VIROLOGY 208, 521–530 (1995)

Antigenic and Sequence Variation in the C-Terminal Unique Domain of the Epstein–Barr Virus Nuclear Antigen EBNA-1

MARK N. WRIGHTHAM,*,[1] JAMES P. STEWART,*,[2] NUSRAT J. JANJUA,* STUART DE V. PEPPER,* CLARE SAMPLE,†‡ CLIONA M. ROONEY,†‡ and JOHN R. ARRAND*

Department of Molecular Biology, Cancer Research Campaign Laboratories, Paterson Institute for Cancer Research, Christie Hospital, Wilmslow Road, Manchester M20 9BX, United Kingdom; †Department of Virology and Molecular Biology, St. Jude Children's Research Hospital, Memphis, Tennessee 38103; and ‡Department of Pathology, University of Tennessee at Memphis, Health Sciences Center, 800 Madison Avenue, Memphis, Tennessee 38163

Received November 2, 1994; accepted February 10, 1995

The Epstein–Barr virus (EBV) nuclear antigen EBNA-1 is essential for viral genome maintenance *in vitro* and may be the only EBV protein expressed by the majority of latently infected cells *in vivo*. EBNA-1 may therefore be critical to the evasion of host immunity which allows persistent infection. EBNA-1 includes a polymorphic internal repeat domain of unknown significance and unique regions which mediate all known functional activities and which have hitherto been assumed to be conserved between strains. Monoclonal antibodies were generated using a construct based on EBNA-1 of the prototype B95-8 strain, deleted for the repeat domain. These antibodies showed a limited profile of recognition of EBNA-1 in common laboratory EBV$^+$ cell lines by immunoprecipitation and immunostaining. The observed antigenic heterogeneity also extended to spontaneously transformed B lymphoblastoid cell lines (LCLs) representing viral isolates circulating within US and UK populations. DNA fragments spanning the C-terminal unique domain of EBNA-1 from eleven spontaneous LCLs were amplified by polymerase chain reaction for sequencing, which directly demonstrated extensive and unexpected variability between diverse type 1 EBV isolates. The resulting polymorphism affects most of the putative MHC Class I binding epitopes which could be identified within this region using published sequence motifs, and influences MHC binding by variants of at least one such peptide in the processing mutant cell line T2. These findings could be related to the apparent lack of recognition of EBNA-1 by cytotoxic T lymphocytes.

INTRODUCTION

Epstein–Barr virus (EBV) shows dual tropism for B cells and oropharyngeal epithelium, establishing reservoirs in both cellular compartments which normally persist for the lifetime of the host. Although asymptomatic individuals intermittently shed virus at the epithelial site (Yao *et al.*, 1985), infection of B cells and EBV-associated tumors is essentially latent. EBV-transformed B lymphoblastoid cell lines (LCL) express a limited number of viral genes encoding a total of nine latent proteins (reviewed by Ring, 1994). These include EBV nuclear antigens (EBNAs) 1, 2, 3a, 3b, and 3c, leader protein (EBNA-LP), and membrane antigens designated latent membrane protein (LMP) and terminal proteins TP 1 and 2. Studies of EBV$^+$ Burkitt's lymphoma (BL) cells *in vitro* have, however, revealed a much more constrained pattern of viral gene expression, designated as the Group I phenotype (Rowe *et al.*, 1987). Such cell lines express only EBNA-1, a profile which has been suggested to parallel the majority of latently infected B cells *in vivo* (Klein, 1989). EBNA-1 has been shown to be the only viral protein required for maintenance of the EBV episome in various cellular contexts (Yates *et al.*, 1984, 1985) and also functions as a transcriptional transactivator (Reisman and Sugden, 1986; Sugden and Warren, 1989), both effects depending on interaction with specific cognate recognition sequences within the EBV origin of plasmid replication, ori P (Wysokenski and Yates, 1989; Hearing *et al.*, 1992). Additional lower affinity binding elements occur in the *Bam*HI-Q region of the viral genome, and these appear to mediate feedback repression of EBNA-1 transcription from the *Bam*HI-F promoter in at least some cells of Group I phenotype (Sample *et al.*, 1992). EBNA-1 is thus a functionally pleiotropic housekeeping protein with a major role in virus persistence.

Most of the known EBV latent proteins elicit cytotoxic T lymphocyte (CTL) responses in the context of particular human major histocompatibility complex (MHC) specificities (Murray *et al.*, 1992; Khanna *et al.*, 1992); these responses may serve to contain *in vivo* dissemination of

Sequence data from this article have been deposited with the EMBL/GenBank Data Libraries under Accession Nos. EG: U21195, PA: U21202, LA: U21198, AM: U21193, WW: U21205, PM: U21203, NL: U21200, JF-1: U21196, SB: U21204, JF-2: U21197, MT: U21199, P3HR-1: U21201 and AG876: U21194.

[1] To whom reprint requests should be addressed.

[2] Present address: Department of Veterinary Pathology, Royal (Dick) School of Veterinary Studies, University of Edinburgh, Summerhall, Edinburgh EH9 1QH, UK.

0042-6822/95 $6.00

overtly transformed cells displaying an LCL-like (Group III) phenotype. By contrast, EBNA-1 invokes a pronounced and ongoing serological response (consistent with sustained expression) in most infected individuals (Rowe *et al.*, 1988), but no study has so far detected specific MHC Class I-restricted CTL responses to this antigen. This apparent lack of recognition has been suggested to facilitate persistence and may therefore be critical to the natural history of EBV infection. Previous studies have, however, addressed this issue only in the context of the prototype B95-8 EBV isolate, which has been assumed to be representative of field isolates with regard to EBNA-1 epitope presentation.

The primary structure of EBNA-1 (B95-8 isolate) features a 238-residue internal repeat sequence (IR3) of unknown function, flanked by N- and C-terminal unique regions. The minimal sequence requirements for known activities of EBNA-1 have been extensively mapped by deletion mutagenesis. Amino acid residues 467–583 in the C-terminal unique domain constitute a potential basic helix–loop–helix structure similar to those of other DNA-binding proteins (Inoue *et al.*, 1991). This domain is closely associated with regions defined as essential for dimerisation, transactivation, and ori P binding by functional assays (Polvino-Bodnar *et al.*, 1988; Ambinder *et al.*, 1991; Chen *et al.*, 1993) and is partially protease-resistant, suggesting a degree of conformational autonomy (Shah *et al.*, 1992). EBNA-1 shows considerable size polymorphism between cell lines, ranging from 69 to 94 kDa, but proteolytic cleavage studies map this variability to the nonessential glycine–alanine repeat region (Allday and MacGillivray, 1985). Sequence variation has not previously been reported within the functionally important unique domains of EBNA-1.

In the present study we use monoclonal antibodies (mAb) generated against an EBNA-1 derivative lacking the IR3 sequence to demonstrate antigenic heterogeneity between diverse virus strains. A remarkably high frequency of sequence variation is shown to occur within the C-terminal unique region of EBNA-1 among Western EBV isolates. Such polymorphism could be critical to the apparent evasion of host CTL responses by EBNA-1 and, hence, to the establishment of lifelong persistent infection. The variant forms of EBNA-1 could also differ at the functional level.

MATERIALS AND METHODS

EBV$^+$ clinical samples and cell lines

Spontaneous EBV-transformed lymphoblastoid cell lines were generated from EBV seropositive donors by culture of peripheral blood mononuclear cells in RPMI 1640 with 10% fetal calf serum and 0.1 μg/ml cyclosporin A (Sandoz) or Ultraculture (Biowhittaker). Transformed cell lines appeared after approximately 6 weeks and were tested for the presence of EBV using monoclonal antibodies specific to EBNA-2 and LMP (Dakopatts). Tumor biopsy specimens from patients with nasopharyngeal carcinoma (NPC) and one patient with post-transplant lymphoproliferative disease (PTLD) were obtained as snap-frozen tissue. Serial sections of 7-μm thickness were dried onto sialane-treated glass microscope slides and fixed in acetone/methanol for 3 min at −70°.

Monoclonal antibody production

EBNA-1 (B95-8 strain) was expressed as a glutathione *S*-transferase fusion protein (E1Δ-GST) in *Escherichia coli*. The construct was deleted for the 7 N-terminal codons of EBNA-1 and the coding sequence for amino acids 40–438, which includes the entire glycine–alanine repeat region (IR3). The resulting fusion protein therefore included all regions to which direct functional importance has so far been ascribed, with the exception of the putative nuclear localization signal (residues 379–386; Ambinder *et al.*, 1991). Recombinant protein was purified on glutathione–Sepharose beads (Pharmacia) and used as immunogen in BALB/c mice. Hybridoma supernatants were tested against E1Δ-GST and GST alone (expressed and purified by an identical method).

Enzyme-linked immunosorbent assays (ELISA)

MAb reactivity with recombinant EBNA-1 fragments was determined by a direct binding ELISA. Test antigen was coated onto microtiter plates (Immulon II, Nunc) by overnight incubation in carbonate–bicarbonate buffer, pH 9.6, at 4°. Plates were then blocked with 1% BSA and washed in TBS/0.1% Tween 20 before addition of test supernatants. Synthetic 15-mer peptides (overlapping by six residues) spanning sections of the B95-8 EBNA-1 sequence were obtained from Alta Bioscience (Birmingham, UK), and reactivity was determined by direct binding or inhibition of mAb binding to directly coated recombinant E1Δ-GST. In all assay systems, secondary detection was by goat anti-mouse Ig-alkaline phosphatase (Sigma), with *p*-nitrophenyl phosphate as substrate. Mouse mAb isotypes were also determined by ELISA using subclass-specific reagents (Sigma).

Immunostaining methods

Cell preparations or cryostat sections for immunostaining were air-dried, fixed in acetone:methanol (1:1) at −20° for 10 min, and preblocked with TBS/0.1% Tween 20 with 5% normal rabbit serum. MAbs were then applied in the presence of 5% normal rabbit serum for 2 hr at 37°. Immunofluorescence staining was carried out using biotinylated F(ab′)$_2$ rabbit anti-mouse Ig secondary and FITC-streptavidin tertiary reagents (Dako). Immunocytochemical staining was by the alkaline phosphatase-antialkaline phosphatase (APAAP) technique, with detection by rabbit anti-mouse Ig and APAAP complexes (Dako); staining was enhanced by repetition of the detecting reagent

stages, and was developed using Fast Red (Sigma). Parallel negative controls were isotype matched mAbs of unrelated specificity.

Immunoprecipitation and immunoblotting

Cell lysates were prepared using 50 m*M* Tris–HCl (pH 8.0), 150 mM NaCl, 1% NP-40, 0.5% deoxycholate, and 0.1% SDS and incubated with mAbs for 3 hr at room temperature. Immune complexes were recovered by incubation with protein A–Sepharose beads precoated with rabbit anti-mouse Igs (Dako), which were then washed five times and boiled for 5 min in the presence of 2% SDS. Supernatants were analyzed by gel electrophoresis in 0.1% SDS using 10% polyacrylamide minigels (Bio-Rad), in parallel with immunoprecipitates from EBV-negative cell lines. Western transfers were carried out in cold 0.25 *M* Tris (pH 8.8), 0.19 *M* glycine, 20% methanol, and 0.1% SDS in a Bio-Rad minigel system at 175 mA for 1 hr. Nitrocellulose membranes were then blocked overnight at 4° with TBS/5% dried milk powder and probed with human serum MO, which has been shown to have a broad profile of recognition encompassing EBNA-1 from a diverse range of EBV strains (Rowe *et al.*, 1987). Membranes were then washed, incubated with alkaline phosphatase-conjugated rabbit anti-human Igs (Dako), and developed using 5-bromo-4-chloro-3-indolyl phosphate/nitro blue tetrazolium (Sigma).

Polymerase chain reaction (PCR) and sequencing

DNA fragments corresponding to the C-terminal unique region of EBNA-1 were amplified from spontaneously transformed LCLs by standard PCR methods using primers 5′-CCC AGG AGT CCC AGT AGT CAG T-3′ (sense; EBV 109090–109111) and 5′-AAC AGC ACG CAT GAT GTC TAC-3′ (antisense; EBV 109969–109949), corresponding to flanking sequences as represented in the B95-8 strain. Double-stranded sequencing of PCR products was performed using internal gene-specific primers, by the dideoxy chain termination method with fluorochrome-conjugated dideoxy nucleotide derivatives, and analyzed using an ABI 373A DNA sequencer in accordance with the supplier's protocol.

Peptide-MHC binding assay using the processing mutant cell line T2

HLA-A2.1 binding activity of peptides was assessed using the antigen processing defective cell line LBL 721.174-T2 (T2), which expresses reduced numbers of unstable empty A2.1 molecules at the cell surface in the absence of exogenous specific peptides (Cerundolo *et al.*, 1990). T2 cells were incubated overnight with synthetic peptides at a final concentration of 200 μg/ml, and stabilization of HLA-A2.1 at the cell surface measured by flow cytometry using the anti-HLA-ABC (monomorphic) mAb W6/32 (Dako). Results were expressed as fluorescence index, determined using the formula (mean sample fluorescence − mean background fluorescence/mean background fluorescence).

RESULTS

EBNA-1-specific monoclonal antibodies detect antigenic heterogeneity

EBNA-1 from the prototype B95-8 EBV strain was expressed as a GST fusion protein deleted for the internal gly–ala repeat region (E1Δ-GST; Fig. 1) and used as immunogen for mAb production. An ELISA screening system identified two hybridomas (3E4 and 4D3), from a total of three successful fusions, showing strong reactivity with the immunizing protein but negligible binding to GST alone; both mAbs are of IgG1 subclass. Immunoprecipitation with these mAbs, followed by Western transfer and probing with a human serum of high anti-EBNA-1 titer yielded a major band of apparent MW 82 kDa from B95-8 cell lysates (Fig. 2). Immunoprecipitation from lysates of EBV$^-$ BL41 cells gave a negative result, and an isotype-matched control mAb of unrelated specificity was likewise ineffective (not shown). Both mAbs showed punctate nuclear staining of acetone–methanol-fixed B95-8 but not EBV$^-$ control cells by biotin–streptavidin immunofluorescence, while isotype-matched control mAbs did not produce detectable staining (not shown).

Analysis of a wider panel of established cell lines by immunofluorescence or immunocytochemistry revealed a heterogeneous pattern of recognition by mAbs 3E4 and 4D3 (Table 1). Among commonly maintained EBV-positive and negative cell lines, staining with either mAb was confined to B95-8 and LCLs transformed using this virus strain. Immunoprecipitation from cell lysates with these mAbs, followed by blotting and probing with a human antiserum known to recognize EBNA-1 from diverse virus strains, exactly mirrored this profile of reactivity (Fig. 2; Table 1). Identical bands were precipitated from B95-8 and the derivative LCL Bristol 8, but not from any other cell line tested. These observations therefore suggested the occurrence of unexpected sequence heterogeneity within the unique regions of EBNA-1 among standard laboratory strains of EBV.

Widespread variation among Western field isolates of EBV

The restricted pattern of recognition of common EBV$^+$ cell lines by these mAbs suggested that the prototype B95-8 EBV strain might be unrepresentative of strains occurring within the general population, perhaps as a result of mutations accumulated during extended passage in culture. In order to assess the natural incidence of this antigenic variant, a panel of 16 spontaneous LCLs of U.S. and UK origin was examined by immunofluorescence staining with these mAbs (Table 2). The cell lines

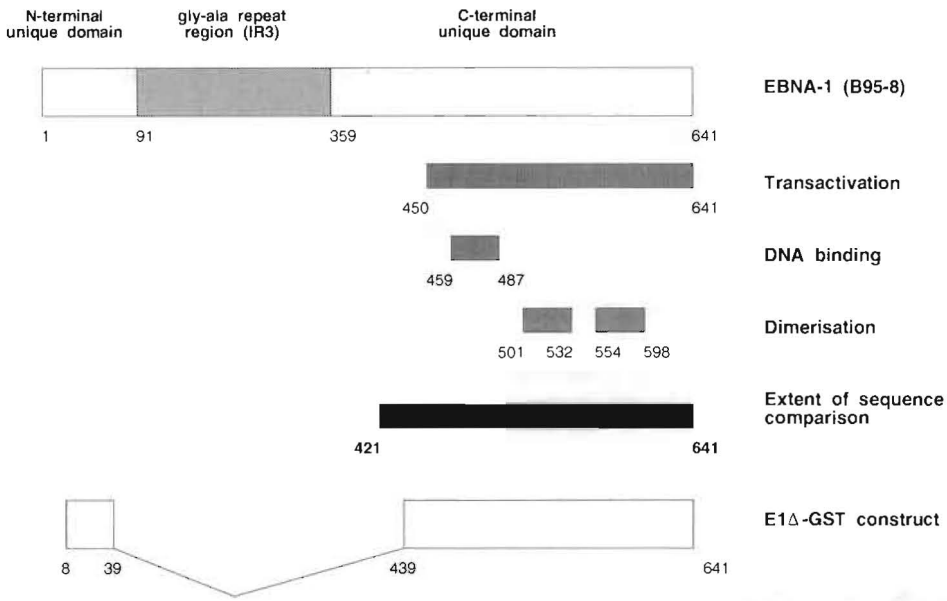

FIG. 1. The domain structure of EBNA-1. Regions of particular importance for specific functions, as defined using deletion mutants (Ambinder *et al*, 1991; Chen *et al*, 1993), are indicated. Regions compared at the sequence level and represented in the E1Δ-GST construct in this study are also shown.

tested were obtained from donors representing a variety of clinical conditions in addition to asymptomatic seropositive individuals. Within this sample, eight lines (50%) showed clear reactivity while the remainder were not recognized. The EBV infection status of all lines examined was confirmed by staining with the LMP-specific mAb pool CS1-4; unrelated control mAbs conversely failed to stain. Cryostat sections of liver biopsy material from a UK patient with post-transplant lymphoma (PTL) were also examined (Table 2). Involved liver tissue from this patient showed focal staining with these mAbs which coincided with nests of positive staining for EBNA-2 and LMP, but not negative control mAbs. Tissue sections from six cases of undifferentiated NPC of Chinese origin were not stained by 3E4 or 4D3, although the presence of EBV in each of the tumors was confirmed by positive staining for LMP (not shown). Thus, the antigenic variation detected within the unique regions of EBNA-1 clearly extends to circulating virus strains, at least within Western populations.

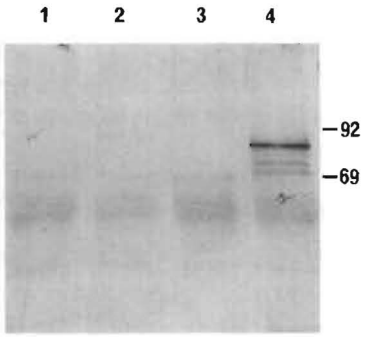

FIG. 2. Immunoprecipitation of EBNA-1 from cell lysates by pooled mAbs 3E4 and 4D3. Western transfers were carried out as described (Materials and Methods) and probed using human serum MO, which predominantly detects EBNA-1, and alkaline phosphatase-conjugated rabbit anti-human Igs. Lane 1, Daudi; 2, Akata; 3, MT; 4, B95-8. No reactivity with EBV⁻ BL41 cells was observed; an isotype-matched control monoclonal antibody was also unreactive (not shown). Molecular weight standards are in kDa.

In order to examine sequence variation directly, DNA fragments spanning the unique regions of EBNA-1 from the UK spontaneous LCL MT, which was not recognized by either mAb (Table 2), were amplified by PCR, cloned, and sequenced. Although the N-terminal unique domain proved to be identical to the prototype B95-8 sequence, multiple amino acid substitutions were identified between residues 471 and 594 of the C-terminal unique domain. DNA fragments corresponding to residues 462–641 of EBNA-1 from a further 10 spontaneous LCLs were therefore amplified by PCR for direct sequence analysis, which revealed a pattern of considerable interstrain diversity such that no two LCLs within the sample were identical in this region (Fig. 3).

Point mutations affecting single amino acids resulted in variation at 17 positions, of which three featured more than two distinct amino acids. The highest level of variability was seen at position 471, which featured a total of four alternative amino acid residues. One isolate (WW) included a tripeptide insertion resulting in an extra iteration of a DDG motif in the C-terminal negatively charged region. Many of the observed point mutations introduced functionally distinct amino acids relative to the B95-8 sequence; a further two codons showed silent nucleotide changes which did not result in altered amino acid residues. Two independent

TABLE 1

Reactivity of 3E4 and 4D3 with Standard EBV-Positive and -Negative Cell Lines

	Immunocytochemistry[a] (Enhanced APAAP)				
	3E4	4D3	LMP	Control	Immpptn[b]
EBV-positive					
B95-8	+	+	++	–	+
Bristol 8[c]	+	+	++	–	+
Kloek[c]	+	+	++	–	NT
Raji	–	–	NT	–	–
Namalwa	–	–	NT	–	NT
Daudi	–	–	–	–	–
Akata	–	–	+/–	–	–
P3HR-1	–	–	NT	–	NT
AW-Ramos	–	–	–	–	NT
AG876	–	–	NT	–	–
EBV-negative					
Ramos	–	–	–	–	–
BL41	–	–	–	–	–

[a] Staining of air-dried acetone/methanol-fixed cell preparations was detected using two cycles of staining with rabbit anti-mouse Igs and mouse APAAP, as detailed in the text. Control was isotype-matched monoclonal antibody of unrelated specificity. NT, not tested. Where tested, similar results were also obtained by biotin/streptavidin immunofluorescence staining.

[b] Immunoprecipitation followed by immunoblotting with human serum MO (Materials and Methods).

[c] LCLs transformed using B95-8 EBV strain.

spontaneous LCLs were available from U.S. donor J.F., of which one was 3E4/4D3$^+$ and the other negative (Table 2), suggesting the presence of at least two distinct antigenically defined EBV variants within this individual; sequence data from these LCLs revealed distinct EBNA-1 species differing at four amino acid positions within the C-terminal unique domain (Fig. 3). The independence of all strains from the prototype B95-8 virus was confirmed by PCR amplification of DNA fragments using primers corresponding to sequences within the deletion characteristic of this isolate as previously described (Lees *et al.*, 1992). In order to determine whether the observed patterns of variation correlated with conventional EBV types, DNA from each cell line was amplified by PCR using primers based on type-specific sequences within EBNA-2A or -2B, which indicated that all strains within the LCL panel were of type 1. Sequencing of PCR products from EBNA-1 of the standard type 2 strains P3HR-1 and AG876, however, revealed a comparable pattern of variation, mostly concentrated on similar residues to those which vary between type 1 strains (Fig. 3).

The determinants recognized by 3E4 and 4D3 could not be localized on the basis of sequence data alone and could not be mapped using synthetic 15-mer peptides or a limited number of expressed fragments representing sequences present in the E1Δ-GST protein (not shown). These observations, together with the failure of these mAbs to recognize partially denatured antigen in immunoblots, suggest that their target epitope or epitopes are highly conformation-dependent.

Comparison of sequence data with published binding motifs for MHC Class I alleles allowed twelve putative MHC-binding peptides to be identified within this region of EBNA-1, of which only three were completely conserved across all of the strains examined (Table 3). Three of the variable putative epitopes showed loss of primary anchor residues predicted to be essential for binding to MHC molecules. The ability of naturally occurring variants of one such peptide (B95-8 amino acid residues 565–574) to bind to the appropriate MHC specificity (HLA-A2.1) was tested *in vitro* using the A2.1$^+$ antigen processing mutant cell line T2 (Fig. 4). The version of this peptide present in the B95-8 strain and a minority (3/11) of spontaneous LCLs clearly bound to A2.1 in this assay, while the homologous sequence as represented in the majority of the LCL panel, which lacks a predicted C-terminal anchor residue, failed to bind.

DISCUSSION

The monoclonal antibodies used in this study were raised against EBNA-1 of B95-8 origin deleted for the

TABLE 2

Reactivity of 3E4/4D3 with Spontaneously Transformed Lymphoblastoid Cell Lines[a]

Donor	Country[b]	Clinical condition[c]	3E4/4D3	LMP	Control
WW	UK	XLP	+	++	–
NL	UK	TR	+	++	–
MT	UK	RA	–	++	–
SB	UK	CA	–	++	–
LA	USA	TR	+	++	–
EG	USA	BL	+	++	–
PM	USA	N	+	++	–
PA	USA	N	+	++	–
AM	USA	N	+	++	–
JF[d]	USA	N	+	++	–
			–	++	–
DG	USA	N	–	++	–
HS	USA	N	–	++	–
JPC	USA	N	–	++	–
PR	USA	N	–	++	–
MP	USA	N	–	++	–
SA[e]	UK	PTLD	+	++	–

[a] Staining of acetone/methanol-fixed cell preparations was detected by immunofluorescence using the biotin/streptavidin system.

[b] UK, United Kingdom; USA, United States of America.

[c] XLP, X-linked lymphoproliferative disorder; TR, transplant recipient; RA, rheumatoid arthritis; CA, non-EBV-related carcinoma; BL, Burkitt's lymphoma; N, normal; PTLD, post-transplant lymphoproliferative disease.

[d] Two independent LCLs established from this donor were examined.

[e] Cryostat sections of a liver biopsy from a patient with post-transplant lymphoproliferative disease were stained by biotin/streptavidin immunofluorescence.

```
B95  421  GGPDGEPDVP PGAIEQGPAD DPGEGPSTGP RGQGDGGRRK KGGWFGKHRG QGGSNPKFEN IAEGLRALLA
EG                                 -                        -  -  --       -  -
PA                                 V                        -  L  -Q       -  T
LA                                 V                        -  -  -Q       -  T
AM                                 -                        -  -  --       -  -
WW                                 -                        -  -  --       -  T
PM                                 -                        -  -  -Q       -  T
NL                                 -                        -  -  -Q       -  T
JF1                                -                        R  E  -Q       -  L

SB                                 -                        -  -  --       D  -
JF2                                -                        -  E  -Q       -  L
MT                                 -                        -  H  -Q       -  L
P3HR1                              -                        R  E  -Q       -  L
AG876                              -                        -  E  SQ       -  L

B95  491  RSHVERTTDE GTWVAGVFVY GGSKTSLYNL RRGTALAIPQ CRLTPLSRLP FGMAPGPGPQ PGPLRESIVC
EG        -       -  -                 -  --                              -
PA        -       ♦  -                 ♦  I-                              ♦
LA        -       ♦  -                 ♦  I-                              ♦
AM        -       ♦  -                 ♦  I-                              ♦
WW        -       ♦  -                 ♦  I-                              ♦
PM        C       ♦  -                 ♦  I-                              ♦
NL        C       ♦  -                 ♦  I-                              ♦
JF1       C       ED N                 ♦  I-                              ♦

SB        -       -  -                 -  --                              -
JF2       C       ED N                 ♦  I-                              ♦
MT        C       ED N                 ♦  I-                              ♦
P3HR1     C       -D N                 ♦  IG                              ♦
AG876     C       ED N                 ♦  IG                              ♦

B95  561  YFMVFLQTHI FAEVLKDAIK DLVMTKPAPT CNIRVTVCSF DDGVDLPPWF PPMVEGAAAE GDDGDDG---
EG          -          -          --       --                                  ---
PA          -          G          --       --                                  ---
LA          I          G          -P       KA                                  ---
AM          I          G          -P       KA                                  ---
WW          I          G          -P       KA                                  DDG
PM          I          G          -P       KA                                  ---
NL          I          G          LL       KA                                  ---
JF1         I          G          LP       --                                  ---

SB          -          -          --       --                                  ---
JF2         -          -          -P       KA                                  ---
MT          I          G          -P       K-                                  ---
P3HR1       I          G          LP       --                                  ---
AG876       I          G          LP       K-                                  ---

B95  628  DEG GDGDEGEEGQ E
EG                -
PA                -
LA                -
AM                -
WW                -
PM                -
NL                -
JF1               -

SB                D
JF2               -
MT                -
P3HR1             -
AG876             -
```

FIG. 3. Comparison of EBNA-1 C-terminal unique region sequences of B95-8, UK/US spontaneous LCLs and standard type 2 isolates P3HR-1 and AG876. Positions at which one or more strains diverged from the B95-8 sequence are indicated; at each such position, – denotes no change relative to B95-8; ♦ denotes a silent base change. All sequences were otherwise identical to B95-8.

IR3 repeat region. While primarily necessary in order to achieve stable expression of the EBNA-1 construct, this modification conferred the additional benefit of removing sequences known to be cross-reactive with host proteins (Birkenfeld *et al.*, 1990; Sabbatini *et al.*, 1993) and in some instances highly immunogenic (Rumpold *et al.*, 1987). The antibodies show a restricted profile of detection of EBNA-1 in EBV$^+$ cell lines by immunoprecipitation and by immunostaining of fixed cells. Interpretation of heterogeneous antibody recognition in both systems is potentially complicated by variation in levels of EBNA-1 expression between cell lines. Published data suggest, however, that levels of this protein are remarkably uniform among diverse EBV$^+$ lymphoid cells, showing a less than 2-fold range of variation among a panel of cell lines in which EBV genome copy number varied by up to 40-fold

TABLE 3

Putative MHC Allele-Specific T Cell Epitopes in EBNA-1[a]

Allele	EBNA-1 amino acid positions (B95-8)	Peptide sequence[b]	Incidence among spontaneous LCLs	
A1	501–510	GTWVAGVFVY	8/11	73%
		GNWVAGVFVY	3/11	27%
A2.1	562–570	FMVFLQTHI	4/11	36%
		FIVFLQTHI	7/11	64%
	565–574	FLQTHIFAEV	3/11	27%
		FLQTHIFAEG	8/11	73%
A3/A11[c]	506–514	GVFVYGGSK	11/11	100%
	567–576	QTHIFAEVLK	3/11	27%
		QTHIFAEGLK	8/11	73%
	571–580	FAEVLKDAIK	3/11	27%
		FAEGLKDAIK	8/11	73%
	577–586	DAIKDLVMTK	3/11	27%
		DAIKDLVLLK	1/11	9%
		DAIKDLVMPK	7/11	64%
A24	477–485	KFENIAEGL	10/11	91%
		KFENIADGL	1/11	9%
	509–517	VYGGSKTSL	11/11	100%
	561–570	YFMVFLQTHI	4/11	36%
		YFIVFLQTHI	7/11	64%
A68	514–522	KTSLYNLRR	11/11	100%
	523–532	GTALAIPQCR	2/11	18%
		GIALAIPQCR	9/11	82%
DR1[d]	515–527	TSLYNLRRGTALA	2/11	18%
		TSLYNLRRGIALA	9/11	82%

[a] Sequences were identified on the basis of published MHC Class I allele-specific motifs defined by primary anchor residues (Guo *et al.*, 1992; Ruppert *et al.*, 1993; Kubo *et al.*, 1994), taking account of secondary anchor interactions where these have been defined.

[b] For each peptide sequence the version occurring in the prototype B95-8 strain is indicated in bold type; primary anchor residues are underlined.

[c] HLA-A3 and A11 motifs are closely related (Kubo *et al.*, 1994).

[d] DR1-restricted epitope detected by EBNA-1-specific CTL (Moss *et al.*, 1994); anchor residues unknown.

(Sternas *et al.*, 1990). The observed pattern of monoclonal antibody recognition of standard EBV$^+$ lines is therefore unlikely to be substantially due to differences in EBNA-1 expression, although the possibility of a contributory effect cannot be formally excluded.

These observations implied the occurrence of interstrain heterogeneity, which was subsequently confirmed directly by sequence analysis, predicting up to 7.5% divergence at the amino acid level between virus isolates over the region examined. Comparison of the observed variation with other regions of the viral genome is complicated by the occurrence of other patterns of interstrain variation arising from insertions and deletions particularly, though not exclusively, within repeat regions, and also because comprehensive sequence data are only available for the B95-8 strain (Baer *et al.*, 1984). Overall homology at the nucleotide level across the *Eco*RI J regions of multiple EBV strains has been reported to be 99%, with the slight variation largely concentrated in noncoding regions (Arrand *et al.*, 1989). A high overall level of conservation is also apparent within most of the open reading frames of the *Bam*HI L region, including the late glycoprotein gp340/220 and dUTPase (97.3 and 98.9% amino acid identity respectively; Lees *et al.*, 1993). Substantial type-specific sequence variation characterizes EBNAs 3a, 3b, 3c, and EBNA-2, of which the latter shows the greatest divergence (Dambaugh *et al.*, 1984; Sample *et al.*, 1990). Both EBNA-2 types also show some sporadic variation in individual amino acid residues, but available partial sequence data suggest a maximum of 5% amino acid dissimilarity between Type 1 strains of extremely diverse geographical origin, and rather less in the case of Type 2 strains (Aitken *et al.*, 1994). The observed variation in EBNA-1, which does not correlate with the conventional virus types, is therefore at least comparable to the highest levels of sporadic amino acid variation known to occur in other EBV proteins.

Extensive EBNA-1 unique region heterogeneity was detected among Type 1 field isolates of EBV from UK and U.S. populations. We were unable to detect the B95-8-associated antigenic marker in six Chinese NPC biopsies; within the limits of small experimental numbers, this result is consistent with a reduced frequency or absence of such variants in the Chinese population. The simultaneous isolation of two EBV strains with differing EBNA-1 unique sequences from donor J.F. is of particular interest, as the occurrence of two or more totally distinct transforming EBV strains in immunocompetent carriers

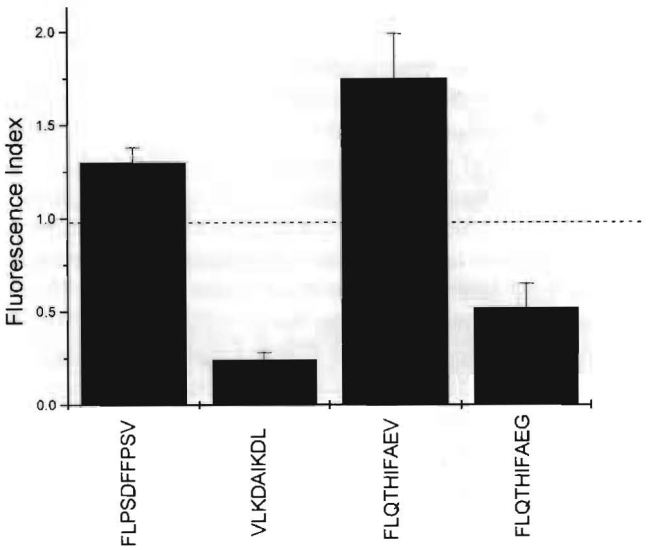

FIG. 4. Differential binding of EBNA-1 variant peptides to HLA-A2.1 using the processing mutant cell line T2. FLPSDFFPSV (HBcAg 18–27) is an authentic A2-binding peptide; VLKDAIKDL is an EBNA-1-derived peptide (574–582; B95-8), predicted to bind poorly using the extended motif of Ruppert *et al.* (1993). Peptide FLQTHIFAEV represents EBNA-1 residues 565–574 of B95-8 and a minority of EBV field isolates; FLQTHIFAEG is the equivalent occurring in most natural strains, which failed to bind in this assay. The threshold fluorescence index of 1 corresponds to 2× background mean linear fluorescence.

is thought to be an extremely rare event (Yao *et al.,* 1991). Coinfection of a minority of such individuals by apparently related strains with EBNA-1 species of differing molecular weights has been reported (Gratama *et al.,* 1994), but these variants were attributed to heterologous recombination within the IR3 repeat sequences during viral replication. Dual infection with independent strains appears to increase in frequency in the context of transplantation or HIV-associated immune dysfunction (Katz *et al.,* 1988; Sculley *et al.,* 1990; Cen *et al.,* 1991; Gratama *et al.,* 1994), but this donor did not show clinical evidence of immunosuppression at the time of sampling. These observations therefore suggest that EBNA-1 unique region polymorphism may define a greater diversity of virus strains within normal donors than might necessarily be detected by analyses based solely on molecular weight.

EBNA-1 is known to vary in the glycine–alanine repeat region, leading to a spectrum of molecular weight variants occurring in different isolates (Allday and MacGillivray, 1985). These findings, together with similar recent observations in tumor biopsy material by another group (Snudden *et al.,* 1995), constitute the first published reports of sequence divergence within the unique domains of EBNA-1. The functional properties of the numerous EBNA-1 variants need to be examined in order to determine the relevance of this diversity to models of EBV persistence and pathogenesis. The observed variation falls largely within the region (ca. aa 450–600) shown to be essential for most of the known functional activities of the molecule (Polvino-Bodnar, 1988; Ambinder *et al.,* 1991; Chen *et al.,* 1993). EBNA-1 is known to be phosphorylated on serine residues (Hearing and Levine, 1985), and some strains show replacement of serine by cysteine at position 492, removing one such site. The functional consequences of this and other nonconservative amino acid changes are unknown at present.

These findings have important implications with regard to host immunological control of differing virus strains at the population level. CTL directed against one or more of the EBV latent proteins have been detected in the context of various common MHC class I specificities, but no such recognition of EBNA-1 has been observed (Murray *et al.,* 1992; Khanna *et al.,* 1992). The apparent lack of response was suggested to facilitate virus persistence in the host, as EBNA-1 expression defines a basal state of virus latency *in vitro* and may therefore be the only viral protein essential for long term maintenance of the EBV genome *in vivo.* Most existing studies of EBNA-1 function have, however, used derivatives of the protein as represented in the B95-8 cell line, this isolate being implicitly assumed to be representative of strains circulating in human populations. By contrast, the panel of spontaneous LCLs examined in the present study included only one virus strain completely identical to B95-8 within the region sequenced. The high level of variability in the C-terminal unique domain results in strain-specific variation in most of the putative Class I-restricted epitopes which can be identified on the basis of published sequence motifs (Table 3). Observed changes included the introduction or loss of primary anchor residues, which in at least one instance clearly determines the capacity of the corresponding peptide to bind to MHC Class I molecules. Conservative and nonconservative amino acid substitutions in positions of potential importance for T cell recognition were also apparent. These findings therefore suggest that experimental reactivation of memory CTL responses may be critically dependent on the host/strain combination used for restimulation. This conclusion thus reopens the possibility that CTL recognition of EBNA-1 may be a major determinant of the virus–host relationship during EBV latency, with MHC haplotype-specific responses to particular epitopes of the protein determining which virus variants can establish persistent infection in a given individual.

Immunologically driven interstrain variation in a dominant HLA A11-restricted CTL epitope has already been documented in EBNA-3b in populations of high A11 prevalence (de Campos-Lima *et al.,* 1993, 1994), suggesting that distinct EBV strains can readily arise in the face of appropriate selection pressure. The relatively pronounced and sustained serological response to EBNA-1 in infected individuals (Rowe *et al.,* 1988) is consistent with continuous expression of this antigen *in vivo.* Recognition of this target by classical MHC Class I-restricted CTLs would therefore be expected to exert significant immunological selection pressure on EBNA-1 variant strains, either at the time of or subsequent to primary infection. CTL recognition of EBNA-1 in the context of MHC Class II has recently been demonstrated (Moss *et al.,* 1994), further suggesting that this antigen is indeed susceptible to immune surveillance. The epitope identified by these workers is restricted through the HLA-DR1 allele and also shows sequence variation within the panel of LCLs in the present study (Table 3), suggesting that Class II-restricted CTL recognition of EBNA-1 could be subject to similar constraints. The incidence of Class I-restricted EBNA-1-specific CTL in diverse strain/haplotype combinations is currently under investigation in our laboratory.

ACKNOWLEDGMENTS

This work was supported by the Cancer Research Campaign, UK, by U.S. Public Health Research Grant CA-56645, Cancer Center CORE Grant CA-21765, and American Lebanese Syrian Associated Charities (ALSAC).

REFERENCES

Aitken, C., Sengupta, S. K., Aedes, C., Moss, D. J. and Sculley, T. B. (1994). Heterogeneity within the Epstein-Barr virus nuclear antigen 2 gene in different strains of Epstein–Barr virus. *J. Gen. Virol.* **75,** 95–100.

Allday, M. J., and MacGillivray, A. J. (1985). Epstein–Barr virus nuclear

antigen (EBNA): Size polymorphism of EBNA-1. *J. Gen. Virol.* **66,** 1595–1600.

Ambinder, R. F., Mullen, M., Chang, Y-N., Hayward, G. S., and Hayward, S. D. (1991). Functional domains of Epstein–Barr virus nuclear antigen EBNA-1. *J. Virol.* **65,** 1466–1478.

Arrand, J. R., Young, L. S., and Tugwood, J. D. (1989). Two families of sequences in the small RNA-encoding region of Epstein–Barr virus (EBV) correlate with EBV types A and B. *J. Virol.* **63,** 983–986.

Baer, R., Bankier, A. T., Biggin, M. D., Deininger, P. L., Farrell, P. J., Gibson, T. J., Hatfull, G., Hudson, G. S., Satchwell, S. C., Seguin, C., Tufnell, P. S., and Barrell, B. G. (1984). DNA sequence and expression of the B95-8 Epstein–Barr virus genome. *Nature (London)* **310,** 207–211.

Birkenfeld, P., Haratz, N., Klein, G., and Sulizeanu, D. (1990). Cross-reactivity between the EBNA-1 p107 peptide, collagen and keratin: Implications for the pathogenesis of rheumatoid arthritis. *Clin. Immunol. Immunopathol.* **54,** 14–25.

Cen, H., Breinig, M. C., Atchison, R. W., Ho, M., and McKnight, J. L. C. (1991). Epstein–Barr virus transmission via the donor organs in solid organ transplantation: polymerase chain reaction and restriction fragment length polymorphism analysis of IR2, IR3 and IR4. *J. Virol.* **65,** 976–980.

Cerundolo, V., Alexander, J., Anderson, K., Lamb, C., Cresswell, P., McMichael, A., Gotch, F. and Townsend, A. (1990). Presentation of viral antigen controlled by a gene in the major histocompatibility complex. *Nature (London)* **345,** 449–452.

Chen, M-R., Middledorp, J. M., and Hayward, S. D. (1993). Separation of the complex DNA-binding domain of EBNA-1 into DNA-recognition and dimerisation subdomains of novel structure. *J. Virol.* **67,** 4875–4885.

Dambaugh, T., Hennessy, K., Chamnankit, L., and Kieff, E. (1984). U2 region of Epstein-Barr virus DNA may encode Epstein–Barr virus nuclear antigen 2. *Proc. Natl. Acad. Sci. USA* **81,** 7632–7636.

de Campos-Lima, P-O., Gavioli, R., Zhang, Q-J., Wallace, L. E., Dolcetti, R., Rowe, M., Rickinson, A. B., and Masucci, M. G. (1993). HLA-A11 epitope loss isolates of Epstein–Barr virus from a highly $A11^+$ population. *Science* **260,** 98–100.

de Campos-Lima, P-O., Levitsky, V., Brooks, J., Lee, S. P., Hu, L. F., Rickinson, A. B., and Masucci, M. G. (1994). T cell responses and virus evolution: Loss of HLA A11-restricted CTL epitopes in Epstein–Barr virus isolates from highly A11-positive populations by selective mutation of anchor residues. *J. Exp. Med.* **179,** 1297–1305.

Gratama, J. W., Oosterveer, M. A. P., Weimar, W., Sintnicolaas, K., Sizoo, W., Bolhuis, R. L. H., and Ernberg, I. (1994). Detection of multiple 'Ebnotypes' in individual Epstein–Barr virus carriers following lymphocyte transformation by virus derived from peripheral blood and oropharynx. *J. Gen. Virol.* **75,** 85–94.

Guo, H-C., Jardetsky, T. S., Garrett, T. P. J., Lane, W. S., Strominger, J. L., and Wiley, D. C. (1992). Different length peptides bind to HLA-Aw68 similarly at their ends but bulge out in the middle. *Nature (London)* **360,** 364–366.

Hearing, J. C., and Levine, A. J. (1985). The Epstein–Barr virus nuclear antigen (*Bam*H1 K antigen) is a single-stranded DNA binding phosphoprotein. *Virology* **145,** 105–116.

Hearing, J., Mulhaupt, Y., and Harper, S. (1992). Interaction of Epstein–Barr virus nuclear antigen 1 with the viral latent origin of replication. *J. Virol.* **66,** 694–705.

Inoue, N., Harada, S., Honma, T., Kitamura, T., and Yanagi, K. (1991). The domain of Epstein–Barr virus nuclear antigen 1 essential for binding to oriP region has a sequence fitted for the hypothetical basic-helix-loop-helix structure. *Virology* **182,** 84–93.

Katz, B. Z., Neiderman, J. C., Olson, B. A., and Miller, G. (1988). Fragment length polymorphisms among independent isolates of Epstein–Barr virus from immunocompromised and normal hosts. *J. Infect. Dis.* **157,** 299–308.

Khanna, R., Burrows, S. R., Kurilla, M. G., Jacob, C. A., Misko, I. S., Sculley, T. B., Kieff, E., and Moss, D. J. (1992). Localisation of Epstein–Barr virus cytotoxic T cell epitopes using recombinant vaccinia: implications for vaccine development. *J. Exp. Med.* **176,** 169–176.

Klein, G. (1989). Viral latency and transformation: The strategy of Epstein–Barr virus. *Cell* **58,** 5–8.

Kubo, R. T., Sette, A., Grey, H. M., Appella, E., Sakaguchi, K., Zhu, N-Z., Arnott, D., Sherman, N., Shabanowitz, J., Michel, H., Bodnar, W. M., Davis, T., and Hunt, D. F. (1994). Definition of specific peptide motifs for four major HLA-A alleles. *J. Immunol.* **152,** 3913–3924.

Lees, J. F., Goodeve, A. C., Arrand, J. E., Ghosh, A. K., Jones, P. H., and Arrand, J. R. (1992). Detection of EBV DNA in post-nasal space biopsy tissue from asymptomatic EBV-seropositive individuals. *J. Med. Virol.* **37,** 30–38.

Lees, J. F., Arrand, J. E., Pepper, S., Stewart, J. P., Mackett, M., and Arrand, J. R. (1993). The Epstein–Barr virus candidate vaccine antigen gp340/220 is highly conserved between virus types A and B. *Virology* **195,** 578–586.

Moss, D. J., Burrows, S. R., Suhrbier, A., and Khanna, R. (1994). Potential antigenic targets on Epstein–Barr virus-associated tumours and the host response. *In* "Vaccines against Virally Induced Cancers" *CIBA Foundation Symposium 187,* pp. 4–20. Wiley, New York.

Murray, R. J., Kurilla, M. G., Brooks, J. M., Thomas, W. A., Rowe, M., Kieff, E., and Rickinson, A. B. (1992). Identification of target antigens for the human cytotoxic T cell response to Epstein–Barr virus (EBV): Implications for the immune control of EBV-positive malignancies. *J. Exp. Med.* **176,** 157–168.

Polvino-Bodnar, M., Kiso, J., and Schaffer, P. A. (1988). Mutational analysis of Epstein–Barr virus nuclear antigen 1 (EBNA-1) *Nucleic Acids Res.* **16,** 3415–3435.

Reisman, D., and Sugden, B. (1986). *trans*-activation of an Epstein–Barr viral (EBV) transcriptional enhancer by the EBV nuclear antigen-1. *Mol. Cell. Biol.* **6,** 3838–3846.

Ring, C. J. A. (1994). The B cell immortalising functions of Epstein–Barr virus. *J. Gen. Virol.* **75,** 1–13.

Rowe, M., Rowe, D. T., Gregory, C. D., Young, L. S., Farrell, P. J., Rupani, H., and Rickinson, A. B. (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.

Rowe, M., Finke, J., Szigeti, R., and Klein, G. (1988). Characterisation of the serological response in man to the latent membrane protein and the six nuclear antigens encoded by Epstein–Barr virus. *J. Gen. Virol.* **69,** 1217–1228.

Rumpold, H., Rhodes, H., Bloch, P. L., Cardon, D. A., and Vaughan, J. H. (1987). The glycine-alanine repeating region is the major epitope of the Epstein–Barr nuclear antigen-1 (EBNA-1). *J. Immunol.* **138,** 593–599.

Ruppert, J., Sidney, J., Celis, E., Kubo, R. T., Grey, H. M., and Sette, A. (1993). Prominent role of secondary anchor residues in peptide binding to HLA-A2.1 molecules. *Cell* **74,** 929–937.

Sabbatini, A., Bombardieri, S., and Migliorini, P. (1993). Autoantibodies from patients with systemic lupus erythematosus bind a shared sequence of SmD and Epstein–Barr virus-encoded nuclear antigen EBNA-1. *Eur. J. Immunol.* **23,** 1146–1152.

Sample, J., Young, L., Martin, B., Chatman, T., Kieff, E., Rickinson, A., and Kieff, E. (1990). Epstein–Barr virus types 1 and 2 differ in their EBNA-3a, EBNA-3b and EBNA-3c genes. *J. Virol.* **64,** 4084–4092.

Sample, J., Henson, E. B. D., and Sample, C. (1992). The Epstein–Barr virus nuclear protein 1 promoter active in Type 1 latency is autoregulated. *J. Virol.* **66,** 4654–4661.

Sculley, T. B., Apolloni, A., Hurren, L., Moss, D. J., and Cooper, D. A. (1990). Coinfection with A- and B-type Epstein–Barr virus in human immunodeficiency virus-positive subjects. *J. Inf. Dis.* **162,** 643–648.

Shah, W. A., Ambinder, R. F., Hayward, G. S., and Hayward, S. D. (1992). Binding of EBNA-1 to DNA creates a protease-resistant domain that encompasses the DNA recognition and dimerization functions. *J. Virol.* **66,** 3355–3362.

Snudden, D. K., Smith, P. R., Jones, M. D., Hara, T., Lai, D., Ng, M-H.,

and Griffin, B. E. (1995). Variations in the structure of the EBV nuclear antigen EBNA-1 in tumours. *Oncogene*, in press.

Sternas, L., Middleton, T., and Sugden, B. (1990). The average number of molecules of Epstein–Barr nuclear antigen 1 per cell does not correlate with the average number of Epstein–Barr virus (EBV) DNA molecules per cell among different clones of EBV immortalized cells. *J. Virol.* **64,** 2407–2410.

Sugden, B., and Warren, N. (1989). A promoter of Epstein–Barr virus that can function during latent infection can be transactivated by EBNA-1, a viral protein required for viral DNA replication during latent infection. *J. Virol.* **63,** 2644–2649.

Wysokenski, D. A., and Yates, J. L. (1989). Multiple EBNA-1 binding sites are required to form an EBNA-1-dependent enhancer and to activate a minimal replicative origin within oriP of Epstein–Barr virus. *J. Virol.* **66,** 2657–2666.

Yao, Q. Y., Rickinson, A. B., and Epstein, M. A. (1985). A re-examination of the Epstein–Barr virus carrier state in healthy seropositive individuals. *Int. J. Cancer.* **35,** 35–42.

Yao, Q. Y., Rowe, M., Martin, B., Young, L. S., and Rickinson, A. B. (1991). The Epstein–Barr virus carrier state: Dominance of a single growth-transforming isolate in the blood and in the oropharynx of healthy virus carriers. *J. Gen. Virol.* **72,** 1579–1590.

Yates, J. L., Warren, N., Reisman, D., and Sugden, B. (1984). A *cis*-acting element from the Epstein–Barr viral genome that permits stable replication of recombinant plasmids in latently infected cells. *Proc. Natl. Acad. Sci. USA* **81,** 3806–3810.

Yates, J. L., Warren, N., and Sugden, B. (1985). Stable replication of plasmids derived from Epstein–Barr virus in a variety of mammalian cells. *Nature (London)* **313,** 812–815.