Widespread Sequence Variation in Epstein-Barr Virus Nuclear Antigen 1 Influences the Antiviral T Cell Response

Melissa J. Bell, Rebekah Brennan, John J. Miles, Denis J. Moss, Jacqueline M. Burrows, and Scott R. Burrows

Epstein-Barr virus (EBV) nuclear antigen (EBNA) 1 is perhaps the most widely studied EBV protein, because of its critical role in maintaining the EBV episome and its expression in all EBV-associated malignancies. Much of this research has focused exclusively on the EBV wild-type (wt) strain (B95-8). Sequence analysis of the gene encoding for EBNA1 in EBV isolates from 43 Caucasians has now revealed considerable EBNA1 sequence divergence from the EBV wt strain in the majority of isolates from this population group. Importantly, T cell recognition of an endogenously processed HLA-B8–binding EBNA1 epitope was greatly influenced by this sequence polymorphism.

The oncogenic herpesvirus Epstein-Barr virus (EBV) is present at high frequency in all human populations and is associated with several malignancies, including Hodgkin lymphoma, Burkitt lymphoma, and nasopharyngeal carcinoma. There are 2 major types of EBV, designated types 1 and 2, that are distinguished by polymorphisms in the latent-cycle genes encoding the EBV nuclear antigens (EBNAs) 2, 3A, 3B, and 3C [1]. Another latent antigen, EBNA1, is essential for maintenance of the EBV episome and its expression in all EBV-associated tumors. Although EBNA1 does not display any obvious type-specific genetic variation, sequence analysis of limited portions of the gene encoding EBNA1 in multiple viral strains has shown considerable polymorphism, particularly across isolates from different ethnic groups [7–10]; however, little is known about the impact of these sequence differences on protein function and/or T cell recognition. Although studies have suggested that specific EBNA1 sequence differences are linked to the EBV-associated malignancies Burkitt lymphoma and nasopharyngeal carcinoma, this remains a controversial issue because in some of these studies it was not possible to distinguish apparent “disease-associated” polymorphisms from EBV strain variation due to the geographic origin of the virus [7].

Methods. Lymphoblastoid cell lines (LCLs) were established from a panel of 26 healthy EBV-seropositive donors and 17 patients with infectious mononucleosis (Australian Caucasians) by spontaneous outgrowth from peripheral blood mononuclear cells (PBMCs). Healthy donor PBMCs were cultured in the presence of cyclosporin A (0.1 μg/mL) [7]. These studies have been reviewed and approved by an appropriate institutional review committee, and blood samples were donated with informed consent.

EBNA1 gene sequencing. DNA was extracted from the cell lines by means of a Qiagen DNA extraction kit. The complete EBNA1 gene was amplified by polymerase chain reaction (PCR) using the 5′ primer 5′-GTCTGCACCTCCCTGTATTCAGAA (B95-8 coordinates 107881–107900) and the 3′ primer 5′-CACAGCA-CGCATGATGCTCT-3′ (B95-8 coordinates 109970–109951). The resulting PCR product was purified using a MinElute spin column (Qiagen) and was sequenced in both directions with additional internal primers by means of a Prism ready reaction Dyedeoxy terminator cycle sequencing kit (Applied Biosystems).

To examine the impact of EBNA1 sequence polymorphism on T cell recognition, cytotoxic T lymphocyte (CTL) bulk cultures or clones, raised by synthetic peptide stimulation (1 μg/mL; Motope), were tested in triplicate for cytotoxicity by the standard 5-h chromium-release assay. Target cells were autologous 51Cr-labeled phytohemagglutinin blasts that either had been pre-
treated with various concentrations of synthetic peptide (YNLRRGIAL or YNLRRGTLAL) or left untreated.

Peptide–major histocompatibility complex (MHC) multimer staining was also used to examine T cell specificity. PBMCs or T cell cultures were incubated for 30 min at 4°C with a YNLRRGIAL-HLA-B*0801 or a YNLRRGTLAL-HLA-B*0801 aliphycocyanin-labeled multimer (ProImmune). Cells were then washed and labeled for 30 min at 4°C with peridinin chlorophyll protein–labeled anti–human CD8 monoclonal antibody (Becton Dickinson). Cells were again washed and analyzed on a FACSCanto flow cytometer using FACSDiva software (version 4.1.2; BD Biosciences).

Interferon (IFN)–γ–enzyme-linked immunospot (ELISPOT) assays were performed using cytokine capture and detection reagents in accordance with the manufacturer’s instructions (Mabtech). In brief, anti–IFN-γ antibodies were coated on the wells of a 96-well nitrocellulose plate, and triplicate wells were seeded with 200,000 PBMCs/well and peptide at 10, 1, or 0.1 μg/mL. After incubation for 16 h, captured IFN-γ was detected with a biotinylated anti–IFN-γ antibody followed by development with streptavidin–horseradish peroxidase complex and chromogenic substrate; spots were counted using an automated plate counter (AID).

**Results.** To assess sequence polymorphism in a large panel of EBV strains carried by individuals of western European origin, LCLs were established without exogenous EBV addition from 43 Australian Caucasian blood donors. The previously described QIMR-Wil EBV strain (also from an Australian Caucasian) was also included in the analysis [11]. The resident EBV strain from each cell line was first tested for EBV type by means of previously published methods [12], and all were shown to be type 1 EBV (data not shown). The gene encoding for EBNA1 was then sequenced across codons 1–89 and 340–641, thereby covering most of the EBNA1 protein with the exclusion of the large glycine-alanine repeat domain, which, for technical reasons, is difficult to accurately sequence. A total of 22 distinct DNA sequences were identified, which encoded for 19 different EBNA1 amino acid sequences (table 1). The EBV wt (B95-8) EBNA1 amino acid sequence was shared by <7% (3/43) of EBV isolates from the study subjects. The most commonly observed EBNA1 protein sequence was carried by >23% of donors. The amino acid sequence of this antigen (referred to as EBNA1/Cauc) diverges from EBV wt EBNA1 at 15 residues, and there was also an 3-aa insertion (Asp-Asp-Gly) toward the C-terminus of the protein. All 15 of these amino acid changes were observed in more than half (23/44 [52%]) of the isolates.

The threonine to isoleucine substitution at position 524, which was observed in 77% of donors, lies within a region that includes an endogenously processed HLA-B8–binding CTL epitope (YNLRRGTLAL) [13]. To examine the impact of this substitution on T cell recognition, PBMCs from a HLA-B8–positive and EBV-seropositive healthy donor were costained with anti–CD8 and either an HLA-B8–YNLRRGIAL or an HLA-B8–YNLRRGTAL multimer. Interestingly, significant staining was observed only with the former multimer, demonstrating a predominance of T cells specific for the YNLRRGIAL peptide but not for the YNLRRGTAL peptide (figure 1A). Interferon-γ ELISPOT assays with PBMCs from this donor confirmed this conclusion, with only the YNLRRGIAL peptide, and not the EBV wt equivalent, stimulating the release of this cytokine (figure 1B). Furthermore, peptide dose-response cytotoxicity assays revealed that 2 CTL clones from this donor, raised by in vitro stimulation with the YNLRRGIAL peptide, could recognize >100-fold lower concentrations of the YNLRRGIAL peptide than the YNLRRGTAL peptide (data not shown). It is also notable that no CTL clones could be raised from this donor by means of in vitro stimulation with the YNLRRGTAL peptide, suggesting a dominance of YNLRRGIAL–specific CTLs within the memory T cell repertoire of this donor.

PBMCs from a further 6 HLA-B8–positive and EBV-seropositive healthy donors were screened for recognition of the 2 MHC-peptide multimers, but the frequency of staining T cells

<table>
<thead>
<tr>
<th>Amino acid position</th>
<th>EBV wt sequence</th>
<th>Residue difference</th>
<th>Frequency (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>16</td>
<td>E (GAG)</td>
<td>Q (CAQ)</td>
<td>27/44 (61.4)</td>
</tr>
<tr>
<td>29</td>
<td>G (GGA)</td>
<td>E (GAQ)</td>
<td>27/44 (61.4)</td>
</tr>
<tr>
<td>70</td>
<td>V (GTC)</td>
<td>A (GCC)</td>
<td>13/44 (29.5)</td>
</tr>
<tr>
<td>85</td>
<td>T (ACC)</td>
<td>A (GCT)</td>
<td>27/44 (61.4)</td>
</tr>
<tr>
<td>86</td>
<td>V (GTC)</td>
<td>M (ATG)</td>
<td>36/44 (81.8)</td>
</tr>
<tr>
<td>471</td>
<td>Q (CAA)</td>
<td>H (CAT)</td>
<td>3/44 (6.8)</td>
</tr>
<tr>
<td>476</td>
<td>P (CCG)</td>
<td>Q (CAQ)</td>
<td>34/44 (77.3)</td>
</tr>
<tr>
<td>487</td>
<td>A (GCT)</td>
<td>T (ACT)</td>
<td>30/44 (68.2)</td>
</tr>
<tr>
<td>252</td>
<td>S (AGT)</td>
<td>C (TGT)</td>
<td>33/44 (75.0)</td>
</tr>
<tr>
<td>563</td>
<td>M (ATG)</td>
<td>I (ATT)</td>
<td>34/44 (77.3)</td>
</tr>
<tr>
<td>574</td>
<td>V (GTT)</td>
<td>G (GGT)</td>
<td>34/44 (77.3)</td>
</tr>
<tr>
<td>585</td>
<td>T (ACA)</td>
<td>P (CCA)</td>
<td>32/44 (72.7)</td>
</tr>
<tr>
<td>594</td>
<td>R (AGG)</td>
<td>K (AGK)</td>
<td>34/44 (77.2)</td>
</tr>
<tr>
<td>596</td>
<td>V (GTC)</td>
<td>A (GCA)</td>
<td>31/44 (70.5)</td>
</tr>
</tbody>
</table>

a From the EBV wt (B95-8) sequence.

b Amino acid residue with the nucleotide codon shown in parentheses.

c Residue changes occurring in only 1 or 2 donor strains were excluded.

d Residue changes found within the most common EBV isolate, designated EBNA1/Cauc.

---

**Table 1. Nonsynonymous Epstein-Barr virus (EBV) nuclear antigen (EBNA) 1 sequence differences at amino acid positions 1–89 and 340–641, compared with EBV wild type (wt).**
was too low to confidently draw conclusions on T cell specificity. Therefore, the PBMCs were stimulated in vitro with either YNLRGIAL (figure 1C and 1E) or YNLRRGTAL (figure 1D and 1F) peptide and were expanded with interleukin-2 for 10 days. The bulk CTL cultures generated from this procedure were then tested for recognition of these peptides by cytotoxicity assays (figure 1C and 1D) and MHC-peptide multimer staining (figure 1E and 1F). Overall, T cells from 6 of 7 donors showed specificity for the EBNA1/Cauc peptide YNLRRGIAL, whereas only 1 donor (donor 4) showed specificity for the EBV wt peptide. Although the resident EBV strains carried by these 7 donors were not characterized, it seems likely that these T cell specificity data reflect the frequency with which the threonine to isoleucine substitution at position 524 occurs in EBV isolates carried by Caucasians.
Discussed. This report represents the most extensive EBNA1 sequence analysis conducted to date on viral strains from Caucasians and demonstrates considerable genetic diversity, with the dominant strains not represented well by EBV wt. Frequently observed sequence polymorphisms included 5 substitutions (at residues 16, 18, 24, 27, and 85) within the N-terminal domain that were observed in ~60% of isolates. The common change at residue 85, which is previously undescribed, lies within a region of EBNA1 essential for driving transcription of EBV’s transforming genes after infection of primary B lymphocytes [14]. A previous sequence analysis of a smaller N-terminal EBNA1 region (codons 1–60) also identified 4 of these coding changes in EBV isolates from East Africans, New Guineans, and Caucasians [7]. Notably, the frequency of type 1 isolates from the Caucasian group with these signature changes within the EBNA1 N-terminal domain (14/23 [60%]) was almost identical to that described in the present study [7].

Sequencing data for residues 340–641 revealed 10 amino acid differences, relative to EBV wt, that were present at very high frequency. More than 68% of isolates carried changes at residues 429, 476, 487, 492, 524, 563, 574, 585, 594, and 595, which lie within the EBNA1 functional domain involved in DNA binding and dimerization [15]. A 3-aa insertion (Asp-Asp-Gly) after residue 621 within the negatively charged C-terminus was also observed in 34% of isolates.

These 10 changes have been described previously in East African patients with Burkitt lymphoma in combination with some, but not all, of the N-terminal changes described above in EBNA1/Cauc [7]. It is notable that EBNA1 sequences in type 1 EBV isolates from East Africa are more closely related to EBNA1/Cauc than the EBV wt sequence and that the EBV wt EBNA1 sequence has not been found in East Africa [7]. This suggests that EBV wt EBNA1 has evolved more rapidly than EBNA1/Cauc from ancestral EBV strains, which are presumed to have originated in Africa. Supporting this proposal, EBNA1 sequence analysis of 2 type 2 EBV strains [9] has revealed many of the residue differences observed in EBNA1/Cauc. Consistent with our results, a subset of these 10 common C-terminal EBNA1 changes has been described previously in Caucasians [7–9]. Some of these previous studies have used sequence polymorphism at residue 487 to classify EBNA1 variants because several different residues have been found at this position in EBV isolates across different ethnic groups. Our Caucasian donor panel carried isolates with either alanine, threonine, or leucine at this position, with the latter variant observed in only 2 EBV isolates.

T cell assays described here provide proof of concept that the extensive genetic diversity within EBNA1 can have a significant impact on immune recognition of this antigen. One of the few regions of EBNA1 that is efficiently processed and presented by class I MHC molecules for CD8+ T cell recognition includes a common amino acid polymorphism, and our data suggest that most HLA-B*0801–positive and EBV-exposed individuals recognize the variant sequence but not the previously defined EBV wt epitope [13]. Thus, any future EBV vaccine that delivers the EBV wt EBNA1 sequence is likely to stimulate YNLRRGAL-specific CTLs, which may have a limited capacity in recognizing the YNLRRGIAL sequence encoded by the most common EBV strains carried by Caucasians. Such problems, which could limit vaccine efficacy, could be circumvented by utilizing antigenic sequences from the most common EBV strains infecting the target population when developing a vaccine. Taken together, these results indicate that a comprehensive investigation into the functional and immunological impact of EBNA1 sequence polymorphism is required, particularly in light of its universal expression in EBV-associated malignancies.

References