Peripheral B cells latently infected with Epstein–Barr virus display molecular hallmarks of classical antigen-selected memory B cells

Tatyana A. Souza*, B. David Stollar†, John L. Sullivan‡, Katherine Luzuriaga‡, and David A. Thorley-Lawson*§

*Department of Pathology, Tufts University School of Medicine, 150 Harrison Avenue, Boston, MA 02111; †Department of Biochemistry, Tufts University School of Medicine, 136 Harrison Avenue, Boston, MA 02111; and ‡Department of Pediatrics and Molecular Medicine, University of Massachusetts Medical School, Biotech II, Suite 318, 373 Plantation Street, Worcester, MA 01605

Communicated by John M. Coffin, Tufts University School of Medicine, Boston, MA, October 26, 2005 (received for review October 6, 2005)

Epstein–Barr virus (EBV) establishes a lifelong persistent infection within peripheral blood B cells with the surface phenotype of memory cells. To date there is no proof that these cells have the genotype of true germinal-center-derived memory B cells. It is critical to understand the relative contribution of viral mimicry versus antigen signaling to the production of these cells because EBV encodes proteins that can affect the surface phenotype of infected cells and provide both T cell help and B cell receptor signals in the absence of cognate antigen. To address these questions we have developed a technique to identify single EBV-infected cells in the peripheral blood and examine their expressed Ig genes. The genes were all isotype-switched and somatically mutated. Furthermore, the mutations do not cause stop codons and display the pattern expected for antigen-selected memory cells based on their frequency, type, and location within the Ig gene. We conclude that latently infected peripheral blood B cells display the molecular hallmarks of classical antigen-selected memory B cells. Therefore, EBV does not disrupt the normal processing of latently infected cells into memory, and deviations from normal B cell biology are not tolerated in the infected cells. This article provides definitive evidence that EBV in the peripheral blood persists in true memory B cells.

latent | mononucleosis | immunoglobulin | germinal center

Epstein–Barr virus (EBV) is a human γ-herpesvirus that establishes a lifelong persistent infection in >90% of the world's population (1, 2). Although in most cases primary infection is asymptomatic, when transmitted during adolescence EBV infection can manifest as acute infectious mononucleosis (AIM), a self-limiting lymphoproliferative disease (3). Perhaps the most striking property of EBV *in vitro* is that it can efficiently infect resting B lymphocytes and drive them to become continuously proliferating lymphoblasts (4). Because of its growth-transforming properties, EBV is associated with various neoplasias (1, 2). In contrast to its growth-promoting activity *in vitro*, we have shown that EBV persists *in vivo* in resting B lymphocytes in the peripheral blood (PB) that have the surface phenotype of memory cells ($CD27^+$, IgD^-, $CD5^-$) (5, 6).

To address the paradox between the proliferating infected cells seen *in vitro* and the resting cells found *in vivo*, we have proposed a model (7, 8) in which naïve B cells newly activated through EBV infection are driven to differentiate into long-lived memory cells through surrogate signaling provided by viral latent proteins. We have provided evidence that EBV in peripheral memory cells sustains persistence by down-regulating expression of all of the latent proteins so that the infected cells are neither immunogenic nor pathogenic to the host (9). This finding has led us to suggest that the latently infected cells in the periphery are maintained by homeostasis-driven cell division allowing their numbers to remain relatively constant over time (10).

Classical long-lived memory B cells arise from a germinal center (GC) reaction after being selected by cognate antigen and receiving help from primed T helper (T_H) cells (11, 12). Memory cells exiting the GC exhibit two cardinal features: class switch recombination (CSR) in the Ig heavy chain constant region, which determines the effector function of the secreted antibody (13), and somatic hypermutations (SHM) in the Ig variable (V) regions, which diversify the B cell pool and improve the affinity of the B cell receptor (BCR) for its antigen (14). Thus, GC-derived memory B cells display a well characterized pattern of SHM in their Ig V regions that is the result of competition for antigen in the GC.

A key unresolved issue with our model arises from the knowledge that EBV latent proteins can have profound effects on the surface phenotype of infected cells (4) and specifically that the viral latent membrane proteins LMP1 and LMP2a are expressed by latently infected GC B lymphocytes in healthy tonsils (15). These two proteins alone potentially have sufficient signaling capacity to drive an infected naïve cell into the memory B cell compartment in the absence of cognate antigen (16–20). Such a process should presumably produce a cell that resembles a memory cell phenotypically but lacks the characteristic patterns of SHM found in antigen-selected, GC-derived memory B cells. Currently, it is unresolved whether the infected cells in the PB are only phenotypically memory cells because of the action of viral latent proteins or are genotypically memory cells produced solely through surrogate signaling by the virus, or if they are truly products of an antigen-selected GC process. It is critical to establish how and to what extent EBV disrupts or regulates normal B cell function *in vivo* to fully understand the conundrum of persistent infection and the mechanism of EBV-associated B cell lymphoma. In this study we sought to distinguish these possibilities by isolating single infected $CD27^+$ B cells from PB and analyzing the sequences of their expressed Ig genes for the presence/absence of SHM and CSR and evidence of antigen selection.

Materials and Methods

Primary Cells. Adolescents (ages 17–24 years) presenting to the clinic at the University of Massachusetts Amherst Campus Student Health Services with clinical symptoms consistent with AIM were recruited as described in ref. 9. These studies were

Conflict of interest statement: No conflicts declared.

Freely available online through the PNAS open access option.

Abbreviations: AID, activation-induced cytidine deaminase; AIM, acute infectious mononucleosis; BCR, B cell receptor; CSR, class switch recombination; EBV, Epstein–Barr virus; EBER1, EBV-encoded RNA 1; CDR, complementarity-determining region; FWR, framework region; GC, germinal center; LMP, latent membrane protein; PB, peripheral blood; SHM, somatic hypermutation; T_H, T helper; V, variable; V_L, V light; V_H, V heavy; R, replacement; S, silent.

Data deposition: The sequences reported in this paper have been deposited in the GenBank database (accession nos. DQ205136–DQ205184).

§To whom correspondence should be addressed. E-mail: david.thorley-lawson@tufts.edu.

© 2005 by The National Academy of Sciences of the USA

IMMUNOLOGY

approved by the Human Studies Committees at the University of Massachusetts Medical School and Tufts University Medical School. Thirty AIM patients were tested, whose frequencies of infected cells ranged from 1 in every 2 memory B cells to 1 in >100 memory B cells. For this study we chose six patients: four with high frequencies of infected memory B cells (IM1, 1 in 2; IM2, 1 in 3; IM3, 1 in 5; IM4, 1 in 7) and two with lower frequencies (IM5, 1 in 22; IM6, 1 in 43).

B Cell Separations. PB mononuclear cells (2×10^7 cells per ml) were obtained as described in ref. 9 and stained with the following antibodies: anti-human CD20 phycoerythrin (BD Pharmingen) or anti-human CD19 CyChrome (DAKO) and anti-human CD27 FITC (BD Pharmingen). Single memory ($CD20^+CD27^+$) or naive ($CD19^+CD27^-$) B cells were sorted with a Cytomation MoFlo fluorescence-activated cell sorter into 10 μl of 1× first-strand buffer (Invitrogen) in 96-well plates, immediately frozen on dry ice, and stored at −80°C.

Limiting Dilution Analysis. Limiting dilution analysis was used to determine the frequency of EBV-infected cells for each AIM patient as detailed in ref. 21.

cDNA Synthesis. Single-cell cDNA synthesis was performed according to the protocol of Wang and Stollar (22), with the exception of adding 5 pmol EBV-encoded mRNA 1 (EBER1) RNA-specific primer (AGGACCTACGCTGCCCTAGA) and 5 pmol Ig Cα-specific primer (GAGGCTCAGCGGGAAGAC) to the primer mixture already containing specific primers to the Cμ, Cγ, Cκ, and Cλ Ig constant chain regions. Eight wells containing all buffers from the time of sorting, minus the single cells, served as negative controls.

Single-Cell PCR. EBER1 PCR, as described in ref. 21, was used to detect infected PB cells, because it is expressed in ≥90% of them (T.A.S., unpublished observations). Ig gene RT-PCR was performed according to the protocol of Wang and Stollar (22) by using the high fidelity hot-start *Pfu-Turbo* polymerase (Stratagene). PCR products were visualized by 2% agarose gel electrophoresis. Ig V genes from infected cells were amplified with similar or better efficiency than from uninfected B cells. PCR efficiency was 38% for V heavy (V_H) chains (32 amplified from 85 infected cells) and 33% for V light (V_L) chains (17 amplified from 48 infected cells) compared with 25% for V_H chains from uninfected cells (16 amplified from 64 cells).

Analysis of Amplified Ig V Genes. V gene PCR products were excised from agarose gel, and DNA was extracted by using the QIAquick gel extraction kit (Qiagen). DNA was sent for sequencing to the Tufts University Core Facility with corresponding constant region primers. Sequences were aligned by using the VBase database and the IMGT database [the international ImMunoGeneTics information system at http://imgt.cines.fr (23)]. The probability that an excess or scarcity of R mutations in V_H chain complementarity-determining region (CDR) or framework region (FWR) were due to chance alone was calculated by using the multinomial distribution model (24). A total of 294 bp (FR1 through FR3) from each of the 32 V_H gene sequences and 288 bp (FR1 through beginning of CDR3) from each of 17 V_L gene sequences were analyzed for frequency and characteristics of the mutations. Mutations in the primer binding regions were disregarded. Only base substitutions were counted. SHM hotspots, DGYW (AGT/*G*/CT/AT), and its reverse complement, WRCH (AT/AG/*C*/ACT), were located in all sequences and quantitated. Sequences determined in this study can be found in the GenBank database under accession nos. DQ205136–DQ205184.

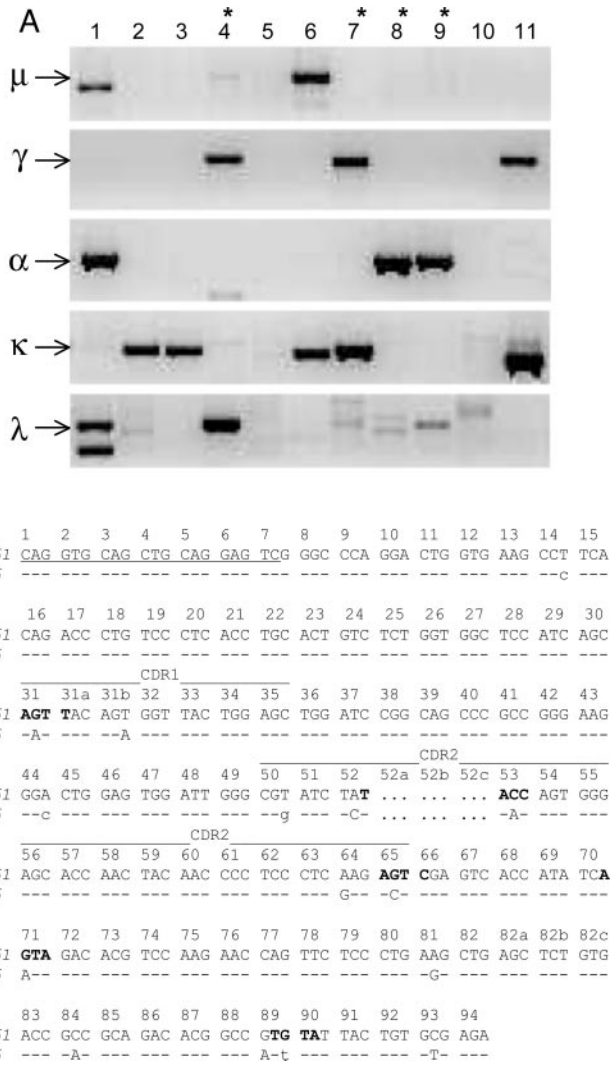

Fig. 1. Amplification and analysis of expressed Ig V regions from single EBV^+ memory B cells. (*A*) Amplification of Ig V regions from 11 single cells. Each cell should result in the 400-bp (arrow) amplification of one heavy chain and one light chain. (The asterisks denote cells that were also EBER1-positive.) (*B*) Sequence alignment of an amplified EBV^+ V_H gene to its germ-line counterpart from AIM patient 4. The codon numbering is according to CHOTHIA. The location of the CDRs is indicated by solid lines above the sequences. R mutations are in uppercase, and S mutations are in lowercase. SHM hotspots are in bold. The location of the forward primer is underlined.

Results

PB B Cells Latently Infected with EBV Are Genotypically Memory Cells. True memory B cells have functional antigen receptors on their surface expressed from Ig genes that have undergone SHM. Therefore, to establish whether or not EBV persists in true memory B cells, we isolated single latently infected cells from the PB and tested for these characteristics. We isolated single cells from six AIM patients by FACS using the pan-B cell marker CD20 and the memory B cell marker CD27 (25). The EBV-infected cells were then identified by using real-time RT-PCR to screen for expression of the abundant EBER1, which is expressed by nearly all of the latently infected cells in the PB (T.A.S., unpublished observations). The Ig genes from the EBER1-positive cells were then amplified (Fig. 1*A*). The resulting 32 V_H regions were sequenced, and the sequences were aligned to their closest germ-line counterparts by using the VBase database and the IMGT database [the international

Table 1. V_H gene sequences of $CD20^+CD27^+$ EBV-infected PB B cells

Cell	Genes				No. of mutations (%)‡
	V_H*	D†	J_H†	C_H	
IM1-1§	VH5-51	NA	JH6	γ	26 (8.8)
IM1-2	VH1-46	D3-9	JH5	α	24 (8.3)
IM1-3	VH1-18	NA	NA	α	17 (5.8)
IM1-4	VH5-51	D3-3	NA	γ	26 (8.8)
IM1-5	VH3-11	D3-10	JH4	γ	13 (4.4)
IM2-1	VH1-2	D6-13	JH5	γ	21 (7.1)
IM2-2	VH3-23	D6-19	JH6	α	24 (8.1)
IM2-3	VH3-7	NA	NA	α	16 (5.4)
IM2-4	VH4-39	D2-2	JH6	α	31 (10.5)
IM2-5	VH4-39	D5-24	JH4	α	9 (3.0)
IM3-1	VH1-69	D1-20	JH3	α	40 (13.6)
IM3-2	VH4-30	D3-22	JH3	γ	11 (3.7)
IM3-3	VH2-5	D3-3	JH4	γ	6 (2.0)
IM3-4	VH3-11	D3-22	JH4	γ	36 (12.2)
IM3-5	VH4-30	D6-13	JH4	γ	13 (4.4)
IM4-1	VH3-74	D5-12	JH4	μ	9 (3.1)
IM4-2	VH2-26	D3-3	JH3	α	24 (8.1)
IM4-3	VH3-49	D2-3	JH2	α	27 (9.1)
IM4-4	VH3-11	D2-8	JH4	α	24 (8.1)
IM4-5	VH4-30	D4-17	JH4	α	14 (4.7)
IM4-6	VH4-61	D3-3	JH4	γ	15 (5.1)
IM5-1	VH1-2	D6-19	JH4	γ	24 (8.2)
IM5-2	VH1-46	D3-3	JH6	α	4 (1.4)
IM5-3	VH3-23	D2-2	JH4	μ	15 (5.1)
IM5-4	VH4-39	D6-19	JH5	α	14 (4.8)
IM5-5	VH4-4	D6-19	JH3	α	5 (1.7)
IM6-1	VH6-1	D6-6	JH6	μ	4 (1.3)
IM6-2	VH1-2	D2-8	JH4	γ	18 (6.1)
IM6-3	VH5-51	D3-16	JH4	γ	13 (4.4)
IM6-4	VH3-48	NA	NA	γ	10 (3.4)
IM6-5	VH1-8	NA	NA	α	15 (5.1)
IM6-6	VH1-69	NA	NA	γ	18 (6.1)

*V_H gene nomenclature by Matsuda and colleagues (55 and 56).
†D and J_H genes, as reported in refs. 57 and 58, respectively. NA, no good match found.
‡% is number of mutations per 100 bp.
§AIM patient 1, clone 1.

ImMunoGeneTics information system at http://imgt.cines.fr (23)], which contain all of the known human polymorphic Ig variants. Fig. 1*B* shows a V_H sequence alignment from AIM patient 4, and Table 1 lists the VDJ gene segment and constant region usage for the 32 sequences. All 32 V_H sequences were mutated and in-frame and did not contain any stop codons, as expected for genotypically normal memory B cells that require a functional Ig molecule to be expressed on the surface. All sequences also had unique VDJ joints (data not shown). One sequence had a 3-bp deletion, which can occur during the SHM process (26, 27). Mutations in the Ig genes from cells latently infected with EBV ranged in number from 4 to 40 per sequence with an overall mutation frequency of 6.0% (Table 2), which is similar to previous reports for human memory B cells (28, 29). As a control, we amplified and analyzed 11 V_H genes of $CD19^+$ $CD27^-$ naïve B cells from AIM patient 4. As expected, EBER1 was never detected in these cells, because EBV does not persist in naïve B cells, and the Ig genes had an average mutation frequency of 0.2% (Table 2). This mutation rate in the naïve B cell population is similar to what was observed previously (30, 31). This validated that the rate of *Pfu* introduced errors during PCR cycling was negligible and that our method for mutation analysis was adequate. From this finding we can conclude that

Table 2. V_H gene somatic mutation summary

Population	Donor	Total no. of sequences	% mutated	Range	Average
$CD27^-$ EBV^-	IM4	11	36	0–2	0.2
$CD27^+$ EBV^+	IM1	5	100	13–26	7.2
	IM2	5	100	9–31	6.8
	IM3	5	100	6–40	7.2
	IM4	6	100	9–27	6.4
	IM5	5	100	4–24	4.2
	IM6	6	100	4–18	4.4
$CD27^+$ EBV^+	IM1-6	32	100	4–40	6.0

the mutations in the V_H genes from the EBV-infected cells were the result of the SHM process and not due to PCR artifacts or analysis errors. These results demonstrate that EBV-infected $CD27^+$ B cells express a functional Ig heavy chain and that their V_H regions carry SHMs at a frequency indistinguishable from normal memory B cells. This finding is definitive evidence that B cells from the PB, latently infected with EBV, are true memory B cells genotypically, and it excludes the possibility that the memory cell surface phenotype of the latently infected cells is an artifact caused by viral gene expression. Furthermore, it suggests that EBV, through LMP2a signaling in lymphoid tissue, does not replace the requirement for an intact antigen receptor for the survival of the latently infected B cells.

EBV^+ $CD27^+$ Memory B Cells Show Characteristics of Positive Antigen Selection. The results described above show that EBV^+ $CD27^+$ B cells bear the hallmark, SHM, of true GC-derived memory cells. The question arises, therefore, whether these cells were rescued from the GC through the signaling activities of the viral latent proteins or whether they were positively selected by cognate antigen. Analysis of somatic mutations can indicate whether a cell has experienced positive selection. During the mutation process, base substitutions can occur in the FWR or in the antigen-binding CDR of the V genes. Base substitutions that lead to amino acid replacements are selected against in the FWRs, to preserve the proper Ig fold, but are favored in the CDRs through the selection process that increases the affinity of the Ig molecule for its antigen. Thus, by comparing the replacement (R) to silent (S) mutation ratios (R/S ratios) in the FWR and CDR it is possible to detect a signature of positive antigen selection. Random mutagenesis is predicted to result in an R/S ratio of ≈3 in the CDRs and ≈1.5 in the FWRs (32). An R/S ratio in the CDRs that is >3 should occur only through antigen selection (32).

We computed the number of R and S mutations in the FWRs and CDRs of the V_H genes from the EBV-infected peripheral B cells described in the previous section. The mean R/S ratios for all donors are listed in Table 3. The mean ratio for all of the EBV^+ cells was 3.4 in the CDRs and 2.1 in the FWRs, remarkably close to published values for productive rearrangements from normal memory B cells, which are reported to be 3.6 in the CDR and 1.9 in the FWRs (33). Therefore, the sequences in the latently infected cells show higher R/S ratios in the CDRs than in the FWRs ($P = 0.017$, one-tailed paired *t* test), and the ratios are similar to those published for antigen-selected memory B cells. We conclude that the R/S ratios we observe in the expressed V_H genes from the EBV-infected cells are consistent with their having experienced antigen selection pressure to preserve a proper protein fold and to increase antigen-binding affinity.

To establish the influence of antigen selection on memory cell formation more definitively, it is necessary to take into account the intrinsic susceptibility of FWRs and CDRs of V_H genes to

Table 3. Analysis for positive antigen selection of single EBV$^+$ CD27$^+$ PB B cells

Patient	Mean		% $P < 0.05^\dagger$		
	R/S_{FWR}*	R/S_{CDR}*	FWR	CDR	FWR + CDR
IM1	2.2	3.2	60	60	100
IM2	1.9	2.8	80	40	50
IM3	1.7	3.7	80	40	50
IM4	2.3	4.3	83	50	60
IM5	1.5	2.7	60	20	33
IM6	2.4	3.6	67	50	75
Total	2.1	3.4‡	73	43	60

*R/S ratios were calculated by dividing the number of R mutations by the number of S mutations in either FWR or CDR regions. For sequences where the denominator was 0, 1 was used to obtain a ratio value.

†*P* values were calculated by using the multinomial distribution model (24). Percent of V_H sequences having significantly fewer R mutations in FWR ($P_{FWR} < 0.05$) and percent of V_H sequences having significantly more R mutations in CDR ($P_{CDR} < 0.05$) are shown. All V_H sequences with $P_{CDR} < 0.05$ also had $P_{FWR} < 0.05$. The last column shows percent of V_H sequences with $P_{FWR} < 0.05$ also having $P_{CDR} < 0.05$.

‡Mean R/S_{CDR} is significantly greater than mean R/S_{FWR} (one-tailed Student's paired *t* test, $P = 0.017$).

Table 4. V_κ and V_λ gene sequences of CD20$^+$CD27$^+$ EBV-infected B cells

Cell	$V_{\kappa/\lambda}$ gene*	$J_{\kappa/\lambda}$ gene†	No. of mutations (%)	% of mutations, V_H gene‡
IM2-1	Vλ2-14	Jλ1	10 (3.5)	7.1
IM2-3	Vλ3-21	Jλ3	9 (3.2)	5.4
IM2-4	Vλ3-10	Jλ3	30 (10.2)	10.5
IM2-5	Vλ3-9	Jλ3	8 (2.7)	3.0
IM2-6	Vκ3-15	NA	13 (4.6)	-
IM2-7	Vκ1-5	NA	6 (2.2)	-
IM2-8	Vκ2D40	Jκ5	3 (1.0)	-
IM2-9	Vκ3-20	Jκ1	11 (3.8)	-
IM3-6	Vκ1D-39	Jκ3	7 (2.5)	-
IM4-1	Vλ1-47	Jλ3	0 (0)	3.1
IM4-2	Vλ2-11	Jλ3	23 (7.8)	8.1
IM4-7	Vλ1-40	Jλ1	18 (6.3)	-
IM4-8	Vκ3-20	Jκ1	14 (5.0)	-
IM4-9	Vκ1-27	NA	2 (0.7)	-
IM4-10	Vκ3-11	Jκ1	7 (2.5)	-
IM5-6	Vκ3-20	Jκ3	6 (2.1)	-
IM5-7	Vκ1D-39	Jκ1	9 (3.2)	-
Means ($n = 17$)			10.4 (3.6)	6.2

*V_κ and V_λ gene nomenclature in refs. 59 and 60, respectively.

†J_κ and J_λ gene nomenclature in ref. 61 and in refs. 62 and 63, respectively. NA, no good match found.

‡V_H gene sequences, where available, are listed in Table 1.

accumulate R mutations. For example, it has been reported that CDR1 is especially susceptible to R mutations based on codon composition (34). To account for this, Lossos *et al.* (24) developed a program that uses the multinomial distribution and takes into account the intrinsic sequence bias of individual germ-line V genes to estimate the influence of positive antigen selection on the mutation process. This program calculates *P* values to quantitate the significance of increased R mutations in the CDR and decreased R mutations in the FWRs. Using this analysis we determined that the proportion of EBV-infected cells having significantly fewer R mutations ($P < 0.05$) in FWRs ranged from 60% to 83% for each donor, and the proportion of cells having significantly more R mutations in CDRs ranged from 20% to 60% (Table 3). The proportion of sequences from EBV-infected cells that were significant by these criteria was similar to the proportion of published sequences of antigen-selected memory B cells that were subjected to the same test (35). That 100% of memory cells do not meet these criteria reflects the stringent nature of the analysis. We conclude that the somatic mutations in the CDRs and FWRs of the expressed Ig genes of EBV-infected cells show the characteristics of positive antigen selection to the same extent as do bona fide memory B cells.

The Light Chains of EBV-Infected Cells Are In-Frame and Somatically Mutated. The conclusion we derived above, that an intact Ig molecule is expressed by EBV$^+$ cells in the PB, was based entirely on analysis of the heavy chains. However, it was conceivable that the light chains could contain aberrant mutations that would result in the loss of Ig expression on the cell surface. To eliminate this possibility, we also amplified and analyzed the V_L chain regions from single CD27$^+$ EBV-infected B cells isolated from the PB. Seventeen sequences were obtained, all of which were in frame and did not contain stop codons. Both κ and λ light chains were used by the EBV$^+$ cells (10 κ genes and 7 λ genes). This usage is typical for human B cells, where κ has been reported to be used by 60% of B cells (36). All light chains carried SHMs, which ranged in number from 0 to 30 per sequence, with an average rate of 3.6% (Table 4). The κ light chains carried a lower rate of mutation (2.7%) than did the λ light chains (4.9%). This finding is consistent with previously published values that showed lower rates (3%) of mutation in the κ light chains (25). The analysis of the light chains confirms that EBV-infected cells express an intact Ig receptor on their surface that has undergone SHM in both heavy and light chains, as is expected for normal memory B cells.

The Heavy Chains of EBV-Infected Cells Are of Various Isotypes. The two defining features of GC-derived memory B cells are SHM and CSR. We have shown above that EBV-infected cells in the PB express Ig V regions that show the characteristics of antigen-selected SHM. To address the question of CSR we performed RT-PCR for the three most prevalent isotypes found on memory B cells in PB (μ, γ, and α). CSR was detected in EBV$^+$ B cells (Table 1). Each cell expressed one constant region heavy chain domain, and all three domains were represented, with γ and α being most prevalent. These data further confirm that EBV-infected cells in PB are true memory cells genotypically because they have undergone CSR and suggest there may be a bias against viral persistence in IgM bearing memory cells.

Cells Latently Infected with EBV in the PB Do Not Show Evidence of Aberrant Activation-Induced Cytidine Deaminase (AID) Behavior. LMP1 has been shown *in vitro* to turn on AID and CSR (18). Therefore, we tested whether the nucleotide composition of the mutations in the expressed Ig genes from the EBV-positive memory cells was characteristic of GC-produced SHM or aberrant AID expression induced by LMP1. Aberrant AID overexpression in cells not ready for SHM results in biased mutations (37–40), with C and G transitions at hotspots being predominant (70–100% of mutations). As evident in Table 5, all nucleotides were targeted for SHM, with only slightly more than half of the mutations occurring at dC and dG bases (57%). Normal SHM of human V_H genes has been reported to cause 55% of the mutations to occur at dC/dG nucleotides (41). Approximately 60% of all transition mutations occurred at dC and dG bases, similar to what was published for V_H genes from the normal human adult memory repertoire (63%) (42). Additionally, ≈35% of mutations occurred at dC and dG nucleotides within the WRCH/DGYW hotspots (43), in remarkable agreement with the reported 35% of mutations in cells that have undergone antigen selection (44). Thus, we can conclude that SHM in the

Table 5. SHM frequency and pattern in single EBV$^+$ CD27$^+$ PB B cells

	A	C	G	T	Total
A	—	34	72	43	149
C	15	—	33	77	125
G	82	54	—	21	157
T	18	29	29	—	76

Fifty-seven percent of mutations are at dC/dG, 44% of mutations are at dA/dT, 61% of dC/dG mutations are transitions, and 35% of mutations are at hotspots.

expressed Ig genes from EBV-infected memory B cells is as expected for normal GC-derived memory B cells and does not display characteristics of aberrant AID targeting caused by EBV infection.

Discussion

In this article we present definitive evidence that cells latently infected with EBV in the PB, the site of viral persistence, are true antigen-selected memory B cells. In our previous studies we proposed that EBV persists in memory B cells based solely on their surface phenotype, CD20$^+$CD27$^+$IgD$^-$CD5$^-$ (6). However, EBV is a transforming virus, and its latent proteins can have a profound effect on the surface phenotype of infected cells (4). It was conceivable, therefore, that the surface phenotype was caused by the presence of the virus rather than a consequence of true memory B cell differentiation, and thus it was unknown whether EBV persisted in true memory B cells. We have now shown that these latently infected cells express Igs that are isotype-switched and carry SHMs. Furthermore, these mutations do not cause stop codons and display the signature of positive antigen selection based on their frequency, type, and location within the Ig gene. Thus, the EBV-infected cells in the PB are indistinguishable from normal memory B cells. Because we find no genotypic aberrations in the EBV-infected memory B cells, we may further conclude that viral latent proteins are prevented from rescuing defective B cells, and any deviations from normal B cell biology are not tolerated in the latently infected peripheral B cells.

Typically, an antigen-activated B cell must enter a follicle to form a GC before becoming a memory cell (11, 12). While in the GC, cells express AID and undergo SHM and CSR. To survive and exit as memory cells they must bind antigen, receive signals from antigen-specific T_H cells, and down-regulate expression of the GC master transcription factor *bcl-6* (45). LMP2a expression is sufficient to cause B cells to form GCs and sustain SHM in the mucosal lymphoid tissue of transgenic mice lacking a BCR (17). LMP2a can also rescue B cells from apoptosis (16) and augment a weak BCR signal (46). LMP1, on the other hand, can constitutively replace T cell help, turn on AID and CSR *in vitro*, and down-regulate *bcl-6* (18–20). Therefore, LMP1 and LMP2a together could provide all of the necessary signals to cause infected naïve B cells to enter, survive in, and exit from the GC to become circulating memory cells without any requirement for signaling by cognate antigen or T_H cells. However, if this were true, the cells would not display the SHM signature of antigen-selected memory B cells that we have observed. Why then are LMP1 and 2a expressed in infected GC cells (15)? The answer may lie in the fact that LMP2a, although it cannot provide a growth signal, can augment a weak BCR signal (46), and LMP1 provides constitutive T cell help. During an immune response GC cells undergo multiple rounds of antigen-driven proliferation and SHM. Because LMP2a cannot drive proliferation, this would be the stage at which cells latently infected with EBV would need to encounter antigen and would explain why these cells show the hallmarks of antigen selection. Because LMP2a can augment weak BCR signals, its role may be to ensure the survival and differentiation of latently infected GC cells bearing a low-affinity BCR as antigen becomes limiting. Ultimately surviving GC cells exit as plasma or memory cells after a short or long exposure, respectively, to T_H cells (47). The role of LMP1 could be to provide a long-lived T_H cell signal to ensure that the latently infected GC cells preferentially differentiate into memory cells. Plasma cell differentiation would be unfavorable for the virus because it leads to reactivation of EBV (48), preventing the virus from establishing long-term latency. Therefore, the role of LMP1 and 2a may be to preferentially direct the survival and differentiation of the latently infected, antigen-selected GC cells into the memory compartment.

A less likely explanation for the presence of LMP1 and 2a in GCs is that they, perhaps in combination with TLR signaling (17, 49), drive the differentiation of infected cells into memory so precisely that the cells are effectively indistinguishable from normal memory B cells even though they have never seen cognate antigen. This outcome could be the result of the SHM machinery itself. If LMP1 turns on AID in GC B cells, they would undergo normal SHM and CSR and accrue both phase I (at C/G nucleotides) and phase 2 (at A/T nucleotides) mutations (50, 51) so that the mutations in the V genes would be indistinguishable from those in normal memory B cells. The signature of positive antigen selection could be a result of the germ-line sequence of the Ig genes and the cytidine deaminase activity of AID. Recently it has been described that the Ig V genes have evolved to direct the induction of SHM by AID into the CDRs and also to minimize R mutations resulting from direct cytosine deaminations (44). If it is true that a memory B cell can express the signature of antigen selection without seeing cognate antigen, it would have profound implications for our understanding of the GC process and the use of SHM to identify the origins of B cells. Specifically, it would mean that the pattern of SHM could no longer be used reliably to identify "true" memory cells. This would not be the first time that EBV studies *in vivo* have produced unexpected findings related to B cell biology. The notion that IgD$^+$ memory B cells were not classical GC-derived memory cells was first suggested by the observation that they could not sustain persistent EBV infection (5, 6), and the first evidence that mucosal memory cells could circulate came from studies of EBV (52).

One of the key observations in this study is that the 32 V_H gene and 17 V_L gene sequences from EBV-infected cells did not contain aberrant mutations or stop codons, and the number of R mutations in FWRs was consistent with preserving a correct Ig protein fold. Thus, a proper Ig molecule on the surface of the infected cells appears to be necessary for their survival. LMP2a can rescue Ig-less B cells from apoptosis. Because we do not detect EBV in any Ig-less B cells, LMP2a must not be expressed very frequently if at all in the infected PB memory B cells. This finding is consistent with our previous work demonstrating that viral proteins are not expressed in these cells (9). LMP2a is present in latently infected GC cells (15) and could rescue these cells from apoptosis if they had acquired aberrant mutations during SHM. Such mutations could result in the loss of surface Ig expression or expression of an autoreactive BCR (16, 53). However, when the cells exit into the PB memory compartment LMP2a expression is down-regulated, and if the BCR is defective the cells would be deleted before entering the periphery. In this manner EBV ensures that it does not allow aberrant cells to survive in the periphery and potentially cause harm to the host.

One possible reason why we did not detect cells with stop codons or aberrant mutations is that such RNAs tend to get degraded more rapidly (54), making them less abundant and harder to amplify by RT-PCR. Therefore, we cannot rigorously exclude the possibility that there are some such cells with sufficiently low transcript copy numbers that they would be

missed by our analysis. However, if these cells exist, they are relatively rare, because our PCR efficiency was, on average, 38% for the infected cells, which was higher than our PCR efficiency with normal memory cells (25%) analyzed in control experiments (data not shown).

One caveat of our results is that we used PB from AIM patients because their frequency of infected cells is up to 1,000-fold greater than healthy carriers (21). To minimize the risk that our findings pertain only to individuals with the highest frequency of infected cells, we extended our analysis to AIM patients with a lower frequency of infection, with similar results. It was not feasible to extend the study to individuals with frequencies of infection $<1/100$. Therefore, technically, we cannot exclude the possibility that long-term viral persistence, which occurs at the level of 1 in 10^5 memory cells, is in atypical memory B cells. However, we think this is unlikely. Although the patients we studied were experiencing AIM, the virus is just as tightly regulated in acute infection as it is during persistent infection, being restricted in the PB to the resting $CD20^+CD27^+IgD^-CD5^-$ B cells that express no detectable viral proteins (9). It is most likely therefore that the single infected cells in the PB from AIM patients are representative of EBV^+ cells from both the acute and persistent stages of viral infection.

In this study we have shown that EBV persists in the PB in true memory B cells that carry functional SHMs and are class-switched. Our data have implications for understanding the relationship between lifelong EBV persistence and EBV-associated tumors. Because the EBV-infected memory B cells in peripheral blood are normal, we may conclude that they are not *per se* a risk factor for B cell lymphoma. Rather, EBV-associated malignancies must represent abnormal behavior by the virus as a consequence of rare circumstances that allow aberrant cells to survive and escape immunosurveillance and other forms of regulation.

We thank Allen Parmellee for FACS, the Tufts University Core Facility for sequencing, Rosemary Ciccarelli (University Health Services, University of Massachusetts, Amherst) for patient recruitment, and Drs. Srimoyee Ghosh and Theresa Imanishi-Kari for insightful discussions. This work was supported by Public Health Research Grants CA65883, AI18757, AI062989 (to D.A.T.-L.) and AI49320 (to K.L.).

1. Thorley-Lawson, D. A. (2001) in *Sampter's Immunologic Diseases*, eds. Austen, K. F., Frank, M. M., Atkinson, J. P. & Cantor, H. (Williams & Wilkins, New York), Vol. 2, pp. 970–985.
2. Rickinson, A. B. & Kieff, E. (2001) in *Virology*, eds. Knipe, D. M. & Howley, P. M. (Lippincott Williams & Wilkins, New York), Vol. 2, pp. 2575–2628.
3. Niederman, J. C., McCollum, R. W., Henle, G. & Henle, W. (1968) *J. Am. Med. Assoc.* **203,** 205–209.
4. Kieff, E. & Rickinson, A. B. (2001) in *Virology*, eds. Knipe, D. M. & Howley, P. M. (Lippincott Williams & Wilkins, New York), Vol. 2, pp. 2511–2574.
5. Babcock, G. J., Decker, L. L., Volk, M. & Thorley-Lawson, D. A. (1998) *Immunity* **9,** 395–404.
6. Joseph, A. M., Babcock, G. J. & Thorley-Lawson, D. A. (2000) *J. Immunol.* **165,** 2975–2981.
7. Thorley-Lawson, D. A. (2001) *Nat. Rev. Immunol.* **1,** 75–82.
8. Thorley-Lawson, D. A. & Gross, A. (2004) *N. Engl. J. Med.* **350,** 1328–1337.
9. Hochberg, D., Middeldorp, J. M., Catalina, M., Sullivan, J. L., Luzuriaga, K. & Thorley-Lawson, D. A. (2004) *Proc. Natl. Acad. Sci. USA* **101,** 239–244.
10. Khan, G., Miyashita, E. M., Yang, B., Babcock, G. J. & Thorley-Lawson, D. A. (1996) *Immunity* **5,** 173–179.
11. Liu, Y. J. & Arpin, C. (1997) *Immunol. Rev.* **156,** 111–126.
12. MacLennan, I. C. (1994) *Annu. Rev. Immunol.* **12,** 117–139.
13. Stavnezer, J. (1996) *Adv. Immunol.* **61,** 79–146.
14. Rajewsky, K. (1996) *Nature* **381,** 751–758.
15. Babcock, G. J., Hochberg, D. & Thorley-Lawson, A. D. (2000) *Immunity* **13,** 497–506.
16. Caldwell, R. G., Wilson, J. B., Anderson, S. J. & Longnecker, R. (1998) *Immunity* **9,** 405–411.
17. Casola, S., Otipoby, K. L., Alimzhanov, M., Humme, S., Uyttersprot, N., Kutok, J. L., Carroll, M. C. & Rajewsky, K. (2004) *Nat. Immunol.* **5,** 317–327.
18. He, B., Raab-Traub, N., Casali, P. & Cerutti, A. (2003) *J. Immunol.* **171,** 5215–5224.
19. Kilger, E., Kieser, A., Baumann, M. & Hammerschmidt, W. (1998) *EMBO J.* **17,** 1700–1709.
20. Panagopoulos, D., Victoratos, P., Alexiou, M., Kollias, G. & Mosialos, G. (2004) *J. Virol.* **78,** 13253–13261.
21. Hochberg, D., Souza, T., Catalina, M., Sullivan, J. L., Luzuriaga, K. & Thorley-Lawson, D. A. (2004) *J. Virol.* **78,** 5194–5204.
22. Wang, X. & Stollar, B. D. (2000) *J. Immunol. Methods* **244,** 217–225.
23. Lefranc, M. P. (2003) *Nucleic Acids Res.* **31,** 307–310.
24. Lossos, I. S., Tibshirani, R., Narasimhan, B. & Levy, R. (2000) *J. Immunol.* **165,** 5122–5126.
25. Klein, U., Rajewsky, K. & Kuppers, R. (1998) *J. Exp. Med.* **188,** 1679–1689.
26. Goossens, T., Klein, U. & Kuppers, R. (1998) *Proc. Natl. Acad. Sci. USA* **95,** 2463–2468.
27. Wilson, P. C., de Bouteiller, O., Liu, Y. J., Potter, K., Banchereau, J., Capra, J. D. & Pascual, V. (1998) *J. Exp. Med.* **187,** 59–70.
28. Chong, Y., Ikematsu, H., Yamaji, K., Nishimura, M., Kashiwagi, S. & Hayashi, J. (2003) *Clin. Exp. Immunol.* **133,** 59–66.
29. Klein, U., Goossens, T., Fischer, M., Kanzler, H., Braeuninger, A., Rajewsky, K. & Kuppers, R. (1998) *Immunol. Rev.* **162,** 261–280.
30. Klein, U., Kuppers, R. & Rajewsky, K. (1997) *Blood* **89,** 1288–1298.
31. Wang, X. & Stollar, B. D. (1999) *Clin. Immunol.* **93,** 132–142.
32. Shlomchik, M. J., Marshak-Rothstein, A., Wolfowicz, C. B., Rothstein, T. L. & Weigert, M. G. (1987) *Nature* **328,** 805–811.
33. Dorner, T., Brezinschek, H. P., Brezinschek, R. I., Foster, S. J., Domiati-Saad, R. & Lipsky, P. E. (1997) *J. Immunol.* **158,** 2779–2789.
34. Chang, B. & Casali, P. (1994) *Immunol. Today* **15,** 367–373.
35. Ghosh, S., Steere, A. C., Stollar, B. D. & Huber, B. T. (2005) *J. Immunol.* **174,** 2860–2869.
36. Kuppers, R., Zhao, M., Hansmann, M. L. & Rajewsky, K. (1993) *EMBO J.* **12,** 4955–4967.
37. Yoshikawa, K., Okazaki, I. M., Eto, T., Kinoshita, K., Muramatsu, M., Nagaoka, H. & Honjo, T. (2002) *Science* **296,** 2033–2036.
38. Martin, A. & Scharff, M. D. (2002) *Proc. Natl. Acad. Sci. USA* **99,** 12304–12308.
39. Martin, A., Bardwell, P. D., Woo, C. J., Fan, M., Shulman, M. J. & Scharff, M. D. (2002) *Nature* **415,** 802–806.
40. Faili, A., Aoufouchi, S., Flatter, E., Gueranger, Q., Reynaud, C. A. & Weill, J. C. (2002) *Nature* **419,** 944–947.
41. Dorner, T., Foster, S. J., Brezinschek, H. P. & Lipsky, P. E. (1998) *Immunol. Rev.* **162,** 161–171.
42. Imai, K., Slupphaug, G., Lee, W. I., Revy, P., Nonoyama, S., Catalan, N., Yel, L., Forveille, M., Kavli, B., Krokan, H. E., *et al*. (2003) *Nat. Immunol.* **4,** 1023–1028.
43. Rogozin, I. B. & Diaz, M. (2004) *J. Immunol.* **172,** 3382–3384.
44. Zheng, N. Y., Wilson, K., Jared, M. & Wilson, P. C. (2005) *J. Exp. Med.* **201,** 1467–1478.
45. Ye, B. H., Cattoretti, G., Shen, Q., Zhang, J., Hawe, N., de Waard, R., Leung, C., Nouri-Shirazi, M., Orazi, A., Chaganti, R. S., *et al*. (1997) *Nat. Genet.* **16,** 161–170.
46. Swanson-Mungerson, M. A., Caldwell, R. G., Bultema, R. & Longnecker, R. (2005) *J. Virol.* **79,** 7355–7362.
47. Arpin, C., Dechanet, J., Van, K. C., Merville, P., Grouard, G., Briere, F., Banchereau, J. & Liu, Y. J. (1995) *Science* **268,** 720–722.
48. Laichalk, L. L. & Thorley-Lawson, D. A. (2005) *J. Virol.* **79,** 1296–1307.
49. Bernasconi, N. L., Onai, N. & Lanzavecchia, A. (2003) *Blood* **101,** 4500–4504.
50. Chaudhuri, J., Khuong, C. & Alt, F. W. (2004) *Nature* **430,** 992–998.
51. Neuberger, M. S., Di Noia, J. M., Beale, R. C., Williams, G. T., Yang, Z. & Rada, C. (2005) *Nat. Rev. Immunol.* **5,** 171–178.
52. Laichalk, L. L., Hochberg, D., Babcock, G. J., Freeman, R. B. & Thorley-Lawson, D. A. (2002) *Immunity* **16,** 745–754.
53. Mancao, C., Altmann, M., Jungnickel, B. & Hammerschmidt, W. (2005) *Blood*, in press.
54. Hillman, R. T., Green, R. E. & Brenner, S. E. (2004) *Genome Biol.* **5,** R8.
55. Matsuda, F., Shin, E. K., Nagaoka, H., Matsumura, R., Haino, M., Fukita, Y., Taka-ishi, S., Imai, T., Riley, J. H., Anand, R., *et al*. (1993) *Nat. Genet.* **3,** 88–94.
56. Matsuda, F., Ishii, K., Bourvagnet, P., Kuma, K., Hayashida, H., Miyata, T. & Honjo, T. (1998) *J. Exp. Med.* **188,** 2151–2162.
57. Corbett, S. J., Tomlinson, I. M., Sonnhammer, E. L., Buck, D. & Winter, G. (1997) *J. Mol. Biol.* **270,** 587–597.
58. Ravetch, J. V., Siebenlist, U., Korsmeyer, S., Waldmann, T. & Leder, P. (1981) *Cell* **27,** 583–591.
59. Lefranc, M. P. (2001) *Exp. Clin. Immunogenet.* **18,** 161–174.
60. Lefranc, M. P. (2001) *Exp. Clin. Immunogenet.* **18,** 242–254.
61. Hieter, P. A., Maizel, J. V., Jr., & Leder, P. (1982) *J. Biol. Chem.* **257,** 1516–1522.
62. Udey, J. A. & Blomberg, B. (1987) *Immunogenetics* **25,** 63–70.
63. Vasicek, T. J. & Leder, P. (1990) *J. Exp. Med.* **172,** 609–620.