CLASS-SWITCH RECOMBINATION: INTERPLAY OF TRANSCRIPTION, DNA DEAMINATION AND DNA REPAIR

Jayanta Chaudhuri and Frederick W. Alt

Class-switch recombination (CSR) of immunoglobulin heavy chains is the genetic process by which a B cell switches from the production of IgM to the production of IgG, IgE or IgA. Although the general characteristics of CSR have been known for some time, the detailed molecular mechanism of this process is only now emerging. CSR is unique, in that it seems to involve transcription-generated, higher-order RNA–DNA structures, specific DNA deamination and several DNA-repair pathways. In this review, we discuss our current knowledge of the mechanism of CSR and highlight the important unanswered questions.

NON-HOMOLOGOUS END JOINING (NHEJ). The process that joins broken DNA ends without depending on extended homology. Components of this pathway include the DNA-end-binding proteins Ku70 and Ku80, the endonuclease ARTEMIS, X-ray repair cross-complementing protein 4 (XRCC4), DNA ligase IV and the catalytic subunit of DNA-dependent protein kinase (DNA-PKcs).

Howard Hughes Medical Institute, Center for Blood Research and Department of Genetics, Harvard University Medical School, Boston, Massachusetts 02115, USA.
Correspondence to F.W.A. e-mail: alt@enders.tch.harvard.edu
doi:10.1038/nri1395

The great diversity of antigen receptors that is produced by the mammalian immune system depends on the unique ability of B and T cells to somatically alter their genomes. B cells are particularly remarkable in that they can undergo three distinct genetic alterations. Developing B-lineage cells in the bone marrow or foetal liver assemble the exons that encode immunoglobulin heavy-chain (IgH) and light-chain (IgL) variable regions from component variable (V), diversity (D) and joining (J) segments, through a process known as V(D)J recombination. This site-directed, antigen-independent recombination event is initiated by the sequence-specific recombination-activating gene 1 (RAG1)–RAG2 endonuclease complex and is completed by components of the NON-HOMOLOGOUS END-JOINING (NHEJ) machinery[1,2]. Productive assembly of IgH and IgL V-region exons allows the expression of IgH and IgL chains as cell-surface IgM by newly generated B cells. IgM$^+$ B cells migrate to secondary lymphoid organs (for example, the spleen and lymph nodes), where they can undergo further antigen-driven immunoglobulin-gene diversification through somatic hypermutation (SHM) and class-switch recombination (CSR) (FIG. 1).

SHM introduces point mutations at high rates, into the V-region exons of both IgH and IgL genes, and this enables the selection of B cells that produce higher-affinity antibodies[3,4]. By contrast, CSR exchanges the initially expressed IgH constant (C)-region (Cμ) exons for an alternative set of downstream IgH C-region exons, such as Cγ, Cα or Cε, known as C_H genes. Therefore, CSR allows the expression of antibodies that have the same antigen specificity but are of a secondary IgH isotype (IgG, IgA or IgE) and thereby have a different effector function[5,6]. Despite being distinct processes targeted to distinct immunoglobulin regions (TABLE 1), CSR and SHM are similar in that they both occur in antigen-stimulated B cells, require transcription through the target region and require the activated B-cell-specific enzyme activation-induced cytidine deaminase (AID)[3–7].

The mouse immunoglobulin locus contains eight C_H genes, which encode proteins that are capable of different effector functions. Each C_H gene, except Cδ, is preceded by specialized DNA sequences called switch (S) regions. CSR involves a recombination event between two S regions, with the intervening sequences, including Cμ, being deleted. As a result, the VDJ exon is juxtaposed with a different downstream C_H gene[5,6] (FIG. 1). Mechanistically, CSR is a deletional recombination reaction that, similar to V(D)J recombination, could best be explained by a 'cut and paste' mechanism, in which breaks are introduced to the DNA of the two participating S regions followed by their fusion. However, CSR differs from

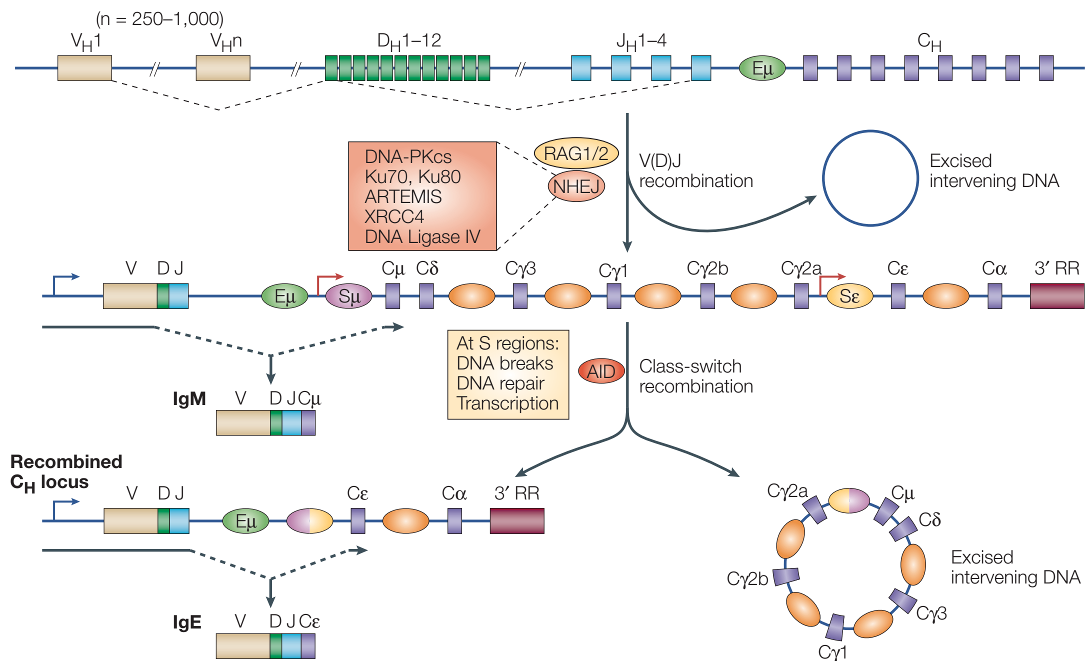


Figure 1 | **Rearrangement at the immunoglobulin heavy-chain locus.** The variable region of the immunoglobulin heavy chain is assembled from component variable (V_H), diversity (D_H), and joining (J_H) gene segments by V(D)J recombination. The process of rearrangement involves cleavage of the recombination signal sequences in the DNA, which flank the rearranging gene segments, which is carried out by the recombination-activating gene 1 (RAG1)–RAG2 complex. Joining of the DNA ends requires non-homologous end-joining (NHEJ) proteins, including Ku70, Ku80, ARTEMIS, X-ray repair cross-complementing protein 4 (XRCC4), DNA ligase IV and the catalytic subunit of DNA-dependent protein kinase (DNA-PKcs). Transcription across the locus is driven by a promoter upstream of the rearranged VDJ segment (blue arrow), which facilitates the synthesis of a μ heavy chain. This then associates with a light chain, thereby forming an IgM molecule, which is displayed on the cell-surface of a B cell. Subsequently, secondary isotypes are produced by class-switch recombination (CSR), a process that exchanges the constant region of the heavy chain (C_H) with a set of downstream constant-region genes (CSR to IgE is shown). This deletional-recombination reaction, which requires the enzyme activation-induced cytidine deaminase (AID), involves the generation of DNA breaks at switch (S) regions, which precede the constant-region genes, followed by the repair of DNA. This leads to a rearranged C_H locus and deletion of the intervening sequence as an episomal circle. Cytokines stimulate transcription (red arrows) through the C_H gene and determine the immunoglobulin isotype that the B cell will switch to. The rearranged variable regions of both the heavy and light chains also undergo a high rate of point mutation through the process of somatic hypermutation (SHM) (not shown). The Eμ and 3′-regulatory-region (3′ RR) enhancers influence V(D)J recombination and CSR, respectively.

V(D)J recombination in several ways: first, CSR does not require RAG[8], and second, CSR has no well-defined consensus target sequence[9].

Recent studies have indicated a working model for the initiation of CSR, in which transcription through the S regions renders these regions substrates for modification by AID[10,11]. Other work has implicated several DNA-repair pathways in the processing and joining of AID-initiated S-region breaks that complete CSR[12]. In this context, several lines of evidence indicate that CSR uses a DNA double-stranded-break (DSB) intermediate[5,6]. In this review, we discuss the current understanding of CSR and the role of S regions, transcription, AID, DSBs and DNA-repair proteins.

Transcription-generated S-region structures

S regions consist of highly repetitive 1–12-kilobase (Kb) sequences with G-rich non-template strands[13]. The S-region repeats can be divided into two general categories: Sμ, Sα and Sε are composed of variations of pentameric sequences, whereas Sγ1, Sγ2a, Sγ2b and Sγ3 mostly contain 49–52-base-pair repeats (FIG. 2a). Deletion of most of the Sμ region has been shown to severely impair CSR to downstream C_H genes[14], whereas deletion of the entire Sγ1 region essentially blocks switching to IgG1 (REF. 10), indicating that the S regions are specialized targets for the CSR process. However, there are certain caveats associated with this interpretation, as neither S-region-deletion study replaced the deleted sequences with an unrelated sequence of equal size.

Transcription through a particular S region is required for CSR to the corresponding C_H gene[5]. In response to specific activators and CYTOKINES (TABLE 2), transcription is initiated from an activation/cytokine-responsive promoter upstream of an exon (known as the I exon) that precedes all C_H genes that undergo CSR (FIG. 2b). Transcription proceeds through the S region and terminates downstream of the corresponding C_H gene. The primary transcript is spliced to remove the intronic S region and to join the I exon to the C_H exons, producing a seemingly non-coding transcript[5] (FIG. 2b). This requirement for transcription and transcription-control elements for CSR has been demonstrated directly in gene-targeting experiments, in which deletion of I-exon

CYTOKINES
Biologically active molecules that are released by cells and affect the function of other cells.

Table 1 | **Main differences between CSR and SHM**

Parameters	SHM	CSR
Target	Variable-region coding DNA in both the heavy and light chains	Switch-region non-coding DNA in the constant region of heavy chains
Transcription-generated structures	Transient transcription bubbles, other structures?	Stable RNA–DNA hybrids, other structures?
AID activity	Single deaminated residue can lead to SHM	High density of deamination might be required, at least in one strand
Completion residue can lead to SHM;	Replication across deaminated intermediates probably does not require DSB intermediates	Probably requires DSB

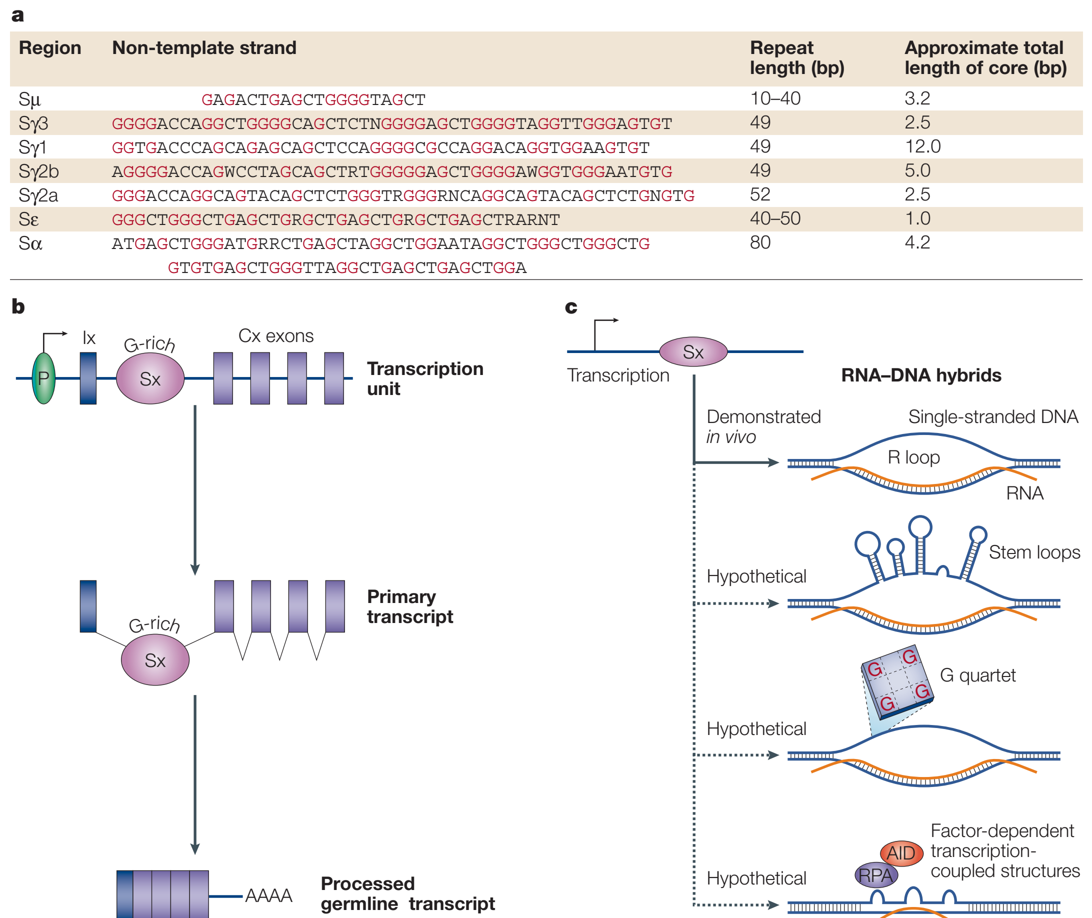
AID, activation-induced cytidine deaminase; CSR, class-switch recombination; DSB, double-stranded break; SHM, somatic hypermutation.

promoters or certain 3′ IgH-enhancer elements abolished or greatly reduced germline transcription of the corresponding C_H genes and CSR to these genes[5,15]. Furthermore, constitutively active heterologous promoters inserted in place of endogenous I-region promoters can also drive CSR[5,16]. Therefore, transcription directly influences CSR.

The role of germline transcription in CSR has long been a subject of speculation. Transcription has been postulated to function at several non-mutually exclusive levels, including rendering S-region chromatin structures accessible to the enzymes involved in CSR[17,18] and generating DNA structures able to be substrates of a CSR recombinase[19–22]. Support for the DNA-structure model arises from the notable feature of all mammalian S regions — that the non-template strand is G-rich. Moreover, when S regions are transcribed *in vitro* in the normal physiological orientation, the transcripts stably associate with the template strand of the DNA to form RNA–DNA hybrids[19–24] (FIG. 2c). More recently, these RNA–DNA hybrids were demonstrated to form R loops, in which the displaced non-template strand exists as single-stranded (ss) DNA, both *in vitro*[22,25] and *in vivo*[25]. Theoretically, the non-template strand of the RNA–DNA hybrid could assume other structures in addition to being single-stranded in R loops: for example, stem loops, from the abundant palindromic motifs in S regions[26], or parallel four-stranded guanosine quartets, stabilized by Hoogstein pairing between G residues[27,28] (FIG. 2c). However, although these latter structures have been postulated to have a role in CSR[26,28], there is no direct evidence to support this hypothesis, and the existence of stem loop and G-quartet structures during *in vivo* CSR has not been demonstrated.

Thermodynamic properties predict that stable R loops should form only when mammalian S regions are transcribed in the physiological orientation[29], and this prediction has been confirmed by *in vitro*-transcription studies[19–25]. Therefore, if RNA–DNA hybrids have a mechanistic role in CSR, the transcriptional orientation of S regions should influence CSR considerably. Studies carried out using extrachromosomal substrates have provided the initial evidence that recombination at S regions is orientation dependent[21]; although this idea has been challenged by other studies that used chromosomally integrated, short S-region substrates[30]. Recent studies directly examined orientation *in vivo* and found that inversion of the 12-Kb Sγ1 significantly reduced CSR to IgG1, although it was not completely abolished[10]. In addition, a random synthetic 1-Kb sequence that was G-rich on the non-template strand and formed R loops *in vitro*, but lacked the characteristic S-region repeat elements, could replace native Sγ1 in supporting CSR *in vivo*, although at a reduced level. Importantly, the activity of this synthetic sequence was again dependent on transcriptional orientation, because it targeted almost no CSR activity when in the inverted (C-rich), non-R-loop forming configuration[10]. These findings provided evidence for the orientation dependence of S-region transcription *in vivo* and led to the proposal that transcription-generated DNA structures, in particular ssDNA, might be substrates for the enzymes involved in CSR[10]. In addition, the finding that CSR can be supported by a synthetic G-rich sequence with no repeat motifs characteristic of endogenous Sγ1 provides further evidence that not only primary S-region structures but also higher-order secondary structures have an important role in targeting CSR.

The finding that inverted Sγ1 supports some CSR (25% of the normal level), even though this orientation does not lead to readily detectable R-loop formation, indicates that other mechanisms might exist to reveal ssDNA structures during CSR[10]. Consistent with this, *Xenopus* S regions are (A+T)-rich[31] so would not be predicted to show extensively displaced non-template ssDNA strands. However, similar to mammalian S regions, *Xenopus* S regions are rich in palindromic repeats, which if rendered single-stranded, could form stem loops. Indeed, recombination breakpoints in *Xenopus* S regions are observed frequently at positions where a computer program, which models ssDNA folding, predicts a transition from a stem to a loop structure[31]. Consistent with this, sequences with the potential to form palindromes supported low-level recombination within substrates introduced into cell lines[26]. Notably, other work has implicated transcriptionally generated ssDNA substrates as having a role in targeting SHM to sequences that do not form R loops[32–34], indicating additional, potentially factor-dependent mechanisms for generating substrates that do not readily reveal extensive ssDNA structures. Given that the initiating enzyme, AID, is the same for SHM and CSR, this finding might also have implications for this aspect of CSR. In support of



Region	Non-template strand	Repeat length (bp)	Approximate total length of core (bp)
Sμ	GAGACTGAGCTGGGGTAGCT	10–40	3.2
Sγ3	GGGGACCAGGCTGGGGCAGCTCTNGGGGAGCTGGGGTAGGTTGGGAGTGT	49	2.5
Sγ1	GGTGACCCAGCAGAGCAGCTCCAGGGGCGCCAGGACAGGTGGAAGTGT	49	12.0
Sγ2b	AGGGGACCAGWCCTAGCAGCTRTGGGGGAGCTGGGGAWGGTGGGAATGTG	49	5.0
Sγ2a	GGGACCAGGCAGTACAGCTCTGGGTRGGGRNCAGGCAGTACAGCTCTGNGTG	52	2.5
Sε	GGGCTGGGCTGAGCTGRGCTGAGCTGRGCTGAGCTRARNT	40–50	1.0
Sα	ATGAGCTGGGATGRRCTGAGCTAGGCTGGAATAGGCTGGGCTGGGCTG GTGTGAGCTGGGTTAGGCTGAGCTGAGCTGGA	80	4.2

Figure 2 | **Switch-region sequence, transcription and transcription-generated structures in CSR. a** | The switch (S)-region repeat elements, which are similar but not identical to each other, are shown. All S regions are G-rich in the non-template (top) strand. N, any nucleotide; R, a purine; W, A or T; Y, a pyrimidine. The table has been compiled from previous studies[15]. **b** | The heavy-chain constant-region (C_H) genes are transcription units, in which transcription initiates from cytokine-inducible I-promoters (P) upstream of the I-exon (I) and proceeds through S regions and the C_H exons. The primary transcript is processed to generate mature germline transcripts that do not code for proteins. **c** | S-region transcripts can stably associate with the template strand to form RNA–DNA hybrids, in which the non-template strand can theoretically assume several structures (including G quartets or stem loops) or can remain single stranded (R loops). Although R loops have been detected *in vivo* and are strongly implicated to have a mechanistic role, at present, there is no direct experimental evidence for the formation of stem loops and G-quartet structures *in vivo*. AID, activation-induced cytokine deaminase; bp, base pairs; RPA, replication protein A; UNG, uracil-DNA glycosylase; x, any constant-region gene.

this, in *Xenopus* S regions, the areas where most CSR breakpoints occur are rich in RGYW MOTIFS, which are also 'hot spots' for SHM[31]. Therefore, these sequences might function somehow to target AID activity in both processes. In this regard, recent studies have implicated replication protein A (RPA), a ssDNA-binding protein, as a factor that can target AID to transcribed sequences in the context of RGYW motifs (Chaudhuri, J., Khong, C. & Alt, F. W., unpublished observations). Overall, current findings indicate that transcription might function in the context of various sequence motifs, and several mechanisms might exist to provide ssDNA substrates that are probably the targets of AID activity.

RGYW MOTIFS
The nucleotide sequence RGYW (where R denotes A or G, Y denotes C or T, and W denotes A or T) is considered a 'hot spot' for somatic hypermutation, as it is considerably more mutable than other random sequences.

The Function of AID in CSR

AID and DNA deamination. The *AID* cDNA was cloned from a subtractive screen for genes that were upregulated in a B-cell line induced to undergo CSR[35]. AID is only expressed by activated B cells[35] and is essential for CSR and SHM[36,37]. Ectopic AID expression by non-lymphoid cells enabled CSR and SHM on artificial substrates[38–40], implicating AID as the only lymphoid-specific factor required for these processes. However, the efficiency of CSR was low in these ectopic-expression studies, and whereas SHM is restricted primarily to V genes in B cells, several transcribed DNA sequences were shown to serve as AID substrates during SHM in non-lymphoid cells, implying that additional factors might be required for the efficiency and specificity of CSR and SHM. AID shares sequence homology with the RNA-editing cytidine deaminase APOBEC1 (apolipoprotein B mRNA-editing enzyme, catalytic polypeptide 1)[35], which led to initial speculation that AID edits an mRNA such that it encodes a novel protein required to activate CSR and SHM[36,37] (FIG. 3). Although this is still considered a possibility by some[41], no direct experimental

Table 2 | **In vitro activation of CSR**

Cytokine	Ig isotype
LPS	IgG2b, IgG3
LPS and IL-4	IgG1, IgE
CD40-specific antibody and IL-4	IgG1, IgE
LPS and IFN-γ	IgG2a
LPS, TGF-β and IL-5	IgA

CSR, class-switch recombination; IFN, interferon; IL, interleukin; LPS, lipopolysaccharide; TGF, transforming growth factor

evidence has been obtained to support the RNA-editing activity of AID. Moreover, it now seems that APOBEC1 is an outlier in a family of DNA-cytidine deaminases that evolved to mediate innate immunity[42]. Therefore, for the remainder of the review, we focus on AID as a DNA-modifying enzyme, because this activity has been clearly documented.

Several lines of evidence indicate that AID acts directly on DNA as a DNA-cytidine deaminase. First, enforced expression of AID in *Escherichia coli*, which are not expected to contain any putative mRNA substrates of AID, induced mutations in several bacterial genes[43]. Second, *in vitro* studies showed that purified AID is a ssDNA deaminase that converts deoxycytidine residues to deoxyuridine on target DNA[11,33,44,45]. Last, chromatin-immunoprecipitation experiments demonstrated that ectopically expressed AID is directly associated with the transcribed S-region DNA in cells that are undergoing CSR[46]. The finding that AID deaminates only ssDNA *in vitro* supports the argument that CSR might be enhanced by ssDNA targets, such as those provided by R-loop formation. Indeed, the G-rich DNA that supported CSR *in vivo*[10] was efficiently deaminated by AID *in vitro*, mainly on the non-template strand (that is the ssDNA of the R loop), when the reaction was coupled to transcription. Conversely, C-rich DNA that did not target CSR *in vivo* was a poor AID substrate when transcribed *in vitro*[11]. Overall, the current findings provide a plausible model by which transcription through mammalian S regions could generate ssDNA R-loop substrates for the cytidine deaminase activity of AID and thereby generate the initiating DNA lesion that is ultimately processed into a DNA break in the S region. In addition, it is probable that there are factors that function in the context of transcription of non-R-loop-forming sequences to generate ssDNA structures that are also AID targets, potentially focusing somehow on RGYW motifs for SHM. Moreover, such a mechanism might also function in CSR, in conjunction with R-loop formation in mammalian S regions (perhaps to target the template strand), or in place of R loops for *Xenopus* S regions.

LIGATION-MEDIATED PCR (LM-PCR). This method, which involves the ligation of unphosphorylated adaptors to broken DNA followed by PCR, can detect DNA double-stranded breaks in defined regions of the genome. It has been used extensively to analyse recombination-activating gene (RAG)-dependent breaks during V(D)J recombination, and more recently, in variable (V) gene segments and switch (S) regions during somatic hypermutation and class-switch recombination, respectively.

DNA DSBs in CSR. The generation of an episomal circle in CSR strongly suggests that CSR proceeds through DNA DSB intermediates. Indeed, in LIGATION-MEDIATED PCR (LM-PCR) assays, DSBs have been detected in S regions[47,48], and these appear to be AID dependent[48]. In addition, the foci of the phosphorylated histone H2AX (γ-H2AX), which form around DNA DSBs[49], occur at the IgH locus in an AID-dependent fashion in B cells that are activated for CSR[50]. A deficiency in H2AX also leads to S-region translocations[51]. Furthermore, a deficiency in p53-binding protein 1 (53BP1), an early component of the DSB response, severely impairs CSR[52]. Finally, some components of the NHEJ pathway, Ku and the catalytic subunit of DNA-dependent protein kinase, DNA-PKcs, are required for effective CSR *in vivo*, which also indicates a role for DSB intermediates[53–55]. Although such evidence points strongly to the involvement of DSB intermediates in CSR, this interpretation requires some caution. LM-PCR studies have detected DSBs in V genes, both in the presence and absence of AID[56–58], and although γ-H2AX foci clearly form at the IgH locus in B cells undergoing CSR, it has not been formally proven that this occurs in response to DSBs in S regions, rather than other regions of the IgH locus[59]. However, despite these potential caveats, all available data are consistent with the idea that CSR proceeds through DSB intermediates.

Given that DSBs are essential intermediates in CSR, how does AID-mediated DNA deamination lead to DSBs? It was proposed that during CSR, the AID-introduced deoxyuridine in S-region DNA is removed by the base-excision-repair (BER) enzyme uracil-DNA glycosylase (UNG) to generate an abasic site, the processing of which, by the apurinic/apyrimidinic endonuclease 1 (APE1), creates a nick[43]. A closely spaced, similarly created nick on the opposite strand could lead to a staggered DSB. In addition, the deoxyuridine–deoxyguanosine mismatch could also be processed by components of the mismatch repair (MMR) machinery to generate staggered DSBs[43]. Accordingly, both UNG deficiency[60,61] and MMR deficiency[62–64] result in CSR defects. Therefore, deaminated DNA might be processed by several DNA-repair pathways, which ultimately lead to the generation of a DNA DSB (FIG. 4).

The AID–UNG–APE1 pathway for the generation of DSBs predicts that AID needs to deaminate both the template and the non-template strands during CSR. *In vitro* studies using a G-rich synthetic substrate showed that although the non-template ssDNA is the preferred AID substrate, the template strand is also deaminated[11]. How the template strand is rendered a substrate is not known, but several mechanisms have been considered. One possibility is that AID acts at ssDNA exposed by the transition between duplex DNA and the R-loop structure[22]. In addition, it has been proposed that collapse of R loops, perhaps secondary to RNA removal by RNase H, could lead to misalignment of repeats on opposite strands, leading to single-stranded loops on both the template and the non-template strands of the DNA[65]. Furthermore, there must be mechanisms other than R loops that allow AID to have access to the DNA (discussed earlier); this includes the possibility that additional factors, such as RPA, target AID to non-R-loop-forming sequences in the context of transcription-coupled mechanisms. In theory, several mechanisms could work together to generate lesions on both S-region strands. For this to occur, a mechanism that generates a

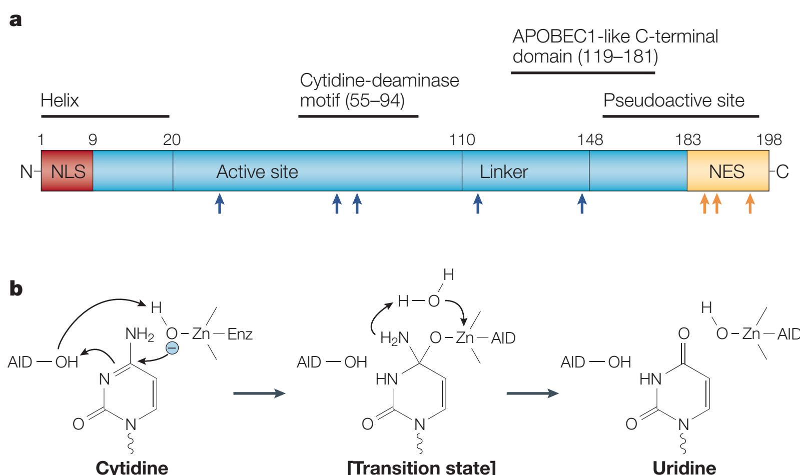


Figure 3 | **AID and DNA deamination. a** | The primary structure of activation-induced cytidine deaminase (AID) is shown, depicting the nuclear-localization sequence (NLS), nuclear-export sequence (NES) and other predicted domains based on the structure of apolipoprotein B mRNA-editing enzyme, catalytic polypeptide 1 (APOBEC1)[103]. Some of the mutations that affect AID function are shown: mutations in the active site and linker region (blue arrows) impair both class-switch recombination (CSR) and somatic hypermutation (SHM), whereas those at the carboxyl (C)-terminus (orange arrows) and a ten amino-acid C-terminal deletion (not shown) impair CSR without affecting SHM. N, amino terminus. **b** | Mechanism of cytidine deamination, based on a bacterial cytidine deaminase that is homologous to APOBEC1 and AID. The reaction proceeds through a direct nucleophilic attack at position 4 of the pyrimidine ring by zinc ions (Zn^{2+}) coordinated with AID[12].

high density of nicks on one strand (for example, by R-loop formation) would only require a mechanism that generates a low density of nicks on the other strand, to allow formation of a set of closely opposing nicks that would create a DSB (FIG. 5).

Processing of DSBs during CSR. Irrespective of the process that generates AID substrates, the deamination model predicts that most DSBs would result from staggered cleavages. Indeed, studies using LM-PCR assays have indicated that most AID-dependent Sμ lesions are staggered single-stranded breaks[66]. Such lesions might need to be filled or resected to generate blunt DSBs or overhangs short enough for efficient ligation[6]. This could be achieved by either nuclease-mediated recession or fill-in through short-patch DNA synthesis, using error-prone polymerases, and could potentially account for the deletions and mutations observed at S-region junctions[9] (FIG. 4). Consistent with this, the absence of the error-prone polymerase η has been shown to affect the mutation profile of S-region junctions[67]. The MMR protein mutS homologue 2 (MSH2) has also been implicated in end processing during CSR[63,64,68,69], and because yeast MSH2 homologues can recruit nucleases, such as the RAD1–RAD10 complex, to cleave branched intermediates[70], it is possible that MSH2 could similarly recruit the mammalian homologues of RAD1 and RAD10, the ERCC1–XPF complex (excision repair cross-complementing protein 1–xeroderma pigmentosum, complementation group F), in B cells to process DNA ends. Both ERCC1–XPF and XPG (xeroderma pigmentosum, complementation group G) have been speculated to participate in CSR as a result of their ability to cleave ssDNA in transcribed S regions *in vitro*[22]. However, mutations in XPF or in XPG do not have a readily detectable effect on the rate of CSR[71,72]; although this finding does not rule out redundancy with other repair pathways.

SYNAPSIS
Non-covalent juxtaposition of two non-adjacent stretches of DNA.

Cell-cycle regulation of CSR

V(D)J recombination occurs in the G1 phase of the cell cycle where NHEJ is the main DSB-repair pathway[1], and understanding the point at which CSR occurs in the cell cycle is crucial for elucidating its mechanism. CSR requires proliferation[73], and inhibitors of DNA synthesis block CSR, indicating a role for progression into S phase and/or replication[16]. In this context, H2AX deficiency on a *p53*-mutant background leads to B-lineage tumours with S-region translocations that might have been initiated in S phase[51]. By contrast, on the basis of the point in the cell cycle at which AID-dependent, IgH-associated γ-H2AX foci appear in B cells activated for CSR, it has been argued that AID deamination and generation of CSR DSBs occurs in G1 (REF. 50). We note that none of the existing pieces of evidence relevant to the cell-cycle dependence of CSR are mutually exclusive, and it is not impossible that processing of lesions initiated in G1 might occur in S phase. In terms of the mechanism, replication, if involved, might function to convert single-stranded S-region nicks into DSBs[74]. Alternatively, CSR might use pathways or factors that are only active, or mainly active, in a particular phase of the cell cycle. Additional work on this important aspect of CSR is required.

Switch region synapsis

Internal deletions in Sμ in the absence of CSR. CSR involves recombination between two S regions, and an efficient recombination reaction probably requires close juxtaposition, or SYNAPSIS, between the two recombining sequences. During V(D)J recombination, the RAG1–RAG2 endonuclease complex cleaves its recognition sites only in the context of a preformed synaptic complex involving two complementary RAG-recognition signal sequences[1]. For CSR, coordinate interaction between two different donor and acceptor S regions does not seem to be required to target AID activity, as indicated by the observation that constitutively transcribed Sμ or Sγ2b regions integrated randomly into a pro-B cell line undergo a high rate of internal deletions following AID expression[75].

In splenic B cells and in B-cell lines activated to undergo CSR, large deletions occur frequently within the endogenous Sμ region[75–77] (FIG. 5). These deletions are AID-dependent and therefore seem to represent the same general process as CSR[75]; however, there has been no direct demonstration that the Sμ breaks that lead to internal deletions are the same as those channelled into CSR with a downstream S region. In theory, unique structural or transcriptional properties of Sμ (for example, proximity to the Eμ intronic enhancer) render it a preferred AID substrate, with AID-targeting to downstream S regions constituting the rate-limiting step in CSR[78]. Therefore, in the

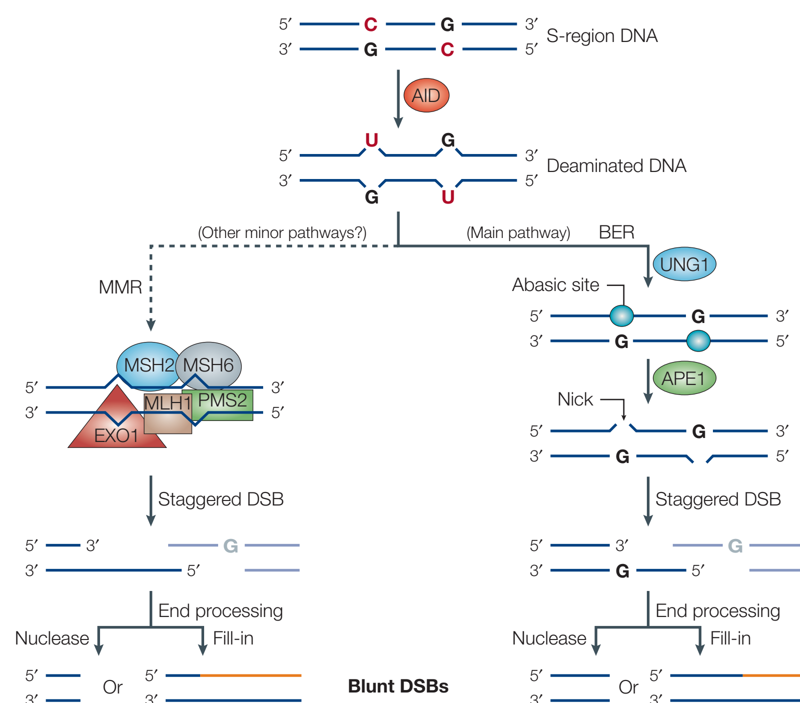


Figure 4 | **Generation of DNA double strand breaks in S regions.** Activation-induced cytidine deaminase (AID) deaminates cytidine residues in DNA converting them to uridine residues. The G–U mismatch can then be processed by either the base-excision repair (BER) pathway or the mismatch-repair (MMR) machinery — which includes the mutS homologue 1 (MSH1), MSH6, exonuclease 1 (EXO1), mutL homologue 1 (MLH1) and post-mitotic segregation (PMS) proteins — to introduce gaps or nicks on opposite strands of the switch (S)-region DNA. The nicks induced by the BER pathway are thought to be generated by the following process: uracil-DNA glycosylase (UNG) removes the AID-introduced deoxyuridine in S-region DNA, creating an abasic site (blue circles) that is processed by the apurinic/apyrimidinic endonuclease 1 (APE1), which creates the nick. Processing of the staggered ends by unknown exonucleases or by error-prone end-filling reactions (orange lines) can lead to blunt double-stranded breaks (DSBs) that can then be ligated to similarly created breaks on downstream S-region DNA to complete class-switch recombination (CSR).

absence of an acceptor S region, DSB Sμ lesions would be resolved by intra-S-region recombination, leading to internal Sμ deletions. Alternatively, when the acceptor S region DSBs are available, Sμ breaks could be efficiently channelled into a genuine CSR reaction. This scenario is consistent with the observation that downstream S-region deletions, in the absence of productive CSR, have rarely been observed[75], unless Sμ was effectively inhibited from being a donor[77].

DSB response and other factors potentially involved in synapsis. Several proteins have been potentially implicated in S-region synapsis. Deficiency in H2AX reduces CSR by 50–80%[50], but mutations and deletions within Sμ occur at normal levels, indicating that targeting of AID to Sμ and assembly of DNA-repair components is not affected and that the CSR defect might be due to an inability to synapse Sμ breaks with downstream S regions[79]. In this regard, γ-H2AX has been proposed to function as an anchor to recruit and assemble lattices of DNA-repair factors at sites of DSBs and thereby to hold broken chromosomal ends in close proximity to allow proper NHEJ[80]. In theory, H2AX might work in a similar way to facilitate the joining of two S regions that contain DSBs[80]. Correspondingly, deficiency for 53BP1, an H2AX-associated DSB response factor nearly abrogates CSR, due to the impairment of CSR steps downstream of AID induction and germline transcription; this is consistent with a role for this factor in the joining phase of CSR[52]. Although the idea that γ-H2AX and its associated factors might function in the context of S-region synapsis during CSR is intriguing, other interpretations are possible. To fully evaluate the synapsis model, it will be important to demonstrate that Sμ deletions and mutations represent a true CSR-initiation reaction, rather than an SHM-like process promoted by proximity to the elements that specify SHM.

The ataxia telangiectasia mutated (ATM) protein, which is a master regulator of the DSB response, has been suggested to have a role in CSR, as ataxia telangiectasia patients, who have mutations in ATM, have modestly reduced levels of IgA, IgE and IgG subclasses[81] and their Sμ–Sα junctions show significantly increased microhomology[82,83]. The precise role of ATM in CSR is unknown, and it is not clear if it has a direct role in CSR or whether the decreased CSR potential of ATM-deficient B cells is secondary to other defects in B-cell development. It is to be noted, however, that H2AX and 53BP1 are activated in the DSB response by ATM and other phosphatidylinositol 3-kinase (PI3K)-related kinases, such as ATM-related protein (ATR) and DNA-PKcs[84–86]. The CSR defect of 53BP1-deficient cells seems more severe than in ATM-deficient cells[52,87], potentially indicating roles for other PI3K-related kinases in CSR. Finally, DNA-PKcs has the ability to synapse DNA ends *in vitro*[88] and, theoretically, could have such a function in CSR.

The heterodimer PMS2–MLH1 (post-mitotic segregation 2–MutL homologue 1) participates in MMR downstream of MSH2-mediated recognition of DNA mismatches. Mice deficient in PMS2 and MLH1 show a 2–5-fold decrease in the efficiency of CSR, and their S junctions have increased microhomology relative to normal mice[68,89]. These findings, along with the observation that yeast homologues of PMS2 and MLH1 can bind directly to two different DNA molecules[90], led to the proposal that these proteins might participate in S-region synapsis during CSR[91]. MLH1 interacts with ATM both *in vitro* and *in vivo*, and the interaction is stabilized by DSBs[92], providing a further potential link. Finally, the DNA binding protein lipopolysaccharide-responsive factor 1 (LR1) has been proposed to participate in S-region synapsis[28]. LR1 is a heterodimer of nucleolin and heterogeneous nuclear ribonucleoprotein D (HNRD), and each subunit can individually bind to G-quartet DNA *in vitro*, leading to the proposal that a single heterodimer of LR1 can bind to two separate S regions — to juxtapose donor and acceptor S regions during CSR[28]; however, this has not yet been formally tested.

In summary, although several factors have been suggested to be involved in the synapsis of donor and acceptor S regions, none of the factors tested have

been found to be absolutely required for CSR, and their precise roles remain to be determined. So, elucidation of the S-region-synapsis mechanism remains an important challenge.

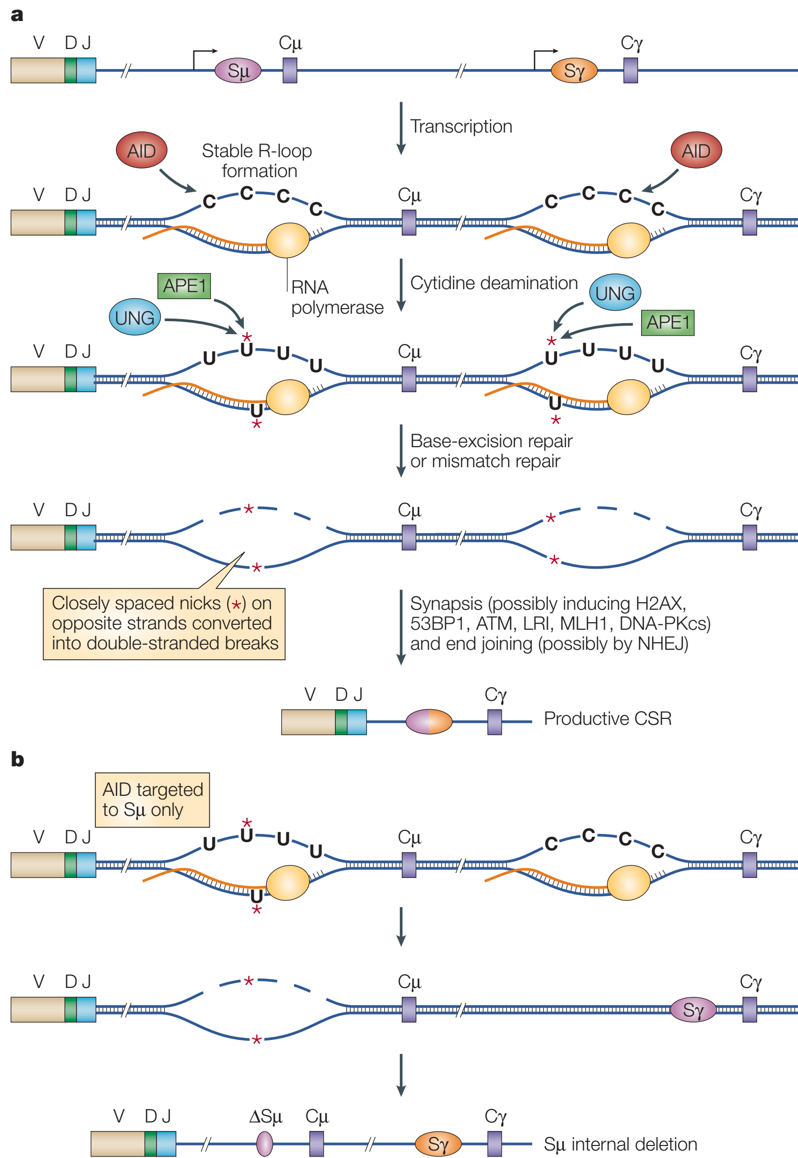


Figure 5 | **Current model for CSR. a** | Transcription through the two participating switch (S) regions generates R loops (or other secondary structures, not shown) that provide single-stranded DNA substrates for activation-induced cytidine deaminase (AID). Subsequently, the activities of uracil-DNA glycosylase (UNG) and apurinic/apyrimidinic endonuclease 1 (APE1) can introduce a high density of nicks on the non-template strand, and this, combined with a closely spaced nick on the opposite strand (*), can generate double-stranded breaks (DSBs) in the S regions. The breaks are synapsed in a process that requires the putative participation of several proteins, including histone 2A family-member X (H2AX), p53 binding protein 1 (53BP1), lipopolysaccharide-responsive protein1 (LR1), ataxia telangiectasia mutated (ATM), mutL homologue 1 (MLH1) and DNA-PKcs (catalytic subunit of DNA-dependent protein kinase). Class-switch recombination (CSR) is completed by fusion of two S regions DSBs, possibly by the non-homologous end-joining pathway (NHEJ)— which has been strongly implicated to have a role in CSR, although this has not been unequivocally proven. The pink and orange oval at the recombination site represents deletions and mutations that are found at S junctions. **b** | In addition to productive CSR, the Sμ region can undergo a high rate of internal deletions (ΔSμ) when this region is targeted by AID but the downstream S region is not. It is not clear if such intraswitch recombination reactions have the same requirements as productive CSR. Cμ, constant region of IgM heavy chain; D, diversity gene segment; J, joining gene segment; V, variable gene segment.

DNA end joining

Completion of CSR seems to involve the joining of two broken S regions. The two main pathways that join broken DNA ends in mammalian cells are homologous recombination and NHEJ. The limited or complete lack of homology at S-region junctions rules out the possibility that the former pathway has a main role, leaving NHEJ as the most probable candidate. A role for NHEJ in CSR was indicated by the observation that Ku70- or Ku80-deficient B cells reconstituted with rearranged IgH and IgL chains had severely impaired CSR[53,54,79]. However, such an effect on CSR could be indirect as Ku-deficient B cells had proliferative defects and died when induced to undergo CSR[54,79]. The latter effect might result from an inability to repair DSBs generated during CSR, because cells activated to divide, but not undergo CSR, proliferated relatively normally[54]. It is also possible that Ku-deficient cells could die before undergoing productive CSR owing to abortive Sμ breaks that are uncoupled from CSR. Although this argument has been countered by the finding that Ku80-deficient B cells that have undergone several rounds of cell division still show impaired CSR[79], such cells could represent those that have escaped activation. Therefore, although Ku activity is required for CSR, it has not been formally proven that its role is through NHEJ.

Unlike Ku-deficient B cells, B cells from mice with a targeted disruption of DNA-PKcs, which completely eliminates expression of any DNA-PKcs protein, showed no proliferation defects[55]. Yet DNA-PKcs deficiency severely impaired CSR to all IgH isotypes except IgG1 (REF. 55), strongly indicating a direct role for DNA-PKcs in CSR. However, analysis of B cells derived from mice homozygous for the severe combined immunodeficiency (SCID) mutation (which inactivates the DNA-PKcs kinase domain but leaves most of the protein intact) has complicated this interpretation. In one study, B cells from SCID mice were found to undergo nearly wild-type levels of CSR[93], whereas another study reported reduced, but significant, levels of CSR to most C_H genes[94]. There could be several possible interpretations of these results: an intriguing one being that the kinase-deficient DNA-PKcs protein in B cells from SCID mice could provide a necessary function for CSR. In this scenario, the role of DNA-PKcs in CSR would be separate from its role in NHEJ, which requires a functional kinase domain. Accordingly, one of the main roles of DNA-PKcs, in the NHEJ component of V(D)J recombination, is end processing through phosphorylation of the ARTEMIS endonuclease[95], and ARTEMIS is dispensable for CSR (Rooney, S., Manis, J. P. & F.W.A. unpublished observations).

One potential kinase-independent role for DNA-PKcs could be its previously mentioned function as a synaptic factor, because the DNA-bridging property of DNA-PKcs *in vitro* is independent of its kinase activity[88]. Other potential DNA-PKcs roles could be the activation of other factors involved with AID recruitment or in the response to DSBs downstream of AID. The observed DNA-PKcs-independent CSR activity of Sγ1 might reflect the differential ability of this S region,

which is considerably larger than the others, to be synapsed or targeted for AID activity. Finally, genetic experiments have provided clear evidence that Ku and DNA-PKcs have NHEJ-independent functions in DNA repair, which might overlap with ATM[96]. So, in this way, even if the CSR defects observed in Ku and DNA-PKcs reflect a true requirement of the proteins in CSR, this does not formally prove that their effects on CSR are mediated through the NHEJ pathway. So far, X-ray repair cross-complementing protein 4 (XRCC4) and DNA Ligase IV are the only proteins required for all types of NHEJ reactions that have no reported roles outside NHEJ. Therefore, although most available evidence points to a role for NHEJ factors in CSR, elucidation of the role of XRCC4 or DNA ligase IV would provide the most unequivocal proof.

Potential Roles of AID in CSR downstream of deamination. Recently, it has been argued that AID itself might participate in the joining phase of CSR. Mutational analyses revealed the C-terminus of AID to be crucial for CSR but dispensable for SHM and gene conversion[97,98] (FIG. 3). Mechanistically, CSR and SHM pathways diverge with respect to targeting — V genes versus S regions — and in the requirement of proteins downstream of deamination — for example, H2AX, 53BP1, and DNA-PKcs — for CSR[50,52,55] but not SHM[52,79,99], probably because of the mechanistic requirement for DSBs in CSR but not SHM (TABLE 3). C-terminal AID mutations do not impair Sμ mutations, indicating that targeting of AID, at least to Sμ, is not affected, leading to the suggestion that the C-terminus of AID is involved in the recruitment or assembly of CSR DNA-repair proteins or in the synapsis of the acceptor and donor S regions[97,98]. However, an alternative interpretation of these interesting findings could be that these AID mutants, owing to an altered conformation or an inability to interact with CSR-specific cofactors, cannot target R loops or other CSR-associated DNA structures with a frequency high enough to generate DSBs in S regions. In this scenario, Sμ mutations might not represent a CSR-initiation reaction, as they might occur through an SHM-like mechanism (for example, because of their proximity to the IgH intronic enhancer). Unlike CSR, SHM probably does not proceed through DSB intermediates, and therefore might not require a high density of AID-mediated target-DNA deaminations. Concerning the hypothesis that CSR might require a higher incidence of deamination relative to that required for SHM, the observation that V genes have a much higher rate of SHM mutation than S regions would need to be accounted for. It is conceivable, given that the proteins and pathways required downstream of deamination are clearly different between CSR and SHM, that during CSR, but not SHM, most deaminated residues in S regions are converted into abasic sites and then into nicks.

Table 3 | **Role of proteins involved in CSR and SHM**

Protein	SHM phenotype in absence of activity	CSR phenotype in absence of activity	Activity in CSR	References
B-cell specific				
AID	Blocked	Blocked	Single-stranded cytidine deaminase, role in recruiting DNA-repair protein?	11,32–37, 44,45,97,98
Base-excision repair				
UNG	Reduced transversion mutations at G–C base pairs	Severely impaired	Removes deoxyuridine from S regions to create abasic sites	60,61
Mismatch repair				
MSH2	Decrease in A–T mutations	2–5-fold reduction	Mismatch recognition, end processing	62–64,68, 69,104
PMS2, MLH1	Decrease in A–T mutations	2–5-fold reduction	Mismatch processing, end-region synapsis	68,89,91
EXO1	Decrease in A–T mutations	Severely impaired	Mismatch processing	105
NHEJ				
DNA-PKcs$^{-/-}$	None	Severely reduced except to IgG1	End joining? synapsis?	54,99
DNA-PKcs inactivated (SCID)		Reduced CSR to all isotypes		93,94
Ku70, Ku80	None	Severely impaired	End joining?	53,54,79
Other DNA-repair proteins				
H2AX	None	50–80% reduced, no effect on Sμ mutations and deletions	Synapsis?	50,79
53BP1	None	Severely impaired	Repair, synapsis?	52
ATM	None	S junctions have increased microhomology	End processing? synapsis?	82,83

53BP1, p53 binding protein 1; AID, activation-induced cytidine deaminase; ATM, ataxia telangiectasia mutated; CSR, class-switch recombination; DNA-PKcs, catalytic subunit of DNA-dependent protein kinase; EXO1, exonuclease 1; H2AX, histone 2A family, member X; MLH1, mutL homologue 1; MSH2, mutS homologue 2; NHEJ, non-homologous end-joining; PMS2, post-mitotic segregation 2; S, switch; SCID, severe combined immunodeficiency; SHM, somatic hypermutation; UNG, uracil-DNA glycosylase.

The CSR phenotype of C-terminal AID mutants closely resembles that found in a group of patients with hyper IgM type IV syndrome. B cells from these patients are considerably impaired in CSR, but SHM and the accumulation of mutations in Sμ are not affected[100]. Given that these cells express wild-type AID, it is possible that the mutation or mutations in these patients resides in genes that encode proteins required in the same reaction in which the putative CSR-specific AID cofactors participate. Therefore, the identification of proteins that interact with the C-terminus of AID and/or the cloning of the gene(s) that contribute to this immunodeficiency syndrome should shed light on the potential CSR-specific roles of AID downstream of deamination. Similarly, it could also be postulated that there will be mutations that could affect SHM but leave CSR intact: for example, mutations that affect the interaction of AID with factors, such as RPA, that target non-R-loop forming structures.

A plethora of unresolved issues

Since the discovery of AID, our understanding of CSR has advanced considerably. We can now envision a model that mechanistically links DNA deamination by AID to long-standing observations on the unique primary structure of S regions and the requirement for germline transcription. However, in addition to the mechanism of S-region synapsis and the other outstanding questions outlined earlier, several additional key issues remain unresolved.

It remains to be determined whether AID is specifically targeted to S regions or is passively recruited to any ssDNA. Studies using extrachromosomal CSR substrates indicate that AID targeting to S regions might be influenced by CSR-specific factors[101,102]. Whereas the ability of AID to mediate CSR on artificial substrates, when it is ectopically expressed in fibroblasts, might be viewed as arguing against IgH-isotype-specific CSR factors[40], the efficiency of CSR in these assays is low compared with endogenous efficiency. Therefore, it will be important to determine whether putative isotype-specific factors or other AID-specific cofactors exist and how they might function. Another important question centres on the current evidence indicating that CSR requires the participation of at least three different DNA-repair pathways: namely, BER, MMR and NHEJ. In this regard, the way in which pathways that use distinct sets of proteins are coordinated to mediate CSR is a mystery. Also, it is unknown how the BER pathway downstream of UNG and APE1 would be subverted from repairing an S-region nick before it is converted into a DSB. One possibility would be the introduction of numerous S-region lesions that overwhelm the BER machinery. Another possibility would be that BER activities downstream of UNG and APE1 might somehow be downregulated in cells undergoing CSR or at S lesions.

Perspective

The discovery of AID and the subsequent delineation of its enzymatic activity have provided us with a model by which the initiation of CSR is effected. However, as discussed here, several aspects of CSR both upstream and downstream of AID remain unresolved, including the identity of potential factors that impart specificity to AID, the process that converts an initial AID-generated DNA lesion into DNA breaks and the mechanism that synapses broken S regions to allow completion of the CSR reaction. A combination of new *in vivo*-genetic manipulations and *in vitro*-biochemical studies will probably be required to address these issues.

1. Bassing, C. H., Swat, W. & Alt, F. W. The mechanism and regulation of chromosomal V(D)J recombination. *Cell* **109**, S45–S55 (2002).
2. Jung, D. & Alt, F. W. Unraveling V(D)J recombination: insights into gene regulation. *Cell* **116**, 299–311 (2004).
3. Papavasiliou, F. N. & Schatz, D. G. Somatic hypermutation of immunoglobulin genes: merging mechanisms for genetic diversity. *Cell* **109**, S35–S44 (2002).
4. Martin, A. & Scharff, M. D. AID and mismatch repair in antibody diversification. *Nature Rev. Immunol.* **2**, 605–614 (2002).
5. Manis, J. P., Tian, M. & Alt, F. W. Mechanism and control of class-switch recombination. *Trends Immunol.* **23**, 31–39 (2002).
6. Honjo, T., Kinoshita, K. & Muramatsu, M. Molecular mechanism of class switch recombination: linkage with somatic hypermutation. *Annu. Rev. Immunol.* **20**, 165–196 (2002).
7. Storb, U. *et al*. *Cis*-acting sequences that affect somatic hypermutation of Ig genes. *Immunol. Rev.* **162**, 153–160 (1998).
8. Lansford, R., Manis, J. P., Sonoda, E., Rajewsky, K. & Alt, F. W. Ig heavy chain class switching in Rag-deficient mice. *Int. Immunol.* **10**, 325–332 (1998).
9. Dunnick, W., Hertz, G. Z., Scappino, L. & Gritzmacher, C. DNA sequences at immunoglobulin switch region recombination sites. *Nucleic Acids Res*. **21**, 365–372 (1993).
10. Shinkura, R. *et al*. The influence of transcriptional orientation on endogenous switch region function. *Nature Immunol.* **4**, 435–441 (2003).
Provides the first evidence *in vivo* that the transcriptional orientation of S regions is important in CSR. In addition, the authors show that a synthetic G-rich sequence that lacks S-region motifs can replace an endogenous S region during CSR *in vivo*.
11. Chaudhuri, J. *et al*. Transcription-targeted DNA deamination by the AID antibody diversification enzyme. *Nature* **422**, 726–730 (2003).
Demonstrates biochemically that AID is a ssDNA deaminase. The authors also show a direct link between transcription and DNA deamination.
12. Neuberger, M. S., Harris, R. S., Di Noia, J. & Petersen-Mahrt, S. K. Immunity through DNA deamination. *Trends Biochem. Sci.* **28**, 305–312 (2003).
13. Stavnezer, J. Immunoglobulin class switching. *Curr. Opin. Immunol.* **8**, 199–205 (1996).
14. Luby, T. M., Schrader, C. E., Stavnezer, J. & Selsing, E. The μ switch region tandem repeats are important, but not required, for antibody class switch recombination. *J. Exp. Med.* **193**, 159–168 (2001).
15. Stavnezer, J. Antibody class switching. *Adv. Immunol*. **61**, 79–146 (1996).
16. Stavnezer, J. Molecular processes that regulate class switching. *Curr. Top. Microbiol. Immunol.* **245**, 127–168 (2000).
17. Yancopoulos, G. D. *et al*. Secondary genomic rearrangement events in pre-B cells: V_HDJ_H replacement by a LINE-1 sequence and directed class switching. *EMBO J.* **5**, 3259–3266 (1986).
18. Stavnezer-Nordgren, J. & Sirlin, S. Specificity of immunoglobulin heavy chain switch correlates with activity of germline heavy chain genes prior to switching. *EMBO J.* **5**, 95–102 (1986).
19. Reaban, M. E. & Griffin, J. A. Induction of RNA-stabilized DNA conformers by transcription of an immunoglobulin switch region. *Nature* **348**, 342–344 (1990).
20. Reaban, M. E., Lebowitz, J. & Griffin, J. A. Transcription induces the formation of a stable RNA. DNA hybrid in the immunoglobulin α switch region. *J. Biol. Chem.* **269**, 21850–21857 (1994).
21. Daniels, G. A. & Lieber, M. R. Strand specificity in the transcriptional targeting of recombination at immunoglobulin switch sequences. *Proc. Natl Acad. Sci. USA* **92**, 5625–5629 (1995).
22. Tian, M. & Alt, F. W. Transcription-induced cleavage of immunoglobulin switch regions by nucleotide excision repair nucleases *in vitro*. *J. Biol. Chem.* **275**, 24163–24172 (2000).
23. Daniels, G. A. & Lieber, M. R. 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 (1995).
24. Mizuta, R. *et al*. Molecular visualization of immunoglobulin switch region RNA/DNA complex by atomic force microscope. *J. Biol. Chem.* **278**, 4431–4434 (2003).
25. Yu, K., Chedin, F., Hsieh, C. L., Wilson, T. E. & Lieber, M. R. R-loops at immunoglobulin class switch regions in the chromosomes of stimulated B cells. *Nature Immunol.* **4**, 442–451 (2003).
Uses chemical probes to demonstrate that S regions form R loops *in vivo* in B cells undergoing CSR.
26. Tashiro, J., Kinoshita, K. & Honjo, T. Palindromic but not G-rich sequences are targets of class switch recombination. *Int. Immunol.* **13**, 495–505 (2001).
27. Sen, D. & Gilbert, W. Formation of parallel four-stranded complexes by guanine-rich motifs in DNA and its implications for meiosis. *Nature* **334**, 364–366 (1988).

28. Dempsey, L. A., Sun, H., Hanakahi, L. A. & Maizels, N. 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 (1999).
29. Ratmeyer, L., Vinayak, R., Zhong, Y. Y., Zon, G. & Wilson, W. D. Sequence specific thermodynamic and structural properties for DNA.RNA duplexes. *Biochemistry* **33**, 5298–5304 (1994).
30. Kinoshita, K., Tashiro, J., Tomita, S., Lee, C. G. & Honjo, T. Target specificity of immunoglobulin class switch recombination is not determined by nucleotide sequences of S regions. *Immunity* **9**, 849–858 (1998).
31. Mussmann, R., Courtet, M., Schwager, J. & Du Pasquier, L. Microsites for immunoglobulin switch recombination breakpoints from *Xenopus* to mammals. *Eur. J. Immunol.* **27**, 2610–2619 (1997).
32. Ramiro, A. R., Stavropoulos, P., Jankovic, M. & Nussenzweig, M. C. Transcription enhances AID-mediated cytidine deamination by exposing single-stranded DNA on the nontemplate strand. *Nature Immunol.* **4**, 452–456 (2003).
Shows that AID targets the non-template strand of DNA when transcription is carried out in bacteria.
33. Sohail, A., Klapacz, J., Samaranayake, M., Ullah, A. & Bhagwat, A. S. Human activation-induced cytidine deaminase causes transcription-dependent, strand-biased C to U deaminations. *Nucleic Acids Res.* **31**, 2990–2994 (2003).
34. Pham, P., Bransteitter, R., Petruska, J. & Goodman, M. F. Processive AID-catalysed cytosine deamination on single-stranded DNA simulates somatic hypermutation. *Nature* **424**, 103–107 (2003).
Shows that AID preferably deaminates cytidine residues within SHM motifs in an ordered fashion.
35. Muramatsu, M. *et al*. 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 (1999).
36. Muramatsu, M. *et al*. Class switch recombination and hypermutation require activation-induced cytidine deaminase (AID), a potential RNA editing enzyme. *Cell* **102**, 553–563 (2000).
37. Revy, P. *et al*. Activation-induced cytidine deaminase (AID) deficiency causes the autosomal recessive form of the Hyper-IgM syndrome (HIGM2). *Cell* **102**, 565–575 (2000).
38. Yoshikawa, K. *et al*. AID enzyme-induced hypermutation in an actively transcribed gene in fibroblasts. *Science* **296**, 2033–2036 (2002).
39. Martin, A. & Scharff, M. D. Somatic hypermutation of the *AID* transgene in B and non-B cells. *Proc. Natl Acad. Sci. USA* **99**, 12304–12308 (2002).
40. Okazaki, I. M., Kinoshita, K., Muramatsu, M., Yoshikawa, K. & Honjo, T. The AID enzyme induces class switch recombination in fibroblasts. *Nature* **416**, 340–345 (2002).
References 38 to 40 provide evidence that AID is probably the only activated B-cell factor required for SHM and CSR.
41. Doi, T., Kinoshita, K., Ikegawa, M., Muramatsu, M. & Honjo, T. *De novo* protein synthesis is required for the activation-induced cytidine deaminase function in class-switch recombination. *Proc. Natl Acad. Sci. USA* **100**, 2634–2638 (2003).
42. Harris, R. S., Petersen-Mahrt, S. K. & Neuberger, M. S. RNA editing enzyme APOBEC1 and some of its homologs can act as DNA mutators. *Mol. Cell* **10**, 1247–1253 (2002).
Shows that other APOBEC1-family members have DNA-deamination activity, indicating that the RNA-editing enzyme APOBEC1 is probably an outlier in this family of proteins.
43. Petersen-Mahrt, S. K., Harris, R. S. & Neuberger, M. S. AID mutates *E. coli* suggesting a DNA deamination mechanism for antibody diversification. *Nature* **418**, 99–103 (2002).
The first paper to present the DNA-deamination model for CSR and SHM. In addition, it provides evidence that the target of AID is not RNA but DNA.
44. Bransteitter, R., Pham, P., Scharff, M. D. & Goodman, M. F. Activation-induced cytidine deaminase deaminates deoxycytidine on single-stranded DNA but requires the action of RNase. *Proc. Natl Acad. Sci. USA* **100**, 4102–4107 (2003).
45. Dickerson, S. K., Market, E., Besmer, E. & Papavasiliou, F. N. AID mediates hypermutation by deaminating single stranded DNA. *J. Exp. Med.* **197**, 1291–1296 (2003).
46. Nambu, Y. *et al*. Transcription-coupled events associating with immunoglobulin switch region chromatin. *Science* **302**, 2137–2140 (2003).
Shows that AID binds to chromatin that is associated with transcribed S regions, thereby providing further evidence that AID acts directly on DNA.
47. Wuerffel, R. A., Du, J., Thompson, R. J. & Kenter, A. L. Ig Sγ3 DNA-specific double strand breaks are induced in mitogen-activated B cells and are implicated in switch recombination. *J. Immunol.* **159**, 4139–4144 (1997).
48. Catalan, N. *et al*. The block in immunoglobulin class switch recombination caused by activation-induced cytidine deaminase deficiency occurs prior to the generation of DNA double strand breaks in switch μ region. *J. Immunol.* **171**, 2504–2509 (2003).
49. Rogakou, E. P., Pilch, D. R., Orr, A. H., Ivanova, V. S. & Bonner, W. M. DNA double-stranded breaks induce histone H2AX phosphorylation on serine 139. *J. Biol. Chem.* **273**, 5858–5868 (1998).
50. Petersen, S. *et al*. AID is required to initiate Nbs1/γ-H2AX focus formation and mutations at sites of class switching. *Nature* **414**, 660–665 (2001).
Shows that H2AX has a role in CSR and that the role of AID in CSR is probably upstream of S-region breaks.
51. Bassing, C. H. *et al*. Histone H2AX: a dosage-dependent suppressor of oncogenic translocations and tumors. *Cell* **114**, 359–370 (2003).
Reports that haploinsufficiency of H2AX can lead to transclocations that involve S regions.
52. Manis, J. P. *et al*. 53BP1 links DNA damage response to immunoglobulin class switch recombination. *Nature Immunol.* **5**, 481–487 (2004).
Describes the effect of 53BP1 mutation in CSR and SHM.
53. Casellas, R. *et al*. Ku80 is required for immunoglobulin isotype switching. *EMBO J.* **17**, 2404–2411 (1998).
54. Manis, J. P. *et al*. Ku70 is required for late B cell development and immunoglobulin heavy chain class switching. *J. Exp. Med.* **187**, 2081–2089 (1998).
55. Manis, J. P., Dudley, D., Kaylor, L. & Alt, F. W. IgH class switch recombination to IgG1 in DNA-PKcs-deficient B Cells. *Immunity* **16**, 607–617 (2002).
Together with references 93 and 94, this paper indicates that the role of DNA-PKcs in CSR might extend beyond the role it has in NHEJ.
56. Bross, L., Muramatsu, M., Kinoshita, K., Honjo, T. & Jacobs, H. DNA double-strand breaks: prior to but not sufficient in targeting hypermutation. *J. Exp. Med.* **195**, 1187–1192 (2002).
57. Papavasiliou, F. *et al*. V(D)J recombination in mature B cells: a mechanism for altering antibody responses. *Science* **278**, 298–301 (1997).
58. Faili, A. *et al*. AID-dependent somatic hypermutation occurs as a DNA single-strand event in the BL2 cell line. *Nature Immunol.* **3**, 815–821 (2002).
59. Woo, C. J., Martin, A. & Scharff, M. D. Induction of somatic hypermutation is associated with modifications in immunoglobulin variable region chromatin. *Immunity* **19**, 479–489 (2003).
60. Rada, C. *et al*. Immunoglobulin isotype switching is inhibited and somatic hypermutation perturbed in UNG-deficient mice. *Curr. Biol.* **12**, 1748–1755 (2002).
61. Imai, K. *et al*. Human uracil-DNA glycosylase deficiency associated with profoundly impaired immunoglobulin class-switch recombination. *Nature Immunol.* **4**, 1023–1028 (2003).
References 60 and 61 show that UNG has a crucial role in CSR and thereby provide evidence for a DNA-deamination step in CSR.
62. Vora, K. A. *et al*. Severe attenuation of the B cell immune response in Msh2-deficient mice. *J. Exp. Med.* **189**, 471–482 (1999).
63. Ehrenstein, M. R. & Neuberger, M. S. Deficiency in Msh2 affects the efficiency and local sequence specificity of immunoglobulin class-switch recombination: parallels with somatic hypermutation. *EMBO J.* **18**, 3484–3490 (1999).
64. Schrader, C. E., Edelmann, W., Kucherlapati, R. & Stavnezer, J. Reduced isotype switching in splenic B cells from mice deficient in mismatch repair enzymes. *J. Exp. Med.* **190**, 323–330 (1999).
65. Yu, K. & Lieber, M. R. Nucleic acid structures and enzymes in the immunoglobulin class switch recombination mechanism. *DNA Repair (Amst.)* **2**, 1163–1174 (2003).
66. Rush, J. S., Fugmann, S. D. & Schatz, D. G. Staggered AID-dependent DNA double strand breaks are the predominant DNA lesions targeted to Sμ in immunoglobulin class switch recombination. *Int. Immunol.* **16**, 549–557 (2004).
67. Faili, A. *et al*. DNA Polymerase η is involved in hypermutation occurring during immunoglobulin class switch recombination. *J. Exp. Med.* **199**, 265–270 (2004).
68. Schrader, C. E., Vardo, J. & Stavnezer, J. Role for mismatch repair proteins Msh2, Mlh1, and Pms2 in immunoglobulin class switching shown by sequence analysis of recombination junctions. *J. Exp. Med.* **195**, 367–373 (2002).
69. Min, I. M. *et al*. The Sμ tandem repeat region is critical for Ig isotype switching in the absence of Msh2. *Immunity* **19**, 515–524 (2003).
70. Paques, F. & Haber, J. E. Multiple pathways of recombination induced by double-strand breaks in *Saccharomyces cerevisiae*. *Microbiol. Mol. Biol. Rev*. **63**, 349–404 (1999).
71. Tian, M., Jones, A. M., Smith, M., Shinkura, R. & Alt, F. W. Deficiency in nuclease activity of xeroderma pigmentosum G in mice leads to hypersensitivity to UV radiation. *Mol. Cell. Biol.* **24**, 2237–2242 (2004).
72. Tian, M., Shinkura, R., Shinkura, N. & Alt, F. W. Growth retardation, early death, and DNA repair defects in mice deficient for the nucleotide excision repair enzyme XPF. *Mol. Cell. Biol*. **24**, 1200–1205 (2004).
73. Hodgkin, P. D., Lee, J. H. & Lyons, A. B. B cell differentiation and isotype switching is related to division cycle number. *J. Exp. Med.* **184**, 277–281 (1996).
74. Kuzminov, A. Single-strand interruptions in replicating chromosomes cause double-strand breaks. *Proc. Natl Acad. Sci. USA* **98**, 8241–8246 (2001).
75. Dudley, D. D. *et al*. Internal IgH class switch region deletions are position-independent and enhanced by AID expression. *Proc. Natl Acad. Sci. USA* **99**, 9984–9989 (2002).
Shows that S-region internal deletions are AID dependent and provides evidence that synapsis between two S regions is not required for targeting AID.
76. Alt, F. W., Rosenberg, N., Casanova, R. J., Thomas, E. & Baltimore, D. Immunoglobulin heavy-chain expression and class switching in a murine leukaemia cell line. *Nature* **296**, 325–331 (1982).
77. Gu, H., Zou, Y. R. & Rajewsky, K. Independent control of immunoglobulin switch recombination at individual switch regions evidenced through Cre–loxP-mediated gene targeting. *Cell* **73**, 1155–1164 (1993).
78. Schrader, C. E. *et al*. Mutations occur in the Ig Sμ region but rarely in Sγ regions prior to class switch recombination. *EMBO J.* **22**, 5893–5903 (2003).
79. Reina-San-Martin, B. *et al*. H2AX is required for recombination between immunoglobulin switch regions but not for intra-switch region recombination or somatic hypermutation. *J. Exp. Med.* **197**, 1767–1778 (2003).
Indicates that H2AX might have a role in the synapsis of two S regions during CSR.
80. Bassing, C. H. & Alt, F. W. H2AX may function as an anchor to hold broken chromosomal DNA ends in close proximity. *Cell Cycle* **3**, 149–153 (2004).
81. Shiloh, Y. Ataxia-telangiectasia and the Nijmegen breakage syndrome: related disorders but genes apart. *Annu. Rev. Genet.* **31**, 635–662 (1997).
82. Pan, Q. *et al*. Alternative end joining during switch recombination in patients with ataxia-telangiectasia. *Eur. J. Immunol.* **32**, 1300–1308 (2002).
83. Pan-Hammarstrom, Q. *et al*. ATM is not required in somatic hypermutation of V_H, but is involved in the introduction of mutations in the switch μ region. *J. Immunol.* **170**, 3707–3716 (2003).
84. Burma, S., Chen, B. P., Murphy, M., Kurimasa, A. & Chen, D. J. ATM phosphorylates histone H2AX in response to DNA double-strand breaks. *J. Biol. Chem.* **276**, 42462–42467 (2001).
85. Stiff, T. *et al*. ATM and DNA-PK function redundantly to phosphorylate H2AX after exposure to ionizing radiation. *Cancer Res.* **64**, 2390–2396 (2004).
86. Ward, I. M. & Chen, J. Histone H2AX is phosphorylated in an ATR-dependent manner in response to replicational stress. *J. Biol. Chem.* **276**, 47759–47762 (2001).
87. Xu, Y. *et al*. Targeted disruption of ATM leads to growth retardation, chromosomal fragmentation during meiosis, immune defects, and thymic lymphoma. *Genes Dev.* **10**, 2411–2422 (1996).
88. DeFazio, L. G., Stansel, R. M., Griffith, J. D. & Chu, G. Synapsis of DNA ends by DNA-dependent protein kinase. *EMBO J.* **21**, 3192–3200 (2002).
89. Ehrenstein, M. R., Rada, C., Jones, A. M., Milstein, C. & Neuberger, M. S. Switch junction sequences in PMS2-deficient mice reveal a microhomology-mediated mechanism of Ig class switch recombination. *Proc. Natl Acad. Sci. USA* **98**, 14553–14558 (2001).
90. Hall, M. C., Wang, H., Erie, D. A. & Kunkel, T. A. High affinity cooperative DNA binding by the yeast Mlh1–Pms1 heterodimer. *J. Mol. Biol.* **312**, 637–647 (2001).
91. Schrader, C. E., Vardo, J. & Stavnezer, J. Mlh1 can function in antibody class switch recombination independently of Msh2. *J. Exp. Med.* **197**, 1377–1383 (2003).
92. Brown, K. D. *et al*. The mismatch repair system is required for S-phase checkpoint activation. *Nature Genet.* **33**, 80–84 (2003).
93. Bosma, G. C. *et al*. DNA-dependent protein kinase activity is not required for immunoglobulin class switching. *J. Exp. Med.* **196**, 1483–1495 (2002).

94. Cook, A. J. *et al*. Reduced switching in SCID B cells is associated with altered somatic mutation of recombined S regions. *J. Immunol.* **171**, 6556–6564 (2003).
95. Ma, Y., Pannicke, U., Schwarz, K. & Lieber, M. R. Hairpin opening and overhang processing by an Artemis/DNA-dependent protein kinase complex in nonhomologous end joining and V(D)J recombination. *Cell* **108**, 781–794 (2002).
96. Sekiguchi, J. *et al*. Genetic interactions between ATM and the nonhomologous end-joining factors in genomic stability and development. *Proc. Natl Acad. Sci. USA* **98**, 3243–3248 (2001).
97. Ta, V. T. *et al*. AID mutant analyses indicate requirement for class-switch-specific cofactors. *Nature Immunol.* **4**, 843–848 (2003).
98. Barreto, V., Reina-San-Martin, B., Ramiro, A. R., McBride, K. M. & Nussenzweig, M. C. C-terminal deletion of AID uncouples class switch recombination from somatic hypermutation and gene conversion. *Mol. Cell* **12**, 501–508 (2003).
References 97 and 98 provide evidence that the carboxyl terminus of AID could have specific roles in CSR that might be downstream of DNA deamination.
99. Bemark, M. *et al*. Somatic Hypermutation in the absence of DNA-dependent protein kinase catalytic subunit (DNA-PKcs) or recombination-activating gene (RAG) 1 activity. *J. Exp. Med.* **192**, 1509–1514 (2000).
100. Imai, K. *et al*. Hyper-IgM syndrome type 4 with a B lymphocyte-intrinsic selective deficiency in Ig class-switch recombination. *J. Clin. Invest.* **112**, 136–142 (2003).
Reports on a newly characterized human immunodeficiency, in which the defect in CSR is probably downstream of AID-mediated DNA deamination.
101. Shanmugam, A., Shi, M. J., Yauch, L., Stavnezer, J. & Kenter, A. L. Evidence for class-specific factors in immunoglobulin isotype switching. *J. Exp. Med.* **191**, 1365–1380 (2000).
102. Ma, L., Wortis, H. H. & Kenter, A. L. Two new isotype-specific switching activities detected for Ig class switching. *J. Immunol.* **168**, 2835–2846 (2002).
103. Ito, S. *et al*. Activation-induced cytidine deaminase shuttles between nucleus and cytoplasm like apolipoprotein B mRNA editing catalytic polypeptide 1. *Proc. Natl Acad. Sci. USA* **101**, 1975–1980 (2004).
104. Martin, A. *et al*. Msh2 ATPase activity is essential for somatic hypermutation at A–T basepairs and for efficient class switch recombination. *J. Exp. Med.* **198**, 1171–1178 (2003).
105. Bardwell, P. D. *et al*. Altered somatic hypermutation and reduced class-switch recombination in exonuclease 1-mutant mice. *Nature Immunol.* **5**, 224–229 (2004).

Acknowledgements
We thank John Manis, Ming Tian, Ali Zarrin, Sheila Ranganath, Craig Bassing and Chan Khuong and other members of the laboratory for careful evaluation of this manuscript. Work on CSR in the Alt laboratory is supported in part by the National Institutes of Health. F.W.A. is an Investigator of the Howard Hughes Medical Institute.

Competing interests statement
The authors declare that they have no competing financial interests.

Online links

DATABASES
The following terms in this article are linked online to:
Entrez Gene:
http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=gene
53BP1 | AID | APE1 | APOBEC1 | ARTEMIS | ATM | DNA Ligase IV | DNA-PKcs | ERCC1 | H2AX | MLH1 | MSH2 | PMS2 | RAG1 | RAG2 | UNG | XPF | XPG | XRCC4
Access to this interactive links box is free online.