EARLY HUMAN IgH GENE ASSEMBLY IN EPSTEIN-BARR VIRUS-TRANSFORMED FETAL B CELL LINES

Preferential Utilization of the Most J_H-proximal D Segment (DQ52) and Two Unusual V_H-related Rearrangements

By KATHERINE G. NICKERSON,* JEFFREY BERMAN,‡ EVA GLICKMAN,* LEONARD CHESS,* AND FREDERICK W. ALT‡§||

*From the Departments of *Medicine, ‡Biochemistry, and §Microbiology, and the ||Howard Hughes Medical Institute, Columbia University College of Physicians and Surgeons, New York, New York 10032*

The genes encoding IgH and L chains are assembled in B cells by somatic recombination of component gene segments (1). Insight into the mechanisms regulating this process has been gained from studies of Abelson murine leukemia virus (AMuLV)[1]-transformed murine pre-B cells that undergo Ig gene assembly in culture (reviewed in references 2 and 3). From this system we know that H chain variable (V_H), diversity (D), and joining (J_H) elements are brought together in an orderly sequence, with D to J_H joining occurring before V_H to DJ_H joining. This process is mediated by conserved heptamer and nonamer recognition sequences that are separated by either 12- (D) or 23-bp spacers (V_H and J_H). Recombination is restricted to segments flanked, respectively, by recognition sequences with 12- and 23-bp spacers (12/23 rule; 4, 5). According to these restrictions, pre-existing DJ_H joins serve as substrates for V_H to DJ_H joining but their replacement by rearrangements of upstream D segments to downstream J_H segments is also permitted (6). One recognized exception to the 12/23 rule at the H chain locus may occur in recently described V_H replacement events in which one V_H gene replaces another in the context of an existing V_HDJ_H rearrangement (7, 8). In this case, internal V_H heptamers that are found at the 3' end of most V_H genes can apparently mediate site-specific recombination in the absense of nonamer or spacer elements.

There seems to be random representation of the possible V_H, D, and J_H elements in combination with one another in the adult murine Ig repertoire, but in both AMuLV-transformed pre-B cells (9) and in murine fetal liver hybridomas (10), there is biased rearrangement and expression of the most 3' (J_H-proximal) V_H genes. It is not known whether or not this 3' bias holds true for early D segment utilization, because AMuLV-transformed cells rapidly undergo secondary D-J_H and V_H-DJ_H rearrangements (6).

This work was supported by the Howard Hughes Medical Institute, grants from the Lupus Foundation (to K. G. Nickerson), and National Institutes of Health (AI-24748 to L. Chess, and AI-20047 and CA-40427 to F. W. Alt). K. G. Nickerson was supported by a Public Health Service Training Grant (AI-07132). J. E. Berman was a fellow of the Damon Runyon/Walter Winchell Cancer Fund and currently is a fellow of the Howard Hughes Medical Institute.

[1] *Abbreviation used in this paper:* AMuLV, Abelson murine leukemia virus.

To establish a system in which to study early events in human IgH gene assembly, we have immortalized human fetal B lineage cells using EBV transformation. Some of these cells are at the initial stages of IgH gene assembly with some alleles apparently still in the unrearranged, germline configuration. While similar cell lines have previously been described, we have for the first time characterized in detail their IgH rearrangements. Of the rearranged alleles the majority represent joining of the most 3′ D segment (DQ52) to various J_H segments, suggesting that this most J_H-proximal D segment is a preferential target for the initiation of IgH gene rearrangements. In addition, we have identified and characterized three rearrangements involving V_H segments. One is a normal in-frame V_HDJ_H, but the other two have unanticipated structures. One has a $V_HDJ_HJ_H$-DJ_H-like structure, while the other consists of the heptamer, nonamer, and 3′ flanking sequences of a V_H4 gene joined in inverted orientation to J_H4.

Materials and Methods

Establishment of Cell Lines. Liver or bone marrow was obtained from 9–16-wk-old fetuses with maternal informed consent at the time of elective abortion. Single cell suspensions were prepared and centrifuged over a Ficoll-diatrazoate gradient to obtain mononuclear cells. Cells were plated in 24-well tissue culture plates, and EBV was obtained from the supernatant of B95-8 marmoset leukocytes (American Type Culture Collection, Rockville, MD) was added. The cell lines were fed twice weekly with Iscove's modified Dulbecco's medium (Life Technologies, Inc., Grand Island, NY) supplemented with 10% FCS (HyClone Laboratories, Logan, UT) and 1% penicillin-streptomycin as described (11). After the cell lines were well established (5–8 wk), cells were simultaneously harvested for phenotypic analysis and DNA extraction.

Phenotypic Analysis. Expression of cell surface markers was examined by indirect immunofluorescence as previously described (11) using a cytofluorograph (model 30-H; Ortho Diagnostic Systems, Inc., Westwood, MA). The following panel of mouse mAbs: anti-HLA-DR, OKB4(12), OKB7(CD21), OKT1(CD5), OKT3(CD3), OKT11(CD2), OKM2 (Ortho Pharmaceutical, Raritan, NJ), B1(CD20), anti-IgM, anti-IgG, anti-κ, and anti-λ (Coulter Electronics Inc., Hialeah, FL) were used. Cells were counterstained using FITC-conjugated affinity-purified $F(ab')_2$ goat anti-mouse IgG + IgM (The Jackson Laboratory, Bar Harbor, ME). For intracytoplasmic staining cells were fixed on microscope slides and examined by indirect immunofluorescence as above with mouse anti-human IgM, IgG, κ, or λ. Secreted Ig was detected in an ELISA using standard techniques (13). Briefly, supernatants were added to 96-well round-bottomed plates (Costar, Cambridge, MA) previously coated with goat anti-human IgA-G-M (Kirkegaard & Perry Laboratories, Inc., Gaithersburg, MD) and blocked with 1% BSA in PBS. Dilutions of purified human monoclonal myeloma proteins of known concentration (a gift of E. F. Osserman, Columbia University, College of Physicians and Surgeons NY, NY) were plated as controls. Plates were developed with alkaline phosphatase-labeled goat anti-human IgM or IgG (Kirkegaard & Perry Laboratories, Inc.) and α-nitrophenyl phosphate. The optical densities were read at 405 nm on an ELISA reader (model EL307; Bio-Tek Instruments, Inc., Burlington, VT). This assay was sensitive to ~10 ng of Ig.

Southern Blotting. High molecular weight DNA was extracted, digested with restriction endonucleases, subjected to electrophoresis through 0.8% agarose gels, transferred to nitrocellulose membranes, and hybridized with ^{32}P nick-translated probes as described (14). Blots were washed in 2× or 0.2× SSC with 0.1% SDS at 68°C for 1 h. The probes used are shown in Fig. 1. Probe A is a 6.5-kb Bam HI–Hind III fragment spanning the entire human germline J_H region (15). This fragment has six functional J_H segments, three pseudo J_H segments, and one D segment, DQ52, which is located between $J\psi1$ and J1. Probe B is a 0.65-kb Pvu II fragment upstream from DQ52, and probe C is a 0.32-kb Sma I–Pst I fragment including only the germline J1 and J2 segments.

Genomic Cloning of Rearranged IgH Genes. Hind III fragments of selected cell lines were cloned into Charon 35 or Charon 21A (a gift of Fred Blattner, University of Wisconsin, Madison, WI) bacteriophage as described (14). Plaques were screened with probe A. J_H hybridizing human insert DNA was subcloned into pUC 13 and extensively mapped by restriction enzyme digestion and agarose gel electrophoresis. Both strands of relevant restriction fragments were sequenced according to the methods of Maxam and Gilbert (16). Sequences were compared with Genbank human Ig sequences as well as additional recently available human D and V_H sequences (14, 17–24) using Microgenie IBM software.

Results

Phenotypic Characteristics of Cell Lines. Cell lines were established from four human fetal livers (FL-1, FL-2, FL-3, and FL-4) and two fetal bone marrows (FBM-1 and FBM-2). All of the lines were similar in growth characteristics and morphology to other EBV-transformed B cells from adult peripheral blood, and they expressed comparable levels of the B cell markers DR, CD20, OKB4, as well as the EBV receptor CD21 (not shown). Although fetal tissues are enriched in the subset of B cells bearing low levels of the T cell antigen CD5 (25), our transformed populations did not express detectable levels of CD5 or other T cell or monocyte markers.

To define the B cell differentiation stage of the cell lines, we first characterized expression of surface and cytoplasmic Ig. Three lines were entirely surface Ig^- (Table I). Two of these, FL-1 and FL-3, also completely lacked cytoplasmic μ chains, while FL-2 had only rare (<1%) cytoplasmic μ-staining cells, indicating that the majority of cells in these cultures were at a very early stage of B cell differentiation. When reexamined after 6 mo in culture, 25% of the cells in FL-2 were surface Ig^+ (FL-2A) and 32% had cytoplasmic μ. Both κ and λ L chains were present, indicating that these cell populations were not clonal, at least at the level of their Ig gene expression. As the original fetal liver population was polyclonal, we could not determine whether IgH rearrangement had occurred in vitro or whether small numbers of mature B cells present in the initial cultures had exhibited a growth advantage over Ig^- cells.

The other three cell lines, FL-4, FBM-1, and FBM-2, were heterogeneous with respect to Ig expression. They contained 25–60% surface Ig^+ cells and had levels of IgM in the supernatants roughly comparable with those of normal adult EBV cell lines. Thus, the fetal B cell lines described here include cells at a spectrum of stages of B cell differentiation ranging from those lacking any expression of Ig to mature B cells expressing surface and secreted H and L chains.

Analysis of DQ52-related Rearrangements. To define the configuration of IgH genes in these cell lines, we analyzed J_H-associated fragments from genomic DNA samples. Hind III digests were assayed on Southern blots for hybridization to probes specific for various regions of the J_H locus (Fig. 1). Probe A, which spans the entire J_H locus and so detects all J_H-associated rearrangements, hybridized to two Hind III fragments from FL-3 and FBM-1, consistent with a clonal population of cells that was not actively undergoing Ig assembly in culture. The other cell lines had three or more J_H-hybridizing fragments, reflecting either oligoclonal populations or rearrangements occurring in culture. Five of the six cell lines (all except FL-3) had a fragment corresponding in size to the germline, unrearranged configuration at 10.5 kb. Consistent results were also found with Eco RI and Bam HI digests (not shown). DNA from both HeLa cells and the neuroblastoma cell line LAN5 were used as

TABLE I

Phenotypic Characteristics and IgH Gene Rearrangements of Fetal B Lineage Cell Lines

	Surface Ig				Cytoplasmic Ig				Secreted Ig		Identity of IgH rearrangements on Southern blots
	μ	γ	κ	λ	μ	γ	κ	λ	μ	γ	
FL-1	–	–	–	–	–	–	–	–	–	–	G DQ52-J1 (DQ52-J2)
FL-2	–	–	–	–	1%	–	1%	1%	–	–	(G) DQ52-J1 DQ52-J3
FL-2A	25	–	11	9	32	–	8	8	+	–	G DQ52-J1 DQ52-J2 DQ52-J3* (*FL2-1*) DQ52-J4 & VH3DJ4* (*FL2-2*)
FL-3	–	–	–	–	–	–	–	–	–	–	DQ52-J1 VH3DJJ-DJ4* (*FL3-1*)
FL-4	52	–	43	24	ND	ND	ND	ND	+	–	G VH4 3'Flank-J4* (*FL4-1*) DQ52-J3 (multiple R)
FBM-1	59	–	–	32	ND	ND	ND	ND	+	–	G or DQ52-J1 VHDJ1 or J2
FBM-2	27	–	13	19	ND	ND	ND	ND	+	–	G DQ52-J1 (multiple R)

The expression of surface and cytoplasmic Ig is shown as percentage of positively staining cells. Secreted Ig is shown as either + or –. Those cell lines that were + all had concentrations of Ig in their supernatants between 500 ng and 1 mg/ml. The cell line FL-2 was initially Ig^-, however, when reexamined 6 mo later, surface, cytoplasmic, and secreted Ig were present and the cell line was designated FL2A. The last column shows the probable identity of IgH rearrangements on Southern blots based on their size and hybridization to specific J_H probes (see also Fig. 1). Parentheses indicate minor bands.

* These rearrangements were isolated and their nucleotide sequences are shown in Figs. 2–4.

nonlymphoid germline controls. It is possible that some of these actually represented rearrangements of the most J_H-proximal D segment, DQ52, to J_H1, as such rearrangements would be only 90 bp smaller than the germline. In either case, the results were indicative that many of the cells in these cultures were at a very early stage of Ig assembly with some alleles still in the germline configuration or at most having undergone a DQ52-J_H1 rearrangement.

To assess the possibility that DQ52 rearrangements to other J_H segments might be present, we probed Hind III–digested DNA with probe B (Fig. 1). This fragment would be deleted in all rearrangements except those involving joining of DQ52 to various J_H segments. All cell lines except for FBM-1 had one or more novel Hind III fragments that hybridized with this probe. All of these fragments were in the range of 0.1–1 kb smaller than the germline fragment, consistent with joining of DQ52 to J_H1-J_H4 (Fig. 1). To confirm this point, we probed blots with probe C (Fig. 1), which under high stringency washing conditions hybridizes specifically to fragments

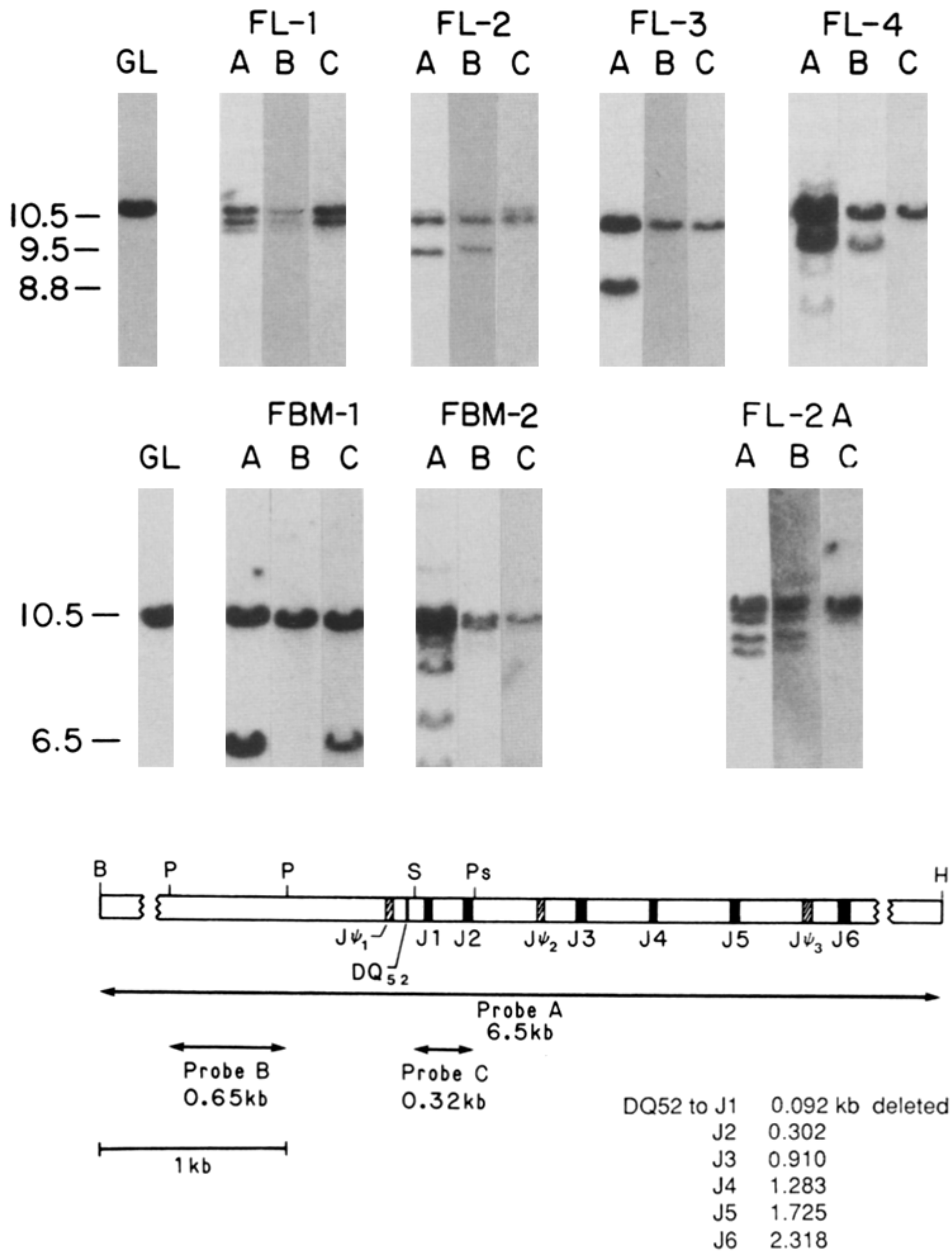

FIGURE 1. Southern blot analysis of J_H-associated rearrangements. DNA from six cell lines and a nonlymphoid, germline (*GL*) control are shown hybridized to three probes (*A*, *B*, and *C*) as indicated above each lane. FL-2, when re-examined after 6 mo in culture, had novel rearrangements and is designated FL-2A. The map of the germline J_H region (15) shows the relative position of the J_H and DQ52 segments and indicates the relevant restriction enzyme sites used to generate the probes. Also shown are the predicted number of basepairs deleted for given DQ52 to J_H related rearrangements. Abbreviations of restriction enzymes are as follows: B, Bam HI; H, Hind III; P, Pvu II; S, Sma I; Ps, Pst I.

bearing J_H1 and/or J_H2. As predicted, of the probe B-hybridizing fragments, only germline-sized fragments and those ~100 or 300 bp smaller were identified with probe C. Thus, five J_H^+ fragments were present in FL-2A (Fig. 1). All of these hybridized to probe B, but only the upper three hybridized with probe C (Fig. 1). This result is consistent with the five fragments from top to bottom, corresponding to germline, DQ52-J_H1, DQ52-J_H2, DQ52-J_H3, and DQ52-J_H4 configurations. The other cell lines, except possibly FBM-1, have comparable DQ52-J_H rearrangements. FBM-1 has one allele in a DQ52-J_H1 or possibly germline configuration, while the other 6.5-kb allele hybridizes to probe C but not probe B (Fig. 1). This rearrangement is most likely a productive V_HDJ_H using J_H1 or J_H2, accounting for the Ig$^+$ phenotype of the cell line.

To confirm that the genomic blotting analysis had correctly identified DQ52-related rearrangements, we isolated one of these fragments from a Hind III library prepared from genomic DNA of FL-2A. The nucleotide sequence of the relevant portion of this clone confirms that this rearrangement represents a normal DQ52-J_H3 join (Fig. 2). Six bases between the DQ52 and J_H3 coding regions correspond to a probable N region addition that occurred during the process of rearrangement.

V_H-associated Rearrangements. Although most rearrangements in these cell lines were identifiable as DQ52-J_H joins, several were candidates to be rearrangements using other D or V_H segments. By molecular cloning we identified three V_H-associated rearrangements from these cell lines (Figs. 3 and 4). Two of them used members of the largest human V_H family, V_H3 (14). One of these, *FL2-2*, was isolated from a genomic Hind III library of FL2A. The cloned fragment was not evident on Southern blots because it comigrated with the rearrangement identified as a DQ52-JH4. It was clearly distinct, however, because on Southern blots of cloned DNA, it did not hybridize with probe B (not shown). The V_H3 gene used by *FL2-2* differs by only 1 bp from that of a germline V_H3 gene previously isolated (V_H1.9III,14). The D

FIGURE 2. Nucleotide sequence of *FL2-1* compared with germline sequences of J_H3, DQ52, and 5' flanking regions including Jψ1; differences between the two sequences are starred; recognition heptamer and nonamer sequences are enclosed in boxes. These sequence data have been submitted to the EMBL/GenBank Data Libraries under the accession number Y00798.

FL2-2 **a**

```
Leader -->
G TTT GGG CTG AGC TGG GTT TTC CTC GTT GCT CTT TTA AGA G
Intron -->
GTGATTCATGGAGAAATAGAGAGACTGAGTGTGAGTGAACATGAGTGA

GAAAAACTGGATTTGTGTGGCATTTTCTGATAACGGTGTCCTTCTGTTT
                             Coding Region -->
GCAG GT GTC CAG TGT CAG GTG CAG CTG GTG GAG TCT GGG GGA
                    Gln Val Gln Leu Val Glu Ser Gly Gly

GGC GTG GTC CAG CCT GGG AGG TCC CTG AGA CTC TCC TGT GCA
Gly Val Val Gln Pro Gly Arg Ser Leu Arg Leu Ser Cys Ala

GCC TCT GGA TTC ACC TTC AGT AGC TAT GGC ATG CAC TGG GTC
Ala Ser Gly Phe Thr Phe Ser Ser Tyr Gly Met His Trp Val
             C
CGC CAG GCT CTA GGC AAG GGG CTG GAG TGG GTG GCA GTT ATA
Arg Gln Ala Leu Gly Lys Gly Leu Glu Trp Val Ala Val Ile

TCA TAT GAT GGA AGT AAT AAA TAC TAT GCA GAC TCC GTG AAG
Ser Tyr Asp Gly Ser Asn Lys Tyr Tyr Ala Asp Ser Val Lys

GGC CGA TTC ACC ATC TCC AGA GAC AAT TCC AAG AAC ACG CTG
Gly Arg Phe Thr Ile Ser Arg Asp Asn Ser Lys Asn Thr Leu

TAT CTG CAA ATG AAC AGC CTG AGA GCT GAG GAC ACG GCT GTG
Tyr Leu Gln Met Asn Ser Leu Arg Ala Glu Asp Thr Ala Val
                             DQ52                JH4
TAT TAC TGT GCG AAA GAT CGA AAC TGG GGT TTT GAC TAC TGG
Tyr Tyr Cys Ala Lys Asp Arg Asn Trp Gly Phe Asp Tyr Trp
                  C
GGC CAA GGA ACG CTG GTC AC
Gly Gln Gly Thr Leu Val
```

FL3-1 **b**

```
Leader -->
G TTT TGG CTG AGC TGG GTT TTC CTT GTT GCT ATT TCA AAA G
Intron -->
GTGATTCATGGAGAACCAGAGATACCGAGTGTGAGTGAATACGAGTGA

GAGAAACAGTGGATTATGTGTGACAGTTCCAACCAATGTCTCTGTGTTT
                             Coding Region -->
GCAG GT GTC CAG TGT GAG GTG TAG CTG GTG GAG ACT GGA GGA
                    Glu Val END Leu Val Glu Thr Gly Gly

GGC TTG ATG CAG CCT GGG GGG TCC CTG AGA CTC TCC TGT GCA
Gly Leu Met Gln Pro Gly Gly Ser Leu Arg Leu Ser Cys Ala

GCC TCT GGG TTC ACC GTC AGT AGC AAT CAC ATG AGC TGG GTC
Ala Ser Gly Phe Thr Val Ser Ser Asn His Met Ser Trp Val

CGC CAG GCT CCA GGG AAG GGG CTG GAG TGG GTC TCA GTT ATT
Arg Gln Ala Pro Gly Lys Gly Leu Glu Trp Val Ser Val Ile

TAT AGC GGT GGT GAC ACA TAC TAC GCA GAC TCC GTG AAG AAC
Tyr Ser Gly Gly Asp Thr Tyr Tyr Ala Asp Ser Val Lys Asn

CGA TTC ACC ATC TCC AGA GAC AAT TCC AAG AAC ACG CTG TTT
Arg Phe Thr Ile Ser Arg Asp Asn Ser Lys Asn Thr Leu Phe

CTT CAA ATG AAC AGC CTG AGA GCC GAG GAC ACG GCC GTG TAT
Leu Gln Met Asn Ser Leu Arg Ala Glu Asp Thr Ala Val Tyr
                          D segment
TAC TGT GCG GGG GGA TCG GAT ATG GCG GCG TCA ACT GGT TTG
Tyr Cys Ala Gly Gly Ser Asp Met Ala Ala Ser Thr Gly Leu

ACT ATT GGG GCC AGG CAA CTG GTT TGA CTA TTG GGG CCA AGG
Thr Ile Gly Ala Arg Gln Leu Val END Leu Leu Gly Pro Arg
        D segment       JH4 or JH5
AAC CGG GTA TAG CAG CAA CTG GTT TGA CTA CTG GGG CCA AGG
Asn Arg Val End Gln Gln Leu Val END Leu Leu Gly Pro Arg

AAC CCT GGT CA
Asn Pro Gly
```

FIGURE 3. (*a*) Nucleotide sequence of *FL2-2* along with its derived amino acid sequence. Leader, intron, and coding regions are indicated according to Kabat et al. (37). Sequences homologous to DQ52 are underlined. Nucleotide differences from the corresponding germline V_H (14) and J_H sequences are indicated in the two places where they occur. (*b*) Nucleotide sequence of *FL3-1* along with its derived amino acid sequence. Stop codons occur as indicated by END. Nucleotides in common with germline D segments (21) are underlined as are the duplicated J_H-related sequences. These sequence data have been submitted to the EMBL/GenBank Data Libraries under the accession number Y00798.

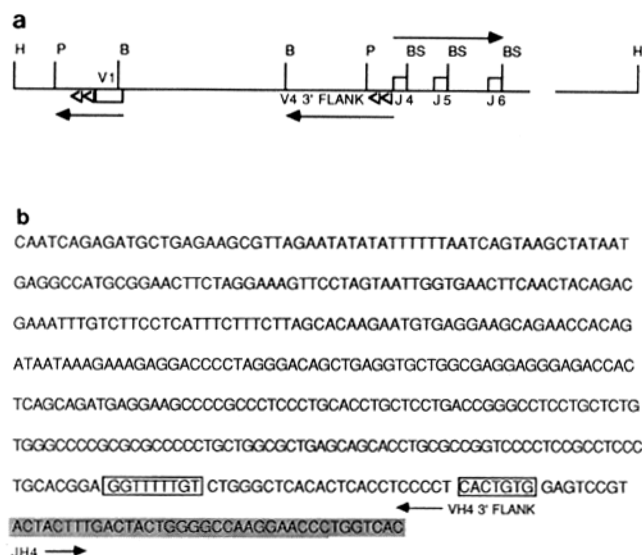

FIGURE 4. (*a*) Map of the Hind III phage insert *FL4-1* showing some of the restriction sites used for mapping and sequencing analysis. Arrows indicate the sense orientation of given fragments. Restriction enzymes are designated as follows: H, Hind III; B, Bam H1; P, Pvu II; BS, Bst EII. (*b*) Nucleotide sequence of the aberrant rearrangement *FL4-1*. J_H4 and the 3′ flanking region of a V_H4 gene are shown and their orientation indicated by arrows. The heptamer and nonamer are enclosed in boxes. The origin of the eight nucleotides between the V_H4 heptamer and J_H4 is unknown. These sequence data have been submitted to the EMBL/GenBank Data Libraries under the accession number Y00798.

segment could be DQ52 as its shares 7 bp with the germline DQ52, and the J_H differs by just 1 bp from J_H4. The base changes from published sequences could be polymorphisms, or may represent limited somatic mutation. This rearrangement appears to be productive, and so, may account for the Ig expression among a subpopulation of cells in FL-2A.

A V_HDJ_H rearrangement isolated from FL-3 also uses a V_H3 gene (Fig. 3 *b*). This V_H is 94% homologous to a previously isolated functional germline gene, V_H 8-1B (14), but has a translation termination codon at the third amino acid. The DJ_H portion of the rearrangement is curious, consisting of what appears to be a D segment with a 6-bp sequence in common with three previously described D segments (21), followed by two nearly identical, tandemly repeated portions of a J_H, in turn, followed by another 11-bp D segment corresponding exactly to part of a recently identified germline D segment (21) appended to J_H4 or J_H5. Our restriction map of the 3′ flanking sequences of this rearrangement is consistent with the J_H being J_H4. Whether the unusual structure of this rearrangement reflects a polymorphism in the germline JH locus, or whether it arose during the Ig assembly process is not clear.

We isolated another very unusual rearrangement from FL-4, *FL4-1*, consisting of the heptamer and nonamer recognition elements and 3′ flanking regions of a V_H4 gene joined in inverted orientation to J_H4 (Fig. 4). The sequence of the heptamer, nonamer, and 23-bp spacer is identical to comparable regions of two previously identified V_H4 genes (V_H58 and $V_H71.4$; reference 23). 3′ flanking regions are not available from these two genes for comparison, but flanking region sequences from two other germline V_H4 genes (14) extending 220 and 110 bp show 88 and 90% homology, respectively. 8 bp between the J_H4 coding region and the V_H4 heptamer are of unknown origin.

Approximately 4 kb upstream of the J_H4-associated part of the phage insert there is a V_H1 gene, which by mapping and partial sequencing, we determined to be in

the same orientation as the 3′ flanking region of the V_H4 gene (Fig. 4). This is another example of human V_H genes of different families being found in close proximity (14, 22), and it also shows that the mechanism giving rise to this rearrangement resulted in a relatively large piece of the germline V_H region being brought into contact with the J_H region in an inverted orientation.

Discussion

We have used EBV to transform early B lineage cells from human fetal liver and bone marrow. As others have shown, EBV can transform B cells at all stages of differentiation, including before any IgH rearrangements (26–29). Our cell lines are oligoclonal, but have predominant populations at the earliest stages of IgH gene assembly, with some alleles still apparently in the germline configuration. We have, for the first time, characterized in detail the rearranged alleles from such early EBV-transformed cell lines.

Comparable murine B lineage cells have not been available because AMuLV-infected pre-B cells usually have DJ_H rearrangements on both alleles, and often rapidly undergo secondary rearrangement events consisting of V_H-DJ_H joining or rearrangement of an upstream D to a downstream J_H (6). By far the most common rearrangements we detected are those involving DQ52 joined to various J_H segments. In both the mouse and human genomes, DQ52 is uniquely situated immediately 5′ of the J_H segments, while the next nearest identified D segment in the mouse is 17 kb upstream (30), and in the human, 22 kb upstream (31). Our data suggest that initial rearrangement events at the IgH locus in B lineage cells may preferentially use DQ52 and the adjacent J_H segments. Recent analyses of IgH rearrangements occurring in normal murine T cells (32), as well as in human leukemic T and B cells (33) also indicate that DQ52-J_H rearrangements are common crosslineage or tumor-associated rearrangements, further implicating DQ52 as a preferred initiation site for rearrangement activity upstream of the J_H cluster.

The significance of DQ52 in specifying antigen binding is not known. Although D segments cannot always be identified with certainty, both because of *N* region additions and because some human germline D segments probably remain to be characterized, DQ52 does appear to be involved in productive V_HDJ_H rearrangements. One example may be *FL2-2*, in which 7 of 17 nucleotides of the D segment are shared with DQ52. Similarly, a possible increased frequency of DQ52 usage among expressed V_HDJ_H genes from a fetal liver sample was noted, with 8 of 14 D segments sharing between 5 and 9 bp with DQ52 (24). An early bias in DQ52 utilization, however, could also be masked either by rearrangements of upstream D segments to downstream J_H segments (6), or by selection for other expressed D segments at the cellular level. Such is the case for murine V_H expression in which the newborn liver repertoire is dramatically biased in favor of J_H-proximal V_H elements, while in the adult spleen, V_H expression is normalized across the entire locus (34).

Our sample of V_H-related rearrangements was too small to draw conclusions about the primary human V_H repertoire. Both V_HDJ_H rearrangements isolated use members of V_H3, the largest human V_H family, which includes members widely dispersed across the V_H locus (14). The chromosomal location of the V_H most closely related to the V_H rearranged in FL-3 is not known, but the V_H gene used by *FL2-2* is not

among the most J_H-proximal V_H genes, as its germline counterpart does not hybridize to the restriction fragment linking V_H and J_H loci (14). As this gene is involved in a productive rearrangement, the V_H gene could have been selected for L chain association or other characteristics. An apparently restricted V_H repertoire was observed in a sampling of 14 expressed V_H genes from a single fetal liver sample (24). The chromosomal location of those genes has not been determined, nor are they represented in our small sampling of cultured cells.

One of the isolated J_H-associated rearrangements, *FL4-1*, consisted of 3' flanking sequences of a VH4 gene inverted and joined to J_H4 coding sequences. The consequence of this join was to replace J_H4 heptamer and nonamer recognition sequences with the recognition sequences of the V_H4 segment; in addition, 8 bp of unknown origin were inserted between JH and the VH heptamer, which could reflect either the contribution of a D segment or could correspond to an N region addition. In this regard, the rearrangement could have arisen in several ways. It could have involved joining of V signal sequences to a preexisting DJ_H substrate, as has recently been found to occur in V_HDJ_H recombination substrates (35). Alternatively, the join could have involved direct joining of V and J elements. In the latter case, such a joining, if mediated by the normal flanking recognition sequences, would necessarily be in violation of the 12/23 joining rule; on the other hand, the join conceivably could have been mediated by the internal VH heptamer (analogous to a V_H to V_HDJ_H join; 7, 8). At this time we have no indication as to whether this unusual joining event involved segments oriented for direct (deletional) or inverted joining. Most, if not all, murine H chain V_H gene segments are oriented for deletional joining (1), whereas human and murine V_κ segments occur in both direct and inverted orientations (36). As at least one human V_H gene is known to be in the same orientation as J_H (14) and others are in the same orientation as their nearest neighbors (14, 22), it seems likely that the V_H and J_H or DJ_H elements involved in this rearrangement were originally in the same transcriptional orientation in the germline (e.g., oriented for direct normal joining). If so, sequences between V_H and J_H or DJ_H substrates would have been inverted by this unusual join. Whatever the orientation of the participating segments, it is notable that the product we isolated retains recombination signal sequences adjacent to J_H or DJ_H sequences, possibly permitting its use as a substrate in further rearrangement events. Thus, even if infrequent, such joining events could be selected by the immune system and represent yet another mechanism for the generation of diversity.

Summary

We have analyzed the phenotypic characteristics and IgH gene rearrangements in a panel of EBV-transformed B lineage cell lines from human fetal liver and bone marrow. Some lines contained only populations of immature, Ig^- B cells, while others contained mixed populations of mature and immature B cells. The majority of identifiable IgH rearrangements involved joining of the most J_H-proximal D segment, DQ52, to various J_H segments, implying that DQ52 is a preferred target for initial DJ_H rearrangements. Three other rearrangements involving V_H-related sequences were also characterized. Two involved V_HDJ_H joining using V_H3 genes, although one of these had a very unusual DJ_H structure. The third consisted of inverted 3' signal sequences and flanking regions of a V_H4 gene appended to a J_H.

The mechanisms by which the later rearrangement could have occurred and its potential physiological significance are discussed.

Received for publication 24 October 1988.

References

1. Tonegawa, S. 1983. Somatic generation of antibody diversity. *Nature (Lond.).* 302:575.
2. Yancopoulos, G. D., and F. W. Alt. 1986. Regulation of the assembly and expression of variable region genes. *Annu. Rev. Immunol.* 4:339.
3. Alt, F. W., T. K. Blackwell, and G. D. Yancopoulos. 1988. Development of the primary antibody repertoire. *Science (Wash. DC).* 238:1079.
4. Early, P., H. Huang, M. Davis, K. Calame, and L. Hood. 1980. An immunoglobulin heavy chain variable region gene is generated from three segments of DNA: V_H, D and J_H. *Cell.* 19:981.
5. Sakano, H., R. Maki, Y. Kurosawa, W. Roeder, and S. Tonegawa. 1980. Two types of somatic recombination are necessary for the generation of complete immunoglobulin heavy chain genes. *Nature (Lond.)* 286:676.
6. Reth, M. G., S. Jackson, and F. W. Alt. 1986. V_HDJ_H formation and DJ_H replacement during pre-B differentiation: non-random usage of gene segments. *EMBO (Eur. Mol. Biol. Organ.) J.* 5:2131.
7. Reth, M., P. Gehrmann, E. Petrac, and P. Wiese. 1986. A novel V_H to V_HDJ_H joining mechanism in heavy-chain-negative (null) pre-B cells results in heavy-chain production. *Nature (Lond.).* 322:840.
8. Kleinfield, R., R. R. Hardy, D. Tarlinton, J. Dangl, L. A. Herzenberg, and M. Weigert. 1986. Recombination between an expressed immunoglobulin heavy-chain gene and a germline variable gene segment in a $Ly1^+$ B-cell lymphoma. *Nature (Lond.).* 322:843.
9. Yancopoulos, G. D., S. V. Desiderio, M. Paskind, J. F. Kearney, D. Baltimore, and F. W. Alt. 1984. Preferential utilization of the most J_H proximal V_H gene segments in pre-B cell lines. *Nature (Lond.).* 311:727.
10. Perlmutter, R. M., J. F. Kearney, S. P. Chang, and L. E. Hood. 1985. Developmentally controlled expression of immunoglobulin V_H genes. *Science (Wash. DC).* 227:1597.
11. Kabat, E. A., K. G. Nickerson, J. Liao, L. Grossbard, E. F. Osserman, E. Glickman, L. Chess, J. B. Robbins, R. Schneerson, and Y. Yang. 1986. A human monoclonal macroglobulin with specificity for $\alpha(2\rightarrow8)$-linked poly-*n*-acetyl neuraminic acid, the capsular polysaccharide of group B meningococci and *Escherichia coli* K1, which crossreacts with polynucleotides and with denatured DNA. *J. Exp. Med.* 164:642.
12. Mittler, R. S., M. A. Talle, K. Carpenter, P. E. Rao, and G. Goldstein. 1983. Generation and characterization of monoclonal antibodies reactive with human B lymphocytes. *J. Immunol.* 131:1754.
13. Atkinson, P., B. Bennett, and R. L. Hunter. 1985. Direct measurement of antibody production in cell suspensions using an enzyme-linked immunosorbent assay. *J. Immunol. Methods.* 76:365.
14. Berman, J. E., S. J. Mellis, R. Pollock, C. L. Smith, H. Suh, B. Heinke, C. Kowal, U. Surti, L. Chess, C. R. Cantor, and F. W. Alt. 1988. Content and organization of the human Ig V_H locus: definition of three new V_H families and linkage to the Ig C_H locus. *EMBO (Eur. Mol. Biol. Organ.) J.* 7:727.
15. Ravetch, J. V., U. Siebenlist, S. Korsmeyer, T. Waldmann, and P. Leder. 1981. Structure of the human immunoglobulin μ locus: characterization of embryonic and rearranged J and D genes. *Cell.* 27:583.
16. Maxam, A., and W. Gilbert. 1980. Sequencing end-labeled DNA with base specific chem-

ical cleavages. *Methods Enzymol.* 65:499.

17. Siebenlist, U., J. V. Ravetch, S. Korsmeyer, T. Waldmann, and P. Leder. 1981. Human immunoglobulin D segments encoded in tandem multigenic families. *Nature (Lond.).* 294:631.
18. Zong, S. Q., S. Nakai, F. Matusda, K. H. Lee, and T. Honjo. 1988. Human immunoglobulin D segments: isolation of a new D segment and polymorphic deletion of the D_1 segment. *Immunol. Lett.* 17:329.
19. Bakshi, A., J. J. Wright, W. Graninger, M. Seto, J. Owens, J. Cossman, J. P. Jensen, P. Goldman, and S. J. Korsmeyer. 1987. Mechanism of the t (14;18) chromosomal translocation: structural analysis of both derivative 14 and 18 reciprocal partners. *Proc. Natl. Acad. Sci. USA.* 84:2396.
20. Ichihara, Y., H. Matsuoka, I. Tsuge, J.-I. Okada, S. Torii, H. Yasui, and Y. Kurosawa. 1988. Abnormalities in DNA rearrangements of immunoglobulin gene loci in precursor B cells derived from an X-linked agammaglobulinemia patient and a severe combined immunodeficiency patient. *Immunogenetics.* 27:330.
21. Ichihara, Y., M. Abe, H. Yasui, H. Matsuoku, and Y. Kurosawa. 1988. At least five D_H genes of human immunoglobulin heavy chains are encoded in nine kilobase DNA fragments. *Eur. J. Immunol.* 18:649.
22. Kodaira, M., T. Kinashi, I. Umemura, F. Matsuda, T. Noma, Y. Ono, and T. Honjo. 1986. Organization and evolution of variable region genes of the human immunoglobulin heavy chain. *J. Mol. Biol.* 190:529.
23. Lee, K. H., F. Matsuda, T. Kinashi, M. Kodaira, and T. Honjo. 1987. A novel family of variable region genes of the immunoglobulin heavy chain. *J. Mol. Biol.* 195:761.
24. Schroeder, H. W., J. L. Hillson, and R. M. Perlmutter. 1987. Early restriction of the human antibody repertoire. *Science (Wash. DC).* 238:791.
25. Bofill, M., G. Janossy, M. Janossa, G. H. Burford, G. J. Seymour, P. Wernet, and E. Keleman. 1985. Human B cell development. II. Subpopulations in the human fetus. *J. Immunol.* 134:1531.
26. Katamine, S., M. Otsu, K. Tada, S. Tsuchiya, T. Sato, N. Ishida, T. Honjo, and Y. Ono. 1984. Epstein-Barr virus transforms precursor B cells even before immunoglobulin gene rearrangements. *Nature (Lond.).* 309:369.
27. Ernberg, I., K. Falk, and M. Hansson. 1987. Progenitor and pre-B lymphocytes transformed by Epstein-Barr virus. *Int. J. Cancer.* 39:190.
28. Gregory, C. D., C. Kirchgens, C. F. Edwards, L. S. Young, M. Rowe, A. Forster, T. H. Rabbitts, and A. B. Rickinson. 1987. Epstein-Barr virus-transformed human precursor B cell lines: altered growth phenotype of lines with germline or rearranged but non-expressed heavy chain genes. *Eur. J. Immunol.* 17:1199.
29. Kubagawa, H., P. B. Burrows, C. E. Grossi, J. Mestecky, and M. D. Cooper. 1988. Precursor B cells transformed by Epstein-Barr virus undergo sterile plasma cell differentiation: J-chain expression without immunoglobulin. *Proc. Natl. Acad. Sci. USA.* 85:875.
30. Wood, C., and S. Tonegawa. 1980. Diversity and joining segments of mouse IgH genes are closely linked and in the same orientation. *Proc. Natl. Acad. Sci. USA.* 80:3030.
31. Matsuda, F., K. H. Lee, S. Nakai, T. Sato, M. Kodaira, S. Q. Zong, H. Ohno, S. Fukihara, and T. Honjo. 1988. Dispersed localization of D segments in the human immunoglobulin heavy chain locus. *EMBO (Eur. Mol. Biol. Organ.) J.* 7:1047.
32. Born, W., J. White, J. Kappler, and P. Marrack. 1988. Rearrangement of IgH genes in normal thymocyte development. *J. Immunol.* 140:3228.
33. Mitzutani, S., A. M. Ford, L. M. Wiedemann, L. C. Chan, A. J. W. Furley, M. F. Greaves, and H. V. Molgaard. 1986. Rearrangement of immunoglobulin heavy chain genes in human leukemic cells shows preferential utilization of the D segment (DQ52) nearest to the J region. *EMBO (Eur. Mol. Biol. Organ.) J.* 5:3467.

34. Yancopoulus, G. D., B. A. Malynn, and F. W. Alt. 1988. Developmentally regulated and strain-specific expression of murine V_H gene families. *J. Exp. Med.* 168:417.
35. Morzycka-Wroblewska, E., F. E. H. Lee, and S. V. Desiderio. 1988. An unusual type of immunoglobulin gene rearrangement results in replacement of recombinational signal sequences. *Science (Wash. DC).* 242:261.
36. Lorenz, W., B. Straubinger, and H. G. Zachau. 1987. Physical map of the human immunoglobulin K locus and its implications for the mechanisms of V_K-J_K rearrangement. *Nucleic Acids Res.* 15:9667.
37. Kabat, E. A., T. T. Wu, M. Reid-Miller, H. M. Perry, and K. Gottesman. 1987. Sequences of Proteins of Immunological Interest. National Institutes of Health, Bethesda, MD. Publication No. 80-2008.