NOVEL REARRANGEMENTS AT THE IMMUNOGLOBULIN D LOCUS

Inversions and Fusions Add to IgH Somatic Diversity

By KATHERYN D. MEEK, CHARLES A. HASEMANN,* AND J. DONALD CAPRA

*From the Department of Microbiology and *Program in Immunology, University of Texas Southwestern Medical Center, Dallas, Texas 75235*

Ig V regions are the precise molecular locations through which antibodies interact with immunogens. The ability to recognize an infinite array of antigens is a consequence of the chromosomal rearrangement of the various V region gene segments. At the Igh locus unique receptors are generated by several mechanisms: (*a*) selection of particular V_H, D, and J_H segments to form a complete V_H region; (*b*) junctional diversity generated at the joints of V_H-D-J_H; (*c*) the addition of nucleotides (N segments) at the junctions of rearranged gene segments (much of the third hypervariable region [HVR3][1] can be composed solely of novel amino acid sequences encoded by "junctional" sequences); (*d*) the somatic mutation of these gene segments; and (*e*) pairing of H chains with κ or λ chains to form heterodimeric molecules (reviewed in references 1-3).

The DNA elements involved in the rearrangement process are fairly well known (reviewed in reference 4). Immune receptor genes have conserved heptamer and nonamer sequences immediately adjacent to their coding sequences. The heptamer is separated from the nonamer by a nonconserved spacer of either 12 (+/− 1) or 23 (+/− 1) bp, i.e., approximately one or two turns of the helix, respectively. This highly conserved sequence motif is found flanking the rearranging elements in the immune receptor genes of all species that have been studied. These sequences are evidently the binding and/or recognition site of the enzyme or enzymes involved in the rearrangement process.

During rearrangement, two new DNA joints are formed (5). The coding sequences are joined to form a rearranged gene, and the recombination signal sequences join and form what has been termed the reciprocal joint. In the reciprocal joints (most well studied at the κ locus) the joining of the two signal sequences is, with rare exception, precise, such that the heptamer sequences from each gene segment are joined exactly, head to head, without addition or deletion of any nucleotides. In contrast, joining of the coding sequences appears to be more random in that the ligation of the two segments occurs at variable nucleotide positions. This sometimes results in

This work was supported by grants from the National Institutes of Health (GM-31689 and AI-12127) and the Robert Welch Foundation. Katheryn Meek is an Investigator of the Arthritis Foundation.

[1] *Abbreviations used in this paper:* HVR3, third hypervariable region; PCR, polymerase chain reaction; Tdt, terminal deoxynucleotidyltransferase.

deletions of significant portions of the coding regions. For example, it is not unusual for the first 12–15 nucleotides of J_H4 to be deleted, generating a shortened HVR3. Furthermore, apparently random additional nucleotides are frequently added at the joints, probably by the enzyme terminal deoxynucleotidyltransferase (Tdt). Thus, due to the imprecisions in these joints, much more somatic diversity is achieved with a particular V_H, D, and J_H than if joining were precise (1, 2, 6, 7).

The organization of the genes at the Ig loci is such that a recombination signal sequence with a 12-bp spacer always rearranges to a recombination signal sequence with a 23-bp spacer. For example, V region gene segments (the most 5′ gene segments) have signal sequences with 23-bp spacers immediately 3′ of their coding sequences. D segments are flanked by signal sequences with 12-bp spacers, and joining segments (the most 3′ gene segments) have signal sequences with 23-bp spacers 5′ of their coding sequences. Thus, V_H-D (23-12) and D-J_H (12-23) recombinations are allowed, while V_H-J_H (23-23) and D-D (12-12) are precluded by the 12-23 rule. This 12/23-bp recombination rule has been maintained at all seven of the immune receptor loci (2, 4).

At the Igh locus, the arrangement of the gene segments would appear to preclude joining of multiple D segments without breaking the 12/23 rule. However, at the time of the discovery of the Igh D locus, Kurosawa and Tonegawa (8) proposed that D-D fusion might occur in antibodies and could account for the additional diversity at the V_H-D and D-J_H joining sites not explained by germline V_H, D, or J_H sequences. They proposed that D-D fusion might be mediated via alternative signal sequences present in certain D segments. An additional heptamer located in the coding region of DSP2 type D segments has a single nucleotide difference from a concensus heptamer sequence, and is 24 bp away from either the 5′ or 3′ nonamer. While spacers are usually 23 bp in length, somewhat shorter or longer spacers have been reported in functional gene segments (for review, see reference 4). Thus, these cryptic signal sequences could indeed be functional. Moreover, recombination mediated by isolated heptamers is thought to be the mechanism by which both V_H region replacement and C_K deletion occurs (9–11), and there is no obvious reason why isolated heptamer recombination should not occur between other genes as well.

Several instances of potential D-D fusions have been reported in expressed antibodies (12–15), some of which are depicted in Fig. 1. So, for example, in antibody 6B1, the 5′ portion of the D segment seems to derive from DSP2.8 and the 3′ portion appears to derive from DSP2.2. Furthermore, D-D fusion has been proposed as the most likely mechanism for the generation of certain conserved N segments (16). Still, D-D fusion has never been definitively proven and is not considered a major source of diversity in Igs. A more accepted mechanism for the generation of N segment diversity is the addition of extra nucleotides by Tdt at the time of rearrangement (17).

We have performed several experiments to address the existence or extent of D-D fusion in the IgH complex. Using polymerase chain reactions (PCR) (18, 19), we have isolated a variety of incomplete Ig H chain rearrangements from murine bone marrow DNA involving the fusion of two D segments without joining to either J_H or V_H. We have demonstrated both direct and inverted rearrangements that have apparently been mediated either by the normal heptamer/nonamer signal sequences or by isolated heptamers. In addition, we present data showing that D segments may generate additional diversity by inverted recombination to J_H segments. This

theoretically allows single D segments to be used in either transcriptional orientation. Collectively, these data suggest new mechanisms for generating diversity in the HVR3 of Ig H chains, probably the most influential region of the antibody molecule.

Materials and Methods

Oligonucleotides. Oligonucleotides used in PCR amplifications are shown in Fig. 2. Oligonucleotides were synthesized on a DNA synthesizer (Applied Biosystems, Inc., Foster City, CA). Sequences (and the restriction endonuclease sites created at their 5′ ends) of the oligonucleotides depicted in Fig. 2 are as follows: 5S, 5′TCCTGCAGAATCTGTTACCTTACTTGGCAG3′ (PST I, primes all DSP2 D segments from the 5′ end); 5F, 5′GATCTAGAGACCATACTGGCTAGGGCTTTT3′ (XBA I, primes DFL16.1 from the 5′ end); 5Q, 5′GCTCTAGAGTCCCTGTGGTCTCTGACTGGT3′ (XBA I, primes DQ52 from the 5′ end); 3X, 5′AGCTGCAGAGTCTGCTGGGCATACTGGGTT3′ (PST I, primes all D segments except DQ52 from the 3′ end); 3Q, 5′CCTCTAGAGCTCCAAACAGAGGGTTTTTG3′ (XBA I, primes DQ52 from the 3′ end); 3JH1, 5′GTTCTAGAATGGAATGTGCAGAAAGAAAAA3′ (XBA I, primes from the 3′ end of JH1); 3JH2, 5′TGTCTAGAAGAGAGAGGTTTTAAGGACTCA3′ (XBA I, primes from the 3′ end of JH2); 3JH3, 5′CATCTAGAATGGGAGAAGTTAGGACTCACC3′ (XBA I, primes from the 3′ end of JH3); 3JH4, 5′GACCTGCAGAGGCCATTCTTACCTGAGGAG3′ (PST I, primes from the 3′ end of JH4); 5′ DFL16 and DSP2 screening, 5′GATTTTTGTCAAGGGATCTACTACTGTG3′ (hybridizes to the 5′ recombination sequence of all D segments except DQ52); 3′ DFL16 and DSP2 screening, 5′GGTTTTTGTTGCTGGATATATCACTGTG3′ (hybridizes to the 3′ recombination sequence of all D segments except DQ52); 5′ DQ52 screening, 5′GGTTTTGACTAAGCGGAGCACCACAGTG3′ (hybridizes to the 5′ recombination sequence of DQ52); 3′ DQ52 screening, 5′CACGGTGACGCGTGGCTCAACAAAAACC3′ (hybridizes to the 3′ recombination sequence of DQ52); JH screening, 5′TCACXGT$^{C}/_{T}$TC$^{T}/_{C}{}^{T}/_{G}$CAGGT, (hybridizes to all JH coding segments).

DNA Isolation. Bone marrow cells were obtained from the femur, tibia, and pelvis of BALB/c mice. Bone marrow and liver DNA was isolated as described previously (20).

PCRs. PCRs were done essentially via the method recommended by the supplier (Perkin-Elmer Corp., Norwalk, CT). Typically, 1 μg of DNA is added to a 200-μM solution of each dATP, dCTP, dGTP, and dTTP, with 100 pmol of each primer and 5 U of the Taq DNA polymerase. Extension time was optimized for each type of rearrangement analyzed. In most instances, PCR cycles were as follows: 3 min at 98°C (denaturation), 2 min at 58°C (annealing), and 3 min at 72°C (extension), controlled in a DNA thermal cycler (Perkin-Elmer Corp.). Certain amplifications worked better with either 5-min (any D to J2, J3, or J4) or 10-min extension times (any D to J1).

Typically, D-J_H rearrangements were detected easily after 40 amplification cycles. Although certain D-D rearrangements could be detected after 40 cycles, detection was simplified after 60 cycles of amplification.

Southern Hybridization Analyses. Amplified DNA could be loaded directly onto a 1.5% agarose gel for Southern filter hybridization analysis. After electrophoresis, DNA was transferred in 0.4 N NaOH onto a zeta probe membrane. Southern filter hybridization was done in 6× SSC, 0.5% SDS, 5× Denhardts. ^{32}P end-labeled oligonucleotides (as depicted in Fig. 2) were used as probes. Washing conditions are as described in each figure legend.

Isolation, Cloning, and Sequencing of Amplified DNA. Alternatively, the resulting DNA was cleaved with the appropriate restriction endonuclease (as shown in Fig. 2) and then loaded onto a 1.5% low-melt agarose gel for size purification. Typically, one fourth of a single DNA amplification was analyzed by Southern filter hybridization and the remainder was gel purified for cloning. In some experiments, gel-isolated DNA was amplified for an additional 60 cycles to facilitate cloning of the DNA. Gel-purified DNA was ligated into PTZ18U, which was then transformed into $CaCl_2$ competent BSJ72 bacteria. Recombinant colonies were identified by transfer to nitrocellulose and subsequently screened with the same ^{32}P end-labeled oligonucleotides used for Southern analysis.

Plasmid DNA for restriction analysis was isolated from each recombinant clone via alka-

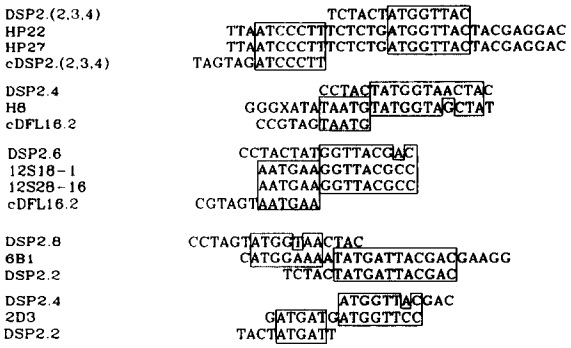

FIGURE 1. Nucleotide sequences of the D segments of several antibodies that were possibly derived via D-D fusion. The D segments have been defined as any nucleotide not obviously derived from V_H or J_H. Boxes indicate areas of homology between expressed D segments and germline D elements (*top line* and *bottom line*). The reverse complement of a particular sequence is indicated by a *c* before the name. Thus, antibodies HP22, HP27, H8, 12S18-1, and 12S28-16 possibly derived their D segments via inverted D-D fusions, while those from 2D3 and 6B would be direct D-D fusions. HP22 and HP27 are from reference 12 the other sequences are from reference 13.

line lysis. ss-DNA for sequencing was isolated from each clone by superinfection with M13K07 and selection with both ampicillin and kanamycin as described previously.

Sequencing was accomplished via dideoxy chain termination as described by Sanger (21), except using T7 DNA polymerase (Sequenase).

Results

Detection of Direct D-J_H Rearrangements. Rearrangement at the Igh locus in B lymphocytes is an ordered process, in that D-J_H rearrangement precedes V_H-D rearrangement (3, 22). It is thought that most rearrangements occur in the pre-B cell pool of the fetal liver or in the bone marrow of adults. We have designed a series of oligonucleotides complementary to the coding or noncoding strands of the mouse Ig D and J_H gene segments, as illustrated in Fig. 2. PCR amplifications utilizing

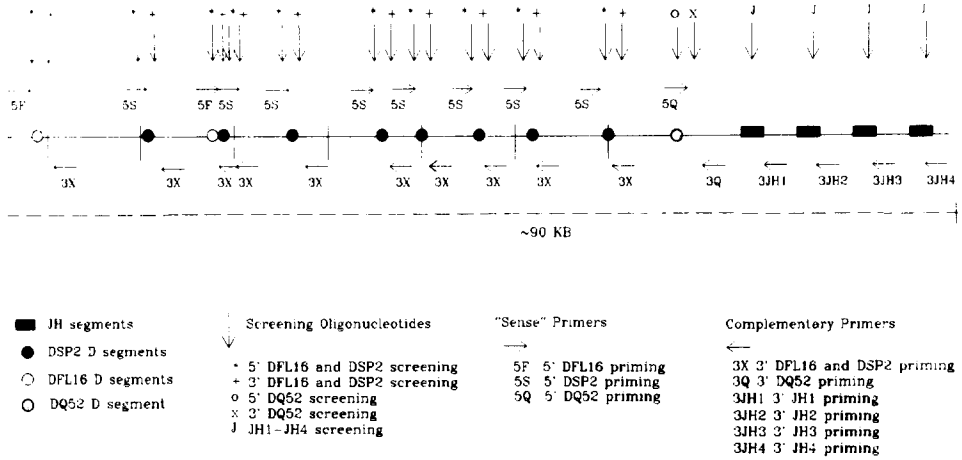

FIGURE 2. Genomic organization of IgH D and J_H regions (8, 31). Sequences complementary to the various oligonucleotide probes are indicated by vertical arrows. Sequences complementary to various oligonucleotides used in PCRs are indicated by horizontal arrows. Coding sequences are depicted as circles (D segments) or squares J_H segments).

oligonucleotide primers 5′ of the D segments and a second primer 3′ of the J_H segments should specifically amplify D-J_H rearrangements. D-J_H rearrangements can subsequently be detected by Southern filter hybridization analyses of the amplification products using oligonucleotide probes representing the D and/or J_H segment coding regions. In the experiment shown in Fig. 3, the 5S primer and each of the four (lanes *1–4*) 3′J_H primers were used in PCR amplifications of mouse bone marrow DNA. We had no difficulty demonstrating rearrangements involving all types of known mouse D elements (DSP2, DFL16, and DQ52) and all four mouse H chain J_H elements (J_H1–4) (data not shown).

Rearrangements of this type are easily detected in bone marrow DNA after 40 cycles of amplification. Surprisingly, these same rearrangements can also be detected in DNA from nonlymphoid tissues. The most straightforward explanation for this result is lymphocyte contamination of the tissue source since a single cell per million with a particular DNA rearrangement can be detected via this technique (19). This prompted the use of DNA from clonal cell lines as a nonrearranging tissue source. DNA from two different mouse fibroblast cell lines (3T3, C2C12) were used in similar experiments and again some D-J_H rearrangement is consistently detected, albeit at a significantly lower level. Fig. 4 demonstrates the extent of D-J_H rearrangement detected in bone marrow, liver, and fibroblast cell line DNAs. Fig. 4, lanes *1–3* represent amplification of germline DSP2 D segments (using 5S primer and 3X primer) from bone marrow, liver, and C2C12 DNA, and lanes *4–6* represent amplifications of D-J_H rearrangements (using 5S primer and 3JH4 primer) from the same DNA samples. Though the extent of D-J_H rearrangement detected in liver DNA (Fig. 4, lane *5*) and especially C2C12 DNA (lane *6*) is much lower than what is detected in bone marrow DNA, it is consistently present. With longer exposures the extent of D-J_H rearrangement is easily detected in liver and C2C12 (Fig. 4, lanes *7–9*). It is unlikely that these results are artifactual in that some D-J_H rearrangement can be detected in multiple different samples of fibroblast DNA, while PCRs with no DNA added are negative. Since the level of D-J_H rearrangement detected from bone marrow DNA is at least an order of magnitude higher than in the other DNA samples (though germline amplifications are similar [Fig. 4, lanes *1–3*]), it is reasonable to conclude that the PCR amplifications detect normal D-J_H rearrangement in B cells. However, it seems likely that a low level of D-J_H rearrangement occurs in nonlymphoid cells as well. The observation that the activity of the lymphoid recombinase system is not absolutely tissue specific is provocative and is being separately studied in our laboratory.

Detection of Inverted D-J_H Rearrangements Suggest that D Segments Can Be Used in Both Orientations. Using the 3X primer and 3′ J_H (3JH1, 3JH2, 3JH3, 3JH4) segment primers (see Fig. 2) in PCR amplifications, inverted D-J_H rearrangements were detected by Southern filter hybridization analysis with oligonucleotide probes complementary to either D or J_H as shown in Fig. 5 *a* and schematically in Fig. 5 *b*. So, in Fig. 5, lane *1*, inverted DSP to J_H1 rearrangements are detected, in lane *2*, inverted DSP to J_H2, etc. As with direct D-J_H joining these rearrangements can be detected involving a variety of different D and J_H gene segments. However, these rearrangements are considerably more difficult to detect, requiring several hours of exposure instead of the several minutes that is required to detect direct D-J_H rearrangement. This would suggest (though not prove) that inverted D-J_H rearrangements are not as common as direct D-J_H rearrangements.

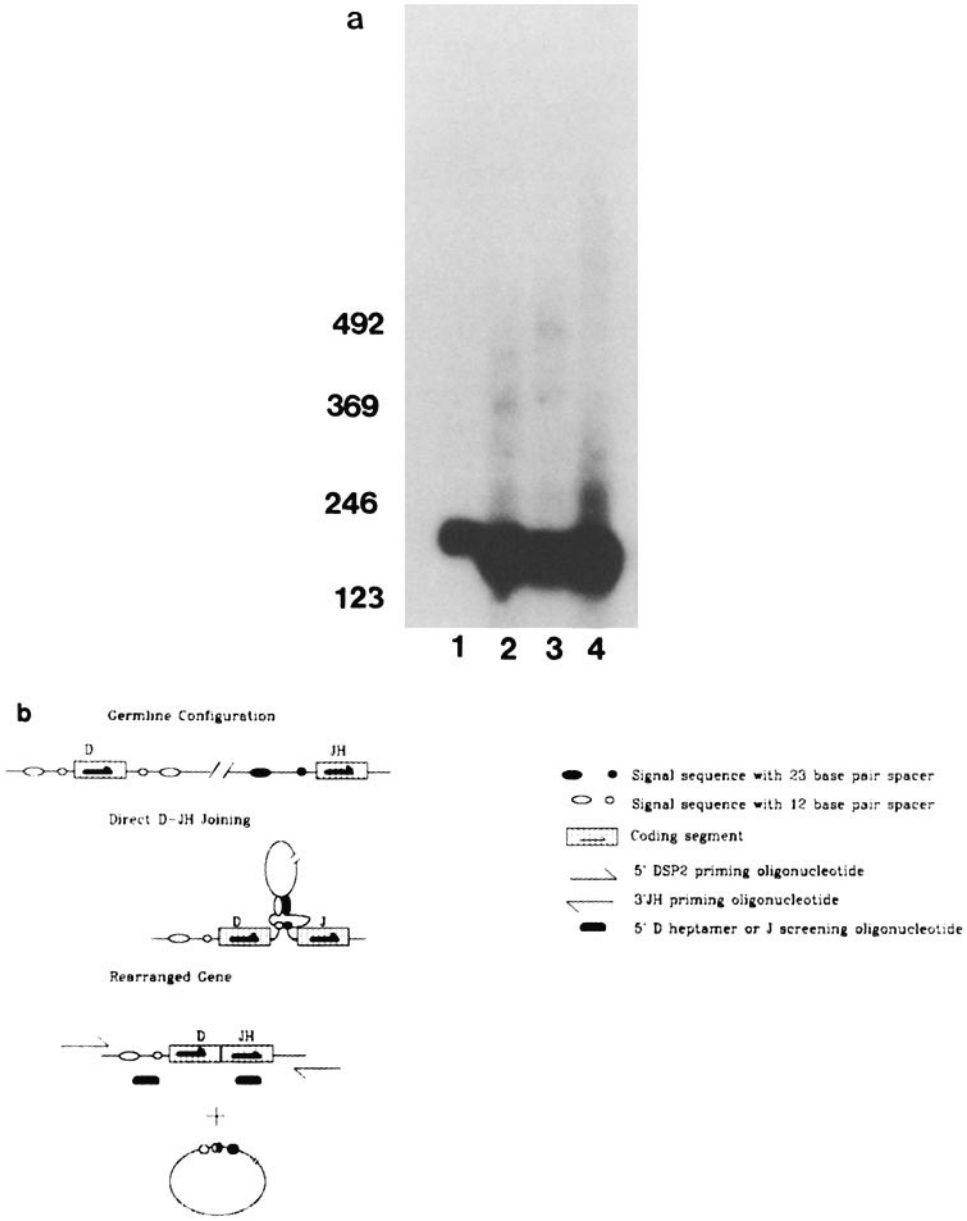

FIGURE 3. Analysis of direct D-J_H joining in bone marrow DNA. (*a*) Southern filter analysis of PCR products from bone marrow DNA using the following oligonucleotides as primers: lane *1*, 5S and 3JH1; lane *2*, 5S and 3JH2; lane *3*, 5S and 3JH3; and lane *4*, 5S and 3JH4. DNA was run on a 1.5% agarose gel, transferred to zeta probe, and hybridized to a ^{32}P-labeled oligonucleotide (J screening). Post washing conditions were 2× SSC, 0.5% SDS, at 45°C. The positions of size markers are shown in base pairs. (*b*) Schematic representation of rearrangements detected with this technique. Low levels of D-J_H1, D-J_H2, and D-J_H3 rearrangements can be detected when amplifying with the 3JH4 primer, since the J_H segments are so closely linked (as depicted in Fig. 2). The 3JH4 primer amplifies the entire (rearranged) J_H locus if a D segment is rearranged to any of the four J_H segments. Similarly, the 3JH3 primer amplifies D-J_H1 and D-J_H2 rearrangements, and the 3JH2 primer amplifies D-J_H1 rearrangements.

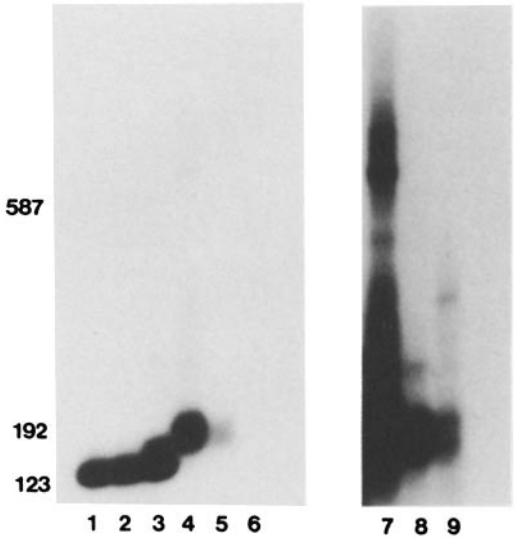

FIGURE 4. Analysis of direct D-J_H joining in bone marrow DNA, liver DNA, and DNA from the mouse fibroblast cell line C2C12. Southern filter analysis of PCR products using the following oligonucleotides as primers with DNA from various sources: lane *1*, 5S and 3X using bone marrow DNA; lane *2*, 5S and 3X using liver DNA; lane *3*, 5S and 3X using C2C12 DNA; lane *4*, 5S and 3JH4 using bone marrow DNA; lane *5*, 5S and 3JH4 using liver DNA; and lane *6*, 5S and 3JH4 using C2C12 DNA. Exposure was for 3 min. Lanes *7–9* are duplicates of lanes *4–6* but exposed five times as long (15 min). DNA was run on a 1.5% agarose gel, transferred to zeta probe, and hybridized to a 32^P-labeled oligonucleotide (5′ DSP2 and DFL16 screening). Post-washing conditions were 2× SSC, 0.5% SDS, at 55°C. The positions of size markers are shown in base pairs. Similar amounts of germline D segments have been amplified from all three tissue sources (lanes *1–3*), however, the degree of D-J_H rearrangement (lanes *4–6*, or lanes *7–9*) in the three tissue sources varies.

To further analyze inverted D-J_H rearrangement, one of these PCR products was cloned and sequenced, as shown in Fig. 6. Clone 4-1 was generated in a PCR amplification using the 3JH1 primer (3JH1) and the 3X primer as depicted in Fig. 2. This clone represents the inverted rearrangement of the 3′ noncoding sequence of a germline D element to the 5′ portion of J_H1. A 6-bp N segment (GTCCCT), typical of Tdt activity, can be found at the joint between D and J_H1. In this rearrangement a large portion of the coding sequence of the D segment was deleted, and it is not possible to tell which germline D segment has been utilized (DSP2.2, 4, 5, and 6 are possibilities). Even though nearly all of the D coding sequence was deleted during this rearrangement, a functional recombination sequence still remains at the appropriate end of this D-J_H1 rearrangement. Therefore, a functional antibody could still have been created by a subsequent V_H rearrangement to this D-J_H. In such an event, the resulting antibody would have a very short, nongermline D segment predominantly derived from N segment additions. It is also likely that inverted D-J_H rearrangements occur in which more of the D segment coding region is preserved. This is supported by the observation that the material detected in hybridization analyses is a smear, representing various sizes of inverted D-J_H rearrangements.

Demonstration of Direct D-D Fusion. Using the 5S and 3Q oligonucleotide primers for PCR amplification of bone marrow DNA, we have been able to detect by Southern hybridization analysis evidence of direct D-D joining, as shown in Fig. 7 *a* and schematically in Fig. 7 *b*. Fig. 7 *a*, lane *1* represents amplification of germline DSP2 D segments using the 5S and 3X primers compared in lane *2* with amplification of D-D rearrangements using the 5S primer and the 3Q primer (Fig. 2). The filter, in this case, has been probed with the 5′ DFL16 and DSP2 screening oligonucleotide. As predicted, the 3′ DFL16 and DSP2 screening oligonucleotide does not hy-

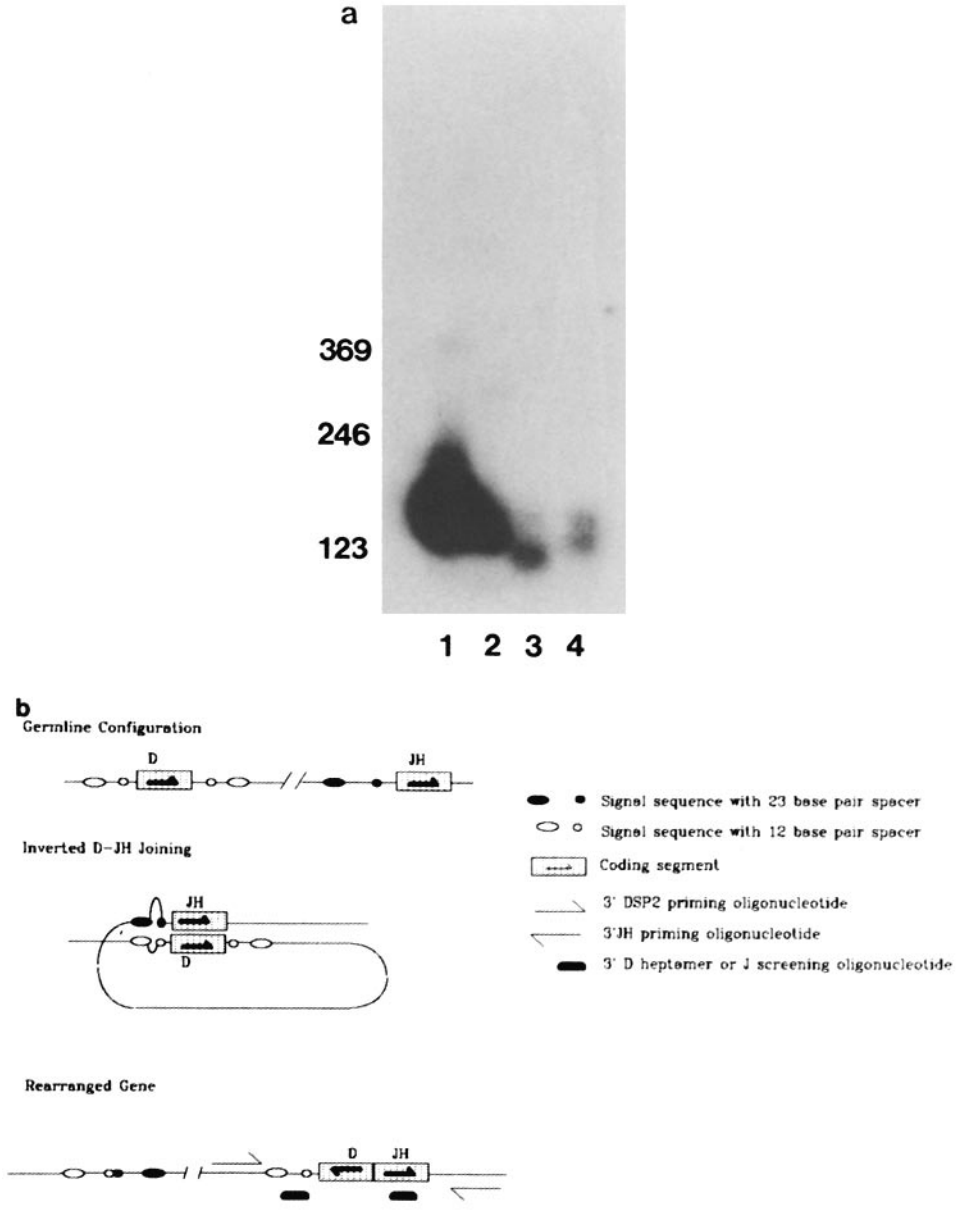

FIGURE 5. Analysis of indirect D-J_H joining in bone marrow DNA. (*a*) Southern filter analysis of PCR products from bone marrow DNA using the following oligonucleotides as primers: lane *1*, 3X and 3JH1; lane *2*, 3X and 3JH2; lane *3*, 3X and 3JH3; and lane *4*, 3X and 3JH4. DNA was run on a 1.5% agarose gel, transferred to zeta probe, and hybridized to a ^{32}P-labeled oligonucleotide (3' DSP2 and DFL16 screening). Post-washing conditions were 2× SSC, 0.5% SDS, at 55°C. The positions of size markers are shown in base pairs. (*b*) Schematic representation of rearrangements detected with this technique.

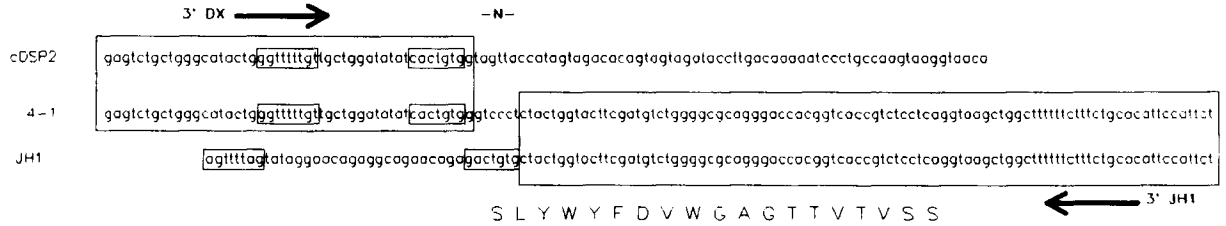

FIGURE 6. Nucleotide sequence of clone 4-1, an inverted D-J_H1 rearrangement. Solid boxes indicate nucleotide homology between 4-1 and germline J_H1 and DSP2 D segments (8, 31). Broken boxes indicate heptamer and nonamer sequences. Primers used in the PCR are as indicated. Extra nucleotides at the point of joining (N segments) are indicated with an N above the sequence.

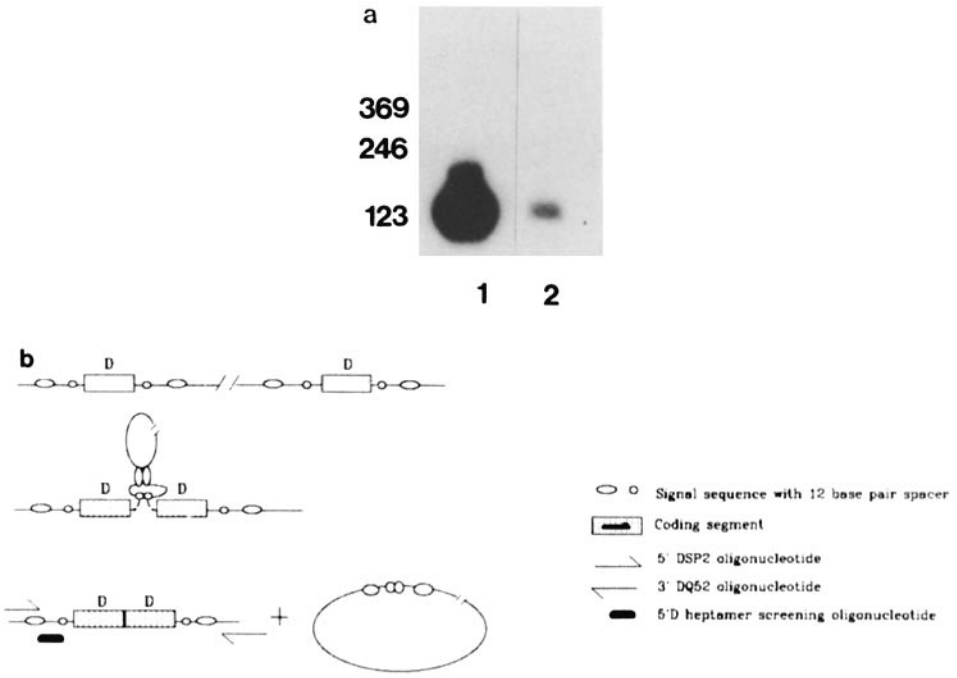

FIGURE 7. Analysis of direct D-D joining in bone marrow DNA. (*a*) Southern filter analysis of PCR products (60 amplification cycles) from bone marrow DNA using the following oligonucleotides as primers: lane *1*, 5S and 3X (amplifying germline DSP2 type D segments); and lane *2*, 5S and 3Q (amplifying direct D-D rearrangements). DNA was run on a 1.5% agarose gel, transferred to zeta probe, and hybridized to a ^{32}P-labeled oligonucleotide (5′ DSP2 and DFL16 screening). Post-washing conditions were 2× SSC, 0.5% SDS, at 55°C. The positions of size markers are shown in base pairs. (*b*) Schematic representation of rearrangements detected with this technique.

bridize to lane *2* (Fig. 7), whereas the 3′ DQ52 screening oligonucleotide does (data not shown). Thus, in multiple experiments we have been able to routinely detect and then characterize amplification products representing D-D fusions by hybridization before isolating clones.

Three clones were isolated from bone marrow DNA amplifications that represent direct D-D fusions, as shown in Fig. 8. Clones 1-3, 1-6, and 1-8 represent rearrangements of the 5′ portion of DSP2.8, DSP2.3, and DSP2.7, respectively, to the 3′ portion of DQ52. The 5′ recombination signal sequence of each DSP2 D segment has been retained, along with the majority of the coding regions (14–17 nucleotides). However, the entire coding segment and 3′ heptamer of DQ52 were deleted during each of these D-D fusions. As discussed earlier, deletion of a significant portion of J_H coding sequences is not uncommon during D-J_H rearrangement. Alternatively, rearrangement in these fusions may have been aberrantly mediated by the 3′ heptamer (instead of the 5′ heptamer) of the DSP2 D segments and the 3′ heptamer of DQ52, since in each instance the breakpoint in DQ52 is just 3′ of this spacer. These rearrangements are nonfunctional in that it is unlikely that any of these re-

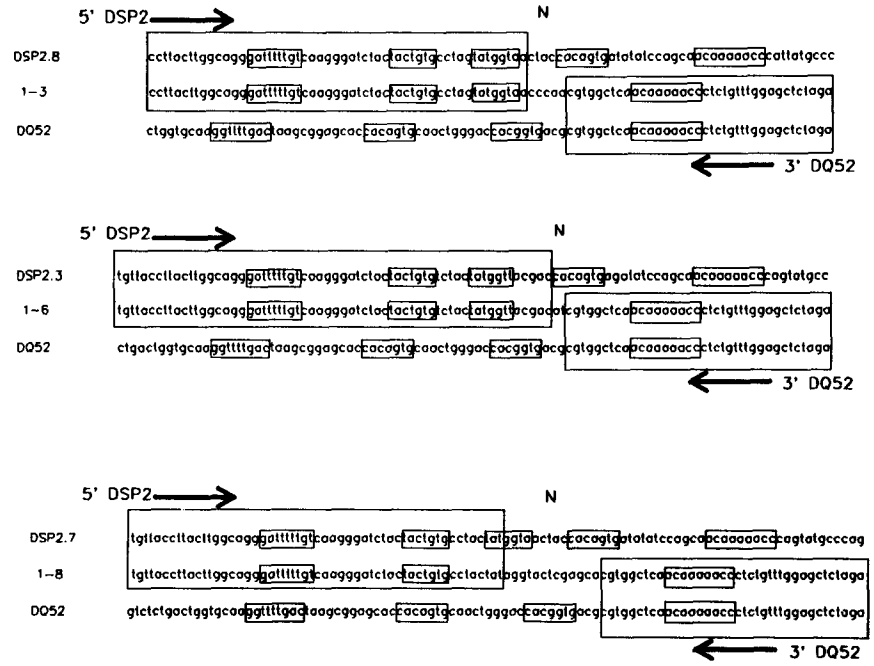

FIGURE 8. Nucleotide sequence of clones 1-3, 1-6, and 1-8, direct D-D rearrangements. Solid boxes indicate nucleotide homology between the three clones and germline DQ52 and DSP2 D segments (8). Broken boxes indicate heptamer and nonamer sequences. Primers used in the PCRs are as indicated. Extra nucleotides at the point of joining (N segments) are indicated with an N above the sequence.

arrangements could have proceeded in a subsequent D-J_H rearrangement due to their lack of a complete 3' signal sequence.

Fig. 9 shows the sequence of clone 5-6-1, which was generated by amplification of bone marrow DNA with the 5Q and 3X priming oligonucleotides and represents the reciprocal joint from a direct D-D fusion. A schematic of this reciprocal joint is shown in Fig. 7 *b*. In this clone, the 5' heptamer from DQ52 has been fused to the 3' heptamer from one of the DSP2 type D segments in the opposite orientation.

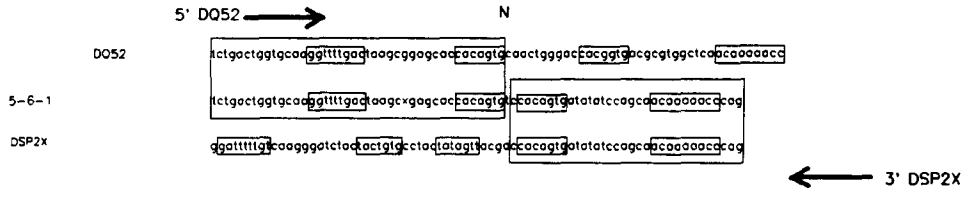

FIGURE 9. Nucleotide sequence of clone 5-6-1, the reciprocal joint from a direct D-D rearrangement. Solid boxes indicate nucleotide homology with germline DSP2 and DQ52 D segments (8). Broken boxes indicate heptamer and nonamer sequences. Primers used in the PCRs are as indicated. Extra nucleotides at the point of joining (N segments) are indicated with an N above the sequence.

Since this reciprocal joint must have been generated from a direct D-D fusion, it most likely existed on closed circular DNA, again emphasizing the sensitivity of PCR amplifications. In contrast to the reciprocal joints described from κ rearrangements, the joining of these heptamers was not precise; two extra nucleotides (TC) having been inserted between the two heptamers. While the C may derive from the DSP2 coding segment, the T cannot be accounted for in either germline sequence and thus must be considered an N segment. In TCR-β reciprocal joints, extra nucleotides from the coding segments have been observed between the heptamers, but additional random nucleotides have rarely been observed (23, 24).

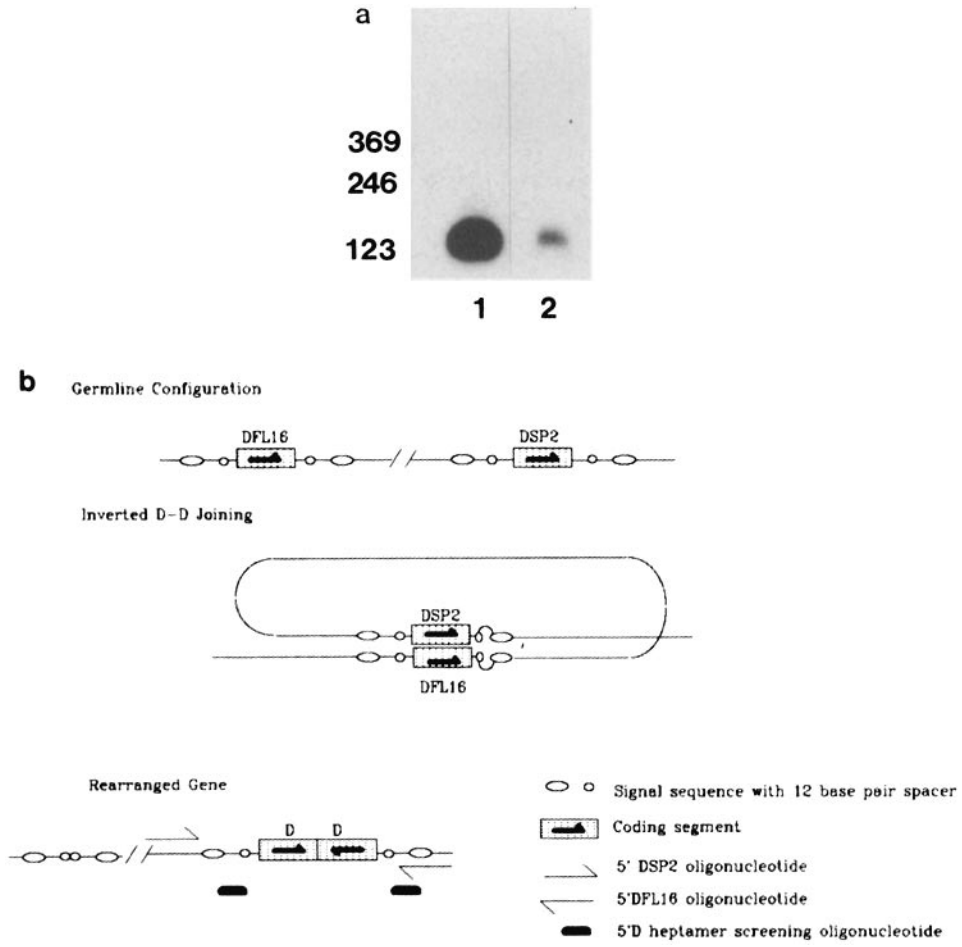

FIGURE 10. Analysis of indirect D-D joining in bone marrow DNA. (*a*) Southern filter analysis of PCR products (60 amplification cycles) from bone marrow DNA using the following oligonucleotides as primers: lane *1*, 5S and 3X (amplifying germline DSP2 type D segments); and lane *2*, 5F and 5S (amplifying indirect D-D rearrangements). DNA was run on a 1.5% agarose gel, transferred to zeta probe, and hybridized to a ^{32}P-labeled oligonucleotide (5′ DSP2 and DFL16 screening). Post-washing conditions were 2× SSC, 0.5% SDS, at 55°C. The positions of size markers are shown in base pairs. (*b*) Schematic representation of rearrangements detected with this technique.

Potentially Functional Inverted D-D Fusions Occur in the Ig H Chain Locus. A Southern filter hybridization analysis of PCR products obtained using the 5F primer and the 5S primer to amplify bone marrow rearrangements is shown in Fig. 10 *a*. The strategy for these amplifications is to detect the inverted rearrangement of two D segments, as illustrated in Fig. 10 *b*. A band of ~130 nucleotides is detected with a 5'D heptamer oligonucleotide probe (5' DFL16 and DSP2 screening; Fig. 2), but at similar stringency, there is no hybridization of this band with a 3' D heptamer oligonucleotide (3' DFL16 and DSP2 screening; Fig. 2), consistent with a tail to tail D segment fusion.

Three PCR-amplified D segments of this type were cloned and sequenced, as depicted in Fig. 11. Clone 9-11-2 is an inverted fusion involving DFL16.1 and DSP2.5. In this case, 13 bases of the coding sequence of DFL16.1 are joined to 14 bases of the coding sequence of DSP2.7 in the opposite orientation, with a five-base N segment (CGGGG) separating the fused D segments. The 5' heptamer/nonamer sequences of the fused D segments now flank the fused coding sequence and should still function in subsequent D-J_H and V_H-D rearrangements. In this case a 10-amino acid D-N-D element quite different in sequence from either parent D segment would be generated.

Clone 9-7 represents an inverted fusion between DFL16.1 and DSP2.7. In this instance, an even larger portion of the coding segment of DFL16.1 has been preserved (23 bp), fused to 13 nucleotides of the coding segment of DSP2.7 in the opposite

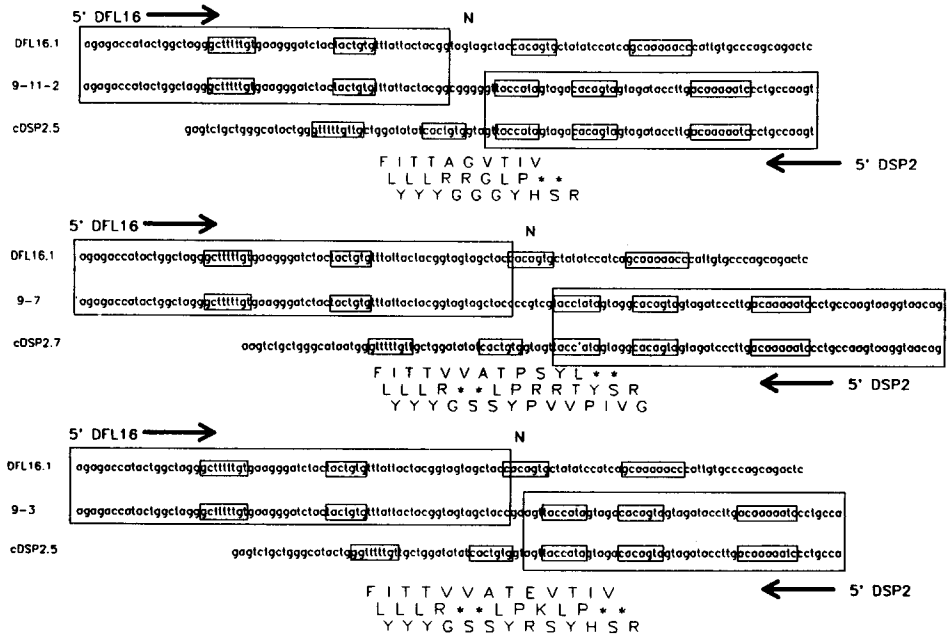

FIGURE 11. Nucleotide sequence of clones 9-11-2, 9-7, and 9-3, indirect D-D rearrangements. Solid boxes indicate nucleotide homology between the three clones and germline DFL16.1 and DSP2 D segments (8). Broken boxes indicate heptamer and nonamer sequences. Primers used in the PCRs are as indicated. Extra nucleotides at the point of joining (N segments) are indicated with an N above the sequence.

orientation with a 6-bp N segment (CCGTCG). There is a single nucleotide insertion in the portion of the sequence derived from the coding segment of DSP2.7. It is not known whether this occurred during D-D joining or if it is a PCR artifact. Taq I DNA polymerase has been shown to have less fidelity than other DNA polymerases.

Clone 9-3 represents a third example of inverted D-D fusion, again involving DFL16.1 and DSP2.5. In this case only a 2-bp N segment (GA) is present between the two coding regions. As in the other two examples, the 5′ signal sequences of each D segment remain intact and now flank the newly generated D coding sequences. Thus, as before, this rearrangement should still be able to undergo normal D-J_H and V_H-D rearrangement.

Discussion

This study demonstrates that rearrangement by inversion occurs at the Igh locus in cells from normal murine tissue, resulting in functional rearrangements. Rearrangement by inversion has been demonstrated at several of the immune receptor loci (κ, and 3′ TCR-β V regions); if the two rearranging genes are in the same transcriptional orientation, then rearrangement is by deletion, but if they are in opposite transcriptional orientation, then rearrangement is by inversion. At the H chain locus, complete V_H-D-J_H rearrangement by deletion has almost exclusively been reported (24), so it is likely that most of the V_H segments are in the same transcriptional orientation as J_H. However, since the D segments are flanked by recombination signal sequences, there is no obvious reason why inverted D-J_H joining should not occur. Still, there have been only two reports of inverted rearrangements involving D and J_H segments. In each case an inverted rearrangement occurred between a rearranged D-J_H segment to a second J_H segment to produce a nonfunctional J_H-D-J_H rearrangement. In one case the rearrangement was from a myeloma cell line (SP2/0-Ag14) and in the other an Abelson MuLV-transformed cell line (14, 15, 17).

There is limited evidence from work with synthetic substrates that rearrangement by deletion is more efficient than rearrangement by inversion (24–26). However, in vivo, rearrangement by inversion is common at the κ locus. In contrast, even though the signal sequences flanking the D segments are structurally equivalent, inverted D-J_H joining has been rarely observed.

Antigenic selection or constraints on antibody structure could be invoked to explain why inverted D-J_H rearrangements are not observed more frequently in expressed antibodies. However, our studies examined events before functional selection occurs, and the observation is the same. Of the many D-J_H rearrangements cloned previously using a variety of different strategies, none contain inverted D segments. Additionally, although PCR amplifications as described here cannot be considered quantitative, there is a significant difference in the ease with which direct vs. inverted D-J_H rearrangements can be demonstrated via Southern filter hybridization. Thus, these data confirm previous reports suggesting that inverted D-J_H rearrangement is rare. While we have shown that inverted D-J_H rearrangement can occur, direct D-J_H rearrangement is clearly favored.

As discussed previously, rearrangement in pre-B cells is thought to be an ordered process (D-J_H, precedes V_H-D-J_H, which precedes rearrangement at the L chain loci). The D-D fusions demonstrated here would indicate that D-D rearrangements could

occur before D-J_H, leaving a functional D-D fusion to rearrange to J_H. Since the D-D fusions obtained in this study arise from pooled tissue, it is impossible to assess the state of the remainder of the IgH loci from which they arose. Thus, there can be no formal proof that D-D fusion preceded any other D-J_H rearrangements. However, the observation that V-D rearrangements occur in the absence of D-J rearrangements at the TCR locus establishes a precedent that might apply to D-D fusion in Igh (27).

We have isolated D-D fusions involving 6 of the 10 known germline D elements. They have included D segments from each D segment family, so it is apparent that this phenomenon is not limited to certain D segments. Though all of the clones described involve a DSP2 type D segment, D-D fusion is not limited to this type of D segment, as we have evidence from Southern filter hybridization experiments for D-D fusions between DQ52 and DFL16.1 (data not shown). The alternative signal sequences do not appear to mediate these fusions, as the cryptic heptamer is found in the coding segments of the fused D segments. Finally, since neither DQ52 nor DFL16.1 contain alternative signal sequences, it seems clear that normal signal sequences suffice to mediate D-D joining.

The N segments at the joints of the seven rearrangements described here (six D-D and one D-J_H) are similar to those previously described to be due to the activity of Tdt. They range in length from 2 to 13 nucleotides and G + Cs predominate 2:1 over A + Ts. Of particular interest are the N segments of clones 1-8 and 9-7. Both of these N segments contain the codon TCG in one of their reading frames. TCG is an exceedingly rare codon in mammalian DNA, probably reflecting a bias against CG pairs. In these seven N segments, only 18 complete codons were generated, and 2 of the 18 were TCG. We have previously reported a conserved TCG in the N segments of antiarsonate antibodies from two different mouse strains (16). We postulated at that time that this TCG could have been generated by a variety of different mechanisms, including D-D fusion or a specific terminal transferase. These results suggest (though we have no formal evidence), that the terminal transferase involved in generating N segments may have a nucleotide preference, such that when it does add a T, it prefers to add C next.

There are numerous reports demonstrating that the 12/23-bp rule is not absolute, including J-J joining of TCR-Jα segments (28), V region replacement (9, 10), and C_κ deletion (11). The D-D fusions described in this report demonstrate another type of rearrangement in which this rule is broken. In this instance, these rearrangements occur at a significant frequency in normal tissue and can generate potentially functional products.

There are several possible mechanisms through which the D-D fusions herein described could have been generated. In the direct D-D fusions, it is unclear whether the 5′ or 3′ heptamer of DQ52 mediated the fusions since both are absent from the coding joints. Since the point of joining in all three clones is adjacent to the position of the 3′ heptamer, these recombination events probably represent "heptamer only"-mediated recombinations. However, it is clear that the 5′ heptamer of DQ52 can also mediate direct D-D fusion since the 5′DQ52 heptamer is joined directly to the 3′DSP2 heptamer in the reciprocal joint presented. This finding also suggests that functional direct D-D fusions (i.e., with recombination sequences flanking the fused D segments) may also be generated by this mechanism.

In each of the inverted D-D fusions, it is apparent that the recombinations were mediated by the 3′ heptamers of the two D segments involved. Whether or not the nonamers of the signal sequences were involved in the joining process is uncertain. The mechanistic details of heptamer-mediated recombination, while impacting on these results, are beyond the scope of this paper.

Though one cannot ignore the fact that Ig H chain rearrangement usually does follow the 12/23 rule, it is clear from these data that D-D fusion does occur. The enzymes that mediate rearrangement work on substrates in various orientations with or without complete recombination signal sequences. Based on the observed frequency of different types of rearrangement at this locus, the recombinase system can be said to have a hierarchical preference for many substrates. Thus, gene segments that conform to the 12/23 rule are rearranged most efficiently, followed by genes in opposite transcriptional orientations with normal signal sequences. Least efficiently rearranged are genes with similar spacers and/or heptamer-mediated recombination.

Although inverted D-J_H joining is not a frequent rearrangement, due to antigenic selection and clonal expansion, the effect of even a low rate of inverted D-J_H joining on the potential antibody repertoire could be very significant. Therefore, in response to an antigen where a simple V_H-D-J_H rearrangement does not provide an antibody of high affinity, it is quite possible that antibodies utilizing D segments in the opposite orientation may provide an antibody of the appropriate specificity. Such a B cell could easily become the predominant clone in response to such an antigen. If each H chain D segment can actually be used in either orientation, the calculated amount of antibody diversity should in effect be doubled.

In an analysis of the N segments of a series of expressed antibodies, some, but not all, N segments fit the simple description of N segments, as proposed by Alt and Baltimore (17) (i.e., CG rich, 1–10 nucleotides). However, in a composite analysis of expressed antibodies that have been reported over the years, many D and N segments are not easily explained by known germline elements or by N segment addition by terminal transferase (12, 13). As Kurosawa and Tonegawa (8) initially proposed, D-D fusion might offer an explanation for the extra nucleotides observed in the HVR3 of these antibodies. However, our data indicate that the alternative signal sequences do not mediate these fusion events. Thus, we extend their original claim to say that any combination of D segments, in any orientation, might be able to generate diversity in the HVR3 of Igs. Thus, the unexplainable nucleotides in many expressed antibodies may be a result of the action of TdT activity and/or D-D fusion. Additionally, in instances where D-D fusion occurs, an added dimension of diversity is added since there are three positions for N segment additions (V_H-N-D-N-D-N-J_H), as is the case in the TCR-δ locus (29).

Even though rearrangements that break the 12/23 rule most likely occur at a lower frequency than normal joining, given the power of antigen selection and clonal expansion, the potential effect on the generation of diversity in Ig H chains could be very significant. Though we have not been able to determine an absolute frequency of D-D joining vs. D-J_H joining, some information can be derived from our cloning experiments. The three unique clones, 9-3, 9-7, and 9-11-2, were all derived from a single amplification. The amount of DNA in that amplification was derived from $\sim 10^7$ unselected bone marrow cells. Pre-B cells comprise 5–10% of total bone

marrow cells. Thus, if our cloning efficiency was 100%, the frequency of D-D joining in pre-B cells must be at least 1 in 33,000. In most somatic cells, an event that occurs at a frequency of 1 in 33,000 would in almost all cases be biologically insignificant. However, in the immune system these rare events may not be without biologic consequence. For example, B cell precursors that have the particular V_H-D-J_H/V_κ-J_κ combination that give rise to the A/J strains' dominant crossreactive idiotype in an antiarsonate response have a frequency of only 1–10 in 10^6 cells (30). Thus, a B cell with a receptor phenotype that occurs only 1–10 times in 10^7 B cells dominates a response. Therefore, even though D-D fusion and inverted D-J_H rearrangement may only occur once out of every 10^7 rearrangements, they likely have an important impact on the generation of antibody diversity.

Summary

IgH rearrangements (V_H-D, D-J_H) are central to the generation of antibody diversity. The majority of the diversity seen in the third hypervariable region is generated by the D segment and at the joints formed by both junctional and N segment variation during D-J_H and V_H-D rearrangements. The mechanisms that regulate rearrangement are thought to obey the 12/23 rule, wherein D-D or V_H-J_H rearrangements are precluded. Here, we present evidence that D-D fusions do in fact occur, either as direct or inverted rearrangements. The fused D segments so generated may be fully capable of proceeding in subsequent D-J_H and V_H-D rearrangements. The resultant V_H-D-D-J_H recombinations add another dimension to the potential repertoire of IgH V regions by increasing the level of combinatorial diversity and by providing additional sites for N region variation.

We thank our colleagues, Drs. R. Baer, W. Garrard, J. Goldstein, C. Slaughter, and P. Tucker, for critical reading of the manuscript and many helpful discussions.

Received for publication 30 November 1988 and in revised form 28 March 1989.

References

1. Honjo, T. 1983. Immunoglobulin genes. *Annu. Rev. Immunol.* 1:499.
2. Tonegawa, S. 1983. Somatic generation of antibody diversity. *Nature (Lond.).* 302:575.
3. Alt, F. W., T. K. Blackwell, R. A. DePinho, M. G. Reth, and G. D. Yancopoulos. 1986. Regulation of genome rearrangement events during lymphocyte differentiation. *Immunol. Rev.* 89:5.
4. Rathbun, G. A., and P. W. Tucker. 1986. Conservation of sequences necessary for V gene recombination. *In* Evolution of the Immune Response. G. Kelsoe and D. Shultz, editors. Academic Press, New York. 75–115.
5. Lewis, S., A. Gifford, and D. Baltimore. 1985. DNA elements are asymmetrically joined during the site specific recombination of kappa immunoglobulin genes. *Science (Wash. DC).* 228:677.
6. Stenzel-Poore, M. P., and M. B. Rittenberg. 1987. Immunoglobulin variable region heptamer-nonamer recognition sequence joined to rearranged D-J segment: implications for the immunoglobulin recombinase mechanism. *J. Immunol.* 138:3055.
7. Desiderio, S. V., and K. R. Wolff. 1988. Rearrangement of exogenous immunoglobulin V_H and DJ_H gene segments after retroviral transduction into immature lymphoid cell lines. *J. Exp. Med.* 167:372.

8. Kurosawa, Y., and S. Tonegawa. 1982. Organization, structure, and assembly of immunoglobulin heavy chain diversity DNA segments. *J. Exp. Med.* 155:201.
9. 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.
10. 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 Ly 1^+ B-cell lymphoma. *Nature (Lond.).* 322:843.
11. Siminovitch, K. A., A. Bakhski, P. Goldman, and S. J. Korsmeyer. 1985. A uniform deleting element mediates the loss of κ genes in human B cells. *Nature (Lond.).* 316:260.
12. Ollier, P., J. Rocca-Serra, G. Somme, J. Theze, and M. Fougereau. 1985. The idiotypic network and the internal image: possible regulation of a germline network by paucigene encoded Ab2 (anti-idiotypic) antibodies in the GAT system. *EMBO (Eur. Mol. Biol. Organ.) J.* 3:3681.
13. Meek, K., C. Hasemann, B. Pollok, S. S. Alkan, M. Brait, M. Slaoui, J. Urbain, and J. D. Capra. 1989. Structural characterization of antiidiotypic antibodies: evidence that Ab2s are derived from the germline differently than Ab1s. *J. Exp. Med.* 169:519.
14. Liu, Z., A. F. Wu, and T. T. Wu. 1987. Short gene inversion involving two adjacent heavy chain joining minigenes and one heavy chain diversity minigene in the nonsecretor SP2/O-Ag14 myeloma cell line. *Nucleic Acids Res.* 15:4688.
15. Liu, Z., C. Wood, and T. T. Wu. 1987. Nucleotide sequence of an anti-fluorescyl hapten antibody heavy chain variable region gene from a BALB/c mouse hybridoma cell line. *Nucleic Acids Res.* 15:6296.
16. Milner, E. C. B., K. Meek, G. Rathbun, P. W. Tucker, and J. Capra. 1986. Are antiarsonate antibody N-segments selected at both the protein and the DNA level? *Immunol. Today.* 7:36.
17. Alt, F., and D. Baltimore. 1982. Joining of immunoglobulin heavy chain gene segments: implications from a chromosome with evidence of three D-J_H fusions. *Proc. Natl. Acad. Sci. USA.* 79:4118.
18. Saiki, R. K., S. Scharf, F. Faloona, K. B. Mullis, G. T. Horn, H. A. Erlich, and N. Arnheim. 1985. Enzymatic amplification of beta-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. *Science (Wash. DC).* 230:1350.
19. Saiki, R. K., T. C. Bugawan, G. T. Horn, K. B. Mullis, and H. A. Erlich. 1986. Analysis of enzymatically amplified beta-globin and HLA-DQa DNA with allele-specific oligonucleotide probes. *Nature (Lond.).* 324:163.
20. Maniatis, T., E. F. Fritsch, and J. Sambrook. Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY. 545 pp.
21. Sanger, F., S. Nicklen, and A. R. Coulson. 1977. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA.* 74:5463.
22. Coleclough, C., R. P. Perry, K. Karjalainen, and M. Weigert. 1981. Aberrant rearrangements contribute significantly to the allelic exclusion of immunoglobulin gene expression. *Nature (Lond.).* 290:372.
23. Malissen, M., C. McCoy, D. Blanc, J. Trucy, C. Devaux, A. Schmitt-Verhulst, F. Fitch, L. Hood, and B. Malissen. 1986. Direct evidence for chromosomal inversion during T-cell receptor beta gene rearrangements. *Nature (Lond.).* 319:28.
24. Hesse, J. E., M. R. Lieber, M. Gellert, and K. Mizuuchi. 1987. Extrachromosomal DNA in pre-B cells undergo inversion or deletion at immunoglobulin V-(D)-J joining signals. *Cell.* 49:775.
25. Okazaki, K., D. D. Davis, and H. Sakano. 1987. T cell receptor B gene sequences in the circular DNA of thymocyte nuclei: direct evidence for intramolecular DNA deletion in V-D-J joining. *Cell.* 49:477.

26. Akira, S., K. Okazake, and H. Sakano. 1987. Two pairs of recombination signals are sufficient to cause immunoglobulin V-(D)-J joining. *Science (Wash. DC).* 238:1134.
27. Chien, Y-h., M. Iwashimi, D. A. Wettstein, K. B. Kaplan, J. F. Elliott, W. Born, and M. M. Davis. 1987. T-cell receptor δ gene rearrangements in early thymocytes. *Nature (Lond.).* 330:722.
28. Baer, R., T. Boehm, H. Yssel, H. Spits, and T. H. Rabbitts. 1988. Complex rearrangements within the human J_δ-C_δ/J_α-C_α locus and aberrant recombination between Ja segments. *EMBO (Eur. Mol. Biol. Organ.) J.* 7:1661.
29. Elliott, J. F., E. P. Rock, P. A. Patten, M. M. Davis, and Y. L. Chien. 1988. The adult T-cell receptor δ-chain is diverse and distinct from that of fetal thymocytes. *Nature (Lond.).* 331:627.
30. Sigal, N. H. 1982. Regulation of azophenylarsonate-specific repertoire expression. *J. Exp. Med.* 156:1352.
31. Sakano, H., Y. Kurosawa, M. Weigert, and S. Tonegawa. 1981. Identification and nucleotide sequence of a diversity DNA segment (D) of immunoglobulin heavy-chain genes. *Nature (Lond.).* 290:561.