An AICDA-independent mechanism of secondary VH gene rearrangement in preimmune human B cells Running title: Secondary rearrangement in human B cells

Nancy S. Longo[*], **Gabrielle J. Grundy**[†], **Jisoo Lee**[*], **Martin Gellert**[†], and **Peter E. Lipsky**[*]

[*] Autoimmunity Branch, National Institute of Arthritis and Musculoskeletal and Skin Diseases, National Institutes of Health, Bethesda, MD 20892

[†] Laboratory of Molecular Biology, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, MD 20892

Abstract

VH replacement is a form of IgH chain receptor editing that is believed to be mediated by recombinase cleavage at cryptic recombination signals (cRSS) embedded in IGHV genes. Whereas there are several reports of IGHV replacement in primary and transformed human B cells and murine models, it remains unclear whether IGHV replacement contributes to the normal human B cell repertoire. We identified VH → VHDJH compound rearrangements from fetal liver, fetal bone marrow and naive peripheral blood, all of which involved invading and recipient IGHV4 genes that contain a cryptic heptamer, 13 base pair (bp) spacer and nonamer in the 5' portion of framework region (FR) 3. Surprisingly, all pseudohybrid joins lacked molecular processing associated with typical VHDJH recombination or nonhomologous end joining. Although inefficient compared to a canonical RSS, the IGHV4 cRSS was a significantly better substrate for *in vitro* RAG-mediated cleavage than the IGHV3 cRSS. It has been suggested that activation induced cytidine deamination (AICDA) may contribute to VH replacement. However, we found similar secondary rearrangements utilizing IGHV4 genes in AICDA-deficient human B cells. The data suggest that IGHV4 replacement in preimmune human B cells is mediated by an AICDA-independent mechanism resulting from inefficient but selective RAG activity.

Keywords

Gene rearrangement; Recombinase (RAG1-2); Activation Induced Cytidine Deaminase (AICDA); Receptor Editing; VH replacement; Human B cell

Introduction

Antigen receptor genes are assembled during lymphoid development by VH(D)JH recombination to generate a diversified immunoglobulin (Ig) repertoire capable of recognizing any foreign antigen(1). During VH(D)JH recombination, BCRs specific for foreign and self antigens are generated and a number of mechanisms have evolved to cull the repertoire of autoreactivity. One mechanism involves secondary gene rearrangement. This mechanism for

changing the specificity of self-reactive receptors involves the products of recombinase activating genes (RAG-1 and RAG-2; RAG1/2) that lead to the replacement of the autoreactive BCR by a secondary Ig gene rearrangement(2,3). Receptor editing was initially thought to be a characteristic property of the L chain(3,4) since the organization of the H chain locus would prohibit secondary rearrangements. In this regard, as a result of the initial H chain rearrangement, extra diversity (D) segments are removed, thereby deleting elements with appropriately spaced heptamer/nonamer recombination signals (RSS) and making it unlikely that secondary rearrangements would occur. However, evidence now indicates that VH replacement can occur at the H chain loci because of the presence of cRSS embedded within IGHV gene segments that can be recognized by RAGs to initiate secondary gene rearrangements(5). Secondary rearrangements of IGHV genes have been reported in transformed murine(6–9) and human(10–13) B cells, murine pro-B cells(14), transgenic models with insertion of an autoreactive V(D)J rearrangement(5), murine strains(15–18), human antibodies(19), human tonsilar germinal center B cells(20), and B cells in rheumatoid arthritis synovium(21).

Although detection of VH replacement in experimental models and *in vivo* during inflammation suggested that H-chain receptor editing occurred frequently, analysis of large VH databases from normal peripheral blood B cells has provided less evidence that this phenomenon occurs frequently in normal human B cells(22,23).

In humans, nearly all VH germline genes contain a 3' cRSS in which a heptamer but no nonamer with an appropriate spacer distance can be identified(24). Despite this, the apparent use of the IGHV1-69 gene and other VH genes with isolated heptamers for secondary rearrangements has been reported in rheumatoid arthritis synovial fluid(21). Other examples of secondary rearrangements using VH genes with minimal cRSS were identified in VH4 family transcripts of IgD$^+$ GC cells from tonsil(20). Moreover, a human B cell line has been shown to undergo secondary VH replacement between similar or different VH gene families, each containing an isolated 3' heptamer, and this process appears to be RAG-mediated(11). More recently, *in vitro* assays using a complete cRSS consisting of a heptamer, a 13 bp spacer and a nonamer identified from a germline IGHV4-34 gene appeared to permit RAG-mediated cleavage(25). Although this cRSS contained more of the essential elements (heptamer/spacer/nonamer sequence) necessary for recombination than those previously reported to be involved in VH secondary rearrangements, there is no 23 bp cRSS and as a result RAG mediated recombination would be expected to be inefficient. Moreover, the coding region of the RAG-mediated recombination that is maintained in the compound rearrangement would be a pseudohybrid join since the cRSS of each VH gene is in the same heptamer/nonamer orientation and the recombined product retains one of the cRSS sequences.

Because of the irregular features of the compound rearrangements, their putative abundance in somatically mutated B cells and the possibility of PCR errors contributing to their apparent identification in some cases, it has been have suggested that another mechanism for initiating double strand breaks, such as activation induced cytidine deaminase (AICDA), might be responsible for secondary VH gene rearrangements(26). AICDA, which is required for gene conversion and class switching(27) may, therefore, also contribute to VH replacement. The recent demonstration that AICDA is expressed during murine B cell ontogeny(28) supports a possible role for this enzyme in mediating VH replacement in developing B cells.

Since most receptor editing in the mouse occurs early during B cell development and the similarity of human VH genes can lead to ambiguity in identification of genes when mutations are present, we generated a database of human fetal and unmutated mature human B cells to explore the occurrence, frequency and possible mechanism of VH replacement in normal B cells. The data provide clear evidence of VH replacement between VH4 genes in developing

fetal B cells as well as in naive peripheral B cells in the adult. Even though examination of sequences suggested that secondary rearrangements might also be possible among members of other VH families since some contained a complete RSS containing a 13 bp spacer, only VH4 hybrids were identified. *In vitro* assays utilizing the IGHV4 and IGHV3 cRSS revealed less efficient RAG binding and cleavage when compared to a consensus RSS, yet site-specific cleavage products were evident in the VH4 substrates, suggesting that VH4 compound rearrangements could be RAG-mediated. Moreover, AICDA did not appear to be required since compound VH4 rearrangements were recovered from AICDA deficient B cells. In summary, secondary replacement of VH genes limited to VH4 family members was detected in fetal and naïve human B cells. The evidence suggests that these secondary rearrangements are likely to arise by AICDA-independent RAG-dependent pseudohybrid joining and may be an additional mechanism for generating repertoire diversity in normal B cells.

Materials and Methods

Cell preparation and sorting

Single cell preparations were made from peripheral blood of 4 normal donors and 7 X-linked hyper-IgM (X-HIgM) patients and three AICDA deficient patients as described previously (22,29,30) and were sorted into CD19+, CD19−, CD19+IgM+, CD19+IgD+CD27+ or CD19+IgD−CD27+ populations. Fetal liver and fetal bone marrow were obtained at 18 wk of gestation. All tissue collections and processing were done in accordance with policies established by the institutional review board for human experimentation at University of Texas Southwestern Medical Center and the National Institutes of Health. Fetal liver cells were mechanically disrupted into tissue fragments followed by filtration through nylon mesh. Fetal bone marrow cells were flushed from long bone specimens. Mononuclear cells were enriched by ficollhypaque density-gradient centrifugation as described(22). The cells were then stained with a phycoerythin-labeled anti-CD19 mAB (Becton Dickinson, San Jose, CA) and FITC labeled anti-IgM (Caltag, Burlingame, CA,), anti-CD27-PE and isotype controls (Becton Dickinson Pharmingen, San Diego, CA) using a FACStar^{Plus} flow cytometer (Becton Dickinson, San Jose, CA) outfitted with an automated single cell deposition unit. One cell was deposited into each well of a 96-well PCR plate assembled on a microAmp base as described previously(22). In some experiments, the CD19+ and CD19− B cells were diluted into aliquots of approximately 1 cell into 96 well microtiter plates. Normal donor cells were sorted into CD19+IgM+, CD19+IgD+CD27+ pre-switch or CD19+IgD−CD27+ post-switch populations. X-HIgM B cells were sorted into CD19+ or CD19− populations and AICDA-deficient B cells were CD19+ populations. The fetal B cells were sorted into CD19+IgM+ and CD19+IgM− populations.

Lysate preparation and PCR amplification

A single cell in 10μl of PCR buffer (10mM NaCl, 5mM Tris-HCl pH 8 at 25°C, 0.1% Triton X-100) containing 0.4mg/ml proteinase K (Sigma Chemical Co., St. Louis) was incubated for 1 hour at 55°C, and the enzyme was inactivated by heating at 95°C for 10 minutes. Primer extension linear preamplification employing random 15-mers and 60 rounds of amplification with Taq polymerase, Promega, Madison, WI) was performed to produce sufficient DNA for multiple subsequent DNA amplifications. Rearranged VHDJH genes were then amplified as described previously(22). Preferential amplification of IGHV4-59 was performed in CD19− cells with external forward 5'-ACATCTGTGGTTCTTCCTTCTC-3' and reverse 5'-ACGGAGGTTTTTGTCTGGGC-3' primers and nested forward 5'-TCACTGTGGGTCTCTCTGTTCA-3' and reverse 5'-TCCCCTCACTGTGTCTCTC-3' primers. The maximum PCR error rate for this method using linear genomic amplification followed by nested PCR is < 0.008%. This rate is based upon finding 2 mutations/25,986 bp in VH4 genes analyzed from non-B cells (30) and 6 mutations/75,000 bp in a non-Ig gene

analyzed from B cells(31). Thus, few, if any, of the nucleotide changes encountered in this analysis can be ascribed to PCR amplification errors.

Sequence analysis and database

PCR products were separated by electrophoresis on a 1.5% Seakem Agarose gel (FMC Bioproducts, Rockland, ME) and purified using GenElute agarose spin columns (Supelco, Bellefonte, PA) or Edge Biosystems (Gaithersburg, MD). Purified products were directly sequenced using the ABI Prism Dye Termination Cycle Sequencing kit (Perkin-Elmer, Foster City, CA) and analyzed with an automated Sequencer (ABI Prism 377; Applied Biosystems, Inc., Foster City, CA). For identification of the germline IGHV gene segments, JoinSolver (32) or the V BASE Sequence Directory was used in conjunction with the software programs Sequencher (Gene Codes Corp., Ann Arbor, MI) and DNASTAR (DNASTAR Inc., Madison, WI). Genbank accession numbers for the normal donor sequences are X87006-X87089, Z80363-Z80770 (excluding Z80606 and Z80679), EF542547-EF542687 and EF542688-EF542796; for X-HIgM are AF077410-AF077525, EF542103-EF542275, EF541488-EF542102 and EF542315-EF542546; and for AICDA-deficiency are EU237493-EU238970). A total of 53 sequences were obtained from fetal liver (AY582384-435) and 72 sequences were obtained from fetal bone marrow.

Identification of VH compound rearrangements

Initially, sequences were considered candidates for H chain secondary rearrangement if there was a 5' region with extensive nucleotide matching to a germline gene and a 3' region with mismatches. When the alignment analysis was repeated using only the 3' region of the VH, the finding of an alignment to another gene in the database without mismatches suggested the presence of a compound rearrangement in which one IGHV gene had invaded another. The accession numbers for all of the VH compound rearrangements are listed in Table I.

DNA and protein

DNA substrates were made by annealing PAGE purified DNA oligonucleotides. Top or bottom strand was labeled at the 3' using cordycepin 5' triphosphate (Perkin Elmer) and TdT (NEB) before annealing. Unincorporated label was removed using BioMicrospin 6 columns (Biorad). The sequences of cRSS within the human IGHV3-23 gene (CCGTGAAGGGCCGGTT<u>CACCATC</u>TCCAGAGACAATT<u>CCAAGAACT</u>CGCTGT) and IGHV4-59 gene (CCCTCAAGAGTCGAGT<u>CACCATA</u>TCAGTAGACACGT<u>CCAAGAACC</u>AGTTCT) and complementary strands were synthesized (predicted heptamer and nonamer sequences are underlined). The sequences of the concensus 12RSS substrate (DAR39 and DAR40), 23RSS substrate (DG61 and DG62) and a non-specific (NS) 50bp substrate with scrambled heptamer and nonamer sequences (DAR81 and DAR82) have been previously described(33).

Truncated HMGB1(34) and recombinant core RAG1 and core RAG2 proteins (co-expressed as fusions with maltose binding protein) were expressed and purified as described elsewhere (33).

DNA binding and cleavage assays

10 µl binding reactions containing 10 ng/µl RAG1/2, 2 nM labeled DNA, 100 nM nonspecific (NS) DNA and 2 ng/ml of high mobility group (HMGB1) were incubated for 10 minutes in a buffer comprising 25 mM MOPS-KOH pH7, 30 mM KCl, 1% glycerol, 0.1 mg/ml BSA, 2 mM $CaCl_2$, 4 mM DTT. Paired complexes were formed by adding 6.25 nM 23RSS to the reaction after 5 minutes. Complexes were loaded on a native 6% polyacrylamide gel in 20% glycerol(35).

Cleavage assays were catalyzed by the addition of 1 mM $MnCl_2$ or 4 mM $MgCl_2$ for 2 h at 37 °C using 40 ng/μl RAG with the same buffer, DNA and HMGB1 concentrations as above, and also included 100 nM NS DNA (DAR81/82). In these reactions the RAG1/2 concentration was 40 ng/μl. Products were separated on 12.5% TBE-urea polyacrylamide gel that was visualized and analyzed using a Molecular Dynamics Typhoon Storm 860 Phosphorimager with ImageQuant 5.1 software (GE Healthcare). An oligonucleotide matching the expected IGHV4-59 (or IGHV4-b) product when it is cleaved 5' of the putative heptamer was synthesized and used as a size marker. The percentage of binding and cleavage was calculated from the signal in the band of expected size divided by the total signal in each lane.

Statistics

The two-tailed Student t-TEST was used to determine differences between the binding and cleavage of NS DNA, IGVH3-23 and IGHV4-59 cRSS substrates and *p* values ≤0.05 were considered significant.

Results

Insertions, deletions, polymorphisms and DNA exchange in normal, mutated peripheral B cells

Because of the reports of VH replacement in tonsillar(20) and synovial B cells(21), we generated a database of 105 nonproductive and 643 productive naive, pre-switch and post-switch B cells from normal donor peripheral blood to identify the frequency of VH replacement. Several normal donor sequences appeared to be candidates for intrachromosomal VH replacement but the presence of mutation made it difficult to identify the VH genes that recombined accurately. Of the 748 sequences from adult naïve and memory B cells examined, we were unable to identify VH replacements such as those described in mutated tonsillar B cell clones(20) or synovial fluid B cells(21). However, irregularities in the variable segments similar to those reported by others(36) were found. Insertions and deletions consisting of one or several nucleotides that have been associated with somatic hypermutation occurred in 16/741 (2%) of the normal donor rearrangements (data not shown). The current data extends the previous reports by noting one unmutated sequence (Z80532) with a single nucleotide deletion. In addition, insertions and deletions, which were found in FRs as well as CDRs, were not predominantly triplets or multiples of triplets, as described in previous reports (36) and thus, were more frequent in nonproductive (8/105, 7.6%) than in productive (8/643, 1.2%) rearrangements.

Several IGHV4-59 sequences with unique CDR3s in the normal donors had 2 identical mismatches in FR3 which were initially presumed to represent a new allelic variant of IGHV4-59. As discussed below, these appeared to be compound rearrangements in which a 5' IGHV4-59 gene invaded a downstream IGHV4-b-DJH rearrangement.

Compound VH rearrangements are found in fetal and X-HIgM sequences

It was apparent that sequences with evidence of VH gene replacement were uncommon in the normal adult B cell repertoire consisting of both naïve and memory B cells. Since most of the murine examples of VH replacement occurred in early progenitor B cells(14) and the similarity between human VH genes made gene identification ambiguous when the sequence was mutated, we created a database from isolated CD19[+]IgM[+] and CD19[+]IgM[−] B cells from fetal liver and fetal bone marrow and CD19[+] B cells from X-HIgM donors in which the B cells are intrinsically normal but unable to undergo SHM.

During the analysis of VHDJH arrangements from fetal B cells, we encountered several potential examples of VH replacement. One sequence from a total of 53 CD19[+]IgM[−] pro/pre

B cells from fetal liver and two different sequences from a total of 72 CD19⁺IgM⁺ immature B cells from fetal bone marrow consisted of a 5' region from one VH gene and the 3' region from a second VH gene (Fig. 1). Using these criteria, the frequency of VH gene replacement was approximately 3/26 (12%) of the VH4 repertoire. It is notable that all of these sequences consisted of secondary rearrangements of IGHV4 gene segments into VH4 rearrangements. One sequence from fetal liver CD19⁺/IgM⁻ pre/pro B cells consisted of a compound rearrangement containing IGHV4-04 and an upstream IGHV4-39 (Fig. 1A). Two compound rearrangements from fetal bone marrow CD19⁺/IgM⁺ immature B cells contained potential pseudohybrid joins between IGHV4-b and IGHV4-59, and consisted of an initial IGHV4-b-DJH rearrangement in which the portion of the sequence upstream of the cRSS was replaced by IGHV4-59 (Fig. 1B). Each of these rearrangements was unique as the CDR3s were distinct.

When the gene locus relationship was analyzed between the VH genes forming the compound rearrangements, the VH gene introduced into the initial VHDJH rearrangement was found to be 5' of the VH gene involved in initial VHDJH recombination in each. IGHV4-39 is located about 400kb upstream of IGHV4-b. In each of these VH4 compound rearrangements, we were able to identify cRSS sequences in the region where the pseudohybrid junction may have potentially occurred. The cRSS sequences consisted of a heptamer (5'-CACCATA-3'), 13 bp spacer, and a nonamer (5'-CCAAGAACC-3') which was embedded in the conserved FR3 region of the VH4 gene family between ImMunoGeneTics (IMGT) codon 76 and IMGT codon 86. We were unable to identify RSS sequences containing a spacer of between 20–25 bps.

Having found compound rearrangements in fetal VH4 sequences, we expanded the search for these rearrangements to B cells from other sources. B cells from X-HIgM patients provide another source of unmutated B cells. When all X-HIgM VH4 rearrangements were examined, 28/327 sequences consisted of IGHV4-59/IGHV4-b and one X-HIgM sequence was an IGHV4-59/IGHV4-34 compound rearrangement (Table II). These compound rearrangements represented 29/327 (8.9%) of the X-HIgM VH4 repertoire.

The IGHV4-59/IGHV4-b compound rearrangement is not an allelic variant of IGHV4-59

The IGHV4-59/IGHV4-b compound rearrangements differ from the IGHV4-59*01 germline sequence by only two nucleotide mismatches in FR3. To exclude the possibility that this was not an allelic variant of IGHV4-59, the VH4 locus was amplified from CD19⁻ non-B cells isolated from the same X-HIgM donor in which the IGHV4-59/IGHV4-b compound rearrangements had been identified in CD19⁺ B cells. A nested PCR approach was chosen to amplify genomic IGHV4-59 from individual cells preferentially. Out of 96 PCR products, 71 matched the IGHV4-59 germline sequence and none of these sequences contained a T > C and G > A substitution in ImMunoGentTics (IMGT) codons 96 and 97, respectively (Table III). The remaining amplifications yielded PCR products of other germline VH4 genes. The IGHV4-59 PCR products from CD19⁻ cells demonstrate that the IGHV4-59 sequence with the 2 single bp substitutions is not an allelic form of IGHV4-59 and indicated that the source of these two FR3 substitutions in IGHV4-59 sequences is from a secondary rearrangement in which IGHV4-59 invaded a primary VHDJH rearrangement utilizing the IGHV4-b gene.

Because we were able to confirm that IGHV4-59 sequences with a T > C and G > A substitution in IMGT codons 96 and 97, respectively, were compound rearrangements, we searched for this specific compound rearrangement in the normal donor database. In total, 16/199 (8.0%) VH4 rearrangements with unique CDR3s had 5' VH alignments to IGHV4-59*01 and 3' VH segment alignments to the IGHV4-b germline sequence. Of these 16 sequences, six had no mismatches other than the two substitutions that are characteristic of the IGHV4-59/4-b compound rearrangement from the unmutated repertoire (X-HIgM), 6 sequences had ≤ 5 mutations and four had 8 or 17 mutations. Therefore, these VH4 replacements can be found in

the normal naive repertoire and B cells expressing these compound rearrangements can undergo SHM.

Because of the absence of any molecular processing at the pseudohybrid join most compound rearrangements were productive. However, 7/56 compound rearrangements (12.5%) were nonproductive (Table IV). Two ND compound rearrangements and one from the X-HIgM database had a productive rearrangement but were nonfunctional because they utilized a D segment reading frame that contained a stop codon. Three other X-HIgM compound rearrangements were nonproductive as a consequence of exonuclease and terminal deoxynucleotidyl transferase (TdT) activities that generated an out of frame JH segment and/ or a stop codon in the DJH junction. The nonproductive compound rearrangements were unmutated. Moreover, since the pseudohybrid join in the compound rearrangement lacked any molecular processing, they appeared to be nonproductive as a result of VHDJH recombination.

Recognition and cleavage of cryptic sequences

The cRSS from IGHV4-59 has 4/7 of the canonical heptamer bases, 6/9 of the canonical nonamer bases and a 13bp spacer instead of one of 12bp. The IGHV3-23 cRSS that was employed as a control has a heptamer and nonamer that is identical to IGHV4-59 except for the last nucleotide in each but only 60% homology in the spacer and flanking sequences. To determine whether these RSSs could be cleaved by RAG1/2s, *in vitro* binding and cleavage assays were carried out. An electrophoretic mobility assay was used to determine the binding of recombinant RAG1/2 proteins to IGHV4-59 and IGHV3-23 cRSS DNA substrates, which also contained 16 bp of the IGHV4-59 or IGHV3-23 sequence 5' of the putative heptamer. The amount of RAG-cRSS complex observed in the native gel was compared to that obtained with an optimal 12bp RSS substrate and a nonspecific (NS) substrate of similar length (50bp) with no heptamer or nonamer sequences. Figure 2 demonstrates that the RAG1/2 protein recognized the IGHV4-59 sequence, albeit with a reduced affinity as compared to the control 12bp RSS substrate. Quantification of the complexes indicated that the IGHV4-59 cRSS bound RAG significantly (p=0.05) more effectively than the IGHV3-23 cRSS.

The ability of RAG proteins to catalyze a double-strand break at the cRSS was also assessed. Physiologically, Mg^{2+} is likely to be the relevant cation, as this ion restricts RAG cleavage to obeying the 12/23 rule by permitting double strand breaks to occur only in the presence of 12bp RSS and 23bp RSS in the RAG-DNA complex(37–39). When one RSS is present, the RAG proteins are only able to catalyze a nick 5' of the heptamer in the presence of Mg^{2+}. When $Mn^{2+}$is used instead of Mg^{2+}, double strand breaks can be detected in the absence of 23bp RSS. In the current study to monitor nicking in the presence of Mg^{2+}, the top strand of the substrate was labeled at the 3' end and the assay was carried out in the presence of an unlabeled 23bp RSS. No significant levels of nicking of the cryptic sequences was detected (p=.06) for both IGHV3-23 and IGHV4-59 cRSS when compared to the control 12bp RSS (Fig. 3B). When the more permissive Mn^{2+} ion was used, more appreciable amounts of nick product of the IGHV4-59 substrate were detected (Fig. 3A). Nicking of the IGHV4-59 substrate was found to be significantly greater than that of NS DNA (p=0.02), whereas nicking of the IGHV3-23 substrate was not (p=0.20). The efficiency of RAG cleavage of the IGHV4-59 substrate varied among replicate assays (17–50% of the control 12RSS). Importantly, however, nicking of the IGHV4-59 cRSS was consistently and significantly (p=.03) greater than that of the IGHV3-23 cRSS. We also found that the nicked cRSS was smaller than the expected size for cleavage at the heptamer border, agreeing with previous reports that nicking on certain cRSSs occurred 1–2 nucleotides (nt) into the cryptic heptamer(25).

Mn^{2+} also allows the formation of hairpins to occur in the absence of the 23bp RSS(37–39). Therefore, we examined the formation of breaks in the bottom strand as well. The control 12bp RSS substrate formed the hairpin product efficiently, whereas the IGHV3-23 and IGHV4-49

substrates yielded several minor products. Of note, the IGHV4-59 substrate resulted in a product of the expected size of a longer hairpin made at the abnormal nick position. No major product resembling a nicked bottom strand (16–18nt) was formed under these conditions, as reported in a previous *in vitro* study of cRSS cleavage(25). Poor hairpinning activity of these cRSSs in Mn^{2+} may have been influenced by inappropriate flanking nucleotides, as it is known that certain bases bordering the heptamer hinder the subsequent hairpin formation catalyzed by Mn^{2+}(40). Indeed, when the coding flank of the optimal 12RSS was replaced with the coding flank of cRSS IGHV3-23 or IGHV4-59 (ending TT or GT, both known to be non-optimal flanks), a reduction of hairpinning was observed, particularly with IGHV4-59 (Fig. 4). Thus, the influence of flanking dinucleotides decreased the capacity of RAG to form the IGHV4-59 hairpin product.

To investigate the action of RAG on cRSS where non-optimal flank effects were operative, we included a 23RSS partner in the presence of Mg^{+2}. However, only small amounts of nick product on the top strand were detected and no hairpin product was detected (Fig. 3B). Similarly, no alternative cleavage site on the bottom strand was detected.

Compound rearrangements are found in AICDA deficient B cell rearrangements

Studies identifying AICDA-targeted motifs (RGYW/WRCY; R=A/G, Y=C/T, W=A/T) near hybrid junctions(26) have led to the suggestion that secondary VH rearrangements may be dependent on AICDA activity. To examine this possibility, we initially determined the frequency of AICDA targeted motifs in the putative region of pseudohybrid joins that corresponds to the region between FR3 IMGT codons 69 and 96. After normalizing the data for the total number of nucleotides in the region in all VH3 versus VH4 genes, we found 23 and 61 RGYW and WRCY motifs, respectively, in the VH3 family compared to 51 and 83 RGYW and WRCY motifs in the VH4 family (data not shown). Thus, the likelihood of AICDA-mediated VH replacement overall may be as much as 1.6-fold greater in VH4 genes than VH3 genes. To determine the role of AICDA in the formation of compound rearrangements in human B cells, we analyzed 284 VH4 rearrangements that were directly amplified from genomic DNA isolated from individual CD19+ AICDA deficient B cells. Twenty-four of these sequences with unique CDR3s had a compound IGHV4-59/IGIGHV4-b rearrangement, providing evidence that these VH replacements were not dependent on AICDA.

Discussion

We identified VH replacement involving VH4 genes from human fetal liver, fetal bone marrow and peripheral blood of normal donors, X-HIgM and AICDA-deficient B cells. Using CD19⁻ non-B cells, we confirmed that the IGHV4-59/4-b pseudohybrid join, which was found in all populations analyzed except fetal liver and differs from germline IGHV4-59*01 by only two single nucleotide mismatches in FR3, was not an allelic variant of IGHV4-59. When these IGHV4-59 rearrangements were originally sequenced, X87091 was in fact submitted to GenBank as a new polymorphism of VH4-59/DP71which was later designated as IGHV4-59*08 by IMGT, the international ImMunoGeneTics information system(41). However, the current study clearly indicates that these rearrangements represent VH replacement. Therefore, the frequency of the VH replacements in the normal repertoire was originally underestimated.

IGHV4-59 and other VH4 genes contain a complete cRSS, which provides the potential for a RAG-mediated pseudohybrid join. In this regard, we found the VH4 cRSS was a substrate for *in vitro* RAG cleavage albeit a less effective one than the canonical RSS. At least part of this inefficiency could be accounted for by the VH4 flanking sequence. It is notable that even though the VH4 cRSS was an inefficient substrate it was cleaved significantly more effectively than a control VH3 cRSS. Sequence specific breaks were detected at the cRSS heptamer

suggesting that RAG could catalyze double strand breaks via the normal nick-hairpin mechanism. Since similar VH replacements were present in AICDA-deficient B cells, we concluded that VH4 replacement was independent of AICDA activity and was likely to result from RAG-dependent pseudohybrid joining or possibly an atypical cleavage and joining mechanism that did not depend on AICDA.

Previous reports suggested that RAG enzymes recognizing cRSS present in VH genes are capable of mediating secondary rearrangements in human B cells that had undergone SHM (20,21). In addition, ligation mediated-PCR products of VH cRSS sites have been detected in human immature bone marrow cells indicating that DNA double strand breaks occur at these sites(11,14). We extended these studies by confirming that VH secondary rearrangements in immature and unmutated mature human B cells can be detected by amplifying genomic DNA from single cells. This approach allowed us to eliminate the possible bias introduced in analyzing the rearranged Ig repertoire acquired by cloning from cDNA(22,42,43). Amplifying Ig genes from genomic DNA of single B cells has been shown to minimize unusual PCR-induced artifacts, such as the creation of chimeric molecules or nucleotide exchanges between the mutational variants, which are known to occur during amplification of cDNA from bulk populations of B cells(44). Importantly, the sequence error rate using this technique is very low (<0.8 x10^{-4}/base pair), largely because the PCR products are directly sequenced(30,31). This makes it possible to distinguish minor differences in sequences and permits a more comprehensive interpretation of the data. Fewer than 10% of the sequences analyzed from X-HIgM and AICDA deficient patients contained mutations, most of which contained a single nucleotide substitution, thereby making it possible to interpret the sequence differences in mature B cells accurately. Compared to the analysis of mature mutated B cells, this was a particular advantage for identifying VH replacement, since VH genes within the same family share more than 80% homology(45).

An important feature of the VH hybrids identified in the current study was that the replacement gene always resided in the heavy chain locus upstream of the VH gene used in the initial rearrangement. This finding supports the conclusion that these compound rearrangements resulted from secondary VH replacement. During the initial rearrangement, VH genes downstream of the primary rearrangement are deleted, leaving only upstream VH genes available for the secondary rearrangement (Fig. 5). Notably, this is different than the VH1 compound rearrangements reported in the rheumatoid arthritis synovium in which only one half of the rearrangements involved the replacement of a downstream gene by an upstream counterpart(21), suggesting that these sequences were either an artifact(26) or that the replacement gene came from the opposite chromosome.

We found VH replacement only among VH4 genes which contrasts with previous reports of RAG-mediated VH replacement involving VH3, VH4 and VH1 genes(20). This finding was of interest since VH3 genes contain a cRSS in FR3 that is very similar to the cRSS in VH4 genes. Importantly, the cRSS in both VH3 and VH4 genes is composed not only of a heptamer 5'-(CACCAT(A/C)-3' but also a 13 bp spacer and a nonamer 5'-CCAAGAAC(C/T). In a canonical RSS, the three nucleotides of the heptamer closest to the recombination site (5'CAC-) are invariant and sequence positions 5 and 6 of the nonamer (5'-AA) are critical for RAG-mediated recombination(46,47). The VH FR3 cRSS meets most of the minimal sequence requirements for a complete RSS, and therefore presents the possibility for RAG binding and cleavage and formation of a pseudohybrid join involving an upstream VH4 gene and a downstream VH4DJH in a secondary rearrangement(4). Each gene in the VH4 family conforms closely to the consensus heptamer and nonamer. The cRSS of VH3 gene family members differ from the VH4 family primarily in the spacer sequence and the first and last position of the nonamer. By comparison, the VH1 cRSS nonamer is even more divergent and lacks the conserved AA dinucleotide. Although the differences in VH4 and VH3 cRSS are apparently

minor, they result in a major loss of efficiency of RAG cleavage. The differential efficiency of VH4 and VH3 cRSS to permit RAG mediated cleavage may contribute to the differential use of these gene families for secondary VH rearrangement. Other considerations such as chromosomal organization (45) and the frequency of sterile transcription (48,49) could contribute to the increased prevalence of IGHV4-59 and IGHV4b genes in secondary rearrangements. Finally, it is known that IGHV4-59 is the only VH4 gene that is rearranged more frequently than expected from random chance(22) and this may contribute to the prevalence of compound rearrangements utilizing IGHV4-59. Other features of VH3 and VH4 genes may contribute to inefficient RAG-mediated cleavage. In this regard, we found that the flanking sequences affect the outcome of RAG-mediated cleavage, in that the hairpinning step was decreased and abnormal when the VH3-23 flank was present. It has been observed that coding sequences ending in TT, such as IGHV3-23, results in lower in vitro recombination frequencies (50). Thus, the presence of the VH3 coding sequence may contribute to the decreased appearance of VH3 containing compound rearrangements. More recently it has been shown that the spacer sequence also contributes to RAG-mediated recombination (51). Such sequence variation in the flank and spacer regions can account for similar RAG binding and cleavage at distinct cRSS but different levels of joining (13,51) and thereby contribute to different levels of VH3 and VH4 containing compound rearrangements. Importantly, neither VH3 nor VH4 family genes contain a cRSS containing a 23 bp spacer. Since the presence of a 23 bp RSS increases the specificity of double stranded breaks (52), the absence of a 23 bp RSS in the FR3 of VH4 and VH3 genes would be anticipated to alter the efficiency of RAG-mediated cleavage.

A discrepancy between our findings and those of others who demonstrated more robust *in vitro* RAG cleavage of the IGHV4-34 cRSS(25) may in part be related to the inclusion of excess non-specific DNA in our reactions, which was added to avoid non-specific cleavage during long incubations. In addition, we avoided alterations in buffer pH or ionic strength and the addition of organic solvents, such as dimethylsulfoxide, which can lead to altered or relaxed specificity of endonucleases in general (53) as well as RAG (54). Others have identified RAG-mediated breaks at CAC sequences under certain conditions (11). In the current study, a low level of semi-specific cleavage was also noted in the NS DNA controls. We concluded that RAG may have contributed to VH replacement but because of the low efficiency, even with permissive conditions, it is possible that another mechanism may be responsible for VH replacement.

Since many of the secondary VH replacements previously reported in human B cells were found in mutated populations, it was suggested that the process was dependent on AICDA or somatic hypermutation (26). Although there are several reports of AICDA expression during early murine B cell development (28,55–58), it is not known whether AICDA may have contributed to the compound rearrangements in human fetal liver and bone marrow. Importantly, the presence of compound rearrangements in AICDA deficient B cells argues against a mechanism that is dependent on AICDA, even though AICDA-targeted RGYW/ WRCY motifs are more frequent in VH4 than VH3 sequences in the region in which the pseudohybrid join could have occurred. The presence of compound VH4 rearrangements in human fetal liver and bone marrow cells also argues against a requirement for SHM. However, we cannot exclude the possibility that there may be more than one mechanism for VH replacement and the compound rearrangements from the mutated populations may have arisen after SHM was initiated.

The data clearly demonstrate that VH4 replacement can occur during early B cell development and in the absence of SHM. However, it remains unclear whether the secondary rearrangement occurs before (IGHV → IGHV) or after a primary rearrangement (IGHV → IGHVDJH) or if both are possible. The finding that compound rearrangements can be found in the

nonproductive repertoire implies that IGHV → IGHV fusions may occur between two VH4 genes before or during VHDJH rearrangement and, moreover, is unlikely to relate to B cell reactivity to self or exogenous antigens. The possibility that IGHV → IGHV fusion may occur before VHDJH rearrangement suggests that the event may be influenced by gene accessibility related to differential sterile transcription of VH genes(48,49) or chromatin remodeling(59–61). This explanation may provide a clue as to why the compound rearrangements preferentially involve VH4 genes.

Interestingly, all of the VH compound rearrangements in the current study exhibited no modification of the pseudohybrid join. In contrast, classic hybrid joins derived from artificial substrates have a considerable degree of end processing(62). The signal end is usually kept intact, whereas the coding end exhibits modifications, with loss or gain of nucleotides. The data suggest that modification enzymes do not have access to the VH4 locus when the compound rearrangement occurs or are less active when the compound rearrangement is initially generated. In this regard, both TdT and exonuclease activities are known to be developmentally regulated(63).

The frequency of compound rearrangements ranged from to 8.6% in X-HIgM to 12% of the VH4 rearrangements in fetal samples. Similarly, 8.3% of the AICDA deficient B cells and 9.4% of the normal donor B VH4 cells contained IGHV4-59/IGHV4-b compound rearrangements. These results indicate that this is not an uncommon event and may be involved in providing additional diversity.

In summary, the current data clearly demonstrate evidence of secondary rearrangement at the heavy chain locus involving VH4 genes in fetal, X-HIgM, AICDA deficient and normal mature human B cells. Our findings suggest that VH replacement in human B cells is likely to involve RAG activity, but does not depend on AICDA.

Acknowledgments

References

1. Tonegawa S. Somatic generation of antibody diversity. Nature 1983;302:575–581. [PubMed: 6300689]

2. Gay D, Saunders T, Camper S, Weigert M. Receptor editing: an approach by autoreactive B cells to escape tolerance. J Exp Med 1993;177:999–1008. [PubMed: 8459227]

3. Tiegs SL, Russell DM, Nemazee D. Receptor editing in self-reactive bone marrow B cells. J Exp Med 1993;177:1009–1020. [PubMed: 8459201]

4. Gellert M. DNA double-strand breaks and hairpins in V(D)J recombination. Sem Immunol 1994;6:125–130.

5. Chen C, Nagy Z, Prak EL, Weigert M. Immunoglobulin heavy chain gene replacement: a mechanism of receptor editing. Immunity 1995;3:747–755. [PubMed: 8777720]

6. Reth M, Gehrmann P, Petrac E, Wiese P. A novel VH to VHDJH joining mechanism in heavy-chain-negative (null) pre-B cells results in heavy-chain production. Nature 1986;322:840–842. [PubMed: 3092105]

7. Kleinfield RW, Weigert MG. Analysis of VH gene replacement events in a B cell lymphoma. J Immunol 1989;142:4475–4482. [PubMed: 2498430]

8. Covey LR, Ferrier P, Alt FW. VH to VHDJH rearrangement is mediated by the internal VH heptamer. Int Immunol 1990;2:579–583. [PubMed: 2085492]

9. Usuda S, Takemori T, Matsuoka M, Shirasawa T, Yoshida K, Mori A, Ishizaka K, Sakano H. Immunoglobulin V gene replacement is caused by the intramolecular DNA deletion mechanism. EMBO J 1992;11:611–618. [PubMed: 1311252]

10. Bende RJ, Aarts WM, Pals ST, van Noesel CJ. Immunoglobulin diversification in B cell malignancies: internal splicing of heavy chain variable region as a by-product of somatic hypermutation. Leukemia 2002;16:636–644. [PubMed: 11960344]

11. Zhang Z, Zemlin M, Wang YH, Munfus D, Huye LE, Findley HW, Bridges SL, Roth DB, Burrows PD, Cooper MD. Contribution of VH gene replacement to the primary B cell repertoire. Immunity 2003;19:21–31. [PubMed: 12871636]

12. Lenze D, Greiner A, Knorr C, Anagnostopoulos I, Stein H, Hummel M. Receptor revision of immunoglobulin heavy chain genes in human MALT lymphomas. Mol Pathol 2003;56:249–255. [PubMed: 14514917]

13. Zhang M, Swanson PC. V(D)J recombinase binding and cleavage of cryptic recombination signal sequences identified from lymphoid malignancies. J Biol Chem 2008;283:6717–6727. [PubMed: 18187418]

14. Davila M, Liu F, Cowell LG, Lieberman AE, Heikamp E, Patel A, Kelsoe G. Multiple, conserved cryptic recombination signals in VH gene segments: detection of cleavage products only in pro B cells. J Exp Med 2007;204:3195–3208. [PubMed: 18056287]

15. Cascalho M, Ma A, Lee S, Masat L, Wabl M. A quasi-monoclonal mouse. Science 1996;272:1649–1652. [PubMed: 8658139]

16. Koralov SB, Novobrantseva TI, Hochedlinger K, Jaenisch R, Rajewsky K. Direct in vivo VH to JH rearrangement violating the 12/23 rule. J Exp Med 2005;201:341–348. [PubMed: 15699070]

17. Kolar GR, Capra JD. Immunoglobulin heavy-chain receptor editing is observed in the NOD/SCID model of human B-cell development. Scand J Immunol 2004;60:108–111. [PubMed: 15238079]

18. Lutz J, Muller W, Jack HM. VH replacement rescues progenitor B cells with two nonproductive VDJ alleles. J Immunol 2006;177:7007–7014. [PubMed: 17082616]

19. Xu C, Sui J, Tao H, Zhu Q, Marasco WA. Human anti-CXCR4 antibodies undergo VH replacement, exhibit functional V-region sulfation, and define CXCR4 antigenic heterogeneity. J Immunol 2007;179:2408–2418. [PubMed: 17675502]

20. Wilson PC, Wilson K, Liu YJ, Banchereau J, Pascual V, Capra JD. Receptor revision of immunoglobulin heavy chain variable region genes in normal human B lymphocytes. J Exp Med 2000;191:1881–1894. [PubMed: 10839804]

21. Itoh K, Meffre E, Albesiano E, Farber A, Dines D, Stein P, Asnis SE, Furie RA, Jain RI, Chiorazzi N. Immunoglobulin heavy chain variable region gene replacement As a mechanism for receptor revision in rheumatoid arthritis synovial tissue B lymphocytes. J Exp Med 2000;192:1151–1164. [PubMed: 11034605]

22. Brezinschek HP, Brezinschek RI, Lipsky PE. Analysis of the heavy chain repertoire of human peripheral B cells using single-cell polymerase chain reaction. J Immunol 1995;155:190–202. [PubMed: 7602095]

23. Brezinschek HP, Foster SJ, Brezinschek RI, Dörner T, Domiati-Saad R, Lipsky PE. Analysis of the human VH gene repertoire. Differential effects of selection and somatic hypermutation on human peripheral CD5(+)/IgM+ and CD5(-)/IgM+ B cells. J Clin Invest 1997;99:2488–2501. [PubMed: 9153293]

24. Radic MZ, Erikson J, Litwin S, Weigert M. B lymphocytes may escape tolerance by revising their antigen receptors. J Exp Med 1993;177:1165–1173. [PubMed: 8459210]

25. Rahman NS, Godderz LJ, Stray SJ, Capra JD, Rodgers KK. DNA cleavage of a cryptic recombination signal sequence by RAG1 and RAG2. Implications for partial V(H) gene replacement. J Biol Chem 2006;281:12370–12380. [PubMed: 16531612]

26. Darlow JM, Stott DI. Gene conversion in human rearranged immunoglobulin genes. Immunogen 2006;58:511–522.

27. Neuberger MS, Harris RS, Di Noia J, Petersen-Mahrt SK. Immunity through DNA deamination. Trends Biochem Sci 2003;28:305–312. [PubMed: 12826402]

28. Han JH, Akira S, Calame K, Beutler B, Selsing E, Imanishi-Kari T. Class switch recombination and somatic hypermutation in early mouse B cells are mediated by B cell and Toll-like receptors. Immunity 2007;27:64–75. [PubMed: 17658280]

29. Dörner T, Brezinschek HP, Foster SJ, Brezinschek RI, Farner NL, Lipsky PE. Comparable impact of mutational and selective influences in shaping the expressed repertoire of peripheral IgM+/CD5- and IgM+/CD5+ B cells. Eur J Immunol 1998;28:657–668. [PubMed: 9521076]

30. Longo NS, Satorius CL, Plebani A, Durandy A, Lipsky PE. Characterization of immunoglobulin gene somatic hypermutation in the absence of activation-induced cytidine deaminase. J Immunol 2008;181:1299–1306. [PubMed: 18606684]

31. Yavuz AS, Lipsky PE, Yavuz S, Metcalfe DD, Akin C. Evidence for the involvement of a hematopoietic progenitor cell in systemic mastocytosis from single-cell analysis of mutations in the c-kit gene. Blood 2002;100:661–665. [PubMed: 12091362]

32. Souto-Carneiro MM, Longo NS, Russ DE, Sun HW, Lipsky PE. Characterization of the human Ig heavy chain antigen binding complementarity determining region 3 using a newly developed software algorithm, JOINSOLVER. J Immunol 2004;172:6790–6802. [PubMed: 15153497]

33. McBlane JF, van Gent DC, Ramsden DA, Romeo C, Cuomo CA, Gellert M, Oettinger MA. Cleavage at a V(D)J recombination signal requires only RAG1 and RAG2 proteins and occurs in two steps. Cell 1995;83:387–395. [PubMed: 8521468]

34. Mo X, Bailin T, Noggle S, Sadofsky MJ. A highly ordered structure in V(D)J recombination cleavage complexes is facilitated by HMG1. Nuc Acids Res 2000;28:1228–1236.

35. Jones JM, Gellert M. Ordered assembly of the V(D)J synaptic complex ensures accurate recombination. EMBO J 2002;21:4162–4171. [PubMed: 12145216]

36. Wilson PC, de Bouteiller O, Liu YJ, Potter K, Banchereau J, Capra JD, Pascual V. Somatic hypermutation introduces insertions and deletions into immunoglobulin V genes. J Exp Med 1998;187:59–70. [PubMed: 9419211]

37. Eastman QM, Leu TM, Schatz DG. Initiation of V(D)J recombination in vitro obeying the 12/23 rule. Nature 1996;380:85–88. [PubMed: 8598914]

38. van Gent DC, Ramsden DA, Gellert M. The RAG1 and RAG2 proteins establish the 12/23 rule in V (D)J recombination. Cell 1996;85:107–113. [PubMed: 8620529]

39. West RB, Lieber MR. The RAG-HMG1 complex enforces the 12/23 rule of V(D)J recombination specifically at the double-hairpin formation step. Mol Cell Biol 1998;18:6408–6415. [PubMed: 9774656]

40. Grundy GJ, Hesse JE, Gellert M. Requirements for DNA hairpin formation by RAG1/2. Proc Natl Acad Sci USA 2007;104:3078–3083. [PubMed: 17307873]

41. Giudicelli V, Duroux P, Ginestoux C, Folch G, Jabado-Michaloud J, Chaume D, Lefranc MP. IMGT/ LIGM-DB, the IMGT comprehensive database of immunoglobulin and T cell receptor nucleotide sequences. Nucl Acids Res 2006;34:D781–784. [PubMed: 16381979]

42. Farner NL, Dörner T, Lipsky PE. Molecular mechanisms and selection influence the generation of the human V lambda J lambda repertoire. J Immunol 1999;162:2137–2145. [PubMed: 9973488]

43. Lee J, Monson NL, Lipsky PE. The V lambda J lambda repertoire in human fetal spleen: evidence for positive selection and extensive receptor editing. J Immunol 2000;165:6322–6333. [PubMed: 11086069]

44. Ford JE, McHeyzer-Williams MG, Lieber MR. Chimeric molecules created by gene amplification interfere with the analysis of somatic hypermutation of murine immunoglobulin genes. Gene 1994;142:279–283. [PubMed: 8194765]

45. Berman JE, Mellis SJ, Pollock R, Smith CL, Suh H, Heinke B, Kowal C, Surti U, Chess L, Cantor CR, et al. Content and organization of the human Ig VH locus: definition of three new VH families and linkage to the Ig CH locus. EMBO J 1988;7:727–738. [PubMed: 3396540]

46. Hesse JE, Lieber MR, Mizuuchi K, Gellert M. V(D)J recombination: a functional definition of the joining signals. Genes Dev 1989;3:1053–1061. [PubMed: 2777075]

47. Ramsden DA, Baetz K, Wu GE. Conservation of sequence in recombination signal sequence spacers. Nucl Acids Res 1994;22:1785–1796. [PubMed: 8208601]

48. Oltz EM, Osipovich O. Targeting V(D)J recombinase: putting a PHD to work. Immunity 2007;27:539–541. [PubMed: 17967406]

49. Cobb RM, Oestreich KJ, Osipovich OA, Oltz EM. Accessibility control of V(D)J recombination. Adv Immunol 2006;91:45–109. [PubMed: 16938538]

50. Gerstein RM, Lieber MR. Coding end sequence can markedly affect the initiation of V(D)J recombination. Genes Dev 1993;7:1459–1469. [PubMed: 8330743]

51. Lee AI, Fugmann SD, Cowell LG, Ptaszek LM, Kelsoe G, Schatz DG. A functional analysis of the spacer of V(D)J recombination signal sequences. PLoS Biol 2003;1:E1. [PubMed: 14551903]

52. Steen SB, Gomelsky L, Roth DB. The 12/23 rule is enforced at the cleavage step of V(D)J recombination in vivo. Genes Cells 1996;1:543–553. [PubMed: 9078384]

53. Nasri M, Thomas D. Relaxation of recognition sequence of specific endonuclease HindIII. Nucl Acids Res 1986;14:811–821. [PubMed: 3003698]

54. Bergeron S, Anderson DK, Swanson PC. RAG and HMGB1 proteins: purification and biochemical analysis of recombination signal complexes. Meth Enzymol 2006;408:511–528. [PubMed: 16793390]

55. Seagal J, Edry E, Naftali H, Melamed D. Generation and selection of an IgG-driven autoimmune repertoire during B-lymphopoiesis in Igmicrodeficient/lpr mice. Int Immunol 2004;16:905–913. [PubMed: 15148286]

56. Macpherson AJ, Lamarre A, McCoy K, Harriman GR, Odermatt B, Dougan G, Hengartner H, Zinkernagel RM. IgA production without mu or delta chain expression in developing B cells. Nat Immunol 2001;2:625–631. [PubMed: 11429547]

57. Ueda Y, Liao D, Yang K, Patel A, Kelsoe G. T-independent activation-induced cytidine deaminase expression, class-switch recombination, and antibody production by immature/transitional 1 B cells. J Immunol 2007;178:3593–3601. [PubMed: 17339456]

58. Seagal J, Edry E, Keren Z, Leider N, Benny O, Machluf M, Melamed D. A fail-safe mechanism for negative selection of isotype-switched B cell precursors is regulated by the Fas/FasL pathway. J Exp Med 2003;198:1609–1619. [PubMed: 14623914]

59. Bates JG, Cado D, Nolla H, Schlissel MS. Chromosomal position of a VH gene segment determines its activation and inactivation as a substrate for V(D)J recombination. J Exp Med 2007;204:3247–3256. [PubMed: 18056289]

60. Maes J, O'Neill LP, Cavelier P, Turner BM, Rougeon F, Goodhardt M. Chromatin remodeling at the Ig loci prior to V(D)J recombination. J Immunol 2001;167:866–874. [PubMed: 11441093]

61. Cherry SR, Baltimore D. Chromatin remodeling directly activates V(D)J recombination. Proc Natl Acad Sci USA 1999;96:10788–10793. [PubMed: 10485904]

62. Lewis SM. The mechanism of V(D)J joining: lessons from molecular, immunological, and comparative analyses. Adv Immunol 1994;56:27–150. [PubMed: 8073949]

63. Souto-Carneiro MM, Sims GP, Girschik H, Lee J, Lipsky PE. Developmental changes in the human heavy chain CDR3. J Immunol 2005;175:7425–7436. [PubMed: 16301650]

A

Fetal Liver

```
                                                            CDR1
VH4-39    GAGACCCTGTCCCTCACCTGCACTGTCTCTGGTGGCTCCATCAGCAGTAGTAGTTACTACTGGGGCTGGATCCGCCAGCCCCCAGGGAAGGGG
FL3-3.4.4G ---------------------------------------------------------------------------------------------
VH4-04    -G-------------------------------------------------***A---GG---A-T---G---------------------

                        CDR2                                            #CACAGTGNNNNNNNNNNNNNNNNACAAA
VH4-39    CTGGAGTGGATTGGGAGTATCTATTATAGTGGGAGCACCTACTACAACCCGTCCCTCAAGAGTCGAGTCACCATATCCGTAGACACGTCCAAG
FL3-3.4.4G -------------------------------------------------------------------------------A--------------
VH4-04    --------------GAA------C--------------------------------------------------------A--------------
                                                                                    △

                                                                                 CDR3
          AACC
VH4-39    AACCAGTTCTCCCTGAAGCTGAGCTCTGTGACCGCCGCAGACACGGCTGTGTATTACTGTGCGAGACA     Junction
FL3-3.4.4G -----------------------------------------G--------C------------------G  GATCGGCTGGGGCGTTCGG
VH4-04    -----------------------------------------G--------C------------------GA
```

```
Germline   ACTACTTTGACTACTGG JH4
FL3-3.4.4G        ------- JH4
```

B

Fetal Marrow

```
                                              CDR1
VH4-59     CTCACCTGCACTGTCTCTGGTGGCTCCATCAGTAGT******TACTACTGGAGCTGGATCCGGCAGCCCCCAGGGAAGGGACTGGAGTGGATT
FM1-2.4.2G  ---------------------------------------------------------------------------------------------
FM1-2.4.12A ---------------------------------------------------------------------------------------------
VH4-b      --------------------------G---------GGT***---------G-----------------------------G-----------

             CDR2                                          #CACAGTGNNNNNNNNNNNNNNNNNACAAAAACC
VH4-59     GGGTATATCTATTACAGTGGGAGCACCAACTACAACCCCTCCCTCAAGAGTCGAGTCACCATATCAGTAGACACGTCCAAGAACCAGTTCTCC
FM1-2.4.2G  ---------------------------------------------------------------------------------------------
FM1-2.4.12A ---------------------------------------------------------------------------------------------
VH4-b      ---AG-------C-T------------T----------G--------------------△--------------------------------------
```

```
                                                                              CDR3
VH4-59     CTGAAGCTGAGCTCTGTGACCGCTGCGGACACGGCCGTGTATTACTGTGCGAGAGA      N        D3-10
FM1-2.4.2G  -----------------------C--A-------------------------      CGGTGG    GGTTCGGGGAGTT
FM1-2.4.12A -----------------------C--A-------------------------      CAAAGACGG  GGGAGCTACTACTAC
VH4-b      -----------------------C--A-------------------------               D1-26
```

```
Germline    ACTACTTTGACTACTGG JH4
FM1-2.4.2G         ----------- JH4
FM1-2.4.12A        --------------- JH4
```

Figure 1. VH compound rearrangements from immature fetal and mature B cells

Three examples of compound rearrangements with pseudohybrid joins found in fetal VH rearrangements are shown. The white arrowheads mark the proposed junction between the two VH genes. At the junction of the two VH genes, cRSS (heptamer-13 bp spacer-nonamer) which matches the canonical RSS is underlined. The canonical RSS is shown (#) for comparison. Asterisks mark deletional differences between the two germline genes (IGHV4-04/IGHV4-39 and IGHV4-b/IGHV4-59). (A) A compound rearrangement (VH FL3-3.4.4G) isolated from 18 week fetal liver CD19[+]IgM[−] B cells. The hybrid molecule consists of IGHV4-39 upstream of the cRSS and IGHV4-04 downstream. (B) Two VH compound rearrangements (FM1-2.4.2G, FM1-2.4.12A) isolated from 18 week fetal bone marrow CD19[+]IgM[+] B cells. Both consist of IGHV4-59 upstream of the cRSS and IGHV4-b downstream.
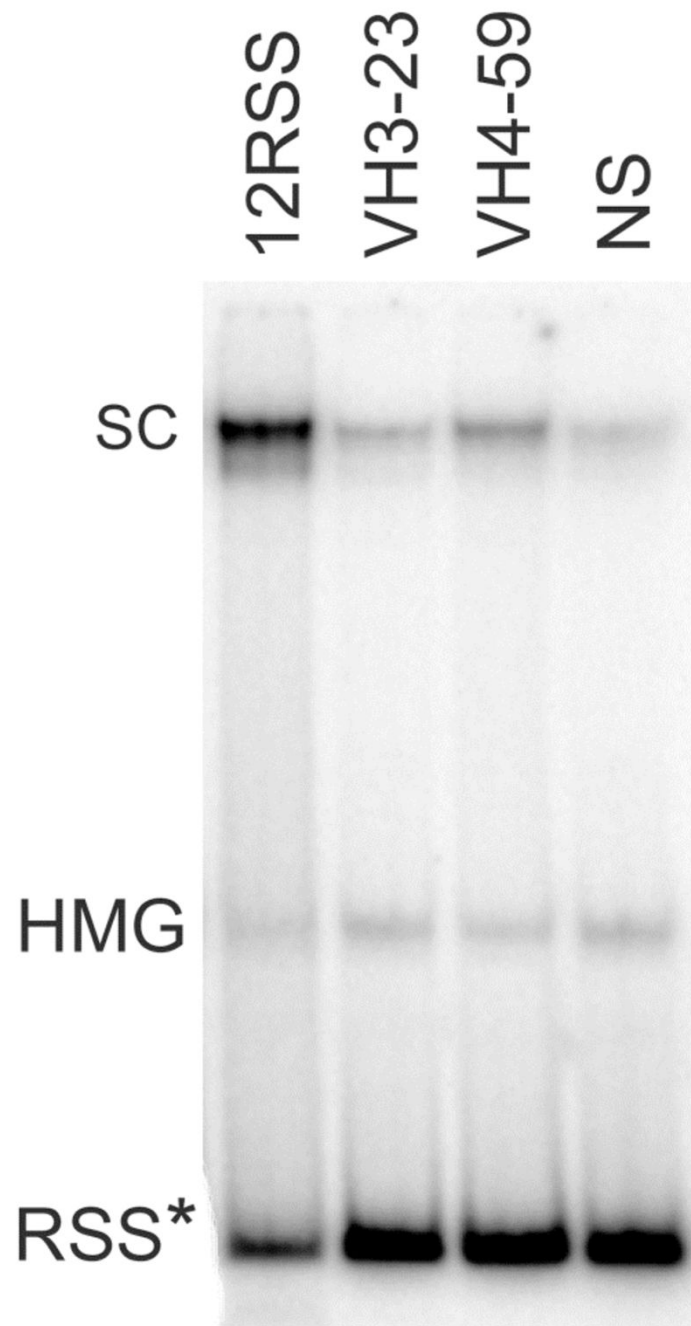
Figure 2. Electrophoretic Mobility Shift Assay of core RAG proteins binding to radiolabeled standard (12RSS), cRSS (from VH3-23 and VH4-59) and non-specific (NS) DNA substrates
Stable RAG-DNA single complexes (SC) migrated more slowly than DNA-HMGB1 (HMG) complexes and free DNA (RSS). The results shown are from one of 3 different binding assays from which the DNA in a single complex was quantified. The mean and standard error (SE) of the percentage of DNA found in a RAG complex was 5.7 ± 1.4 for NS DNA, 6.6 ± 1.4 for VH3-23, 10.3 ± 1.1 for VH4-59 and 40.5 ± 7.6 for 12RSS. Significant differences were found when comparing the SC involving IGHV4-59 (p=.03) versus NS DNA and IGHVH3-23 versus IGHV4-59 (p=.05). RSS* denotes the mobility of each radiolabeled unbound RSS.
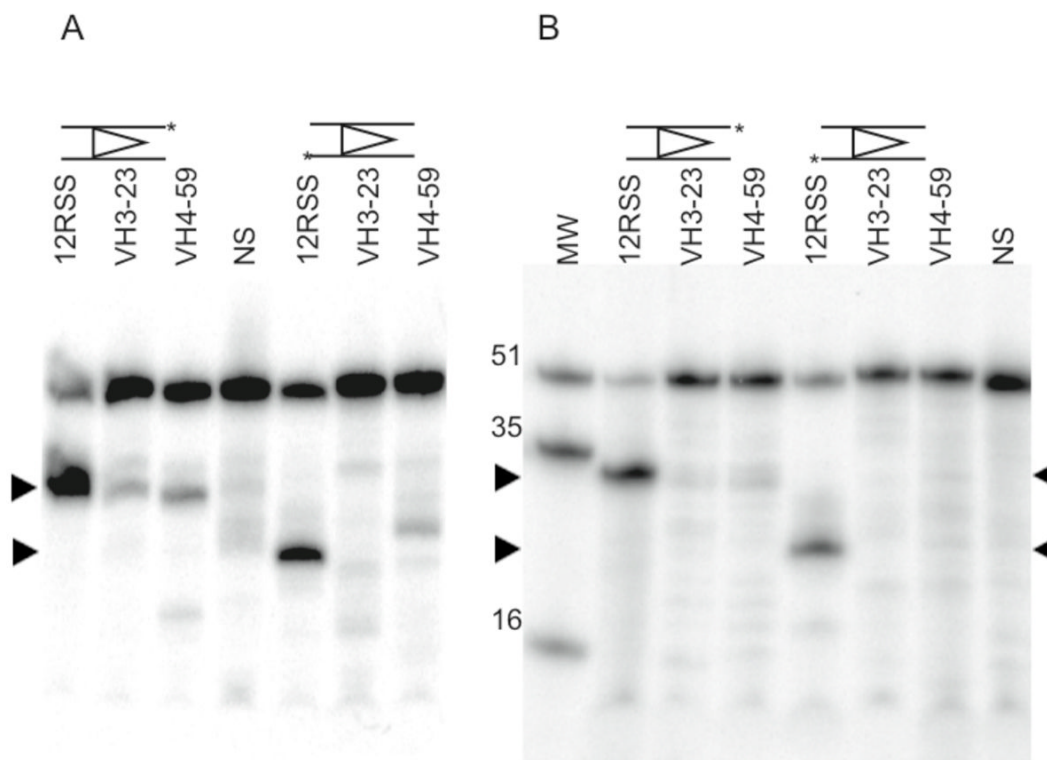
Figure 3. Cleavage of cRSS by RAG proteins
Standard, cryptic and nonspecific DNA substrates were 3' labeled (*) on top or bottom strands. Expected nick (35nt) and nick-nick (16nt) products were run in the MW lane to compare the actual nick sizes. Hairpin products were of various sizes. Reactions included (A) an excess of unlabeled NS DNA and Mn^{2+} and (B) 23RSS and Mg^{2+}. The arrow indicates the nick product. The results are representative of 7 nicking assays in the presence of Mn^{2+} and 4 experiments using Mg^{2+}. In the presence of Mg^{2+}, the mean and SE of the cleavage product expressed as the percentage of total DNA in the lane appearing in the band of expected size (34 bp) was 4.6 ± 1.8 for NS DNA, 6.2 ± 2.2 for VH3-23, 9.1 ± 3.2 for VH4-59 and 62.8 ± 15.7 for 12RSS. In the presence of Mn^{2+}, the mean ± SE of the cleavage product was 7.6 ± 1.9 for NS DNA, 9.4 ± 2.1 for VH3-23, 17.3 ± 4.6 for VH4-59 and 57.0 ± 7.3 for 12RSS. Significant differences were found in RAG-mediated cleavage of IGHV4-59 versus NS DNA (*p*=0.02) and IGHV3-23 versus IGHV4-59 (*p*=.03) in the presence of Mn^{2+}.
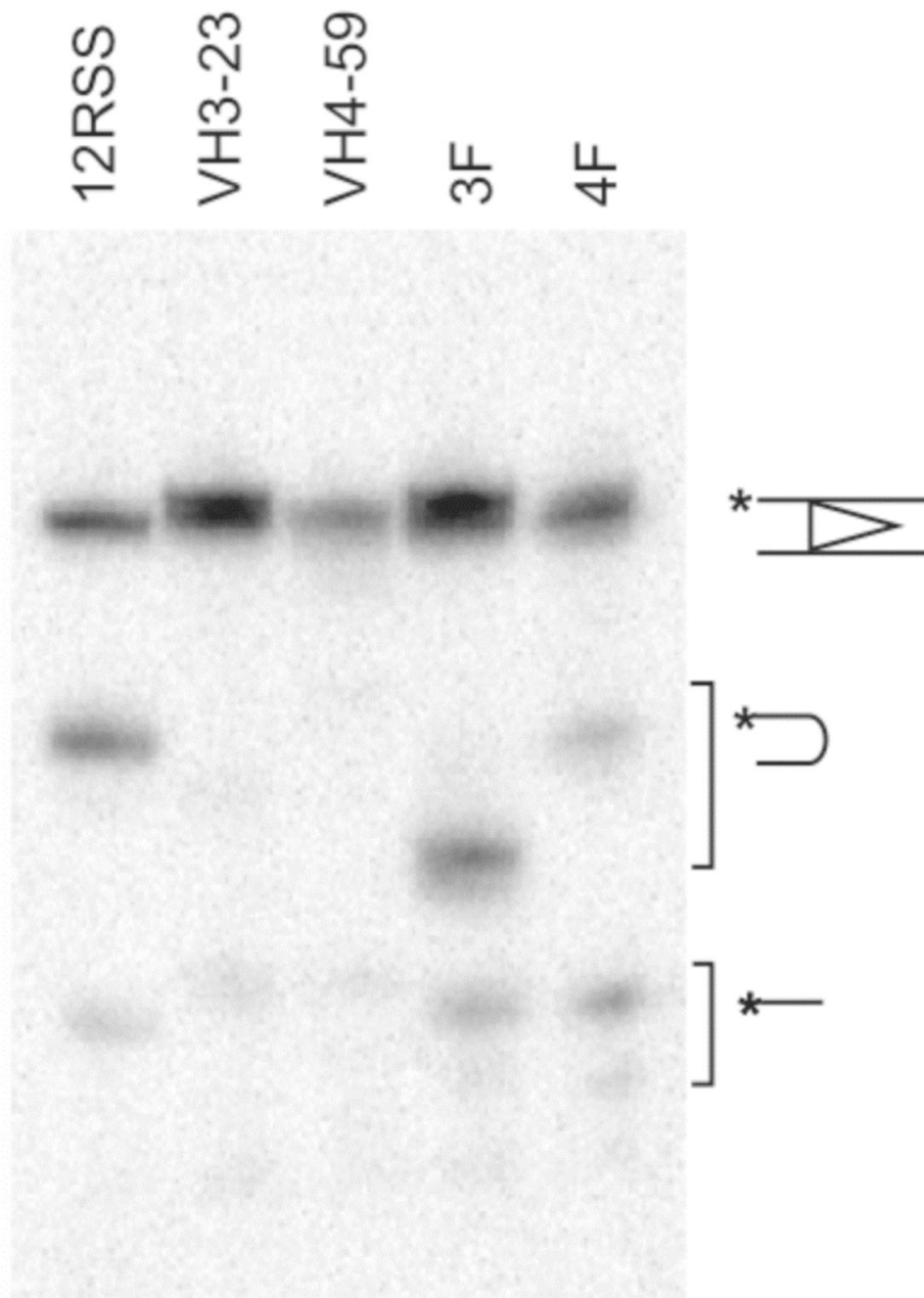
Figure 4. The coding flank reduces hairpin formation
A substrate was prepared in which the 12RSS flanking sequence was replaced with either the IGHV4-59 flank (4F) or the IGHV3-23 flank (3F) to determine the effect of flank residues on RAG cleavage in the presence of Mn^{+2}. A consensus 12RSS and the IGHV3-23 and IGHV4-59 substrates were 5' labeled (*) on the top strand with T4 polynucleotide kinase. A supportive flank sequence ends in YR. The IGHV4-59 flank ending in GT is considered an unsupportive flank (RY) and results in reduced hairpinning. Hairpinning on the IGHV3-23 flank which ends in TT (YY) was less severely inhibited but yielded an unusually short hairpin. The hairpin and nick products are indicated.
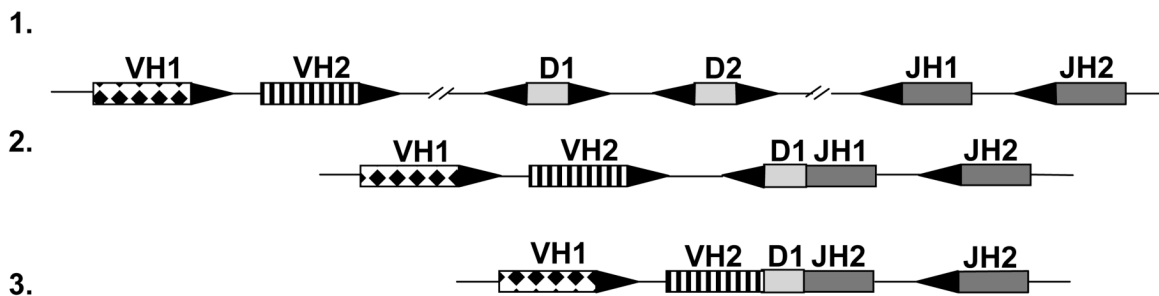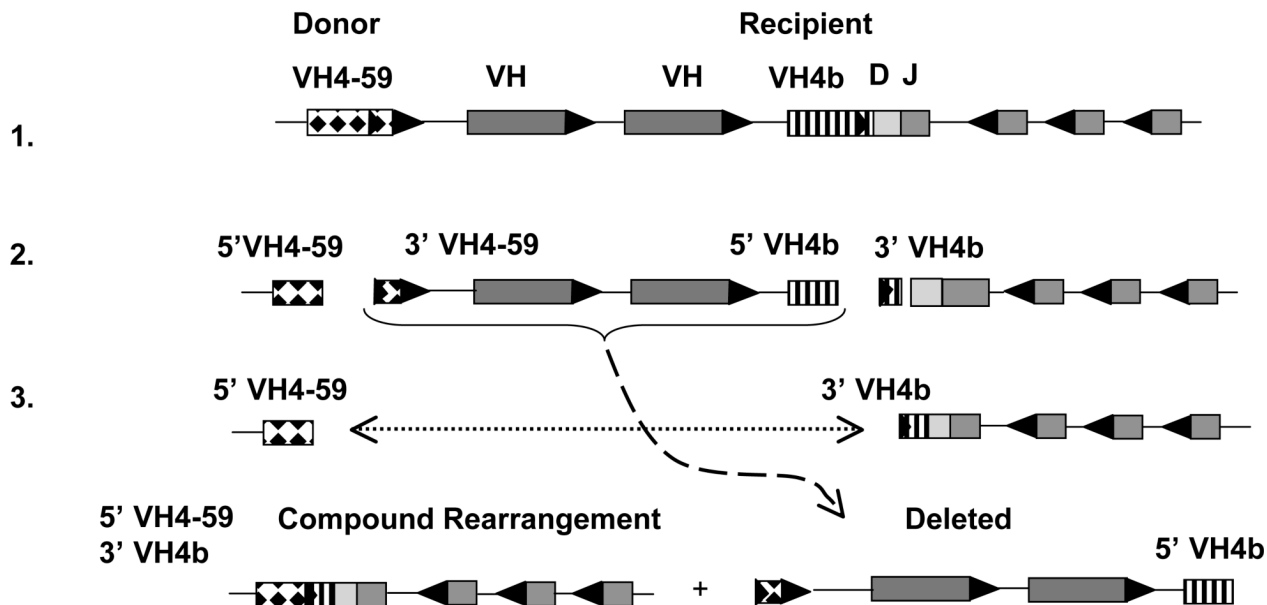
Figure 5. Schematic representation of a primary 12/23 VHDJH recombination and a secondary 12/12 rearrangement between two VH4 genes
A. The heavy chain locus on chromosome 14 is partially represented to demonstrate normal VHDJH recombination. (1) The 23 RSS (black triangles) flank VH and JH genes and the 12RSS (white triangle) flank D segments. Each RSS which is targeted and cleaved by RAG1/2 is oriented with the heptamer adjacent to either the 5' (D and JH) or 3' (D and VH) end of the gene. (2) The D1 gene is joined to JH1 and the intervening D2 segment is deleted. (3) The VH1 gene is joined to the DJH segment and the intervening VH2 gene is deleted. B. A region of the heavy chain locus containing a primary IGHV4-bDJH rearrangement is represented. (1) RAG enzymes bind to the cRSS embedded in a non-rearranged VH4-59 donor gene and the cRSS of the recipient IGHV4-b gene that has undergone prior rearrangement with a D and JH segment. (2) RAG cleaves at the heptamer of the cRSS in IGHV4-59 and IGHV4-b (noted with arrows) (3) After RAG-mediated double strand cleavage at the cRSS heptamers, the intervening gene segment is removed and the 5' portion of the donor gene (coding end) and the recipient VH4 gene containing the cRSS (signal end) are joined together forming a VH pseudohybrid join. Black triangles represent 23 bp RSS; white triangles represent 12bp RSS.

Table I

Accession numbers for VH compound rearrangements

Fetal	X-HIgM	AICDA-/-	Normal Donors
AY013306	AF077496	EU237512	EF542609
AY013307	AF077502	EU237526	EF542612
AY013308	AF077509	EU237977	EF542614
	AF077514	EU237990	EF542628
	AF077518	EU238008	EF542636
	EF541548	EU238013	EF542650
	EF541594	EU238030	EF542652
	EF541595	EU238206	EF542752
	EF541597	EU238230	EF542753
	EF541598	EU238238	EF542758
	EF541613	EU238589	Z80390
	EF541681	EU238597	Z80398
	EF541693	EU238603	Z80558
	EF541708	EU238634	Z80563
	EF541774	EU238650	Z80735
	EF541861	EU238660	Z80741
	EF541884	EU238661	X87083
	EF541983	EU238665	X87081
	EF541985	EU238673	X87077
	EF542024	EU238682	
	EF542038	EU238685	
	EF542104	EU238694	
	EF542108	EU238849	
	EF542122	EU238861	
	EF542172		
	EF542181		
	EF542195		
	EF542203		
	EF542325		

Table II

Human B cell VHDJH compound rearrangements with pseudohybrid joins

Number of Rearrangements	B cell Source	Invading Gene	Recipient Gene
1	Fetal liver pre/pro CD19+IgM-	VH4-39	VH4-4
2	Fetal marrow immature CD19+IgM+	VH4-59	VH4-b
28	X-HIgM mature CD19+	VH4-59	VH4-b
1	X-HIgM mature CD19+	VH4-59	VH4-34
24	AICDA deficient CD19+	VH4-59	VH4-b

A total of 53 fetal liver, 72 fetal bone marrow and 327 X-HIgM rearrangements were analyzed for pseudohybrid joins. An additional 284 AICDA deficient VH4 B cell rearrangements were analyzed specifically for VH4-59/4b pseudohybrid joins.

Table III

Frequency of VH4-59 products from CD19 CD19⁻ cells

Gene	CD19−	CD19+
IGHV4-b	0	0
IGHV4-04	1/96 (1%)	7/55 (13%)
IGHV4-34	24/96 (25%)	12/55 (22%)
IGHV4-39	0	12/55 (22%)
IGHV4-59	71/96 (74%)	12/55 (33%)
VH4 hybrids	0	7/69 (10%)

CD19− and CD19+ cells were isolated from PBMCs from a single donor. Nested PCR reactions were performed on both populations. The nested primers used on the CD19+ cells were designed to amplify all rearranged VH4 genes and the nested primers used on the CD19− cells preferentially amplified unrearranged VH4-59 gene segments in the heavy chain locus.

Table IV

The CDR3 sequence of nonproductive compound rearrangements

Source	Accession Number	Functionality	CDR3 DNA Sequence					
			codon 104	VH	N	D	N	JH
ND	EF542652	Stop codon in D segment	TGT	GCG AGA--	CAT GGA CGC CAC G	------GG TGG **TAA**----	TTC TG	---C TTT GAC TAC
ND	EF542614	Stop codon in D segment	TGT	GCG AGA--	CTT ACT C	---TA TTG**TAG**GAG TAT CAG CTG------	GTA AGT C	---AC TGG TTC GAC CCC
X-HIgM	EF541613	Stop codon in D segment	TGT	GCG AGA--	CGG GGT ACT	AGG ATA TTG**TAG** **TAG**TAC CAG CTGC-	GC	---A TAC GCT GCT TTT GAT ATC
X-HIgM	EF541884	Stop codon in DJH junction	TGT	GCG AGA--	CAA AGA GGG TGG AC	-T ACC AGC TGC T-------	TA **TGA** **TGA** GTG GGG	-TAC TAC TAC TAC TAC GGT ATG GAC GTC
X-HIgM	EF542038	Stop codon in DJH junction and JH is out of frame	TGT	GCG AGA--	CCC	GTA TTA CGA TTT TTG GAG TGG TTA------	GTC TAT**TG**	--**A** CTA CTA CTA CTA CGG TAT GGACGTC
X-HIgM	EF542108	JH is out of frame	TGT	GCG AC---	C CTG	GGG TAT AGC AGT GGC TGG---	CC	---------CTC
AICDA−/−	EU238603	Stop codon in VD junction and JH is out of frame	TGT	------	**GAC AC**	G TAT TAC TAT GGT TCG GGG AGT CAT TAT----	TCC ATT CG	ACTA CTA CTA CTA CAT GGA CGTC

All of the nonproductive compound rearrangements from each database are displayed. The reading frame for each CDR3 is derived from the primary IGHV4bDJH rearrangements. Stop codons are in bold face type, dashes (---) represent nucleotides excised from germline genes during recombination and N= nontemplated nucleotides inserted by terminal deoxynucleotidyl transferase (TdT).