Annu. Rev. Immunol. 1992. 10:359–83

V(D) J RECOMBINATION: Molecular Biology and Regulation

David G. Schatz

Howard Hughes Medical Institute and Section of Immunobiology, Yale University School of Medicine, New Haven, Connecticut 06510

Marjorie A. Oettinger

Department of Molecular Biology, Massachusetts General Hospital, Boston, Massachusetts 02114

Mark S. Schlissel

Departments of Medicine, and Molecular Biology and Genetics, Johns Hopkins University School of Medicine, Baltimore, Maryland 21205

KEY WORDS: V(D)J recombination, lymphoid development, recombination activating gene, RAG-1, RAG-2

THE V(D) J RECOMBINATION REACTION

The vertebrate immune system is capable of specifically recognizing and responding to an enormous number of antigens. The interaction with antigen is mediated by the immunoglobulin (Ig) and T cell receptor (TCR) molecules expressed by B and T lymphocytes, respectively. Ig and TCR polypeptides consist of separate structural domains: a variable domain that forms the antigen binding site, and one or more constant region domains that mediate various effector functions. The exons encoding the variable domains are assembled during lymphocyte development from component gene segments by a site-specific recombination reaction known as V(D) J recombination. This reaction generates much of the required

0732–0582/92/0410–0359$02.00

diversity in the binding specificities of Igs and TCRs. This review discusses the V(D)J recombination reaction, the recombination activating genes, RAG-1 and RAG-2, and the regulation of recombination during lymphoid development.

Seven loci (α, β, γ, and δ for TCRs, and heavy chain and light chain—κ and λ—for the Ig genes) undergo gene rearrangements during lymphocyte development. These loci have a similar organization (for reviews, see 1–3), consisting of multiple variable (V), diversity (D, present only at μ, β and δ loci) and joining (J) gene segments that can be juxtaposed to form unique variable region exons. The constant region is encoded by separate exons that typically lie downstream from the variable region gene segments. In the Ig heavy-chain locus, for example (Figure 1A), V, D, and

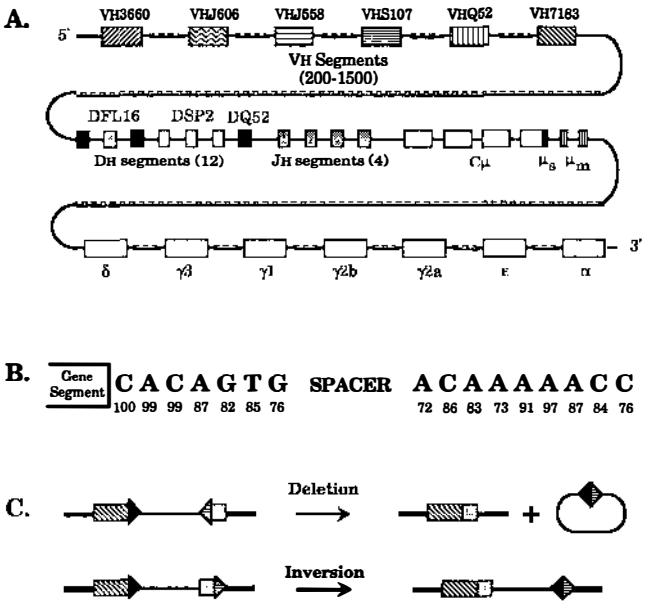

Figure 1 Structure of the murine heavy chain locus and the recombination signal sequence. A. Striped boxes indicate only six of the nine known V_H segments, while D and J segments are represented by smaller, shaded boxes. The separate constant region exons are shown only for Cμ. Also shown are the exons that specify the membrane bound form of μ (μm); the terminal portion of the fourth Cμ exon (μs) specifies secretion. Not drawn to scale. Adapted from (1). B. The recombination signal sequence. The consensus sequence for a RSS lying 3′ of its flanking gene segment is shown. The spacer is 12 or 23 nucleotides long. Below each residue of the heptamer and nonamer is shown the frequency of occurrence of that base at that position, based on a compilation of 226 RSS sequences. Data taken from (7). C. Deletional and inversional V(D)J recombination. Coding segments are represented as rectangles and RSSs as triangles.

J gene segments are present as clusters spread out along the chromosome over distances ranging from a few kilobases (J) to hundreds of kilobases (V), with V and D gene segments grouped into families containing members with closely related sequences. This [V-(D)-J]-C structure of Igs and TCRs is evolutionarily conserved, but the precise number and organization of the component gene segments vary greatly between loci and between species.

The V(D)J recombination reaction that assembles the variable domains is the only site-specific recombination reaction known to occur in vertebrates. Structural analysis of rearrangements at Ig and TCR loci has revealed the types of reactions that occur and various aspects of their developmental and lineage-specific regulation (discussed below). In addition, it has been possible to study artificial recombination substrates that are either integrated into the chromosome or are freely replicating plasmids. These substrates need only contain the recombination signal sequences (see below) to be capable of rearranging, and thus they allow examination of what the V(D)J recombination machinery is capable of in the absence of developmental regulation, selection, or the limitations imposed by the organization of the endogenous loci.

Recombination Signal Sequences

Conserved sequence motifs flank one or both sides of all gene segments known to rearrange. The consensus recombination signal sequence (RSS) (Figure 1B) consists of a dyad-symmetric heptamer sequence directly adjacent to the coding element and an A/T rich nonamer separated from the heptamer by a spacer region of nonconserved nucleotides (4). Two types of RSSs are defined by the length of the spacer, either 12($\pm$1) or 23($\pm$1) nucleotides, with efficient joining requiring one RSS of each type. The germline organization of V, D, and J segments reflects this requirement: all gene-segments of one kind have the same configuration of RSSs, and joining partners carry opposite types of signals. Among all the loci and species known to carry out V(D)J recombination, the joining signals have been conserved and in some cases have been shown to be functionally interchangeable (see references in 5). Studies of artificial recombination substrates have shown that a pair of RSSs is both necessary and sufficient to direct V(D)J recombination (6, 7); no other specific DNA sequence, coding segment DNA, or particular chromosomal context is required for the reaction.

The optimal RSS for directing V(D)J recombination has been defined by mutational analysis (7). Naturally occurring RSSs rarely have this sequence, but those residues most important for function in artificial substrates are also those most highly conserved (Figure 1B). The variability

in naturally occurring signals may have important effects on the rearrangement frequency of individual segments. Furthermore, the range of sequences that can serve as RSSs allows nontraditional events to occur, such as rearrangements directed by one intact RSS and an isolated heptamer. Such events are seen in "V gene replacement" reactions where an upstream V segment with an intact RSS recombines with a previously rearranged VDJ_H segment through a heptamer sequence conserved at the 3′ end of most V_Hs (8, 9). In addition, recombination events between a lone heptamer in the J_κ-C_κ intron and sequence downstream of C_κ resembling an RSS with a 23 base pair spacer have been observed in both human and mouse (10, 11).

Products of the Reaction

The normal V(D) J recombination reaction results in the formation of two new structures: the joining of coding elements to form the "coding joint" and the joining of RSSs, heptamer to heptamer, to form the "signal joint" (Figure 1C). The orientation of RSSs with respect to each other determines whether rearrangement is deletional or inversional. Analysis of artificial recombination substrates has shown only a small inherent preference in the recombination machinery for deletion over inversion, as the frequencies of the two types of joining are similar (12, 13). The structure of the Ig and TCR loci dictates that the majority of rearrangements occur by deletion, but inversion has been shown to occur at the β (14), δ (15), and κ loci (16–18). Signal joint-containing circular products resulting from deletional rearrangement (Figure 1C) have been recovered from thymocytes undergoing V(D)J recombination (19, 20).

The V(D) J recombination reaction is much more permissive with regard to the possible products of recombination than are other site-specific recombination reactions that have been described. In addition to the "normal" recombination products described above, work with artificial substrates has shown that other structures, "open and shut" joints and "hybrid" joints, can be formed (21–23). Hybrid joints arise when RSSs become attached to their reaction partners' coding ends. Up to 30% of the rearrangements of artificial, nonintegrated substrates can result in hybrid joint formation (21). Less commonly seen are "open and shut" events where the RSS is cut away from the coding sequence, base loss and addition occurs at the coding end, and then the original signal and coding ends are rejoined. Putative open and shut joints and hybrid joints have also been reported at endogenous recombining loci (discussed in 21). The finding of these reaction products indicates that the V(D) J recombination machinery does not exclusively join signal ends to signal ends and coding ends to coding ends. Rather, for any given configuration

of RSSs, both inversional and excisional recombination events are possible.

The minimum and maximum distances over which successful recombination can occur have yet to be determined, but recombination between joining signals separated by less than 200 bp has been observed in artificial substrates (7, 12), and chromosomal rearrangements can span hundreds of kilobases. No interplasmid recombination has been reported using extrachromosomal substrates, but recombination between the γ and δ and between the γ and β TCR loci, which lie on different chromosomes, has been detected by PCR (24, 25). Whether these interlocus recombinations were preceded by unrelated chromosomal translocations or were instead the result of bona fide interchromosomal V(D)J recombination is not clear.

JUNCTIONAL DIVERSITY Coding ends and signal ends are treated quite differently by the recombination machinery. The joining of coding ends is generally imprecise, with base loss and base addition of 1–10 nucleotides occurring commonly. This imprecision contributes to the potential diversity of the immune receptor repertoire, but also results in the production of nonfunctional genes due to out-of-frame joining and the introduction of stop codons. Signal joints, in contrast, are much more precise. Base loss is not seen at this junction and nucleotide additions are much less frequent (5–25% of signal joints have base additions) (13, 26). The distinction between signal and coding joint formation is highlighted by the *scid* recombination phenotype (reviewed in 27). The *scid* mutation results in a drastic reduction in proper coding joint formation, but has little effect on the formation of signal joints (22, 28, 29).

There are two classes of nucleotide addition to recombinant junctions: templated and random. Random additions of so-called "N nucleotides" occur at the coding junctions of TCR and Ig heavy chain genes. N nucleotide addition is probably mediated by terminal deoxynucleotidyl transferase (TdT; discussed further below). Templated additions, termed "P nucleotides", have been observed adjacent to some coding segments where base loss has not occurred (30, 31). Strikingly, these typically mono- and di-nucleotide additions are complementary copies of the last nucleotides of the coding segment to which they abut. The addition of complementary nucleotides may be an inherent part of the V(D)J joining reaction that can only be observed in the absence of base trimming (30). P nucleotide additions of up to 3 or 4 bases have been seen in wild type cells. Longer P regions, up to 12 nucleotides in length, have been documented at the boundaries of some of the rare coding joints recovered from *scid* lymphocytes (32), raising the possibility that the *scid* defect affects this step of the joining process.

Reaction Mechanism and Enzymatic Machinery

Steps likely to occur during the joining reaction can be inferred from the substrates and products of the reaction, and a general model has recently been proposed (33). A priori, V(D)J recombination should involve the recognition of the recombination signal sequences, cutting at the coding sequence-heptamer borders, synapsis of the joining complexes, base loss and addition, repair of the ends and ligation. The homology between the two RSSs is probably not indicative of base pairing between them, as mutations in one RSS are not rescued by compensatory mutations in the other (7). However, short sequence homologies present at the 3′ end of D_H coding sequences and the 5′ end of J_H elements may lead to preferential joining in a particular reading frame (34, 35). The existence of hybrid joint structures implies that the four free ends generated by cutting between the RSSs and coding sequences must be relatively stable and that cutting and rejoining are not concerted reactions.

In contrast to our understanding of the substrates and products of V(D)J recombination, little is known about the enzymatic machinery that carries out the reaction. A common V(D)J recombinase is thought to carry out all V(D)J-type rearrangement events in both B cells and T cells. This conclusion is based on three lines of indirect evidence. As noted above, all rearranging loci are flanked by conserved RSSs, and reporter constructs utilizing gene segments from any of the lymphoid rearranging loci will rearrange in any cell line active in gene rearrangement. In addition, D-to-J rearrangement of Ig heavy chain and TCR β chain genes occurs in both B and T cells (36–38, and M. S. Schlissel, unpublished results). Finally, the same two genes (RAG-1 and RAG-2, see below) that induce V(D)J recombination in nonlymphoid cells are expressed in all recombinationally active lymphocytes.

The factor mutated in *scid* mice (39) is strongly implicated as one component of the V(D)J recombinase. The *scid* mutation is recessive and causes a severe immunodeficiency due to the virtual absence of mature B and T cells. V(D)J recombination in *scid* lymphocytes is defective and is characterized by large deletions of coding information, usually resulting in nonfunctional alleles (reviewed in 27). Because nonlymphoid cells from *scid* animals demonstrate increased radiation sensitivity (40–42), the *scid* factor is probably a general factor involved in DNA repair that is recruited for the V(D)J recombination reaction.

A second factor, terminal deoxynucleotidyl transferase (TdT), is probably responsible for at least some of the nucleotide addition seen at recombinant junctions. The insertion of N regions correlates with the level of TdT expression (43 and references therein), and the composition of these

insertions is consistent with the known preference of this enzyme for polymerizing G nucleotides (13 and references therein). TdT expression and N nucleotide addition, however, are not required for V(D)J recombination. Furthermore, short nontemplated base additions are occasionally seen in the apparent absence of TdT (as in fibroblasts transfected with RAG-1 and RAG-2; M. A. Oettinger, D. G. Schatz, and D. Weaver, unpublished results) and are thus likely to be contributed by a different mechanism. Such base additions may be generated by the same mechanism by which nucleotide addition occurs at recombinant junctions formed when transfected, linear DNA is joined in mammalian cells (44).

Candidates for some of the other activities expected to be associated with V(D)J recombination have been identified biochemically. Heptamer-binding (6, 45), nonamer-binding (46), and endonuclease activities (47–49) have been detected in extracts from lymphoid cells. The gene for one heptamer-binding factor (RBP-J_κ) has been cloned, and the predicted amino acid sequence suggests it is a member of the integrase family (50), an intriguing finding since such a homology might be expected if this factor were to cleave DNA. Expression of RPB-J_κ, however, is not lymphoid specific (T. Honjo, personal communication). The relationship of any of these factors and activities to V(D)J recombination remains uncertain.

In contrast, several lines of evidence implicate the products of the recombination activating genes, RAG-1 and RAG-2, as factors critical for V(D)J recombination. In the next section, we review the isolation and properties of these genes.

THE RECOMBINATION ACTIVATING GENES, RAG-1 AND RAG-2

RAG-1 and RAG-2 activate the V(D)J recombinase when introduced together into nonlymphoid cells, and transcripts of the two genes are found in all cell lines and tissues in which V(D)J recombination occurs. While there is as yet no definitive demonstration that the proteins encoded by RAG-1 and RAG-2 participate directly in the recombination reaction, strong indirect evidence suggests that they are essential, lymphoid-specific components of the V(D)J recombinase.

Isolation and Genomic Organization

V(D)J recombinase activity has been detected only in immature lymphoid cell lines (51–53). However, when 3T3 fibroblasts were transfected with human or mouse genomic DNA, approximately 0.1% of transfectants gave rise to cells that stably expressed V(D)J recombinase activity (53).

This surprising result strongly suggested that transfer of a single genetic locus was sufficient to activate the V(D)J recombinase in a fibroblast. This locus was subsequently identified by its co-segregation with V(D)J recombinase activity through successive rounds of genomic transfection (5). Molecular cloning of the locus revealed that it contained two genes, named RAG-1 and RAG-2, that were transcribed in pre-B and pre-T cells and in recombinase-positive (transfected) fibroblasts, but not in a wide variety of other types of cell lines (5, 54).

The structure of the murine RAG-1 and RAG-2 genomic locus (Figure 2) is striking in two respects: The two genes are separated by only approximately 8 kb, and each gene has a similar, compact organization in which the entire coding and 3′ untranslated regions are contained in a single exon (54). The structure of the chicken RAG locus is similar to that of the mouse (55), and RAG-1 and RAG-2 are separated by less than 10 kb in the Xenopus genome (P. Higgins, L. A. Steiner, personal communication). Thus, the genomic organization of RAG-1 and RAG-2 has been highly conserved through vertebrate evolution. The two genes are convergently transcribed and therefore must utilize different promoters. Northern blotting reveals a single RAG-1 transcript of approximately 6.6 kb (5), and a predominant RAG-2 transcript of approximately 2.2 kb as well as several larger species (54). RAG-1 and RAG-2 have been mapped to human chromosome 11p13 and to the syntenic region of mouse chromosome 2 (56). No known human or mouse genetic diseases map precisely to these positions. The *scid* mutation maps to mouse chromosome 16 (57) and is therefore distinct from RAG-1 and RAG-2.

RAG-1 and RAG-2 have no sequence similarity to one another and therefore are unlikely to have arisen by gene duplication. It is unusual to find closely linked genes in the mammalian genome with related functions that lack sequence similarity. This, and the evolutionary conservation of the structure of RAG genomic locus, have led to the hypothesis that the

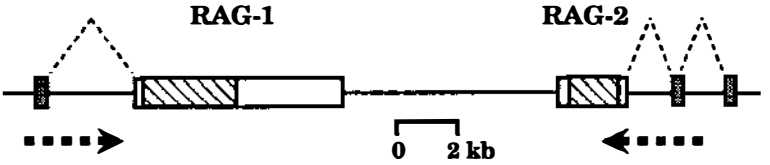

Figure 2 The RAG genomic locus. 5′ untranslated exons are represented as dark boxes, coding regions as hatch-mark boxes and other untranslated regions as unfilled boxes. RAG-1 consists of two exons, while RAG-2 appears to have multiple upstream 5′ untranslated exons (only two are shown; D. Silver, personal communication). Dashed arrows indicate the direction of transcription.

genes arose as part of a viral or fungal recombination system and have co-evolved to play their current role in V(D) J recombination (54).

Evolutionary Conservation and Sequence Homologies

Both the RAG-1 and RAG-2 coding sequences have been conserved through vertebrate evolution. Low stringency filter hybridization analyses have revealed RAG-1 and RAG-2 cross-hybridizing sequences in all vertebrate species examined (5, 54). Human, mouse, and chicken RAG-1 have been sequenced; the predicted mouse (1040 amino acids) and human (1043 amino acids) polypeptides are 90% identical (5), while the mouse and chicken (1041 amino acids) polypeptides are 75% identical (55). The predicted mouse (527 amino acids) and chicken (528 amino acids) RAG-2 polypeptides are 70% identical.

Computer sequence analysis has uncovered two intriguing homologies between the predicted RAG-1 protein and other known proteins. RAG-1, and a variety of other proteins thought to interact with DNA, have in common an approximately 50 amino acid–long, cysteine-rich motif which may represent a novel metal binding domain (58). The C-terminal half of the RAG-1 protein is homologous to the product of the yeast *HPR1* gene (59), which in turn is homologous to yeast DNA topoisomerase I (60). Mutations in *HPR1* cause a dramatic increase in the rate of homologous, intrachromosomal deletional recombination (60). Both RAG-1 (Tyr-998) and HPR1 (Tyr-532) have tyrosine residues in a context similar to that of the active site tyrosines of type I and type II topoisomerases, leading to the hypothesis that these proteins may have topoisomerase activities (59). Mutation of Tyr-532 in HPR1, however, does not result in a hyper-recombination phenotype (H. Klein, personal communication). Thus while RAG-1 may be evolutionarily related to topoisomerases, the relevance of topoisomerase activity to its function is unclear. No homologies have been found between RAG-2 and any known protein or nucleic acid sequence.

Synergistic Activation of the V(D)J Recombinase

Co-transfection of the RAG-1 and RAG-2 cDNAs (in appropriate expression vectors) is sufficient to activate the V(D) J recombinase in 3T3 fibroblasts, P19 embryonal carcinoma cells, and LyD9 pro-B cells (56, and unpublished results). V(D) J recombination is at least several thousand times more efficient with both RAG-1 and RAG-2 than with either gene alone. Rearrangements of recombination substrates in 3T3 fibroblasts expressing RAG-1 and RAG-2 are similar to those seen in lymphoid cell lines, including the deletion of nucleotides at coding junctions and the perfect fusion of the two recombination signal sequences at their heptamer borders (56, and unpublished results). Expression of RAG-1 and RAG-2

in *scid* 3T3 fibroblasts generates a *scid*-type V(D)J recombinase, capable of efficient formation of normal signal joints but not coding joints (unpublished results). Therefore, in conjunction with accessory factors (such as *scid*) already present in nonlymphoid cells, RAG-1 and RAG-2 are necessary and sufficient to activate the V(D)J recombinase.

RAG-1 and RAG-2 may be toxic to cells when co-expressed at high levels. We have been unable to derive fibroblasts stably expressing high levels of both transcripts (unpublished results). The levels of the RAG transcripts as detected by Northern blots decrease dramatically during continuous culture of mature B-cell clones transfected with RAG-1 and RAG-2 expression constructs, or of freshly isolated pre-B cells (F. Alt, G. Rathbun, personal communication). The basis for the growth disadvantage of cells that express high levels of RAG-1 and RAG-2 is unknown, but might be due to the deleterious effects of chromosomal rearrangements caused by the V(D)J recombinase.

The Pattern and Regulation of Expression of RAG-1 and RAG-2

Except for two cases discussed below, RAG-1 and RAG-2 transcripts have been detected only in pre-B and pre-T cell lines, and these lines invariably express both genes (5, 54). V(D)J recombinase activity is found almost exclusively in pre-B and pre-T cell lines (51–53), and there are as yet no examples of V(D)J recombination in the absence of either RAG-1 or RAG-2. As these results would predict, lymphopoietic organs such as the thymus (61), fetal liver (unpublished results) and bone marrow (F. Alt, G. Rathbun, personal communication) express both RAG-1 and RAG-2.

REGULATION OF EXPRESSION DURING T CELL DEVELOPMENT In situ hybridization has revealed that the vast majority of mouse cortical thymocytes express both RAG-1 and RAG-2 (62), a surprising result since many of these cells express TCR on their surface. Sorting of human thymocytes and Northern blot analysis confirmed that double-positive ($CD4^+CD8^+$) TCR^+ thymocytes as well as TCR^- thymocytes express both RAG-1 and RAG-2, and that the more mature, single-positive TCR^+ thymocytes do not express either gene (62). These results imply that surface expression of the TCR is not itself sufficient to terminate RAG transcription, and that cessation of RAG expression in thymocytes coincides approximately with the processes of positive and negative selection (63, 64). In fact, when TCR^+ human thymocytes are treated with antibodies that cross-link the TCR/CD3 complex, or with phorbol 12-myristate 13-acetate (PMA) and a calcium ionophore, RAG transcript levels are reduced dramatically (62). Thus, thymocytes may continue to express V(D)J recombinase activity

after expressing TCR on their surface, and cessation of V(D)J recombination may rely on signals transduced through the TCR and the phospholipase C/protein kinase C signal transduction pathway. These results raise the possibility that thymocytes continue to rearrange their TCR gene segments even in the presence of functionally rearranged TCR loci, an idea consistent with a report of significant endogenous α locus rearrangement in the presence of a rearranged α TCR transgene (65).

REGULATION OF EXPRESSION IN B CELLS Two experiments with immortalized cell lines suggest some parallels between the regulation of RAG expression during B-cell and T-cell development. First, treatment of immortalized mouse pre-B cells with PMA and a calcium ionophore results in an 8-fold decrease in V(D)J recombinase activity (66), raising the possibility that B cells and T cells use similar second messengers to terminate RAG expression. Interestingly, agents that increase the intracellular level of cAMP increase V(D)J recombinase activity and RAG-1 and RAG-2 mRNA levels in pre-B cell lines (66). Second, surface Ig-positive B-cell lines expressing an n-myc or c-myc transgene downregulate RAG-1 and RAG-2 transcript levels in response to treatment with anti-IgM antibodies (F. Alt, A. Ma, personal communication). This result may be relevant to the regulation of RAG expression during normal B-cell development since transcripts for the two genes can be detected in both the spleen (61) and lymph nodes (unpublished results), organs normally devoid of precursor-lymphocytes.

DISCORDANT EXPRESSION OF RAG-1 AND RAG-2 Low levels of the RAG-1 transcript have been detected in the fetal and postnatal murine central nervous system (CNS), with the transcript being most strongly associated with post-mitotic neurons (61). In contrast, no clear evidence for RAG-2 transcripts in the CNS could be found (61). It remains to be demonstrated what function, if any, RAG-1 plays in the CNS. The reciprocal situation of RAG-2 expression in the absence of RAG-1 expression occurs in the chicken bursa of Fabricius (55). Bursal B cells diversify their rearranged Ig heavy and light chain genes through a process of repeated, targeted gene conversion (67, 68). RAG-2 expression is strongly correlated with the gene conversion process in bursal B cells and in immortalized chicken B-cell lines, raising the possibility that RAG-2 plays a role in Ig-specific gene conversion (55).

Thus, while RAG-1 and RAG-2 together are involved in V(D)J recombination, the data suggest that the genes also function independently of one another, in processes as diverse as Ig gene conversion and perhaps neuronal development. In interpreting the data presented above, however, it is worth noting that no information is yet available concerning the

distribution or levels of the RAG proteins, since immunological reagents for their detection have yet to be developed.

Models for the Mechanism of Action of RAG-1 and RAG-2

Two formal possibilities exist for the mechanism by which RAG-1 and RAG-2 activate the V(D)J recombinase: first, that they directly encode lymphoid-specific components of the V(D)J recombinase, and second, that they function indirectly by activating other factors that compose the recombinase. A variety of data, summarized below, lend support to the first model:

1. Both RAG-1 and RAG-2 have been highly conserved through vertebrate evolution, as have the *cis*-acting DNA elements (RSSs) that mediate V(D)J recombination. Chicken RAG-1 is fully functional in mouse fibroblasts when co-transfected with mouse RAG-2, and the same is true of chicken RAG-2 in conjunction with mouse RAG-1 (unpublished results). In addition, it is worth noting that the RAG-1 and RAG-2 proteins have been evolutionarily conserved throughout their entire length (54, 55). The data are consistent with the idea that the RAG proteins are involved in multiple interactions with other components of the recombination reaction (perhaps both proteins and DNA), and that these interactions, like RSSs, have been conserved through evolution.

2. V(D)J recombinase activity is the only lymphoid-specific property found in fibroblasts expressing RAG-1 and RAG-2. If RAG-1 and RAG-2 were regulatory factors, they might be expected to trigger other lymphoid-specific activities in addition to the V(D)J recombinase. No such activities have been detected in fibroblasts expressing both genes (summarized in 5). Thus, RAG-1 and RAG-2 activate the enzyme for V(D)J recombination without activating other aspects of the normal lymphocyte developmental program.

3. The correlation of RAG-2 expression with avian Ig gene conversion provides a connection between RAG-2 and a second recombination process. Interestingly, Ig gene conversion, like V(D)J recombination, may be associated with heptamer-like elements (69).

4. The RAG-1 protein is homologous to the yeast topoisomerase-like factor, HPR1. Topoisomerases are associated with site-specific recombinases in a variety of other systems (70).

THE REGULATION OF GENE REARRANGEMENT

Ig and TCR gene rearrangements are lineage specific, occur in an ordered fashion, show preferences for using certain gene segments, and result in

allelic exclusion of antigen receptor expression. In this section, we describe these regulatory phenomena and assess various hypotheses regarding the mechanisms underlying the regulation of V(D) J recombination. This discussion is confined primarily to murine Ig genes.

The Ordered Nature of Ig Gene Rearrangement

Structural analyses of Ig gene loci in a large number of lymphoid tumors and virally transformed cell lines and immunocytochemical analysis of cells from normal bone marrow and fetal liver have led to the following description of the ordered nature of gene rearrangements (71).

D-TO-J_H REARRANGEMENT The first gene rearrangement during B cell development is heavy chain D-to-J_H rearrangement, which almost invariably occurs on both heavy-chain alleles. Interestingly, this Ig gene rearrangement is also found in a fraction of T cells (36, 37) and TCR β chain gene D-to-J_β rearrangements are found in some B cells (38 and M. S. Schlissel, unpublished results). DJ_H alleles are not rearranged any further in the T-cell lineage. Furthermore, a transgenic TCR gene rearrangement reporter construct undergoes D-to-J_β rearrangement in both lymphoid lineages (72). Two possible explanations for these observations are that D-to-J rearrangement precedes divergence of the B and T cell lineages or that the D-to-J rearrangement step is not lineage-restricted.

D_H gene segments are flanked on both sides by RSSs containing 12-nucleotide spacers, yet D-to-J rearrangement invariably uses the J-proximal RSS (73, 74). Furthermore, there is a strong preference for the first reading frame among D genes in the completely rearranged heavy chain genes of peripheral B cells (75). Rajewsky and colleagues have recently shown that this reading frame preference is due to three mechanisms: (i) short homologies at the ends of D_H and J_H skew the recombination reaction in favor of the first reading frame (34),; (ii) the presence of multiple in-frame stop codons in reading frame three, and (iii) negative section of cells using reading frame two due to the expression of D_μ protein (35). In contrast, there is relatively little bias in J_H-gene utilization (34).

V-TO-DJ_H REARRANGEMENT Precursor B cells then undergo V_H-to-DJ_H rearrangement on one or both DJ_H alleles, generating potentially functional VDJ_H heavy-chain genes. V_H-to-D_H rearrangement does not occur on alleles that have not first undergone D-to-J rearrangement (73, 74). In mice and humans, V_H-to-DJ_H rearrangement is characterized by a clear preference for the most J-proximal V_H gene segments (76, 77). Since V-to-DJ rearrangement is not precise, two thirds of these arrangements are "non-productive" out-of-frame joining events (VDJ^- alleles).

LIGHT-CHAIN GENE REARRANGEMENT Most often, the next rearrangement during B-cell development is light chain V-to-Jκ rearrangement. Ig κ locus rearrangements, however, occasionally precede productive heavy-chain gene rearrangement in normal human bone marrow (78), in Abelson virus-transformed pre-B cell lines (79), and in *scid* mice (80, 81). There is a significant preference for use of the first J gene, Jκ1, but this rearrangement, unlike heavy-chain V_H-to-DJ_H rearrangement, lacks a particular Vκ gene preference (82). A small fraction of B cells, often ones that have deleted one or both of their κ loci, then go on to rearrange some of their four λ light chain alleles (83).

The end-product of this developmental lineage is a mature B cell with a unique antigen receptor complex consisting of IgM and several accessory proteins on its cell surface (84). The observation that in a given cell only single heavy- and light-chain proteins are expressed on the cell surface is termed allelic exclusion (85). It has been proposed (71) that this is due to the feedback regulation of Ig gene rearrangement by the products of those rearranged genes, functional heavy- and light-chain proteins (discussed below).

The Accessibility Hypothesis

Several lines of indirect evidence (reviewed above) have led to the suggestion that a single recombinase catalyzes all V(D)J-type rearrangement events in B and T cells. If this is true, then recombinase activity must be regulated at the level of the template. The "accessibility hypothesis" has been proposed to explain regulated V(D)J recombination—the accessibility of a locus within the chromosome determines whether or when that locus is a template for the recombinase, and transcriptional activation of unrearranged gene segments may either cause or reflect that accessibility (71). Support for this notion comes from several types of experiments, described below.

GERMLINE TRANSCRIPTION AND REARRANGEMENT All rearranging loci in B cells are transcriptionally active at the time of their rearrangement (74, 86). Developmentally regulated "germline" Cμ, DJ_H, V_H, Cκ, and V_κ transcripts (Figure 3) have been characterized in cells carrying out the rearrangement of those genes (74, 87–90), and developmentally regulated transcripts from unrearranged TCR gene segments have also been detected (91).

It has been suggested that inactivation of rearrangement at a particular locus is correlated with inactivation of germline transcription, but there are conflicting data on this point. V_H and DJ_H transcripts either decrease (88, 89) or remain unchanged (74) in Abelson virus-transformed pre-B

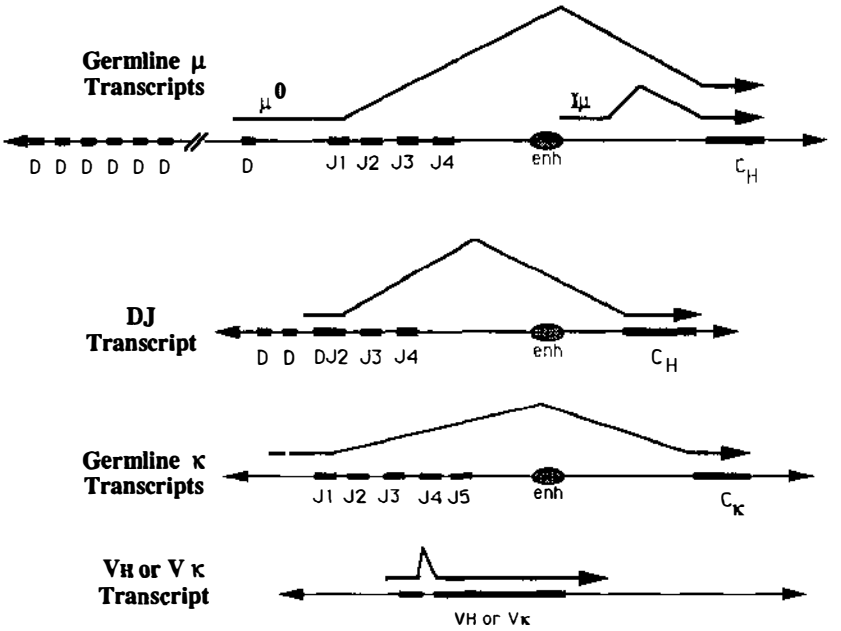

Figure 3 Structures of germline Ig gene transcripts. The locations of exons (thick lines) and introns (thin lines) are displayed above a schematic diagram of the relevant loci. The locations of the intronic μ and κ enhancers are indicated (enh). The figure is not drawn to scale.

cells of increasing maturity (i.e. those that have undergone productive V-to-DJ_H rearrangement).

Two recent studies have shown that treatments which activate germline gene transcription increase the frequency of gene rearrangement. The first study, which involved the κ locus, took advantage of the observation that germline Cκ transcripts are inducible in pre-B cells by bacterial lipopolysaccharide (LPS) (92), a treatment known to activate the κ enhancer binding transcription factor NF-κB (93). Abelson virus-transformed pre-B cells with nonproductive VDJ_H rearrangements on both chromosomes, cultured in the presence of LPS, showed significant increases in the levels of both germline C_κ transcription and V-to-J_κ rearrangement, leading to the suggestion that developmentally regulated activation of germline C_κ transcription might be responsible for the timing of onset of κ gene rearrangement (79).

The second experiment involved stable transfection and over-expression of the gene encoding transcription factor E47 in the pre-T cell line 2017. E47 can bind to two sites, μE2 and μE5, in the heavy-chain gene enhancer.

The Abelson-virus-induced pre-T cell lymphoma 2017 (94) has V(D)J recombinase activity (51) but rarely transcribes or rearranges the Ig heavy-chain locus (74). Overexpression of E47 resulted in the induction of the Iμ but not the μ^0 germline heavy chain gene transcript (Figure 3), and a striking 30 to 100-fold increase in the frequency of D-to-J_H rearrangement (95). The transcription of several other lymphoid genes, including RAG-1, RAG-2 and Oct-2, was also induced. These transfected pre-T cells did not, however, transcribe their DJ_H alleles or go on to rearrange V_H-to-DJ_H. These experiments provide further support for the role of transcription, or transcription factor binding, in the activation of gene rearrangement. These observations also place new limits on the nature of the relationship between transcription and gene rearrangement since the transcription associated with D-to-J_H rearrangement did not encompass the RSSs being rearranged.

Another piece of data in support of the accessibility hypothesis is the behaviour of the endogenous Ig genes in RAG-transfected fibroblasts. Despite the presence of an active V(D)J recombinase, the endogenous Ig genes, which are transcriptionally inert and presumably inaccessible, do not undergo rearrangement (D. Schatz, M. Oettinger, M. Schlissel, unpublished results).

TRANSCRIPTION AND REARRANGEMENT IN REPORTER CONSTRUCTS Transcriptional activity also affects the rearrangement activity of both transfected and transgenic reporter constructs. Selection for active transcription from a linked gene increased the frequency of reporter construct rearrangement (96). A transgenic rearrangement construct consisting of the Ig heavy-chain enhancer and constant region linked to V_β, D_β and J_β sequences underwent D-to-J_β rearrangement in both the thymus and spleen, but underwent V_β- to DJ_β rearrangement only in the thymus (97). The rearrangements detected in the thymus and spleen corresponded to the transcriptional patterns of the gene segments in those organs. Rearrangement was not detected, however, if the heavy chain enhancer was not present in the transgene. In contrast, a transgenic κ gene reporter construct underwent Vκ-to-Jκ rearrangement in the thymus despite not being transcribed there (98). Differences in the behavior of these constructs might be due to the exact nature of the sequences used, since DNA sequences that regulate transcription in vivo can be tens to hundreds of kb away (99).

A MECHANISTIC LINK BETWEEN TRANSCRIPTION AND REARRANGEMENT? One can only speculate about the mechanistic nature of the link between the activation of transcription and the targeting of the V(D)J recombinase. Transcription might reflect another, as yet undefined, event that makes Ig

genes accessible to both RNA polymerase and the V(D) J recombinase. It is possible, however, that the nascent germline transcript is involved in one of the steps of the rearrangement reaction. Alternatively, transcription might alter local structural characteristics of the DNA, for example by melting the helix and introducing compensatory superhelicity, making the RSSs better targets for rearrangement. Finally, transcriptional enhancers might actually function as rearrangement enhancers, able to activate V(D) J recombination over considerable distances by serving as binding sites for essential components of the recombinase.

In summary, a substantial body of evidence supports a role for template accessibility in the regulated activation of Ig and TCR gene rearrangement. The specific structural features of chromatin responsible for this accessibility remain to be determined as does the exact role, if any, of germline transcription in the rearrangement process.

Heavy Chain Protein as a Regulator of Gene Rearrangement

It has been suggested that heavy-chain protein serves as a signal for the cessation of heavy-chain gene rearrangement and the activation of κ light-chain gene rearrangement (71, 100). Likewise, it has been proposed that the complete IgM molecule expressed on the surface of a nascent B cell provides the signal that results in complete inactivation of the recombination machinery (71). We review here recent data regarding the potential role of heavy-chain protein and surface Ig in the temporally regulated activation and inactivation of recombinase activity.

HEAVY-CHAIN PROTEIN AS AN ACTIVATOR OF κ GENE REARRANGEMENT

There are conflicting data on the role of heavy-chain protein in the activation of light-chain gene rearrangement. It was reported that transfection of a functional heavy-chain gene into a pre-B cell line significantly increased the frequency of endogenous κ gene rearrangement (100). A subsequent report described the failure to reproduce this result (81). Several lines of evidence demonstrate that heavy-chain protein is not necessary for κ gene rearrangement. First, Abelson virus–transformed pre-B cells with the heavy-chain genotype VDJ^-/VDJ^- lack heavy-chain protein yet still rearrange their κ genes at a detectable rate (79). Second, Abelson lines from *scid* mice in which both heavy-chain loci have been deleted still activate κ gene rearrangement (80, 81). The transfection of a functional heavy-chain gene into these cells did not affect the frequency of κ gene rearrangement. Finally, human bone marrow cells that stain with anti-κ antibody but not with anti–heavy chain antibody have been detected (78). EBV transformation of these cells and subsequent DNA analysis reveals V-to-Jκ rearrangements in clones lacking VDJ^+ alleles. The absence of a

demonstrable effect of heavy-chain protein on κ gene rearrangement in transformed cell lines does not rule out the possibility that it might modulate the process in vivo.

HEAVY-CHAIN PROTEIN AS AN INHIBITOR OF HEAVY-CHAIN GENE REARRANGEMENT Transgenic experiments have provided the strongest evidence in favor of a role for μ protein as a mediator of heavy-chain allelic exclusion. Experiments involving transgenic Ig (101) and TCR genes (65) have been reviewed recently. This discussion focuses on experiments involving heavy-chain gene transgenics. Transgenic mice that express the membrane (μ_m) but not the secreted (μ_s) form of μ protein show varying but significant extents of endogenous heavy-chain gene allelic exclusion. Between 10% and 90% or more of the heavy-chain loci in peripheral B cells, B-cell hybridomas, and Abelson virus–transformed pre-B cell lines from μ transgenic mice are unrearranged or only partially rearranged (102–104). Analogous results have been obtained with TCR β-chain transgenic mice (105). The extent of allelic exclusion is associated with the level of transgene expression (101).

These experiments have been interpreted as demonstrating a direct feedback-inhibitory role of μ_m on heavy-chain gene rearrangement. Other explanations of these observations are possible, however. For example, the effect of μ_m on heavy-chain allelic exclusion might be due to its function as a positive regulator of cell survival and putative activator of κ gene rearrangement. In support of this notion, Rajewsky and colleagues have recently shown that μ_m is necessary for B-cell development since homologous recombinant mice homozygous for a disruption of the μ membrane exon did not produce mature B cells (106). Productive V-to-DJ_H rearrangement is a relatively rare occurrence in nontransformed bone marrow B-cell progenitors (107). If in a transgenic animal 100% of developing B cells successfully express μ_m, the variable degrees of allelic exclusion of endogenous heavy-chain gene rearrangement might be due to stimulated rates of κ rearrangement and subsequent complete inactivation of recombinase in μ^+/κ^+ cells rather than specific inactivation of heavy-chain locus rearrangement.

Also in support of this alternative explanation for heavy-chain allelic exclusion is the observation that virally transformed pre-B cells do not show feedback inactivation of heavy-chain gene rearrangement. Experiments have shown no difference in the rates of ongoing heavy-chain gene rearrangement in Abelson pre-B cell lines with the heavy-chain locus genotypes VDJ+/DJ and VDJ−/DJ (74). Complicating the interpretation of these experiments, however, is the fact that Abelson virus–transformed pre-B cells rarely express μ protein on the cell surface and that requisite extracellular signals are probably missing in these cell cultures.

If μ protein does play a signalling role during early B-cell development, it likely does so via a cell-surface complex. μ protein is expressed on the cell surface in association with so-called "surrogate" light chains, the products of the V-pre-B and $\lambda 5$ genes (108, 109). These genes are homologous to the variable and constant regions of λ light-chain genes, but do not undergo gene rearrangement (110). It has been proposed that this complex might be involved in the signalling function of μ in cells that have yet to rearrange their light-chain genes (108, 109). Complicating assessments of the function of this cell-surface complex during development is the observation that most bone marrow cells that express it also express normal IgM (109).

The Inactivation of V(D)J Recombination

Mature B and T cell lines lack measurable V(D)J recombinase activity (51) and do not express detectable levels of RAG-1 or RAG-2 mRNA (5, 54). This inactivation of the recombination machinery in mature cells is presumably necessary to prevent the production of unselected antigen receptor specificities and decrease the likelihood of aberrant rearrangements involving oncogenic loci. It has been proposed that allelic exclusion of light-chain gene expression and overall inactivation of the V(D)J recombinase results from feedback inhibition by surface Ig. Data from light-chain only and heavy-plus-light chain transgenic experiments support this hypothesis (101). As discussed above, however, more recent evidence suggests that in both the T- and B-cell lineages there is a lag between surface expression of TCR or Ig and the inactivation of the recombinase, suggesting that other developmental factors might be involved in the inactivation of the recombinase.

Regulation of V(D)J Recombination During B-Cell Development: A Model

What follows is a speculative model for the regulation of gene rearrangement during B-cell development (Figure 4). It incorporates features of the accessibility model, but proposes an alternative explanation for allelic exclusion.

The developmentally regulated activation of unrearranged Ig gene transcription targets the different Ig-gene segments for rearrangement. D-to-J_H rearrangement is unique in that it is very efficient, almost invariably occurring on both chromosomes before the onset of either V-to-DJ_H rearrangement or V-to-Jκ rearrangement, and does not allow for direct V_H-to-D_H or inversional D_H-to-J_H rearrangement. For these reasons, we suggest that a specific accessory protein is involved in this class of

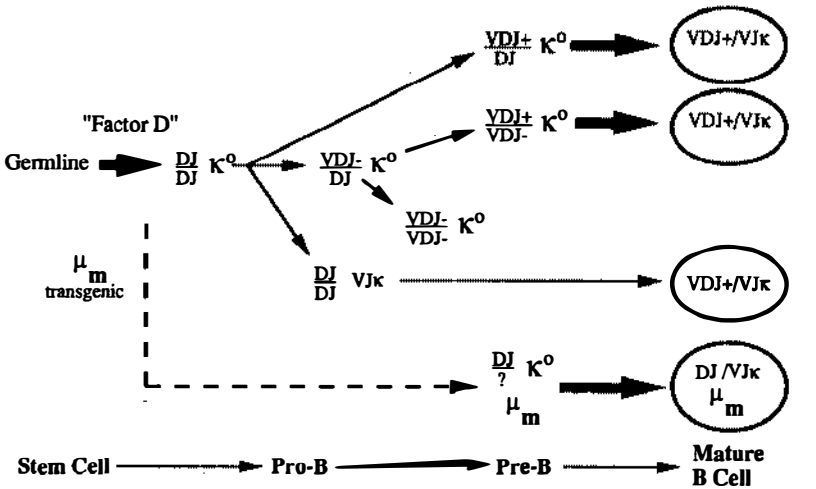

Figure 4 The regulation of Ig gene rearrangement: a model. Inefficient rearrangement is depicted by slender arrows, and efficient rearrangement by the thicker arrows. VDJ^+ and VDJ^- indicate productively and nonproductively rearranged heavy-chain genes respectively, while K^0 represents the unrearranged κ loci. The dashed line indicates what might happen in a μ_m transgenic mouse, where μ protein would induce efficient κ rearrangement in all B-cell progenitors. Factor D is a hypothetical DJ-specific rearrangement factor. The genotypes enclosed within circles are those capable of encoding complete IgM. Only illustrative genotypes are shown. See the text for further explanation.

rearrangements (Factor D in Figure 4). We propose that, initially, both V-to-DJ_H and V-to-Jκ rearrangement are inefficient processes. However, once a cell undergoes productive V-to-DJ_H rearrangement and produces μ_m, a signalling pathway, possibly operating through the μ/surrogate light-chain complex, significantly increases the frequency of κ gene rearrangement. This would account for the seemingly ordered nature of gene rearrangement and the occurrence of $\kappa+/\mu-$ cells. The inefficiency of V-to-DJ_H rearrangement would make it unlikely to occur a second time in a productive fashion before activated V-to-Jκ rearrangement occurs. Light-chain synthesis would then result in expression of cell-surface Ig leading to the complete inactivation of the V(D)J recombinase, accounting for both heavy-chain and light-chain allelic exclusion.

ACKNOWLEDGMENTS

We would like to thank the following people for sharing with us their unpublished data: F. Alt, G. Rathbun, A. Ma, T. Honjo, H. Klein, L. A. Steiner, and P. Higgins. We also thank S. Ghosh and K. Struhl for critical

reading of and helpful comments on the manuscript. D. G. Schatz is supported by the Howard Hughes Medical Institute.

Literature Cited

1. Blackwell, T. K., Alt, F. W. 1988. Immunoglobulin genes. In *Molecular Immunology*, ed. B. D. Hames, D. M. Glover, pp. 1–60. Washington, DC: IRL
2. Davis, M. M. 1988. T cell antigen receptor genes. In *Molecular Immunology*, ed. B. D. Hames, D. M. Glover, pp. 61–79. Washington, DC: IRL
3. Raulet, D. H. 1989. The structure, function, and molecular genetics of the γ/δ T cell receptor. *Annu. Rev. Immunol.* 7: 175–207
4. Toncgawa, S. 1983. Somatic generation of antibody diversity. *Nature* 302: 575–81
5. Schatz, D. G., Oettinger, M. A., Baltimore, D. 1989. The V(D)J recombination activating gene (RAG-1). *Cell* 59: 1035–48
6. Aguilera, R. J., Akira, S., Okazaki, K., Sakano, H. 1987. A pre-B cell nuclear protein which specifically interacts with the immunoglobulin V-J recombination sequences. *Cell* 51: 909–17
7. Hesse, J. E., Lieber, M. R., Mizuuchi, K., Gellert, M. 1989. V(D)J recombination: a functional definition of the joining signals. *Genes Dev.* 3: 1053–61
8. Kleinfield, R., Hardy, R. R., Tarlinton, D., Dangl, J., Herzenberg, L. A., Weigert, M. 1986. Recombination between an expressed immunoglobulin heavy-chain gene and a germline variable gene segment in a Ly 1^+ B-cell lymphoma. *Nature* 322: 843–46
9. Reth, M. G., Jackson, S., Alt, F. W. 1986. V_HDJ_H formation and DJ_H replacement during pre-B differentiation: non-random usage of gene segments. *EMBO J.* 5: 2131–38
10. Durdik, J., Moore, M. W., Selsing, E. 1984. Novel κ light-chain gene rearrangements in mouse λ light chain-producing B lymphocytes. *Nature* 307: 749–52
11. Stavnezer, J., Kekish, O., Batter, D., Grenier, J., Balazs, I., Henderson, E., Zegers, B. J. M. 1985. Aberrant recombination events in B cell lines derived from κ-deficient human. *Nucl. Acids Res.* 13: 3495–514
12. Hesse, J. E., Lieber, M. R., Gellert, M., Mizuuchi, K. 1987. Extrachromosomal DNA substrates in pre-B cells undergo inversion or deletion at immunoglobulin V-(D)-J joining signals. *Cell* 49: 775–83
13. Lieber, M. R., Hesse, J. E., Mizuuchi, K., Gellert, M. 1988. Lymphoid V(D) J recombination: nucleotide insertion at signal joints as well as coding joints. *Proc. Natl. Acad. Sci. USA* 85: 8588–92
14. Malissen, M., McCoy, C., Blanc, D., Trucy, J., Devaux, C., Schmitt-Verhulst, A.-M., Fitch, F., Hood, L., Malissen, B. 1986. Direct evidence for chromosomal inversion during T-cell receptor β-gene rearrangements. *Nature* 319: 28–33
15. Korman, A. J., Maruyama, J., Raulet, D. H. 1989. Rearrangement by inversion of a T-cell receptor δ variable region gene located 3′ of the δ constant region gene. *Proc. Natl. Acad. Sci. USA* 86: 267–71
16. Feddersen, R. M., van Ness, B. G. 1985. Double recombination of a single immunoglobulin κ-chain allele: implication for the joining mechanism. *Proc. Natl. Acad. Sci. USA* 82: 4793–97
17. Lewis, S., Rosenberg, N., Alt, F., Baltimore, D. 1982. Continuing kappa-gene rearrangement in a cell line transformed by Abelson murine leukemia virus. *Cell* 30: 807–16
18. Lorenz, W., Straubinger, B., Zachau, H. G. 1987. Physical map of the human immunoglobulin κ locus and its implications for the mechanisms of Vκ-Jκ rearrangement. *Nucl. Acids Res.* 15: 9667–76
19. Fujimoto, S., Yamagishi, H. 1987. Isolation of an excision product of T-cell-receptor α-chain gene rearrangement. *Nature* 327: 242–43
20. Winoto, A., Baltimore, D. 1989. Separate lineages of T cells expressing the $\alpha\beta$ and $\gamma\delta$ receptors. *Nature* 338: 430–32
21. Lewis, S. M., Hesse, J. E., Mizuuchi, K., Gellert, M. 1988. Novel strand exchanges in V(D)J recombination. *Cell* 55: 1099–1107
22. Lieber, M. R., Hesse, J. E., Lewis, S., Bosma, G. C., Rosenberg, N., Mizuuchi, K., Bosma, M. J., Gellert, M. 1988. The defect in murine severe combined immune deficiency: Joining

of signal sequences but not coding segments in V(D) J recombination. *Cell* 55: 7–16
23. Morzycka-Wroblewska, E., Lee, F. E. H., Desiderio, S. V. 1988. Unusual immunoglobulin gene rearrangement leads to replacement of recombinational signal sequences. *Science* 242: 261–63
24. Lipkowitz, S., Stern, M.-H. 1990. Hybrid T cell receptor genes formed by interlocus recombination in normal and Ataxia-telangectasia lymphocytes. *J. Exp. Med.* 172: 409–18
25. Tycko, B., Palmer, J. D., Sklar, J. 1989. T cell receptor gene trans-rearrangements: chimeric γ-δ genes in normal lymphoid tissues. *Science* 245: 1242–46
26. Lewis, S., Gifford, A., Baltimore, D. 1985. DNA elements are asymmetrically joined during the site-specific recombination of kappa immunoglobulin genes. *Science* 228: 677–85
27. Schuler, W. 1990. The scid mouse mutant: biology and nature of the defect. In *Molecular Basis of B Cell Developments. Cytokines*, ed. C. Sorg, 3: 132–53. Basel: S. Karger
28. Schuler, W., Weiler, I. J., Schuler, A., Phillips, R. A., Rosenberg, N., Mak, T. K., Kearney, J. F., Perry, R. P., Bosma, M. J. 1986. Rearrangement of antigen receptor genes is defective in mice with severe combined immune deficiency. *Cell* 46: 963–72
29. Hendrickson, E. A., Schatz, D. G., Weaver, D. T. 1988. The *scid* gene encodes a *trans*-acting factor that mediates the rejoining event of Ig gene rearrangement. *Genes Dev.* 2: 817–29
30. Lafaille, J. J., DeCloux, A., Bonneville, M., Takagaki, Y., Tonegawa, S. 1989. Junctional sequences of T cell receptor gamma delta genes: implications for gamma delta T cell lineages and for a novel intermediate of V-(D)-J joining. *Cell* 59: 859–70
31. McCormack, W. T., Tjoelker, L. W., Carlson, L. M., Petryniak, B., Barth, C. F., Humphries, E. H., Thompson, C. B. 1989. Chicken IgL gene rearrangement involves deletion of a circular episome and addition of single nonrandom nucleotides to both coding segments. *Cell* 56: 785–91
32. Schuler, W., Ruetsch, N. R., Amsler, M., Bosma, M. J. 1991. Coding joint formation of endogenous T cell receptor genes in lymphoid cells from scid mice: unusual P-nucleotide additions in VJ-coding joints. *Eur. J. Immunol.* 21: 589–96
33. Lewis, S., Gellert, M. 1989. The mechanism of antigen receptor gene assembly. *Cell* 59: 585–88
34. Gu, H., Forster, I., Rajewsky, K. 1990. Sequence homologies, N sequence insertion and J_H gene utilization in V_HDJ_H joining: implications for the joining mechanism and the ontogenetic timing of Ly1 B cell and B-CLL progenitor generation. *EMBO J.* 9: 2133–40
35. Gu, H., Kitamura, D., Rajewsky, K. 1991. B cell development regulated by gene rearrangement: arrest of maturation by membrane-bound Dμ protein and selection of D_H element reading frames. *Cell* 65: 47–54
36. Forster, A., Hobart, M., Hengartner, H., Rabbitts, T. H. 1980. An immunoglobulin heavy chain gene is altered in two T cell clones. *Nature* 286: 897–99
37. Kurosawa, Y., Von Boehmer, H., Hass, W., Sakano, H., Traunecker, A., Tonegawa, S. 1981. Identification of D segments of immunoglobulin heavy chain genes and their rearrangement in T lymphocytes. *Nature* 290: 565–70
38. Cook, W. D., Balaton, A. M. 1987. T-cell receptor and immunoglobulin genes are rearranged together in Abelson virus-transformed pre-B and pre-T cells. *Mol. Cell. Biol.* 7: 266–72
39. Bosma, G. C., Custer, R. P., Bosma, M. J. 1983. A severe combined immunodeficiency mutation in the mouse. *Nature* 301: 527–30
40. Fulop, G. M., Phillips, R. A. 1990. The *scid* mutation in mice causes a general defect in DNA repair. *Nature* 347: 479–82
41. Biedermann, K. A., Sun, J. R., Giaccia, A. J., Tosto, L. M., Brown, J. M. 1991. *scid* mutation in mice confers hypersensitivity to ionizing radiation and a deficiency in DNA double-strand break repair. *Proc. Natl. Acad. Sci. USA* 88: 1394–97
42. Hendrickson E. A., Qin, X.-Q., Bump, E. A., Schatz, D. G., Oettinger, M. A., Weaver, D. T. 1991. A link between double-strand break repair and V(D) J recombination: the *scid* mutation. *Proc. Natl. Acad. Sci. USA* 88: 4061–65
43. Landau, N. R., Schatz, D. G., Rosa, M., Baltimore, D. 1987. Increased frequency of N-regional insertion in a murine pre-B-cell line infected with a terminal deoxynucleotidyl transferase retroviral expression vector. *Mol. Cell. Biol.* 7: 3237–43
44. Roth, D. B., Chang, X. C., Wilson, J.

H. 1989. Comparison of filler DNA at immune, nonimmune, and oncogenic rearrangements suggests multiple mechanisms of formation. *Mol. Cell. Biol.* 9: 3049–57

45. Hamaguchi, Y., Matsunami, N., Yamamoto, Y., Honjo, T. 1989. Purification and characterization of a protein that binds to the recombination signal sequence of the immunoglobulin J kappa segment. *Nucl. Acids Res.* 17: 9015–26
46. Halligan, B. D., Desiderio, S. V. 1987. Identification of a DNA binding protein that recognizes the nonamer recombinational signal sequence of immunoglobulin genes. *Proc. Natl. Acad. Sci. USA* 84: 7019–23
47. Desiderio, S., Baltimore, D. 1984. Double-stranded cleavage by cell extracts near recombinational signal sequences of immunoglobulin genes. *Nature* 308: 860–62
48. Kataoka, T., Kondo, S., Nishi, M., Kodaira, M., Honjo, T. 1984. Isolation and characterization of endonuclease J: a sequence-specific endonuclease cleaving immunoglobulin genes. *Nucl. Acids Res.* 12: 5995–6010
49. Hope, T. J., Aguilera, R. J., Minie, M. E., Sakano, H. 1986. Endonucleolytic activity that cleaves immunoglobulin recombination sequences. *Science* 231: 1141–45
50. Matsunami, N., Hamaguchi, Y., Yamamoto, Y., Kuze, K., Kanagawa, K., Matsuo, H., Kawaichi, M., Honjo, T. 1989. A protein binding to the J kappa recombination sequence of immunoglobulin genes contains a sequence related to the integrase motif. *Nature* 342: 934–37
51. Lieber, M. R., Hesse, J. E., Mizuuchi, K., Gellert, M. 1987. Developmental stage specificity of the lymphoid V(D)J recombination activity. *Genes Devel.* 1: 751–61
52. Desiderio, S. V., Wolff, K. R. 1988. Rearrangement of exogenous immunoglobulin V_H and DJ_H gene segments after retroviral transduction into immature lymphoid cell lines. *J. Exp. Med.* 167: 372–88
53. Schatz, D. G., Baltimore, D. 1988. Stable expression of immunoglobulin gene V(D)J recombinase activity by gene transfer into 3T3 fibroblasts. *Cell* 53: 107–15
54. Oettinger, M. A., Schatz, D. G., Gorka, C., Baltimore, D. 1990. RAG-1 and RAG-2, adjacent genes that synergistically activate V(D)J recombination. *Science* 248: 1517–23
55. Carlson, L. M., Oettinger, M. A., Schatz, D. G., Masteller, E. L., Hurley, E. A., McCormack, W. T., Baltimore, D., Thompson, C. B. 1991. Selective expression of RAG-2 in chicken B cells undergoing immunoglobulin gene conversion. *Cell* 64: 201–8
56. Oettinger, M. A., Stanger, B., Schatz, D. G., Glaser, T., Call, K., Housman, D., Baltimore, D. 1991. The recombination activating genes, RAG-1 and RAG-2, are on chromosome 11p in humans and chromosome 2 in mice. *Immunogenetics.* In press
57. Bosma, G. C., Davisson, M. T., Ruetsch, N. R., Sweet, H. O., Shultz, L. D., Bosma, M. J. 1989. The mouse mutation severe combined immune deficiency (scid) is on chromosome 16. *Immunogenetics* 29: 54–57
58. Freemont, P. S., Hanson, I. M., Trowsdale, J. 1991. A novel cysteine-rich sequence motif. *Cell* 64: 483–84
59. Wang, J. C., Caron, P. R., Kim, R. A. 1990. The role of DNA topoisomerases in recombination and genome stability: a double-edged sword? *Cell* 62: 403–6
60. Aguilera, A., Klein, H. L. 1990. HPR1, a novel yeast gene that prevents intrachromosomal excision recombination, shows carboxy-terminal homology to the *Saccharomyces cerevisiae* TOP1 gene. *Mol. Cell. Biol.* 10: 1439–51
61. Chun, J. J., Schatz, D. G., Oettinger, M. A., Jaenisch, R., Baltimore, D. 1991. The recombination activating gene-1 (RAG-1) transcript is present in the murine central nervous system. *Cell* 64: 189–200
62. Turka, L. A., Schatz, D. G., Oettinger, M. A., Chun, J. J. M., Gorka, C., Lee, K., McCormack, W. T., Thompson, C. B. 1991. Thymocyte expression of the recombination activating genes RAG-1 and RAG-2 can be terminated by T-cell receptor cross-linking. *Science* 253: 778–81
63. Blackman, M., Kappler, J., Marrack, P. 1990. The role of the T cell receptor in positive and negative selection of developing T cells. *Science* 248: 1335–41
64. von Boehmer, H., Kisielow, P. 1990. Self-nonself discrimination by T cells. *Science* 248: 1369–73
65. von Boehmer, H. 1990. Developmental biology of T cells in T cell-receptor transgenic mice. *Annu. Rev. Immunol.* 8: 531–56
66. Menetski, J., Gellert, M. 1990. V(D)J recombination activity in lymphoid cell lines is increased by agents that elevate cAMP. *Proc. Natl. Acad. Sci. USA* 87: 9324–28

67. McCormack, W. T., Thompson, C. B. 1990. Somatic diversification of the chicken immunoglobulin light-chain gene. *Adv. Immunol.* 48: 41–67
68. Reynaud, C. A., Dahan, A., Anquez, V., Weill, J. C. 1989. Somatic hyperconversion diversifies the single V_H gene of the chicken with a high incidence in the D region. *Cell* 59: 171–83
69. McCormack, W. T., Carlson, L. M., Tjoelker, L. W., Thompson, C. B. 1989. Evolutionary comparison of the avian IgL locus: combinatorial diversity plays a role in the generation of the antibody repertoire in some avian species. *Int. Immunol.* 1: 332–41
70. Wang, J. C. 1985. DNA topoisomerases. *Annu. Rev. Biochem.* 54: 665–97
71. Alt, F., Blackwell, T., Yancopoulos, G. 1987. Development of the primary antibody repertoire. *Science* 238: 1079–87
72. Ferrier, P., Krippl, B., Furley, A. J., Blackwell, T. K., Suh, H., Mendelsohn, M., Winoto, A., Cook, W. D., Hood, L., Costantini, F., Alt, F. W. 1989. Control of VDJ recombinase activity. *Cold Spring Harbor Symp. Quant. Biol.* 54: 191–202
73. Alt, F. W., Yancopoulos, G. D., Blackwell, T. K., Wood, C., Thomas, E., Boss, M., Coffman, R., Rosenberg, N., Tonegawa, S., Baltimore, D. 1984. Ordered rearrangement of immunoglobulin heavy chain variable region segments. *EMBO J.* 3: 1209–19
74. Schlissel, M. S., Corcoran, L. M., Baltimore, D. 1991. Virus-transformed pre-B cells show ordered activation but not inactivation of immunoglobin gene rearrangement and transcription. *J. Exp. Med.* 173: 711–20
75. Kaartinen, M., Makela, O. 1985. Reading of D genes in variable frames as a source of diversity. *Immunol. Today* 6: 324–30
76. Malynn, B. A., Yancopoulos, G. D., Barth, J. E., Bona, C. A., Alt, F. W. 1990. Biased expression of the J_H-proximal V_H genes occurs in the newly generated repertoire of neonatal and adult mice. *J. Exp. Med.* 171: 843–59
77. Yancopoulos, G., Desiderio, S., Paskind, M., Kearney, J., Baltimore, D., Alt, F. 1984. Preferential utilization of the most J_H proximal V_H gene segments in pre-B-cell lines. 311: 727–33
78. Kubagawa, H., Cooper, M. D., Carroll, A. J., Burrows, P. D. 1989. Light-chain gene expression before heavy-chain gene rearrangement in pre-B cells transformed by Epstein-Barr virus. *Proc. Natl. Acad. Sci. USA* 86: 2356–60
79. Schlissel, M. S., Baltimore, D. 1989. Activation of immunoglobulin kappa gene rearrangement correlates with induction of germline kappa gene transcription. *Cell* 58: 1001–7
80. Hendrickson, E. A., Schlissel, M. S., Weaver, D. T. 1990. Wild-type V(D) J recombination in scid pre-B cells. *Mol. Cell. Biol.* 10: 5397–407
81. Blackwell, K., Malynn, B., Pollock, R., Ferrier, P., Covey, L., Fulop, G., Phillips, R., Yancopoulos, G., Alt, F. 1989. Isolation of scid pre-B cells that rearrange kappa light chain genes: formation of normal signal and abnormal coding joins. *EMBO J.* 8: 735–42
82. Lawler, A. M., Kearney, J. F., Kuehl, M., Gearhart, P. J. 1989. Early rearrangements of genes encoding murine immunoglobulin κ chains, unlike genes encoding heavy chains, use variable gene segments dispersed throughout the locus. *Proc. Natl. Acad. Sci. USA* 86: 6744–47
83. Selsing, E., Durdik, J., Moore, M. W., Persiani, D. M. 1989. Immunoglobulin λ genes. In *Immunoglobulin Genes*, ed. T. Honjo, F. W. Alt, T. H. Rabbitts, pp. 111–112. London: Academic
84. Reth, M., Hombach, J., Wienands, J., Cambell, K. S., Chien, N., Justement, L. B., Cambier, J. C. 1991. The B-cell antigen receptor complex. *Immunol. Today* 12: 196–201
85. Pernis, B. G., Chiappino, G., Kelus, A. S., Gell, P. G. H. 1965. Cellular localization of immunoglobulins with different allotype specificities in rabbit lymphoid tissue. *J. Exp. Med.* 122: 853–75
86. Lennon, G. G., Perry, R. P. 1990. The temporal order of appearance of transcripts from unrearranged and rearranged Ig genes in murine fetal liver. *J. Immunol.* 144: 1983–87
87. Lennon, G. G., Perry, R. P. 1985. Cμ-containing transcripts initiate heterogenously with the IgH enhancer region and contain a novel 5′-nontranslatable exon. *Nature* 318: 475–78
88. Yancopoulos, G. D., Alt, F. W. 1985. Developmentally controlled and tissue-specific expression of unrearranged V_H gene segments. *Cell* 40: 271–81
89. Reth, M., Alt, F. 1984. Novel immunoglobulin heavy chains are produced from DJ_H gene segment rearrangements in lymphoid cells. 312: 418–23
90. Van Ness, B., Weigert, M., Coleclough, C., Mather, E., Kelley, D., Perry, R. 1981. Transcription of the unre-

arranged mouse Cκ locus: Sequence of the initiation region and comparison of activity with a rearranged Vκ-Cκ gene. *Cell* 27: 593–602

91. Pardoll, D. M., Fowlkes, B. J., Lechler, R. I., Germain, R. N., Schwartz, R. H. 1987. Early genetic events in T cell development analyzed by in situ hybridization. *J. Exp. Med.* 165: 1624–38
92. Nelson, K., Mather, E., Perry, R. 1984. Lipopolysaccharide-induced transcription of the kappa immunoglobulin locus occurs on both alleles and is independent of methylation status. *Nucl. Acids Res.* 12: 1911–23
93. Baeuerle, P. A., Baltimore, D. 1991. The physiology of the NF-κB transcription factor. In *The Hormonal Regulation of Transcription*, ed. P. Cohen, J. G. Foulkes, VI: 409–432. Amsterdam: Elsevier Sci.
94. Spolski, R., Miescher, G., Erard, R., Risser, R., MacDonald, H. R., Mak, T. W. 1988. Regulation of expression of T cell gamma chain, L3T4 and Ly-2 messages in Abelson/Moloney virus-transformed T cell lines. *Eur. J. Immunol.* 18: 295–300
95. Schlissel, M., Voronova, A., Baltimore, D. 1991. Helix-loop-helix transcription factor E47 activates germline immunoglobulin heavy-chain gene transcription and rearrangement in a pre-T cell line. *Genes Dev.* 5: 1367–76
96. Blackwell, T. K., Moore, M. W., Yancopoulos, G. D., Suh, H., Lutzker, S., Selsing, E., Alt, F. W. 1986. Recombination between immunoglobulin variable region gene segments is enhanced by transcription. *Nature* 324: 585–89
97. Ferrier, P., Krippl, B., Blackwell, T. K., Furley, A. J., Suh, H., Winoto, A., Cook, W. D., Hood, L., Costantini, F., Alt, F. W. 1990. Separate elements control DJ and VDJ rearrangement in a transgenic recombination substrate. *EMBO J.* 9: 117–25
98. Goodhardt, M., Babinet, C., Lutfalla, G., Kallenbach, S., Cavelier, P., Rougeon, R. 1989. Immunoglobulin κ light chain gene promoter and enhancer are not responsible for B-cell restricted gene rearrangement. *Nucl. Acids Res.* 17: 7403–15
99. Townes, T. M., Behringer, R. R. 1990. Human globin locus activation region (LAR): role in temporal control. *Trends Genet.* 6: 219–23
100. Reth, M., Petrac, E., Wiese, P., Lobel, L., Alt, F. 1987. Activation of Vκ gene rearrangement in pre-B cells follows the expression of membrane-bound immunoglobulin heavy chains. *EMBO J.* 6: 3299–305
101. Storb, U. 1987. Transgenic mice with immunoglobulin genes. *Annu. Rev. Immunol.* 5: 151–74
102. Muller, W., Ruther, U., Viera, P., Hombach, J., Reth, M., Rajewsky, K. 1989. Membrane-bound IgM obstructs B cell development in transgenic mice. *Eur. J. Immunol.* 19: 923–28
103. Nussenzweig, M. C., Shaw, A. C., Sinn, E., Danner, D. B., Holmes, K. L., Morse, H. C., Leder, P. 1987. Allelic exclusion in transgenic mice that express the membrane form of immunoglobulin M. *Science* 236: 816–19
104. Weaver, D., Reis, M. H., Albanese, C., Costantini, F., Baltimore, D., Imanishi-Kari, T. 1986. Altered repertoire of endogenous immunoglobulin gene expression in transgenic mice containing a rearranged mu heavy chain gene. *Cell* 45: 247–59
105. Uematsu, Y., Ruser, S., Dembic, Z., Borgulya, P., Krimpenfort, P., Berns, A., von Boehmer, H., Steinmetz, M. 1988. In transgenic mice the introduced functional T cell receptor β gene prevents expression of endogenous β genes. *Cell* 52: 831–41
106. Kitamura, D., Roes, J., Kuhn, R., Rajewsky, K. 1991. A B cell-deficient mouse by targeted disruption of the membrane exon of the immunoglobulin μ chain gene. *Nature* 350: 423–26
107. Hardy, R., Carmack, C. E., Shinton, S. A., Kemp, J. D., Hayakawa, K. 1991. Resolution and characterization of pro-B and pre-pro-B cell stages in normal mouse bone marrow. *J. Exp. Med.* 173: 1213–25
108. Melchers, F. 1989. The many roles of immunoglobulin molecules in the development and growth control of the B-lymphocyte lineage. In *Immunoglobulin Genes*, ed. T. Honjo, F. W. Alt, T. H. Rabbitts, pp. 23–44. London: Academic
109. Cherayil, B. J. Pillai, S. 1991. The ω/λ5 surrogate immunoglobulin light chain is expressed on the surface of transitional B lymphocytes in murine bone marrow. *J. Exp. Med.* 173: 111–16
110. Melchers, F., Strasser, A., Bauer, S. R., Kudo, A., Thalmann, P., Rolink, A. 1989. Cellular stages and molecular steps of murine B-cell development. *Cold Spring Harbor Symp. Quant. Biol.* 54: 183–89