Mechanism and control of class-switch recombination

John P. Manis, Ming Tian and Frederick W. Alt

Class-switch recombination (CSR) occurs by an unusual and intriguing mechanism that has not been clearly elucidated as yet. Currently, we know that this mechanism involves recombination between large and highly repetitive switch (S) regions, is targeted by S-region transcription and requires the activity of the newly discovered activation-induced deaminase (AID). In this review, we discuss the potential role of these factors in CSR, discuss potential relationships between CSR and somatic hypermutation, and speculate how CSR and related mechanisms might contribute to genomic instability.

The production of specific antibodies by mature B cells involves up to three separate types of Ig gene alterations. During early B-cell development, Ig heavy-chain (IgH) and Ig light-chain (IgL) variable-region exons are assembled from component germline variable (V), diversity (D) and joining (J) gene segments by V(D)J recombination [1]. In response to antigen, mature B cells can change their expressed IgH constant region (C_H), and associated effector function, by a distinct recombination process, termed IgH class-switch recombination (CSR) [2–4]. At this stage, variable-region gene exon sequences can be further altered by somatic hypermutation (SM), potentially generating higher affinity antibodies [5]. Whereas V(D)J recombination has been well-characterized, much less is known about CSR. In this review, we discuss recent advances in our understanding of the mechanism and control of CSR. In particular, we focus on the crucial, but as yet not fully characterized, roles of transcription and the recently identified activation-induced deaminase (AID), a novel protein that is required absolutely for CSR [6,7]. In addition, we outline the substantial mechanistic contrasts between V(D)J recombination and CSR, discuss potential mechanistic overlaps between CSR and SM, and speculate how CSR and related mechanisms might be linked to genomic instability and cancer. Finally, although this review focuses primarily on CSR in the mouse, much of what we discuss applies generally also to CSR in humans.

John P. Manis
Ming Tian
Frederick W. Alt*
Howard Hughes Medical Institute and Children's Hospital, Center for Blood Research and Dept of Genetics, Harvard Medical School, Boston, MA 02115, USA.
*e-mail: alt@rascal.med.harvard.edu

Class-switch recombination and V(D)J recombination employ distinct mechanisms

V(D)J recombination is initiated in progenitor B lymphocytes by recombinase-activating gene 1 (RAG-1) and RAG-2 proteins, which introduce DNA double-strand breaks (DSBs) precisely between target recombination signal sequences (RSSs) and flanking V, D or J coding segments [1]. Targeted inactivation of *Rag1* or *Rag2* in mice results in a complete lack of mature B and T cells, owing to their inability to initiate V(D)J recombination. Joining of RAG-liberated coding sequences and RSSs is carried out by ubiquitously expressed, nonhomologous end-joining (NHEJ) proteins, which are employed also for the general repair of DSBs [8]. Three of these proteins are sub-units of the DNA-dependent protein kinase (DNA-PK), which is comprised of the Ku70 and Ku80 DNA end-binding complex (Ku), and the catalytic subunit (DNA-PKcs). Two additional NHEJ proteins, DNA ligase IV (Lig4) and XRCC4, are involved in ligation. A deficiency in Ku70, Ku80, Lig4 or XRCC4 leads to a severe combined immunodeficient (SCID) phenotype, owing to an inability to complete RAG-initiated V(D)J recombination, as well as more-general defects, including impaired cellular proliferation. Deficiency of DNA-PKcs causes a SCID phenotype also, owing to defective coding joining, but does not cause severely defective RSS-joining or proliferation defects, suggesting that DNA-PKcs might be involved only in a subset of the functions of NHEJ proteins [8].

Functional rearrangement and expression of IgH and IgL leads to the generation of IgM^+ B lymphocytes. The murine IgH locus contains eight different C_H genes. Differentiating B lymphocytes produce μ heavy chains first and, correspondingly, IgM; but upon activation, they can juxtapose the V(D)J exon to a downstream C_H by CSR, allowing the generation of different antibody classes (e.g. IgG, IgA and IgE) and effector functions (Fig. 1) [3,4]. Some of the crucial early discoveries with respect to the mechanism of CSR have been summarized recently [9]. CSR occurs within the 1–10 kb repetitive switch (S) region sequences located 5′ of each C_H gene, except Cδ (which is expressed through alternative RNA splicing). Thus, in contrast to V(D)J recombination, CSR is carried out by mature B lymphocytes and does not use a consensus target sequence. Although we still do not know the initiating factor for CSR, RAG-1- or RAG-2-deficient B cells, generated by the introduction of functionally rearranged V_HDJ_H and V_LJ_L knock-ins, are capable of CSR [10]. Conversely, AID deficiency blocks CSR completely but does not affect V(D)J recombination, further distinguishing the two reactions [7]. As we outline CSR in more detail, we will further compare this process with V(D)J recombination.

http://immunology.trends.com 1471-4906/02/$ – see front matter PII: S1471-4906(01)02111-1

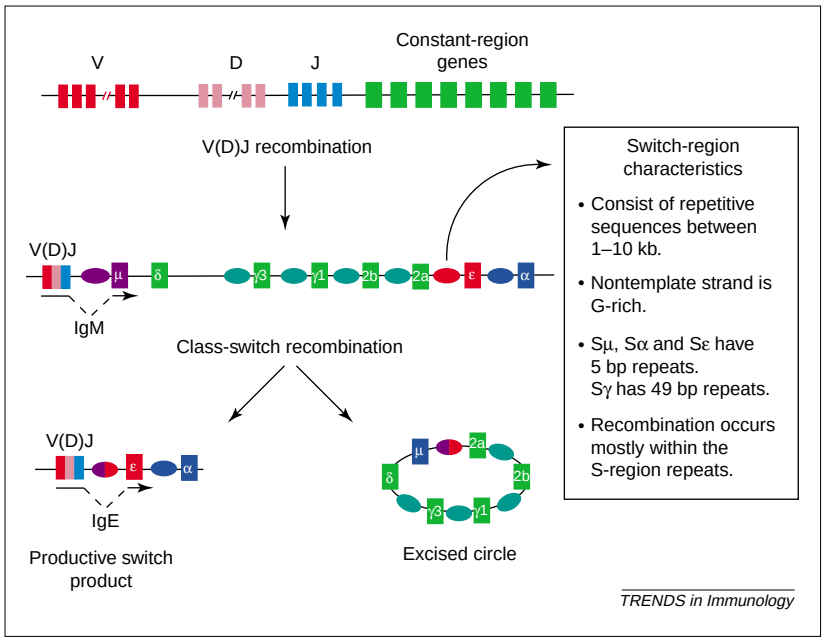

Fig. 1. Class-switch recombination (CSR) occurs by a recombination/deletion mechanism. Ig heavy-chain variable genes are assembled in progenitor B cells by V(D)J recombination. Subsequently, CSR can exchange the constant-region heavy chain (C_H) gene expressed with the assembled V(D)J (see text for details). Abbreviations: D, diversity; J, joining; S, switch; V, variable.

Regulation of class-switch recombination

B lymphocytes migrate from the bone marrow to peripheral lymphoid organs, such as the spleen, lymph nodes or gut-associated lymphoid tissues. In these sites, naive B cells, often in association with T cells, undergo antigen-driven clonal expansion in germinal centers (GCs) [11]. At this time, B cells might undergo CSR, although CSR can occur also outside of the GC and, for IgA, might even occur without prior expression of membrane-bound IgM [12]. In contrast to V(D)J recombination, CSR occurs only in B-lineage cells. Moreover, CSR is directed specifically to particular C_H genes by signals generated from extracellular cytokines in the context of various forms of B-cell activation [3,13]. CSR tends to occur for the same S regions on both productive and nonproductive alleles, indicating that the various external agents that activate specific CSR events generate signals that, ultimately, direct the DNA recombination process [14]. *In vivo*, CSR can occur by T-cell-dependent (TD) mechanisms that involve the interaction of CD40 on B cells with CD40 ligand on T cells, as well as by non-T-cell (non-CD40)-dependent routes, through T-cell-independent (TI) antigens [3,13]. TD responses can be mimicked, *in vitro*, by stimulating B cells with anti-CD40 antibodies in the presence or absence of particular lymphokines, whereas TI responses can be mimicked by treatment with agents such as bacterial lipopolysaccharide (LPS). Moreover, specific CSR events can be induced by the *in vitro* exposure of splenic B cells to particular combinations of activators and cytokines [3,13]. For example, the activation of murine B cells with LPS leads to the generation of IgG2b- or IgG3-secreting cells; whereas simultaneous treatment with LPS plus interleukin-4 (IL-4) suppresses the generation of IgG2b- and IgG3-secreting cells, but leads to the production of IgG1 and IgE (Fig. 2).

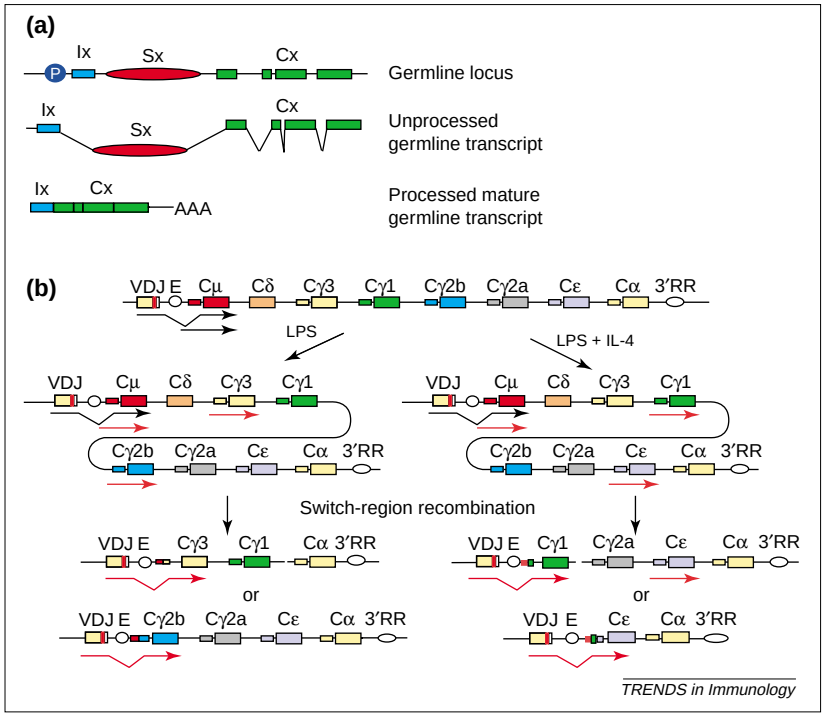

Fig. 2. Class-switch recombination (CSR) is directed by germline transcription. (a) Germline transcripts originate from the promoter (P) upstream of the I exon, which precedes all the S regions (adapted from Ref. [14]). (b) CSR of particular heavy-chain constant-region (C_H) genes is preceded by the induction of germline transcription of the C_H region targeted for CSR. Abbreviations: C, constant region; D, diversity; IL-4, interleukin-4; J, joining; LPS, lipopolysaccharide; RR, regulatory region; S, switch; V, variable.

Germline C_H transcription

Germline C_H genes are organized into germline transcription units, in which transcription initiates from a promoter 5′ of the I exon, runs through the S region and undergoes polyadenylation downstream of the C_H exons [2,15] (Fig. 2). RNA splicing generates a processed germline transcript by fusing the I exon to the C_H exons and deleting intervening S-region-derived sequences. Although such transcripts occur in the cytoplasm, they do not appear to encode proteins. In this context, the activation of CSR by specific cytokines correlates directly with the ability of particular treatments to induce or suppress specific germline C_H transcription units before CSR, strongly supporting a cause–effect relationship [14,15] (Fig. 2). Gene-targeting studies have demonstrated that promoter integrity of the I exon is required for efficient CSR to associated S regions [16–20]. In addition, constitutively transcribed drug-resistance cassettes (e.g. the *pgk-neo*^{r} cassette) can replace I-region promoters in directing CSR to a downstream S region in activated B cells and alter CSR profiles stimulated by particular activation treatments [18–20]. Thus, transcription through the S region plays a primary role in targeting CSR.

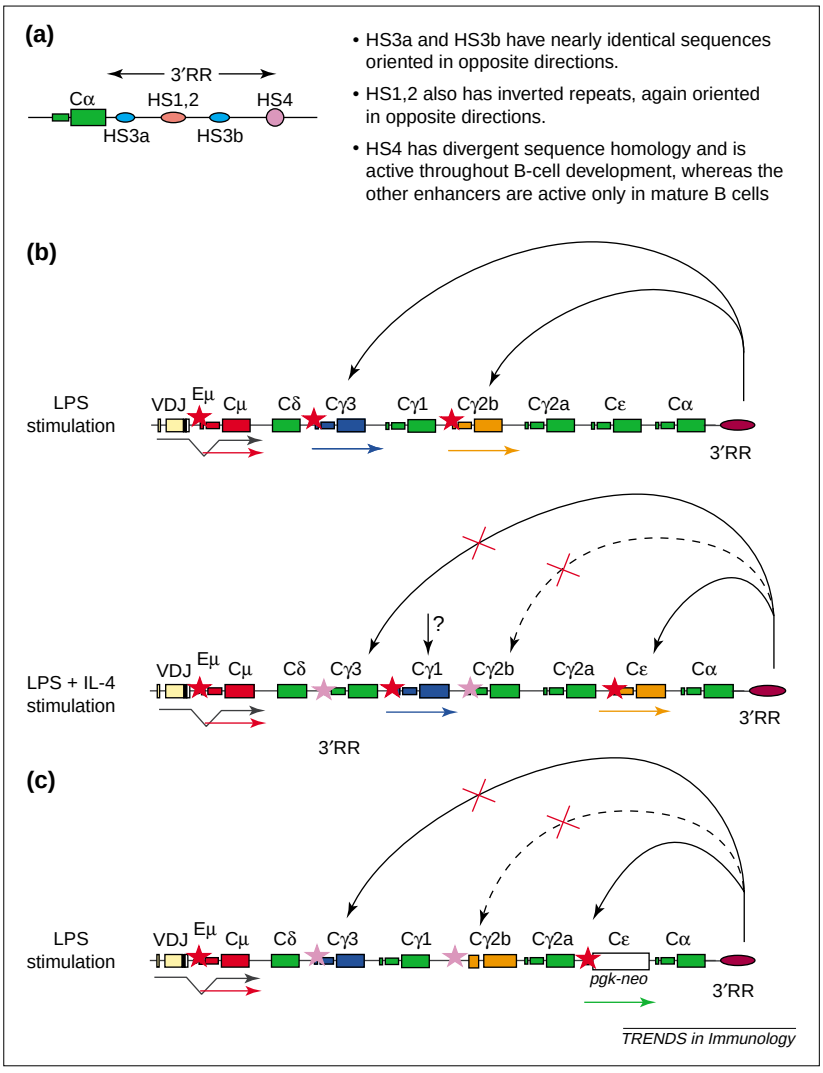

Fig. 3. The 3′ Ig heavy-chain (IgH) regulatory region (3′ RR). (a) The region just downstream of the IgH locus contains several DNAse hypersensitivity sites (HSs), which also have position-independent enhancer activity in B cells [24]. (b) Class-switch recombination (CSR) might be regulated by a promoter competition mechanism. In this case, induction of germline transcription of certain heavy-chain constant-region (C_H) genes would require both activation of a particular germline promoter and also, its interaction with the 3′ IgH RR. In the context of such a model, adding interleukin-4 (IL-4) along with lipopolysaccharide (LPS) would extinguish germline Cγ2b and Cγ3 expression by activating the Cε or Cγ1 promoters, which would compete more effectively for the 3′ IgH RR activity. The question mark indicates that Cγ1 transcription can be activated independently of the 3′IgH RR. (c) Placement of certain exogenous promoters (e.g. the *pgk-neo*^{r} promoter) might, similarly, inactivate expression of upstream germline C_H transcription units dependent on the 3′ IgH RR for expression (for details, see Refs [28,29]). The red stars represent an activated C_H promoter that interacts with the 3′ IgH RR to initiate germline transcription. The pink stars represent activated promoters that are unable to induce germline transcripts owing to promoter competition for the 3′ IgH RR. Abbreviations: D, diversity; J, joining; V, variable.

Regulation of germline C_H transcription

Germline C_H promoters contain sequences capable of conferring LPS- and/or lymphokine-mediated transcriptional regulation. In several cases, the signal-transduction pathways leading from the surface cytokine receptor to the activation of *cis*-acting I-region promoters by specific transcription factors have been elucidated [15]. However, efficient *in vivo* germline C_H transcription requires sequences in addition to germline C_H promoters. In this regard, CSR or CSR-related deletions of downstream S regions occur, albeit at reduced efficiency, in the absence of the B-cell specific transcriptional enhancer element (iEμ) located just downstream of the J_H segments [21,22]. Various observations, including activation of the c-myc oncogene following translocation into the downstream portion of the IgH locus, have suggested that major control elements lie 3′ of the IgH locus [23]. For example, the 40 kb region downstream from Cα harbors four enhancer elements, referred to as HS3a, HS1,2, HS3b and HS4 [24] (Fig. 3a), which are conserved in humans and appear to have locus control region (LCR)-like properties [24,25]. HS1,2, also termed the 3′ IgH enhancer, was thought to be a prime candidate for a major CSR control element [23]. Replacement of HS1,2 with a *pgk-neo*^{r} cassette disrupted germline transcription of, and CSR to, a series of C_H genes, including Cγ3, which lies 120 kb upstream [26]. Although these studies are consistent with a major role for HS1,2, the other possibility is that the *pgk-neo*^{r} cassette interferes somehow with a higher-level CSR regulatory locus 3′ of the IgH locus (IgH 3′RR).

To resolve these models, the effects of *pgk-neo*^{r} insertions in place of HS3a or HS1,2 were compared with 'clean' deletions of these elements (replacement with a loxP site). These studies showed that neither 3′ enhancer element was required for CSR, but insertion of the *pgk-neo*^{r}, in place of either enhancer, inhibited germline transcription of, and CSR to, various upstream C_H genes [27]. Moreover, other studies showed that insertion of transcribed drug-resistance cassettes at additional sites within the C_H genes led to the inhibition of germline transcription and/or CSR in the context of upstream, but not downstream, C_H genes [28,29]. However, insertion of a *pgk-neo*^{r} cassette just downstream of HS4 had no obvious effect on CSR [30]. Together, these insertional mutation analyses point to the existence of a 3′IgH RR with crucial components lying between HS1,2 and HS4. Strikingly, deletion of the HS3a–HS4 region has been shown recently to block germline transcription of, and CSR to, a similar set of C_H genes as those affected by insertion of *pgk-neo*^{r} into HS1,2 or HS3a [31], consistent with the predictions of the insertional mutagenesis studies. In this regard, HS4 is a particularly attractive candidate to play a major role in CSR because of its activity in late-stage B cells [25,32].

Based on the *pgk-neo*^{r} inhibition studies, it has been suggested that the IgH 3′RR might operate by a promoter competition mechanism to help regulate the differential expression of distal C_H genes, although other models are conceivable [24] (Fig. 3b). Moreover, it remains to be demonstrated that the inhibitory effects of the *pgk-neo*^{r} cassette are actually caused by promoter activity. Of note, germline transcription of, and CSR to, certain C_H genes (e.g. Cγ1 and Cα) is not affected markedly by either *pgk-neo*^{r} insertions or deletion of HS3b and/or HS4 [26,31], consistent with the possibility that such C_H genes might be regulated

principally by proximal elements [33,34] or uncharacterized long-range elements. Finally, the precise functions of individual 3′ enhancers remain to be elucidated; conceivably, they might also function in other processes, such as SM [35,36].

Mechanism of class-switch recombination

The large and unusual nature of the S regions has led to various models to describe their role in CSR (see, for example, Refs [9,15,37–39]). A common starting point is the knowledge that CSR results from the fusion of an upstream S region to a downstream S region with the deletion of intervening DNA (Fig. 1). Recombination points are found throughout individual S regions and can occur outside of the main S region [9]. Therefore, unlike V(D)J recombination, CSR is region-specific, occurs within introns and does not influence translational reading frames. S regions are composed of tandem repeats of pentamers (predominantly, GAGCT and GGGGT) or a 49 base pair sequence, and share varying degrees of homology [3]. However, CSR breakpoints lack homology generally and also, consensus junctions, arguing against either a specific CSR signal motif or a primary role for homologous recombination [9,15,40].

The degeneracy of S-region sequences and heterogeneity of CSR sites led to CSR models in which the S-region recognition code might lie in higher-order structures, rather than primary sequences. In this regard, murine and human S regions are GC-rich and one strand is highly G-rich, properties which can contribute to structures such as G quartets [38,41]. However, the S regions of frogs, although repetitive, lack homology to each other and are AT-rich [42]. Assuming a common mechanism, the latter findings might support models that argue for a role for palindromic sequences in targeting a cleavage event [4,43]. At this time, however, the precise function of S regions remains speculative; indeed, recent evidence showed that deletion of a major portion of the tandem repeats of the Sμ sequence reduced greatly, but did not abrogate, CSR in mice [44]. In that case, residual CSR might have been promoted by a small number of S-region elements left by the deletion.

V(D)J-recombination substrates have revealed many insights into the V(D)J-recombination mechanism. Similarly, a variety of CSR substrates have been introduced transiently or stably into mammalian cell lines to study aspects of the CSR mechanism [37,45–49]. Several of these substrates appeared to recombine more efficiently in activated B-lineage cells, suggesting a link to CSR. However, studies of CSR substrates have led to a very wide range of interpretations and models, particularly with regard to potential S-region functions. One potential explanation for the divergent findings is that recombination in some substrates might have occurred by pathways distinct from CSR that were stimulated by the unusual S-region-sequence composition. In this regard, it will be of great importance to demonstrate that recombination events in such substrates actually represent *bona fide* CSR. The ultimate resolution of these issues might be possible by testing model substrates in AID-deficient cells.

Initiation and resolution: unresolved questions

CSR has been shown to generate two products: the rearranged chromosome and an extrachromosomal circle containing deleted intervening sequences [50–53]. Although CSR-derived circles have not been demonstrated unequivocally to be direct reciprocal products of CSR junctions, such recombination products can be explained most easily by a simple cut-and-join mechanism [54] (Fig. 1). In this model, CSR would be initiated by DSBs or staggered single-strand breaks (SSBs) in the S region, and completed by joining DNA ends in the recombinant configuration. Indeed, S-region DSBs have been detected in B cells undergoing CSR [55]. However, a major question has involved the nature of the activity that would generate DSBs or other potential initiating DNA lesions, and whether or not it involves a specific CSR recombinase (analogous to RAG) or a more general factor and/or process associated with S-region transcription. It is, perhaps, notable that S-region instability, as indicated by internal deletions, occurs in association with S-region transcription, independent of CSR [21,56]. If, as hypothesized, such instability were related mechanistically to CSR [56], it might imply that initiation of CSR does not require synapsis of two S regions, in contrast to the requirement for two participating RSSs in V(D)J recombination. Assays of the dependence of the deletion phenomenon on AID should help to reveal its relationship to CSR.

The numerous models that attempt to explain CSR can be divided into two general, but not mutually exclusive, categories, based on a direct versus indirect role for transcription in generating the initiating cleavage event (Fig. 4). First, in terms of an indirect role, the correlation between germline C_H transcription and subsequent CSR has been interpreted in the context of accessibility of the S region to a specific CSR recombinase, expression of which is induced in activated B cells [14,15]. The V(D)J recombination accessibility model was based on the existence of a single V(D)J recombinase [57]; however, it is not yet clear that there is a single CSR recombinase that recognizes all S regions. In fact, S-region-specific CSR factors that direct CSR specifically in certain cell lines have been implicated [49]. However, the nature of such factors and their precise role in CSR requires further clarification. A second model argues that transcription plays a more primary role in the initiation of CSR, perhaps leading directly to the initiating event [14,15].

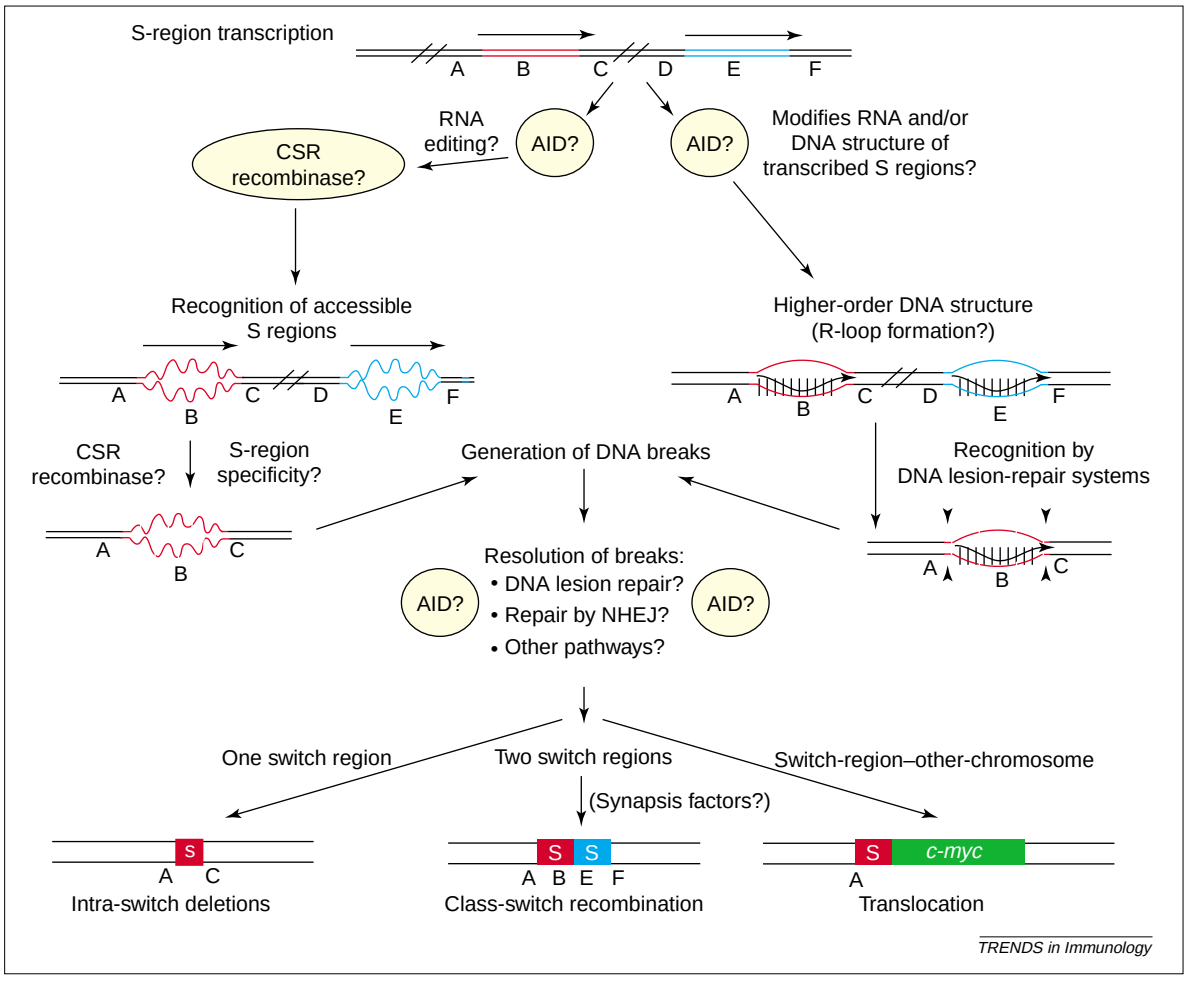

Fig. 4. General models for the mechanism of class-switch recombination (CSR). Two general models for CSR are outlined. Both are based on the original recombination/deletion model [54] and the requirement for germline heavy-chain constant-region C_H transcription (top). The model on the left is based on a common CSR recombinase and a major role for transcription in generating S-region accessibility. The model on the right is based on a primary role for germline transcription in initiation of the CSR reaction. The two models are not mutually exclusive, share overlapping features and remain speculative, as indicated by the question marks (see text for details). Also, additional models or versions of these models have been envisioned (for example, see Refs [4,37–39,49]). Abbreviations: AID, activation-induced deaminase; NHEJ, nonhomologous end-joining; S region, switch region.

In addition to the initiating mechanism, the S-region joining mechanism is unknown also. The heterogeneous nature of CSR junctions has pointed to mechanisms such as NHEJ. Mutations in components of the NHEJ system inhibit CSR [58–60]. Thus, Ku70- and Ku80-deficient B cells, generated by complementation with germline IgH and IgL knock-ins, show little ability to undergo CSR either *in vivo* or *in vitro*, despite the activation and induction of germline transcripts, implying a defect in the terminal phases of CSR. However, as a deficiency of Ku impairs B-cell proliferation also, the effect of the Ku mutation on CSR could be indirect. Intriguingly, DNA-PKcs-deficient B cells show no apparent proliferation defects in response to appropriate activation, support normal induction of germline transcripts and yet, are defective in CSR to all C_H genes, except Cγ1 (J.P. Manis and F.W. Alt, unpublished). These data provide the strongest evidence in favor of a role for DNA-PKcs (and NHEJ) in CSR. If so, however, then efficient CSR to IgG1 might occur also through an NHEJ-independent pathway or DNA-PKcs-independent NHEJ, similar to RSS joining in the absence of DNA-PKcs. It is notable also that V(D)J recombination occurs in the G0 and/or G1 stage of the cell cycle of nonproliferating, developing lymphocytes [1], the stage of the cell cycle in which the NHEJ DSB-repair pathway is most crucial. By contrast, CSR occurs in proliferating cells, has been argued to require DNA replication and might occur in S phase [15,61], suggesting the possibility of a role for DNA replication and associated processes. Another important question is the mechanism by which S regions are synapsed during CSR [62].

AID – a crucial component of the class-switch recombination process

Proteins have been identified that bind to, cleave or recombine S regions *in vitro* [15], but their relevance to CSR remains unclear. Recently, however, a subtractive approach has identified gene sequences, the expression of which was induced in a murine B-cell line stimulated *in vitro* to undergo CSR [6]. One

such sequence encoded a novel protein referred to as activation-induced deaminase (AID), based on sequence homology to a known deaminase. Dramatically, AID-deficient mice, generated by gene-targeted mutation, lack completely the ability to undergo CSR (and, apparently, SM; see below), but show no other obvious phenotype, except enlarged GCs containing IgM-producing activated B cells [7]. Likewise, a group of immunodeficient patients (hyper-IgM II syndrome) with impaired CSR had mutations in the human gene encoding AID [63]. Notably, the defect in CSR in AID-deficient B cells is cell-intrinsic and does not involve known steps in B-cell activation leading to CSR. Overall, the specificity of the AID-deficient phenotype for CSR is strikingly reminiscent of the absolute specificity of RAG mutations for V(D)J recombination. These findings suggest that AID is required for the actual CSR event itself and might either activate or be a primary component of a CSR recombinase. In this context, the enlarged lymph nodes and hyper-IgM phenotype of AID deficiency might result from an accumulation of activated B cells that cannot differentiate further owing to their inability to effect CSR [7,63].

The exact function of AID is speculative. In theory, it might catalyze some stage of CSR (e.g. initiation) directly; however, its sequence suggests other possibilities also [6]. Thus, AID belongs to a family of cytidine deaminases that catalyze the hydrolytic deamination of cytidine to uridine, either as free nucleotide or in the context of RNA. AID bears closest resemblance to the RNA editing enzyme apolipoprotein B mRNA-editing complementation protein 1 (APOBEC-1) [6], which catalyzes a site-specific cytosine to uridine change in APOB mRNA and, thereby, introduces a premature stop codon that leads to a novel truncated protein [64]. Similarly, editing by AID might generate the functional message for a putative component of the CSR recombinase. We envision a CSR model in which antigen stimulation would lead to two crucial events (Fig. 4). The first would be germline transcription to make particular transcribed S regions accessible and guide S-region specificity. The second would include the induction of expression of AID, which would generate a CSR recombinase that leads to cleavage of the S region [7,48]. However, a role for AID downstream of initiation has not been eliminated; thus, it will be of significant interest to determine whether S regions are still cleaved and internal S-region deletions still occur in AID-deficient mice. In any case, following cleavage, S regions could be joined by the general cellular repair machinery. Additional potential roles for AID are outlined below. Whatever its role, assays in AID-deficient cells might provide the gold-standard for determining the accuracy with which various substrates induce a *bona fide* CSR process.

Potential roles for germline transcription and DNA-lesion-repair pathways

Although there is no demonstrated role for transcription *per se* in V(D)J recombination, evidence suggests a more intimate, perhaps mechanistic, role for germline transcription in CSR. Thus, the organization of C_H transcription units is conserved among diverse species. However, I exons contain multiple in-frame stop codons and are not conserved at the level of nucleotide sequence [2,3]. Yet, appropriate splicing of these apparently 'sterile' germline C_H transcripts might be required for efficient CSR [65–67], although more fine-targeting of endogenous splice sites is needed to clarify this potential relationship. In addition, transient transfection assays show that recombination between transcribed S regions proceeds more efficiently when they are oriented in their normal, 'sense' orientation [47]. Thus, current data are consistent with the possibility that germline C_H transcripts play a direct role in promoting CSR, and this role might depend on both general structure and, possibly, sequence composition (e.g. S-region orientation). However, the latter notion has been challenged by studies indicating that S-region orientation plays no role in promoting recombination events between S regions in chromosomally integrated substrates in a B-lymphoma cell line [48]. Ultimately, this issue could be resolved by gene-targeted inversion or replacement of endogenous S regions.

Early studies showed that when S regions are transcribed *in vitro*, transcripts remain on the template DNA, forming RNA–DNA hybrids [68,69]. Although various higher order structures were hypothesized, a recent study showed that the displaced nontemplate strand remains largely single-stranded and the RNA–DNA complex takes the form of an R-loop [39]. Furthermore, this R-loop structure can be recognized and cleaved by structure-specific endonucleases [xeroderma pigmentosum group F (XPF)-ERCC1 and XPG] involved in nucleotide excision repair [39]. Based on these observations, an R-loop model for the initiation of CSR was proposed in which germline transcription would lead to the formation of S-region R-loops, which could then be cleaved by structure-specific endonucleases to yield S-region strand breaks. However, because the R-loop model is based entirely on *in vitro* observations, its validity remains to be evaluated *in vivo*. In the context of such general models, it is of note also that the displaced DNA strand in transcribed S-region DNA would be subject to the formation of secondary structures [9] that could target cleavage activities also. What could be the role of AID in the context of R-loop or higher-order DNA-structure models? Although editing roles could still be envisioned, one could also imagine that AID might modify RNA or DNA structures in some way to make them more stable or better substrates for repair endonucleases (Fig. 4).

Mice deficient in mismatch repair (MMR) genes show reduced CSR and altered CSR junctions focused on consensus repeats [70,71]. Because MMR-deficient B cells appear to activate and proliferate normally, the effect of MMR-deficiency on CSR is probably direct. In this context, the recession of S-region DSBs by strand-specific exonucleases could expose short stretches of microhomology in the form of single-stranded DNA; such microhomology might then be used for annealing cleaved S regions [72]. Because microhomologies might occur internal to a DSB, such annealing intermediates might contain also heterologies in the form of branched DNA structures. Based on the activities of yeast homologs, such branches might be removed by the MMR proteins before joining [73]. In addition, MMR proteins have been proposed to recognize other unusual DNA structures, such as loops and misaligned DNA sequences, that might be generated during CSR [74]. A potential link between MMR and R-loop cleavage might be provided by the observation that Rad1–Rad10 (the yeast homolog of XPF-ERCC1) and MMR proteins collaborate in cutting branched DNA structures [74].

Potential overlaps in the class-switch recombination and somatic hypermutation mechanisms

SM occurs in GC B cells during antigen-induced proliferation at frequencies as high as one event per 1000 nucleotides cell^{-1} generation^{-1} [5]. In contrast to the large deletions involved in CSR, SM comprises point mutations, small deletions and insertions in the V(D)J exons [75]. Intriguingly, however, CSR junctions often contain point mutations and other sequence alterations found in SM [40,76]. In addition, CSR and SM share other features. First, SM relies on transcription through the target V(D)J exons [77]. Second, the rate and nature of SM is influenced also by the MMR system, and is focused on sequences reminiscent of the hypermutation consensus sequence in MMR-deficient cells [5,71,78]. Third, AID appears to be required for SM, although it remains possible that the apparent function of AID in SM is indirect [7,63]. One hypothesis is that AID might serve to generate the functional transcript for a DSB initiator in SM separate from that employed for CSR [7]. An alternative thinking is that similarities between SM and CSR might be indicative of a common mechanism, perhaps involving R-loops or other secondary DNA structures within strongly transcribed V(D)J sequences that are augmented by AID functions [79].

As for CSR, the nature of the SM initiator is unknown, although, again, it is linked closely to transcription [77]. However, SM has been associated with DNA strand breaks [80], which might include both SSBs [81] and DSBs [82,83]. The DSBs appear primarily during S and G2 phases of the cell cycle, as opposed to G1 when NHEJ is most active [82], consistent with the proposal that DSBs in SM might be repaired by homologous recombination [78]. Whatever the initiation mechanism, mutations associated with SM might be introduced subsequent to an initiating DNA lesion resulting from error-prone repair involving low-fidelity DNA polymerases [5]. Finally, important differences might exist between CSR and SM. For example, SCID (DNA-PKcs-deficient) mice reconstituted with transgenic B- and T-cell receptors undergo efficient SM, suggesting that, in contrast to CSR to most C_H genes, DNA-PKcs might not be required for SM [84]. However, because DNA-PKcs is required for only a subset of NHEJ transactions and might have roles outside of NHEJ, clarification of these issues will require assays of SM and CSR in Ku-, XRCC4- or Lig4-deficient B cells that lack apoptotic responses to DNA damage. Finally, it is of note that SM can occur independently of CSR [85,86].

Genomic instability and cancer due to errors in class-switch recombination or somatic hypermutation

Most human B-cell lymphomas arise in GCs. In many cases, lymphoma cells harbor cellular oncogenes, such as Bcl-2 or Bcl-6, translocated into IgH S regions, resulting in deregulated oncogene expression by association with IgH regulatory sequences [87]. Other examples of S-region translocations in B-cell tumors include murine plasmacytomas and sporadic variants of human Burkitt's lymphoma that harbor *c-myc* translocations into the IgH S regions [88]. It seems probable that aberrant CSR plays a major role in the generation of such translocations. Given the potential role of DNA strand breaks in CSR and SM, such breaks might play an important role in promoting translocations associated with the genesis of mature B-cell tumors [87,89]. A major question, however, is the mechanism and factors that influence the joining of such breaks to other chromosomal loci. Finally, because progenitor-B-cell lymphomas harboring specific translocations arise in mouse models as a result of errors in V(D)J recombination in the setting of defective NHEJ and lack of expression of the p53 checkpoint protein [90], it will be of significant interest to attempt to promote the occurrence of mature B-cell lymphomas by, similarly, impairing DNA repair and checkpoint pathways in mature B cells induced to attempt CSR or SM.

As noted previously, transcribed S regions can be unstable in B-lineage cell lines, probably in association with the activity of AID. By analogy, certain non-Ig loci might be similarly unstable owing to inherent sequence properties and this could contribute to translocations. In addition, many genetic diseases are associated with sequence instability; for example, the amplification of triplet repeats, which are linked to chromosomal fragile sites [91]. Moreover, *in vitro* studies have shown that some of these repeats can adopt unusual DNA structures, such as hairpins [91]. Thus, mechanisms related to

those of strand-break formation in S regions might operate also in causing more-general chromosomal instability. The role of AID or related proteins in contributing to the instability of putative non-Ig sequence targets is a question of significant potential interest also. Clearly, understanding the basic mechanisms of CSR and SM might lead to important insights into mechanisms underlying genomic instability and that contribute to B-cell lymphomagenesis.

Perspective

Work on the elucidation of the V(D)J recombination mechanism was aided greatly by the discovery of the RAG proteins, which permitted extensive *in vitro* and cell-based analyses. In contrast to V(D)J recombination, progress in elucidating the CSR mechanism has not advanced rapidly. Although genetic evidence has suggested roles for several DNA repair pathways in CSR, the precise mechanisms leading to the initiation and resolution of this reaction in the participating DNA strands remain relatively uncharacterized. However, we anticipate that discovery of the AID protein should result in rapid progress. In particular, the availability of appropriate cell lines deficient in this protein should allow the firm establishment of CSR-substrate assays and might provide also the basis for future *in vitro* studies. Given the potential involvement of aspects of the basic CSR mechanism in promoting S-region sequence instability, results of such future studies might have substantial implications for understanding the basis of other types of genomic instability in the context of particular diseases.

Acknowledgements
This work was supported by NIH grant AI31541 (to F.W.A.), NIH training grant AI 07512 (to M.T.) and a grant from the Hood Foundation and the Cure for Lymphoma Foundation (to J.P.M.). F.W.A. is an Investigator of the Howard Hughes Medical Institute and M.T. was an Associate of the Howard Hughes Medical Institute and an Irvington Institute Fellow.

References

1 Hesslein, D.G. and Schatz, D.G. (2001) Factors and forces controlling V(D)J recombination. *Adv. Immunol.* 78, 169–232
2 Zhang, J. *et al.* (1995) *Regulation of Class-Switch Recombination of the Immunoglobulin Heavy-Chain Genes*, Academic Press
3 Stavnezer, J. (1996) Antibody class switching. *Adv. Immunol.* 61, 79–146
4 Kinoshita, K. and Honjo, T. (2000) Unique and unprecedented recombination mechanisms in class switching. *Curr. Opin. Immunol.* 12, 195–198
5 Jacobs, H. and Bross, L. (2001) Towards an understanding of somatic hypermutation. *Curr. Opin. Immunol.* 13, 208–218
6 Muramatsu, M. *et al.* (1999) Specific expression of activation-induced cytidine deaminase (AID), a novel member of the RNA-editing deaminase family in germinal-center B cells. *J. Biol. Chem.* 274, 18470–18476
7 Muramatsu, M. *et al.* (2000) Class-switch recombination and hypermutation require activation-induced cytidine deaminase (AID), a potential RNA-editing enzyme. *Cell* 102, 553–563
8 Sekiguchi, J.M. *et al.* (1999) Nonhomologous end-joining proteins are required for V(D)J recombination, normal growth and neurogenesis. *Cold Spring Harbor Symp. Quant. Biol.* 64, 169–181
9 Kinoshita, K. *et al.* (1999) Molecular mechanism of immunoglobulin class-switch recombination. *Cold Spring Harbor Symp. Quant. Biol.* 64, 217–226
10 Lansford, R. *et al.* (1998) Ig heavy-chain class switching in Rag-deficient mice. *Int. Immunol.* 10, 325–332
11 Tarlinton, D. (1998) Germinal centers: form and function. *Curr. Opin. Immunol.* 10, 245–251
12 Macpherson, A.J. *et al.* (2001) IgA production without μ or δ chain expression in developing B cells. *Nat. Immunol.* 2, 625–631
13 Snapper, C.M. *et al.* (1997) The immunoglobulin class switch: beyond 'accessibility'. *Immunity* 6, 217–223
14 Rothman, P. *et al.* (1989) The molecular events in heavy-chain class-switching. *Semin. Immunol.* 1, 65–77
15 Stavnezer, J. (2000) Molecular processes that regulate class switching. *Curr. Top. Microbiol. Immunol.* 245, 127–168
16 Jung, S. *et al.* (1993) Shutdown of class-switch recombination by deletion of a switch-region control element. *Science* 259, 984–987
17 Zhang, J. *et al.* (1993) A selective defect in IgG2b switching as a result of targeted mutation of the Iγ2b promoter and exon. *EMBO J.* 12, 3529–3537
18 Bottaro, A. *et al.* (1994) S-region transcription *per se* promotes basal IgE class-switch recombination but additional factors regulate the efficiency of the process. *EMBO J.* 13, 665–674
19 Harriman, G.R. *et al.* (1996) IgA class switch in Iα exon-deficient mice. Role of germline transcription in class-switch recombination. *J. Clin. Invest.* 97, 477–485
20 Seidl, K.J. *et al.* (1998) An expressed neo(r) cassette provides required functions of the 1γ2b exon for class switching. *Int. Immunol.* 10, 1683–1692
21 Gu, H. *et al.* (1993) Independent control of immunoglobulin switch recombination at individual switch regions evidenced through Cre-loxP-mediated gene targeting. *Cell* 73, 1155–1164
22 Bottaro, A. *et al.* (1998) Deletion of the IgH intronic enhancer and associated matrix-attachment regions decreases, but does not abolish class switching at the μ locus. *Int. Immunol.* 10, 799–806
23 Pettersson, S. *et al.* (1990) A second B-cell-specific enhancer 3′ of the immunoglobulin heavy-chain locus. *Nature* 344, 165–168
24 Khamlichi, A.A. *et al.* (2000) The 3′ IgH regulatory region: a complex structure in a search for a function. *Adv. Immunol.* 75, 317–345
25 Madisen, L. and Groudine, M. (1994) Identification of a locus control region in the immunoglobulin heavy-chain locus that deregulates c-myc expression in plasmacytoma and Burkitt's lymphoma cells. *Genes Dev.* 8, 2212–2226
26 Cogne, M. *et al.* (1994) A class-switch control region at the 3′ end of the immunoglobulin heavy-chain locus. *Cell* 77, 737–747
27 Manis, J.P. *et al.* (1998) Class switching in B cells lacking 3′ immunoglobulin heavy-chain enhancers. *J. Exp. Med.* 188, 1421–1431
28 Seidl, K.J. *et al.* (1999) Position-dependent inhibition of class-switch recombination by PGK-neo^r cassettes inserted into the immunoglobulin heavy-chain constant-region locus. *Proc. Natl. Acad. Sci. U. S. A.* 96, 3000–3005
29 Qiu, G. *et al.* (1999) Iα exon-replacement mice synthesize a spliced HPRT-C(α) transcript which may explain their ability to switch to IgA. Inhibition of switching to IgG in these mice. *Int. Immunol.* 11, 37–46
30 Manis, J.P. *et al.* Elucidation of a downstream boundary of the 3′ IgH regulatory region. *Mol. Immunol.* (in press)
31 Pinaud, E. *et al.* (2001) Localization of the 3′ IgH locus elements that effect long-distance regulation of class-switch recombination. *Immunity* 15, 187–199
32 Michaelson, J.S. *et al.* (1995) Identification of 3′ α-hs4, a novel Ig heavy-chain enhancer element regulated at multiple stages of B-cell differentiation. *Nucleic Acids Res.* 23, 975–981
33 Lin, S.C. and Stavnezer, J. (1996) Activation of NF-κB/Rel by CD40 engagement induces the mouse germ-line immunoglobulin Cγ1 promoter. *Mol. Cell. Biol.* 16, 4591–4603
34 Cunningham, K. *et al.* (1998) Germline transcription and recombination of a murine VDJμδγ1 transgene. *Int. Immunol.* 10, 1027–1037
35 Tumas-Brundage, K.M. *et al.* (1997) Evaluation of the role of the 3′α heavy-chain enhancer [3′α E(hs1,2)] in Vh gene somatic hypermutation. *Mol. Immunol.* 34, 367–378
36 Terauchi, A. *et al.* (2001) A pivotal role for DNase-I-sensitive regions 3b and/or 4 in the induction of somatic hypermutation of IgH genes. *J. Immunol.* 167, 811–820
37 Ballantyne, J. *et al.* (1998) Efficient recombination of a switch substrate retrovector in CD40-activated B lymphocytes: implications for the control of CH gene-switch recombination. *J. Immunol.* 161, 1336–1347
38 Maizels, N. (1999) Immunoglobulin class-switch recombination: will genetics provide new clues to mechanism? *Am. J. Hum. Genet.* 64, 1270–1275
39 Tian, M. and Alt, F.W. (2000) Transcription-induced cleavage of immunoglobulin switch regions by nucleotide excision-repair nucleases *in vitro*. *J. Biol. Chem.* 275, 24163–24172
40 Dunnick, W. *et al.* (1993) DNA sequences at immunoglobulin switch-region recombination sites. *Nucleic Acids Res.* 21, 365–372
41 Sen, D. and Gilbert, W. (1988) Formation of parallel four-stranded complexes by guanine-rich motifs in DNA and its implications for meiosis. *Nature* 334, 364–366

42 Mussmann, R. *et al.* (1997) Microsites for immunoglobulin-switch recombination breakpoints from *Xenopus* to mammals. *Eur. J. Immunol.* 27, 2610–2619
43 Tashiro, J. *et al.* (2001) Palindromic but not G-rich sequences are targets of class-switch recombination. *Int. Immunol.* 13, 495–505
44 Luby, T.M. *et al.* (2001) The μ switch-region tandem repeats are important, but not required, for antibody class-switch recombination. *J. Exp. Med.* 193, 159–168
45 Ott, D.E. *et al.* (1987) Immunoglobulin heavy-chain switch-region recombination within a retroviral vector in murine pre-B cells. *EMBO J.* 6, 577–584
46 Leung, H. and Maizels, N. (1992) Transcriptional regulatory elements stimulate recombination in extrachromosomal substrates carrying immunoglobulin switch-region sequences. *Proc. Natl. Acad. Sci. U. S. A.* 89, 4154–4158
47 Daniels, G.A. and Lieber, M.R. (1995) Strand specificity in the transcriptional targeting of recombination at immunoglobulin-switch sequences. *Proc. Natl. Acad. Sci. U. S. A.* 92, 5625–5629
48 Kinoshita, K. *et al.* (1998) Target specificity of immunoglobulin class-switch recombination is not determined by nucleotide sequences of S regions. *Immunity* 9, 849–858
49 Shanmugam, A. *et al.* (2000) Evidence for class-specific factors in immunoglobulin isotype switching. *J. Exp. Med.* 191, 1365–1380
50 Matsuoka, M. *et al.* (1990) Switch circular DNA formed in cytokine-treated mouse splenocytes: evidence for intramolecular DNA deletion in immunoglobulin class switching. *Cell* 62, 135–142
51 Iwasato, T. *et al.* (1990) Circular DNA is excised by immunoglobulin class-switch recombination. *Cell* 62, 143–149
52 Honjo, T. and Kataoka, T. (1978) Organization of immunoglobulin heavy-chain genes and allelic deletion model. *Proc. Natl. Acad. Sci. U. S. A.* 75, 2140–2144
53 von Schwedler, U. *et al.* (1990) Circular DNA is a product of the immunoglobulin class-switch rearrangement. *Nature* 345, 452–456
54 Kataoka, T. *et al.* (1980) Rearrangement of immunoglobulin γ1-chain gene and mechanism for heavy-chain class switch. *Proc. Natl. Acad. Sci. U. S. A.* 77, 919–923
55 Wuerffel, R.A. *et al.* (1997) Ig Sγ3 DNA-specifc double-strand breaks are induced in mitogen-activated B cells and are implicated in switch recombination. *J. Immunol.* 159, 4139–4144
56 Alt, F.W. *et al.* (1982) Immunoglobulin heavy-chain expression and class switching in a murine leukaemia cell line. *Nature* 296, 325–331
57 Yancopoulos, G.D. and Alt, F.W. (1986) Regulation of the assembly and expression of variable-region genes. *Annu. Rev. Immunol.* 4, 339–368
58 Rolink, A. *et al.* (1996) The SCID but not the RAG-2 gene product is required for S-μ–S-ε heavy-chain class switching. *Immunity* 5, 319–330
59 Manis, J.P. *et al.* (1998) Ku70 is required for late B-cell development and immunoglobulin heavy-chain class switching. *J. Exp. Med.* 187, 2081–2089
60 Casellas, R.N. *et al.* (1998) Ku80 is required for immunoglobulin-isotype switching. *EMBO J.* 17, 2404–2411
61 Lundgren, M. *et al.* (1995) Cell-cycle regulation of immunoglobulin class-switch recombination and germ-line transcription: potential role of Ets family members. *Eur. J. Immunol.* 25, 2042–2051
62 Dempsey, L.A. *et al.* (1999) G4 DNA binding by LR1 and its subunits, nucleolin and hnRNP D. A role for G–G pairing in immunoglobulin-switch recombination. *J. Biol. Chem.* 274, 1066–1071
63 Revy, P. *et al.* (2000) Activation-induced cytidine deaminase (AID) deficiency causes the autosomal recessive form of the hyper-IgM syndrome (HIGM2). *Cell* 102, 565–575
64 Chester, A. *et al.* (2000) RNA editing: cytidine to uridine conversion in apolipoprotein B mRNA. *Biochim. Biophys. Acta* 1494, 1–13
65 Petry, K. *et al.* (1999) An extrachromosomal switch-recombination substrate reveals kinetics and substrate requirements of switch recombination in primary murine B cells. *Int. Immunol.* 11, 753–763
66 Lorenz, M. *et al.* (1995) Switch transcripts in immunoglobulin-class switching. *Science* 267, 1825–1828
67 Hein, K. *et al.* (1998) Processing of switch transcripts is required for targeting of antibody class-switch recombination. *J. Exp. Med.* 188, 2369–2374
68 Daniels, G.A. and Lieber, M.R. (1995) RNA:DNA complex formation upon transcription of immunoglobulin switch regions: implications for the mechanism and regulation of class-switch recombination. *Nucleic Acids Res.* 23, 5006–5011
69 Reaban, M.E. and Griffin, J.A. (1990) Induction of RNA-stabilized DNA conformers by transcription of an immunoglobulin switch region. *Nature* 348, 342–344
70 Schrader, C.E. *et al.* (1999) Reduced isotype switching in splenic B cells from mice deficient in mismatch-repair enzymes. *J. Exp. Med.* 190, 323–330
71 Ehrenstein, M.R. and Neuberger, M.S. (1999) Deficiency in Msh2 affects the efficiency and local sequence specificity of immunoglobulin class-switch recombination: parallels with somatic hypermutation. *EMBO J.* 18, 3484–3490
72 Dunnick, W. and Stavnezer, J. (1990) Copy-choice mechanism of immunoglobulin heavy-chain switch recombination. *Mol. Cell. Biol.* 10, 397–400
73 Sugawara, N. *et al.* (1997) Role of *Saccharomyces cerevisiae* Msh2 and Msh3 repair proteins in double-strand break-induced recombination. *Proc. Natl. Acad. Sci. U. S. A.* 94, 9214–9219
74 Paques, F. and Haber, J.E. (1997) Two pathways for removal of nonhomologous DNA ends during double-strand break repair in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* 17, 6765–6771
75 Goossens, T. *et al.* (1998) Frequent occurrence of deletions and duplications during somatic hypermutation: implications for oncogene translocations and heavy-chain disease. *Proc. Natl. Acad. Sci. U. S. A.* 95, 2463–2468
76 Lee, C.G. *et al.* (1998) Frequent but biased class-switch recombination in the S μ flanking regions. *Curr. Biol.* 8, 227–230
77 Storb, U. (1996) The molecular basis of somatic hypermutation of immunoglobulin genes. *Curr. Opin. Immunol.* 8, 206–214
78 Harris, R.S. *et al.* (1999) Somatic hypermutation and the three R's: repair, replication and recombination. *Mutat. Res.* 436, 157–178
79 Gonzalez-Fernandez, A. *et al.* (1994) Somatic mutation of immunoglobulin λ chains: a segment of the major intron hypermutates as much as the complementarity-determining regions. *Proc. Natl. Acad. Sci. U. S. A.* 91, 12614–12618
80 Sale, J.E. and Neuberger, M.S. (1998) TdT-accessible breaks are scattered over the immunoglobulin V domain in a constitutively hypermutating B-cell line. *Immunity* 9, 859–869
81 Kong, Q. and Maizels, N. (2001) DNA breaks in hypermutating immunoglobulin genes: evidence for a break-and-repair pathway of somatic hypermutation. *Genetics* 158, 369–378
82 Papavasiliou, F.N. and Schatz, D.G. (2000) Cell-cycle-regulated DNA double-stranded breaks in somatic hypermutation of immunoglobulin genes. *Nature* 408, 216–221
83 Bross, L. *et al.* (2000) DNA double-strand breaks in immunoglobulin genes undergoing somatic hypermutation. *Immunity* 13, 589–597
84 Bemark, M. *et al.* (2000) Somatic hypermutation in the absence of DNA-dependent protein kinase catalytic subunit (DNA-PK_{cs}) or recombination-activating gene (RAG)1 activity. *J. Exp. Med.* 192, 1509–1514
85 Sohn, J. *et al.* (1993) Somatic hypermutation of an immunoglobulin μ heavy-chain transgene. *J. Exp. Med.* 177, 493–504
86 Mantovani, L. *et al.* (1993) Human rheumatoid B-1a ($CD5^+$ B) cells make somatically hypermutated high-affinity IgM rheumatoid factors. *J. Immunol.* 151, 473–488
87 Kuppers, R. *et al.* (1999) Cellular origin of human B-cell lymphomas. *New Engl. J. Med.* 341, 1520–1529
88 Klein, G. (1999) Immunoglobulin-gene-associated chromosomal translocations in B-cell derived tumors. *Curr. Top. Microbiol. Immunol.* 246, 161–167
89 Pasqualucci, L. *et al.* (2001) Hypermutation of multiple proto-oncogenes in B-cell diffuse large-cell lymphomas. *Nature* 412, 341–346
90 Ferguson, D.O. and Alt, F.W. (2001) DNA double-strand break repair and chromosomal translocation: lessons from animal models. *Oncogene* 20, 5572–5579
91 Wells, R.D. (1996) Molecular basis of genetic instability of triplet repeats. *J. Biol. Chem.* 271, 2875–2878