



HHS Public Access

Author manuscript

Curr Opin Immunol. Author manuscript; available in PMC 2017 April 01.

Published in final edited form as:

Curr Opin Immunol. 2016 April ; 39: 96–102. doi:10.1016/j.coi.2016.01.008.

AID hits the jackpot when missing the target

Amy L. Kenter*, Satyendra Kumar, Robert Wuerffel, and Fernando Grigera

Department of Microbiology and Immunology, University of Illinois College of Medicine, Chicago, Illinois, USA

Abstract

Activation induced deaminase is the single B cell specific factor mediating class switch recombination and somatic hypermutation. Numerous studies have shown that AID preferentially targets Ig substrates and also attacks non-Ig substrates to create DNA damage that contributes to lymphomagenesis. AID targeting to Ig loci is linked to transcription but the mechanism governing this process has been obscure. Here we discuss research that illustrates the connection between AID targeting to DNA substrates and transcription processes to reveal rules governing the specificity of AID attack. These observations are woven together to provide an integrated view of AID function and a surprising linkage with global regulation of gene expression.

Keywords

B lymphocytes; Class switch recombination; Somatic hypermutation; Activation induced deaminase

Humoral immunity is mediated in B cells by antigen receptors (BCR) that are composed of immunoglobulin (Ig) heavy (H) and light (L) chains. The antigen receptor component of Ig is assembled from multiple gene segments via V(D)J recombination during early B cell development in the bone marrow and is mediated by the RAG1 and RAG2 recombinases [1]. In the peripheral lymphoid organs, B cells become activated by antigen and undergo somatic hypermutation (SHM) and class switch recombination (CSR) and upon terminal differentiation secrete Ig as antibody [2,3]. Antibody repertoires are diversified at high frequency during somatic hypermutation (SHM), by introducing mutations into *Igh* and *Igl* V(D)J exons. SHM occurs in germinal centers of the peripheral lymphoid organs where mutations are selected to produce BCR with increased affinity. Constant (C_H) region genes encode the C-terminal domains of the Ig chains and determine IgH effector activities. IgH effector function is diversified through CSR, while retaining the original antigen binding specificity arising during V(D)J recombination. Activation induced deaminase (AID), a cytosine deaminase is essential for initiating both SHM and (CSR) in mature B cells [4]**.

*Correspondence: Dr. Amy L. Kenter, Department of Microbiology and Immunology (M/C 790), University of Illinois College of Medicine, 835 South Wolcott, Chicago, Illinois, 60612, USA, Phone: 312-996-5293, Fax: 312-996-6415, star1@uic.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

The mouse *Igh* locus includes eight C_H genes, encoding μ , δ , $\gamma 3$, $\gamma 1$, $\gamma 2b$, $\gamma 2a$, ϵ , and α chains and each is paired with repetitive switch (S) DNA (with the exception of C δ). The C_H region genomic area spans 220 kb and is flanked by the intronic E μ and 3'E α enhancers [5]. CSR is focused on S regions and involves intra-chromosomal deletional rearrangements that replace the initial C μ with a downstream C_H region gene (fig. 1). Prior to and during CSR, a S-S synaptosome is formed to facilitate proximity between the donor S μ and a downstream acceptor S region [6–9]. Formation of the S-S synaptosome is critically dependent on long range chromatin interactions that are tethered by key transcriptional elements (reviewed in [10]). AID initiates CSR by deaminating cytosines in S regions in donor S μ and a downstream acceptor S region (Box 1) [11]. AID dependent DNA damage is focused to highly degenerate WRC hotspot motifs (W=A/T, R=A/G) that are found in V genes and at high density in S regions (reviewed in [12]). Conversion of AID induced lesions to DNA double strand breaks by general DNA repair factors has been extensively reviewed [2,3,11–17]. Degeneracy of the AID hotspot motif raises questions regarding the mechanism by which AID is targeted to its Ig substrates.

Box 1

AID introduces DNA damage by converting deoxycytidine (dC) to deoxyuracil (dU) [14]. The AID initiated DNA lesions are processed by engagement with base excision repair (BER) and mismatch repair (MMR) pathways to create mutations required for SHM and DNA double strand break (DSB) intermediates that consumed in CSR (reviewed in [2,3,11,14]). The observations that AID dependent dU residues are detected in the 5'S μ region [99] and that uracil DNA glycosylase (UNG) is required for formation of double strand breaks (DSB) and mutations demonstrate that AID deamination initiates SHM and CSR [100,101]. Conversion of AID induced lesions to DNA double strand breaks is mediated by general DNA repair factors [17].

Transcription is a hallmark of SHM and CSR and transcriptional elements are candidates for providing specificity to AID targeting [5,18]. V region transcription initiates from a promoter 5' proximal to the rearranged V(D)J exon and terminates 3' of the C_H region exons. CSR is focused to specific S regions by differential activation of germline transcription [10,19]. C_H gene transcription units are comprised of the noncoding intervening (I) exon, an S region and a C_H coding region. Germline transcription initiates at a transcription start site (TSS) 5' of each I exon, proceeds through the S region and terminates downstream of the corresponding C_H gene (fig. 1A). The V(D)J mutation profile has a sharp 5' boundary ~120 bp downstream of the TSS and a less defined 3' border ~1 kb downstream of the promoter. Alteration of the V region promoter position displaces transcription initiation and perturbs the mutation distribution [20*,21] thereby linking induction of AID dependent mutations with transcription. Similarly, AID induced DNA lesions in S regions begin ~150 bp downstream of the I exon TSS [22]. A recent study shows that sequence intrinsic features target AID dependent DNA lesions to Ig templates [23]. Although AID targeting to Ig substrates requires transcription, the unique transcriptional features that determine preferential AID attack at Ig templates have been difficult to discern.

AID induced double strand breaks (DSBs) in normal B cells occur at hundreds of non-Ig sites many of which are syntenic with sites of translocations, deletions, and amplifications found in human B cell lymphomas [24]. Physiological levels of AID in GC B cells have been linked to deamination of a large cohort of non-Ig genes [25*–28] where the mutation rate is 20–100 fold lower than at Ig loci [26]**. The findings that SHM and CSR are tightly linked to transcription [20*,29**,30] and AID interacts with RNA Pol II (RNAP II) [27*,31] have led to the notion that transcription and RNAP II-associated proteins might facilitate the binding of AID to target DNA sequences. Based on provocative new studies, we address three important interrelated questions regarding AID targeting. What is the mechanism by which AID preferentially targets Ig substrates and its corollary, how is AID attack at non-Ig substrates directed? Finally, is there an adaptive advantage for AID attack at non-Ig templates? We focus primarily on CSR as this area has been the most intensively investigated. Recent work implies that AID attack on non-Ig substrates may mediate functional outcomes that contribute to a biological “jackpot”.

Transcribed S regions are specialized targets of AID

S regions are repetitive, nonidentical, 1–12 kb long and guanine rich on the nontemplate strand [32]. Deletion, inversion or replacement of S regions reduces CSR frequency indicating that S regions are specialized targets of CSR [33–36]. However, CSR S/S junctions are notable for their lack of consensus sequence or homology originating from site specific or homologous recombination. The degeneracy of the S region repeats and the absence of discernable recombination signal motifs have led to models in which higher order structures provide recognition motifs for the CSR machinery. In vitro transcription studies indicated that the looped out ssDNA nontemplate strand can assume specialized structures including stem loops [37], four-stranded G quartets [38] and R-loops (Box 2) [39,40]. In vivo studies confirm that transcription through S regions in vivo creates G quartets [41,42] and generates long stretches of R-loops that form on the nontemplate strand [33,43*,44] and that enhance CSR efficiency.

Box 2

R-loops are composed of RNA:DNA hybrids in which the nontemplate ssDNA strand is looped out while the template strand is stably annealed to nascent S region RNA transcripts [43,102]. The presence of R-loops within a transcription unit can cause RNAP II stalling [103]. Frequent R-loop mediated RNAP II stalling may account, at least in part, for enrichment of initiating RNAP II occupancy in S regions that in turn mediates the introduction of activating histone modifications and increased chromatin accessibility [45,104]. Accordingly, targeted inversion of an S region leads to loss of R-loop formation and reduction of CSR frequency [33]. Furthermore, S region deletion leads to reduced RNAP II occupancy at flanking sites [45,104] and loss of activating histone marks [45]. The enrichment of initiating RNAP II at promoter distal sites provides a plausible explanation for increased levels of histone activating marks and chromatin accessibility throughout the S region. Biochemical studies indicated that transcription generated R-loops within dsDNA permit AID directed deamination on the ssDNA nontemplate strand

while the template strand remained blocked by the nascent RNA transcript (reviewed in [10,11]).

Epigenetic studies indicate that the initiating form of RNAP II phosphorylated on serine 5 (p-ser5) and activating histone marks decorate the length of transcribed S regions [45*,46]. In contrast, genome wide studies show that transcriptionally active genes are associated with promoter proximal enrichment of RNAP II p-ser5 coupled with activating histone modifications [47–51]. R-loops within a transcription unit can act as a structural impediment for RNAP II elongation and cause stalling (reviewed in [52]). S region inversion led to loss of R-loops and reduced CSR *in vivo* thereby linking R-loops and CSR [33]. However, the contribution of RNAP II stalling in S regions to the CSR reaction has remained unclear.

AID specifically binds to components of the transcription machinery

Several AID binding proteins that facilitate association of AID with transcribed S DNA and modulate CSR have been identified in genetic screens, including the adaptor 14-3-3 protein [53] and the RNA processing and/or splicing factors, polypyrimidine-tract binding protein (PTBP2) [54], RNAP II stalling cofactor, Spt5 [27] and the 11-subunit cellular noncoding RNA 3′–5′ exonucleolytic processing complex, RNA exosome [55]**. SPT5 collaborates with the DSIF complex and negative elongation factor (NEF) to stall RNAP II p-ser5 at promoter proximal sites [56] and links RNAP II to splicing factors [57], capping enzyme [57,58], and the RNA exosome (reviewed in [52]). The RNA exosome contains 3′–5′ exoribonucleases that process structural RNA, degrade improperly processed pre-mRNAs and some long noncoding (lnc) RNAs of which GLT RNA hybridized to the S region in an R-loop is an example [52,59]. The RNA exosome interacts indirectly with RNAP II via SPT5 and SPT6 [55]**. Deletion of PTBP2, SPT5 and RNA exosome components impair CSR and link transcription processes with AID recruitment and function at S DNA (reviewed in [11]). These observations directly demonstrate that the transcriptional machinery associated with RNAP II stalling, the presence of RNA:DNA hybrid structures as impediments to RNAP II elongation and the release of those obstacles by means of the RNA exosome are integral to targeting AID to S region DNA.

A recent provocative study has identified another regulatory feature contributing to AID targeting specificity that is directly dependent on expression of the long noncoding GLT RNA from transcribed S regions. Evidence indicates that AID functions as an RNA binding protein with specificity for S region GLTs through G-quadruplexes structures and is required for CSR [42]. Indeed, AID is one of 12 members of the APOBEC family of DNA/RNA cytidine deaminases and has recently been shown to mutate small RNA genes when expressed in yeast [60] suggesting a role for RNA in the recruitment of AID to S regions. Furthermore, short RNA segments have been shown to function as specific guides for nuclease modification of the genome [61] and to regulate DNA rearrangements in ciliates [62]. The requirement for G rich S region guide-RNAs for AID targeting implies that AID off target genes might also express non-coding RNA that functions in a similar capacity. Thus, AID targeting to Ig loci is dependent on transcription machine components.

AID targeting and mistargeting to genome wide substrates

Mistargeting of AID to non-Ig substrates has been implicated in the pathogenesis of B cell lymphomas and chromosomal translocations [63,64]. In the Western world, the vast majority of lymphomas arise in B cells that actively engage in SHM and CSR [65] that in turn generate recurrent chromosomal translocations [66]. Normal mature B cells are particularly prone to dynamic AID dependent chromosomal translocations that juxtapose Ig genes and proto-oncogenes. The high frequency of these events allows detection in the absence of selection (reviewed in [17]). The preferential focus of AID to Ig loci and the nonrandom aspect of AID attack on non-Ig genes implies a determinant set of rules governing AID targeting. Recent GRO-seq studies indicate that most AID induced off-target translocations occur at defined regions of target genes in which sense and anti-sense transcription converge [67]**. Strikingly, convergent transcription is due to antisense transcription originating from super-enhancers within sense transcribed gene bodies [67]. Super-enhancers are clusters of enhancers that have prominent roles in cell type specific processes [68,69] and express relatively high levels of enhancer RNAs (eRNA) [70]. Super-enhancers have the propensity to associate in three dimensional nuclear space via long range chromatin interactions [71] that are mediated by lineage specific transcription factors (TFs) [70]. Hence, lineage specific TFs indirectly facilitate AID initiated chromosomal translocations [67]**.

The RNA exosome is an RNA surveillance complex that degrades a variety of non-coding RNAs (reviewed in [52]). RNA species arising in response to exosome subunit deficiencies have revealed the identities of non-Ig genes and intergenic regions that are the focus of AID mediated translocations [52,72]. Several types of non-coding exosome substrate RNAs accumulate in the transcriptomes of exosome-deficient B cells including, transcription start site (TSS)-associated antisense transcripts (xTSS-RNAs). The xTSS-RNAs are divergently transcribed from cognate coding gene transcripts that accumulate R-loops, AID-mediated mutations, and/or are frequent translocation partners of *Igh* DNA DSBs in B cells [59]. A subset of xTSS-RNAs originate as overlapping sense and antisense x-eRNAs at super-enhancers and are located at recurrent translocation hotspots [72]. It is not clear whether the characterization of these transcripts as divergent represents a real or semantic difference with the “convergent” transcripts defined by Meng et al. described above. Collectively, these findings begin to define a set of predictive transcriptional features that characterize Ig- and non-Ig substrates of AID.

Noncanonical AID Expression

The prevalence of AID off-target sites implies that this phenomenon may have adaptive advantages. Low levels of AID have been detected in a number of tissues and cell types including early stage B cells [73,74], and in pluripotent cells such as embryonic stem cells [75]* and spermatocytes [76] where it may have an unanticipated function in cellular reprogramming [77–80]. AID expression can be upregulated in some instances of infection [81] and under conditions of chronic inflammation that promote tumorigenesis [81–85]. Intriguingly, inflammatory stimuli (LPS) robustly induce AID in primary pro-B cells and Abelson transformed B cell lines leading to CSR prior to V(D)J joining [86] and mutations associated with acute lymphoblastic leukemia [87,88]. Strikingly, AID expression in

transitional B cells appears to mediate tolerance [89,90] by a B cell intrinsic mechanism [91]. It is intriguing to speculate that by analogy to new gene expression linked to RAG recombinase induced DNA DSBs [92], AID induced DSBs direct new gene expression leading to functional outcomes including immune tolerance.

Role for AID in DNA Demethylation

One provocative explanation for AID expression during development and in early B cells is related to DNA demethylation. DNA methylation is an epigenetic modification that is considered central to the establishment and maintenance of stable cellular identities (reviewed in [93]). The processes by which methylation is removed from cytosine were unclear until recent studies indicated active modes of DNA demethylation that involve modification of the meC base coupled to DNA repair. One pathway proceeds through oxidation catalyzed by the TET (ten eleven translocation) enzymes [93,94]. A second pathway uses AID, which promotes DNA demethylation through direct deamination of meC to thymidine [75]*, and subsequent repair of the resultant T:G mismatch by classical repair pathways [77,79,80,95]. Evidence suggests that AID's demethylation activity is required for reprogramming in zebrafish embryos [80] and in mice [79,96]. AID interacts with and demethylates the promoters of the *OCT4* and *NANOG* genes during reprogramming of human fibroblasts fused to mouse ES cells [78]. In mice, DNA demethylation is mediated by base excision repair (BER) through AID/Apobec deamination of 5meC to thymidine followed by G:T mismatch repair by TDG [75*,79,95] (Box 1). Notably, differentially methylated cytosines between naïve and GC B cells are enriched in non-Ig genes that are targeted by AID for SHM, and these genes form networks required for B cell development and proliferation [97]. Thus, emerging evidence implies that genome wide DNA demethylation by AID may be an important mechanism toward global regulation of gene expression programs [98].

Concluding Thoughts

Recent discoveries indicate that AID targeting to Ig and non-Ig loci is highly integrated with the transcription machinery and non-coding RNA biogenesis and degradation implying that off-target events are regulated and might be adaptive. Parallel studies link AID to active genome wide DNA demethylation. The final link in this chain is a new study that shows that gene loci hotspots for AID dependent mutagenesis are also DNA demethylated and their expression is networked. This raises the intriguing possibility that what we had previously considered AID off-targeting is actually integral to functional gene regulation. More work is necessary to define the molecular mechanism mediating apparently non-canonical AID effects at non-Ig loci.

Acknowledgments

This work was supported in part by the National Institutes of Health (R01AI121286, R21AI117687, R21AI117687) to A.L.K.

Literature Cited

1. Schatz DG, Ji Y. Recombination centres and the orchestration of V(D)J recombination. *Nature reviews Immunology*. 2011; 11:251–263.
2. Peled JU, Kuang FL, Iglesias-Ussel MD, Roa S, Kalis SL, Goodman MF, Scharff MD. The biochemistry of somatic hypermutation. *Annu Rev Immunol*. 2008; 26:481–511. [PubMed: 18304001]
3. Stavnezer J, Guikema JE, Schrader CE. Mechanism and regulation of class switch recombination. *Annu Rev Immunol*. 2008; 26:261–292. [PubMed: 18370922]
4. Muramatsu M, Kinoshita K, Fagarasan S, Yamada S, Shinkai Y, Honjo T. Class switch recombination and hypermutation require activation-induced cytidine deaminase (AID), a potential RNA editing enzyme. *Cell*. 2000; 102:553–563. [PubMed: 11007474]
5. Perlot T, Alt FW. Cis-regulatory elements and epigenetic changes control genomic rearrangements of the IgH locus. *Adv Immunol*. 2008; 99:1–32. [PubMed: 19117530]
6. Wuerffel R, Wang L, Grigera F, Manis J, Selsing E, Perlot T, Alt FW, Cogne M, Pinaud E, Kenter AL. S-S synapsis during class switch recombination is promoted by distantly located transcriptional elements and activation-induced deaminase. *Immunity*. 2007; 27:711–722. [PubMed: 17980632]
7. Feldman S, Achour I, Wuerffel R, Kumar S, Gerasimova T, Sen R, Kenter AL. Constraints Contributed by Chromatin Looping Limit Recombination Targeting during Immunoglobulin Class Switch Recombination. *J Immunol*. 2015
8. Sellars M, Reina-San-Martin B, Kastner P, Chan S. Ikaros controls isotype selection during immunoglobulin class switch recombination. *J Exp Med*. 2009; 206:1073–1087. [PubMed: 19414557]
9. Dong J, Panchakshari RA, Zhang T, Zhang Y, Hu J, Volpi SA, Meyers RM, Ho YJ, Du Z, Robbiani DF, et al. Orientation-specific joining of AID-initiated DNA breaks promotes antibody class switching. *Nature*. 2015; 525:134–139. [PubMed: 26308889]
10. Kenter AL. AID targeting is dependent on RNA polymerase II pausing. *Semin Immunol*. 2012; 24:281–286. [PubMed: 22784681]
11. Matthews AJ, Zheng S, DiMenna LJ, Chaudhuri J. Regulation of immunoglobulin class-switch recombination: choreography of noncoding transcription, targeted DNA deamination, and long-range DNA repair. *Adv Immunol*. 2014; 122:1–57. [PubMed: 24507154]
12. Maul RW, Gearhart PJ. AID and somatic hypermutation. *Adv Immunol*. 2010; 105:159–191. [PubMed: 20510733]
13. Chahwan R, Edelmann W, Scharff MD, Roa S. AIDing antibody diversity by error-prone mismatch repair. *Seminars in immunology*. 2012; 24:293–300. [PubMed: 22703640]
14. Di Noia JM, Neuberger MS. Molecular mechanisms of antibody somatic hypermutation. *Annu Rev Biochem*. 2007; 76:1–22. [PubMed: 17328676]
15. Saribasak H, Rajagopal D, Maul RW, Gearhart PJ. Hijacked DNA repair proteins and unchained DNA polymerases. *Philos Trans R Soc Lond B Biol Sci*. 2009; 364:605–611. [PubMed: 19008198]
16. Boboila C, Alt FW, Schwer B. Classical and alternative end-joining pathways for repair of lymphocyte-specific and general DNA double-strand breaks. *Adv Immunol*. 2012; 116:1–49. [PubMed: 23063072]
17. Alt FW, Zhang Y, Meng FL, Guo C, Schwer B. Mechanisms of programmed DNA lesions and genomic instability in the immune system. *Cell*. 2013; 152:417–429. [PubMed: 23374339]
18. Storb U. Why does somatic hypermutation by AID require transcription of its target genes? *Adv Immunol*. 2014; 122:253–277. [PubMed: 24507160]
19. Chaudhuri J, Basu U, Zarrin A, Yan C, Franco S, Perlot T, Vuong B, Wang J, Phan RT, Datta A, et al. Evolution of the immunoglobulin heavy chain class switch recombination mechanism. *Adv Immunol*. 2007; 94:157–214. [PubMed: 17560275]
20. Peters A, Storb U. Somatic hypermutation of immunoglobulin genes is linked to transcription initiation. *Immunity*. 1996; 4:57–65. [PubMed: 8574852]

21. Bachl J, Carlson C, Gray-Schopfer V, Dessing M, Olsson C. Increased transcription levels induce higher mutation rates in a hypermutating cell line. *J Immunol.* 2001; 166:5051–5057. [PubMed: 11290786]
22. Xue K, Rada C, Neuberger MS. The in vivo pattern of AID targeting to immunoglobulin switch regions deduced from mutation spectra in *msh2*^{-/-} *ung*^{-/-} mice. *J Exp Med.* 2006; 203:2085–2094. [PubMed: 16894013]
23. Yeap LS, Hwang JK, Du Z, Meyers RM, Meng FL, Jakubauskaite A, Liu M, Mani V, Neuberger D, Kepler TB, et al. Sequence-Intrinsic Mechanisms that Target AID Mutational Outcomes on Antibody Genes. *Cell.* 2015; 163:1124–1137. [PubMed: 26582132]
24. Staszewski O, Baker RE, Ucher AJ, Martier R, Stavnezer J, Guikema JE. Activation-induced cytidine deaminase induces reproducible DNA breaks at many non-Ig Loci in activated B cells. *Mol Cell.* 2011; 41:232–242. [PubMed: 21255732]
25. Shen HM, Peters A, Baron B, Zhu X, Storb U. Mutation of BCL-6 gene in normal B cells by the process of somatic hypermutation of Ig genes. *Science.* 1998; 280:1750–1752. [PubMed: 9624052]
26. Liu M, Duke JL, Richter DJ, Vinuesa CG, Goodnow CC, Kleinstein SH, Schatz DG. Two levels of protection for the B cell genome during somatic hypermutation. *Nature.* 2008; 451:841–845. [PubMed: 18273020]
27. Pavri R, Gazumyan A, Jankovic M, Di Virgilio M, Klein I, Ansarah-Sobrinho C, Resch W, Yamane A, Reina San-Martin B, Barreto V, et al. Activation-induced cytidine deaminase targets DNA at sites of RNA polymerase II stalling by interaction with Spt5. *Cell.* 2010; 143:122–133. [PubMed: 20887897]
28. Yamane A, Resch W, Kuo N, Kuchen S, Li Z, Sun HW, Robbiani DF, McBride K, Nussenzweig MC, Casellas R. Deep-sequencing identification of the genomic targets of the cytidine deaminase AID and its cofactor RPA in B lymphocytes. *Nat Immunol.* 2011; 12:62–69. [PubMed: 21113164]
29. Stavnezer-Nordgren J, Sirlin S. Specificity of immunoglobulin heavy chain switch correlates with activity of germline heavy chain genes prior to switching. *EMBO J.* 1986; 5:95–102. [PubMed: 3007121]
30. Yancopoulos GD, DePinho RA, Zimmerman KA, Lutzker SG, Rosenberg N, Alt FW. Secondary genomic rearrangement events in pre-B cells: VHDJH replacement by a LINE-1 sequence and directed class switching. *EMBO J.* 1986; 5:3259–3266. [PubMed: 3028778]
31. Nambu Y, Sugai M, Gonda H, Lee CG, Katakai T, Agata Y, Yokota Y, Shimizu A. Transcription-coupled events associating with immunoglobulin switch region chromatin. *Science.* 2003; 302:2137–2140. [PubMed: 14684824]
32. Gritzmacher CA. Molecular aspects of heavy-chain class switching. *Crit Rev Immunol.* 1989; 9:173–200. [PubMed: 2505810]
33. Shinkura R, Tian M, Smith M, Chua K, Fujiwara Y, Alt FW. The influence of transcriptional orientation on endogenous switch region function. *Nat Immunol.* 2003; 4:435–441. [PubMed: 12679811]
34. Bhattacharya P, Wuerffel R, Kenter AL. Switch region identity plays an important role in Ig class switch recombination. *J Immunol.* 2010; 184:6242–6248. [PubMed: 20427773]
35. Zarrin AA, Tian M, Wang J, Borjeson T, Alt FW. Influence of switch region length on immunoglobulin class switch recombination. *Proc Natl Acad Sci U S A.* 2005; 102:2466–2470. [PubMed: 15684074]
36. Khamlichi AA, Glaudet F, Oruc Z, Denis V, Le Bert M, Cogne M. Immunoglobulin class-switch recombination in mice devoid of any S mu tandem repeat. *Blood.* 2004; 103:3828–3836. [PubMed: 14962903]
37. Honjo T, Kinoshita K, Muramatsu M. Molecular mechanism of class switch recombination: linkage with somatic hypermutation. *Annu Rev Immunol.* 2002; 20:165–196. [PubMed: 11861601]
38. Dempsey LA, Sun H, Hanakahi LA, 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.* 1999; 274:1066–1071. [PubMed: 9873052]

39. Reaban ME, Griffin JA. Induction of RNA-stabilized DNA conformers by transcription of an immunoglobulin switch region. *Nature*. 1990; 348:342–344. [PubMed: 1701219]
40. Reaban ME, Lebowitz J, Griffin JA. Transcription induces the formation of a stable RNA. DNA hybrid in the immunoglobulin alpha switch region. *J Biol Chem*. 1994; 269:21850–21857. [PubMed: 8063829]
41. Duquette ML, Handa P, Vincent JA, Taylor AF, Maizels N. Intracellular transcription of G-rich DNAs induces formation of G-loops, novel structures containing G4 DNA. *Genes Dev*. 2004; 18:1618–1629. [PubMed: 15231739]
42. Zheng S, Vuong BQ, Vaidyanathan B, Lin JY, Huang FT, Chaudhuri J. Non-coding RNA Generated following Lariat Debranching Mediates Targeting of AID to DNA. *Cell*. 2015; 161:762–773. [PubMed: 25957684]
43. Yu K, Chedin F, Hsieh CL, Wilson TE, Lieber MR. R-loops at immunoglobulin class switch regions in the chromosomes of stimulated B cells. *Nat Immunol*. 2003; 4:442–451. [PubMed: 12679812]
44. Zhang ZZ, Pannunzio NR, Han L, Hsieh CL, Yu K, Lieber MR. The strength of an Ig switch region is determined by its ability to drive R loop formation and its number of WGCW sites. *Cell Rep*. 2014; 8:557–569. [PubMed: 25017067]
45. Wang L, Wuerffel R, Feldman S, Khamlichi AA, Kenter AL. S region sequence, RNA polymerase II, and histone modifications create chromatin accessibility during class switch recombination. *J Exp Med*. 2009; 206:1817–1830. [PubMed: 19596805]
46. Wang L, Whang N, Wuerffel R, Kenter AL. AID-dependent histone acetylation is detected in immunoglobulin S regions. *J Exp Med*. 2006; 203:215–226. [PubMed: 16418396]
47. Barski A, Cuddapah S, Cui K, Roh TY, Schones DE, Wang Z, Wei G, Chepelev I, Zhao K. High-resolution profiling of histone methylations in the human genome. *Cell*. 2007; 129:823–837. [PubMed: 17512414]
48. Bernstein BE, Kamal M, Lindblad-Toh K, Bekiranov S, Bailey DK, Huebert DJ, McMahon S, Karlsson EK, Kulbokas EJ 3rd, Gingeras TR, et al. Genomic maps and comparative analysis of histone modifications in human and mouse. *Cell*. 2005; 120:169–181. [PubMed: 15680324]
49. Guenther MG, Levine SS, Boyer LA, Jaenisch R, Young RA. A chromatin landmark and transcription initiation at most promoters in human cells. *Cell*. 2007; 130:77–88. [PubMed: 17632057]
50. Adelman K, Lis JT. Promoter-proximal pausing of RNA polymerase II: emerging roles in metazoans. *Nat Rev Genet*. 2012; 13:720–731. [PubMed: 22986266]
51. Jonkers I, Lis JT. Getting up to speed with transcription elongation by RNA polymerase II. *Nat Rev Mol Cell Biol*. 2015; 16:167–177. [PubMed: 25693130]
52. Pefanis E, Basu U. RNA Exosome Regulates AID DNA Mutator Activity in the B Cell Genome. *Adv Immunol*. 2015; 127:257–308. [PubMed: 26073986]
53. Xu Z, Fulop Z, Wu G, Pone EJ, Zhang J, Mai T, Thomas LM, Al-Qahtani A, White CA, Park SR, et al. 14-3-3 adaptor proteins recruit AID to 5'-AGCT-3'-rich switch regions for class switch recombination. *Nat Struct Mol Biol*. 2010; 17:1124–1135. [PubMed: 20729863]
54. Nowak U, Matthews AJ, Zheng S, Chaudhuri J. The splicing regulator PTBP2 interacts with the cytidine deaminase AID and promotes binding of AID to switch-region DNA. *Nat Immunol*. 2011; 12:160–166. [PubMed: 21186367]
55. Basu U, Meng FL, Keim C, Grinstein V, Pefanis E, Eccleston J, Zhang T, Myers D, Wasserman CR, Wesemann DR, et al. The RNA exosome targets the AID cytidine deaminase to both strands of transcribed duplex DNA substrates. *Cell*. 2011; 144:353–363. [PubMed: 21255825]
56. Glover-Cutter K, Kim S, Espinosa J, Bentley DL. RNA polymerase II pauses and associates with pre-mRNA processing factors at both ends of genes. *Nat Struct Mol Biol*. 2008; 15:71–78. [PubMed: 18157150]
57. Pei Y, Shuman S. Interactions between fission yeast mRNA capping enzymes and elongation factor Spt5. *J Biol Chem*. 2002; 277:19639–19648. [PubMed: 11893740]
58. Wen Y, Shatkin AJ. Transcription elongation factor hSPT5 stimulates mRNA capping. *Genes Dev*. 1999; 13:1774–1779. [PubMed: 10421630]

59. Pefanis E, Wang J, Rothschild G, Lim J, Chao J, Rabadan R, Economides AN, Basu U. Noncoding RNA transcription targets AID to divergently transcribed loci in B cells. *Nature*. 2014; 514:389–393. [PubMed: 25119026]
60. Taylor BJ, Wu YL, Rada C. Active RNAP pre-initiation sites are highly mutated by cytidine deaminases in yeast, with AID targeting small RNA genes. *Elife*. 2014; 3:e03553. [PubMed: 25237741]
61. Sternberg SH, Redding S, Jinek M, Greene EC, Doudna JA. DNA interrogation by the CRISPR RNA-guided endonuclease Cas9. *Nature*. 2014; 507:62–67. [PubMed: 24476820]
62. Nowacki M, Vijayan V, Zhou Y, Schotanus K, Doak TG, Landweber LF. RNA-mediated epigenetic programming of a genome-rearrangement pathway. *Nature*. 2008; 451:153–158. [PubMed: 18046331]
63. Nussenzweig A, Nussenzweig MC. Origin of chromosomal translocations in lymphoid cancer. *Cell*. 2010; 141:27–38. [PubMed: 20371343]
64. Pasqualucci L, Bhagat G, Jankovic M, Compagno M, Smith P, Muramatsu M, Honjo T, Morse HC 3rd, Nussenzweig MC, Dalla-Favera R. AID is required for germinal center-derived lymphomagenesis. *Nat Genet*. 2008; 40:108–112. [PubMed: 18066064]
65. Kuppers R. Mechanisms of B-cell lymphoma pathogenesis. *Nature reviews Cancer*. 2005; 5:251–262. [PubMed: 15803153]
66. Kuppers R, Dalla-Favera R. Mechanisms of chromosomal translocations in B cell lymphomas. *Oncogene*. 2001; 20:5580–5594. [PubMed: 11607811]
67. Meng FL, Du Z, Federation A, Hu J, Wang Q, Kieffer-Kwon KR, Meyers RM, Amor C, Wasserman CR, Neuberg D, et al. Convergent transcription at intragenic super-enhancers targets AID-initiated genomic instability. *Cell*. 2014; 159:1538–1548. [PubMed: 25483776]
68. Hnisz D, Schuijers J, Lin CY, Weintraub AS, Abraham BJ, Lee TI, Bradner JE, Young RA. Convergence of developmental and oncogenic signaling pathways at transcriptional super-enhancers. *Mol Cell*. 2015; 58:362–370. [PubMed: 25801169]
69. Whyte WA, Orlando DA, Hnisz D, Abraham BJ, Lin CY, Kagey MH, Rahl PB, Lee TI, Young RA. Master transcription factors and mediator establish super-enhancers at key cell identity genes. *Cell*. 2013; 153:307–319. [PubMed: 23582322]
70. Hnisz D, Abraham BJ, Lee TI, Lau A, Saint-Andre V, Sigova AA, Hoke HA, Young RA. Super-enhancers in the control of cell identity and disease. *Cell*. 2013; 155:934–947. [PubMed: 24119843]
71. Ing-Simmons E, Seitan VC, Faure AJ, Flicek P, Carroll T, Dekker J, Fisher AG, Lenhard B, Merckenschlager M. Spatial enhancer clustering and regulation of enhancer-proximal genes by cohesin. *Genome Res*. 2015; 25:504–513. [PubMed: 25677180]
72. Pefanis E, Wang J, Rothschild G, Lim J, Kazadi D, Sun J, Federation A, Chao J, Elliott O, Liu ZP, et al. RNA exosome-regulated long non-coding RNA transcription controls super-enhancer activity. *Cell*. 2015; 161:774–789. [PubMed: 25957685]
73. Han JH, Akira S, Calame K, Beutler B, Selsing E, Imanishi-Kari T. Class switch recombination and somatic hypermutation in early mouse B cells are mediated by B cell and Toll-like receptors. *Immunity*. 2007; 27:64–75. [PubMed: 17658280]
74. Mao C, Jiang L, Melo-Jorge M, Puthenveetil M, Zhang X, Carroll MC, Imanishi-Kari T. T cell-independent somatic hypermutation in murine B cells with an immature phenotype. *Immunity*. 2004; 20:133–144. [PubMed: 14975236]
75. Morgan HD, Dean W, Coker HA, Reik W, Petersen-Mahrt SK. Activation-induced cytidine deaminase deaminates 5-methylcytosine in DNA and is expressed in pluripotent tissues: implications for epigenetic reprogramming. *J Biol Chem*. 2004; 279:52353–52360. [PubMed: 15448152]
76. Schreck S, Buettner M, Kremmer E, Bogdan M, Herbst H, Niedobitek G. Activation-induced cytidine deaminase (AID) is expressed in normal spermatogenesis but only infrequently in testicular germ cell tumours. *J Pathol*. 2006; 210:26–31. [PubMed: 16783758]
77. Bhutani N, Brady JJ, Damian M, Sacco A, Corbel SY, Blau HM. Reprogramming towards pluripotency requires AID-dependent DNA demethylation. *Nature*. 2010; 463:1042–1047. [PubMed: 20027182]

78. Bhutani N, Decker MN, Brady JJ, Bussat RT, Burns DM, Corbel SY, Blau HM. A critical role for AID in the initiation of reprogramming to induced pluripotent stem cells. *FASEB J*. 2013; 27:1107–1113. [PubMed: 23212122]
79. Popp C, Dean W, Feng S, Cokus SJ, Andrews S, Pellegrini M, Jacobsen SE, Reik W. Genome-wide erasure of DNA methylation in mouse primordial germ cells is affected by AID deficiency. *Nature*. 2010; 463:1101–1105. [PubMed: 20098412]
80. Rai K, Huggins IJ, James SR, Karpf AR, Jones DA, Cairns BR. DNA demethylation in zebrafish involves the coupling of a deaminase, a glycosylase, and gadd45. *Cell*. 2008; 135:1201–1212. [PubMed: 19109892]
81. Matsumoto Y, Marusawa H, Kinoshita K, Endo Y, Kou T, Morisawa T, Azuma T, Okazaki IM, Honjo T, Chiba T. Helicobacter pylori infection triggers aberrant expression of activation-induced cytidine deaminase in gastric epithelium. *Nat Med*. 2007; 13:470–476. [PubMed: 17401375]
82. Endo Y, Marusawa H, Kinoshita K, Morisawa T, Sakurai T, Okazaki IM, Watashi K, Shimotohno K, Honjo T, Chiba T. Expression of activation-induced cytidine deaminase in human hepatocytes via NF-kappaB signaling. *Oncogene*. 2007; 26:5587–5595. [PubMed: 17404578]
83. Endo Y, Marusawa H, Kou T, Nakase H, Fujii S, Fujimori T, Kinoshita K, Honjo T, Chiba T. Activation-induced cytidine deaminase links between inflammation and the development of colitis-associated colorectal cancers. *Gastroenterology*. 2008; 135:889–898. 898 e881–883. [PubMed: 18691581]
84. Komori J, Marusawa H, Machimoto T, Endo Y, Kinoshita K, Kou T, Haga H, Ikai I, Uemoto S, Chiba T. Activation-induced cytidine deaminase links bile duct inflammation to human cholangiocarcinoma. *Hepatology*. 2008; 47:888–896. [PubMed: 18306229]
85. Kou T, Marusawa H, Kinoshita K, Endo Y, Okazaki IM, Ueda Y, Kodama Y, Haga H, Ikai I, Chiba T. Expression of activation-induced cytidine deaminase in human hepatocytes during hepatocarcinogenesis. *Int J Cancer*. 2007; 120:469–476. [PubMed: 17066440]
86. Kumar S, Wuerffel R, Achour I, Lajoie B, Sen R, Dekker J, Feeney AJ, Kenter AL. Flexible ordering of antibody class switch and V(D)J joining during B-cell ontogeny. *Genes Dev*. 2013; 27:2439–2444. [PubMed: 24240234]
87. Swaminathan S, Klemm L, Park E, Papaemmanuil E, Ford A, Kweon SM, Trageser D, Hasselfeld B, Henke N, Mooster J, et al. Mechanisms of clonal evolution in childhood acute lymphoblastic leukemia. *Nat Immunol*. 2015; 16:766–774. [PubMed: 25985233]
88. Tsai AG, Lu H, Raghavan SC, Muschen M, Hsieh CL, Lieber MR. Human chromosomal translocations at CpG sites and a theoretical basis for their lineage and stage specificity. *Cell*. 2008; 135:1130–1142. [PubMed: 19070581]
89. Meyers G, Ng YS, Bannock JM, Lavoie A, Walter JE, Notarangelo LD, Kilic SS, Aksu G, Debre M, Rieux-Laucat F, et al. Activation-induced cytidine deaminase (AID) is required for B-cell tolerance in humans. *Proceedings of the National Academy of Sciences of the United States of America*. 2011; 108:11554–11559. [PubMed: 21700883]
90. Kuraoka M, Holl TM, Liao D, Womble M, Cain DW, Reynolds AE, Kelsoe G. Activation-induced cytidine deaminase mediates central tolerance in B cells. *Proc Natl Acad Sci U S A*. 2011; 108:11560–11565. [PubMed: 21700885]
91. Cantaert T, Schickel JN, Bannock JM, Ng YS, Massad C, Oe T, Wu R, Lavoie A, Walter JE, Notarangelo LD, et al. Activation-Induced Cytidine Deaminase Expression in Human B Cell Precursors Is Essential for Central B Cell Tolerance. *Immunity*. 2015
92. Bredemeyer AL, Helmink BA, Innes CL, Calderon B, McGinnis LM, Mahowald GK, Gapud EJ, Walker LM, Collins JB, Weaver BK, et al. DNA double-strand breaks activate a multi-functional genetic program in developing lymphocytes. *Nature*. 2008; 456:819–823. [PubMed: 18849970]
93. Bhutani N, Burns DM, Blau HM. DNA demethylation dynamics. *Cell*. 2011; 146:866–872. [PubMed: 21925312]
94. Franchini DM, Schmitz KM, Petersen-Mahrt SK. 5-Methylcytosine DNA demethylation: more than losing a methyl group. *Annu Rev Genet*. 2012; 46:419–441. [PubMed: 22974304]
95. Cortellino S, Xu J, Sannai M, Moore R, Caretti E, Cigliano A, Le Coz M, Devarajan K, Wessels A, Soprano D, et al. Thymine DNA glycosylase is essential for active DNA demethylation by linked deamination-base excision repair. *Cell*. 2011; 146:67–79. [PubMed: 21722948]

96. Palermo A, Doyonnas R, Bhutani N, Pomerantz J, Alkan O, Blau HM. Nuclear reprogramming in heterokaryons is rapid, extensive, and bidirectional. *FASEB J.* 2009; 23:1431–1440. [PubMed: 19141533]
97. Dominguez PM, Teater M, Chambwe N, Kormaksson M, Redmond D, Ishii J, Vuong B, Chaudhuri J, Melnick A, Vasanthakumar A, et al. DNA Methylation Dynamics of Germinal Center B Cells Are Mediated by AID. *Cell Rep.* 2015; 12:2086–2098. [PubMed: 26365193]
98. Illingworth RS, Bird AP. CpG islands--'a rough guide'. *FEBS Lett.* 2009; 583:1713–1720. [PubMed: 19376112]
99. Maul RW, Saribasak H, Martomo SA, McClure RL, Yang W, Vaisman A, Gramlich HS, Schatz DG, Woodgate R, Wilson DM 3rd, et al. Uracil residues dependent on the deaminase AID in immunoglobulin gene variable and switch regions. *Nat Immunol.* 2010; 12:70–76. [PubMed: 21151102]
100. Schrader CE, Linehan EK, Mochegova SN, Woodland RT, Stavnezer J. Inducible DNA breaks in Ig S regions are dependent on AID and UNG. *J Exp Med.* 2005; 202:561–568. [PubMed: 16103411]
101. Rada C, Williams GT, Nilsen H, Barnes DE, Lindahl T, Neuberger MS. Immunoglobulin isotype switching is inhibited and somatic hypermutation perturbed in UNG-deficient mice. *Curr Biol.* 2002; 12:1748–1755. [PubMed: 12401169]
102. Huang FT, Yu K, Balter BB, Selsing E, Oruc Z, Khamlichi AA, Hsieh CL, Lieber MR. Sequence dependence of chromosomal R-loops at the immunoglobulin heavy-chain Smu class switch region. *Mol Cell Biol.* 2007; 27:5921–5932. [PubMed: 17562862]
103. Tous C, Aguilera A. Impairment of transcription elongation by R-loops in vitro. *Biochem Biophys Res Commun.* 2007; 360:428–432. [PubMed: 17603014]
104. Rajagopal D, Maul RW, Ghosh A, Chakraborty T, Khamlichi AA, Sen R, Gearhart PJ. Immunoglobulin switch mu sequence causes RNA polymerase II accumulation and reduces dA hypermutation. *J Exp Med.* 2009; 206:1237–1244. [PubMed: 19433618]

Highlights

AID is guided by non-coding germline transcripts to switch region targets

Integration of transcription and AID targeting with class switch recombination

AID deamination function intersects with global DNA demethylation

AID regulation of genomic DNA methylation may regulate gene expression

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

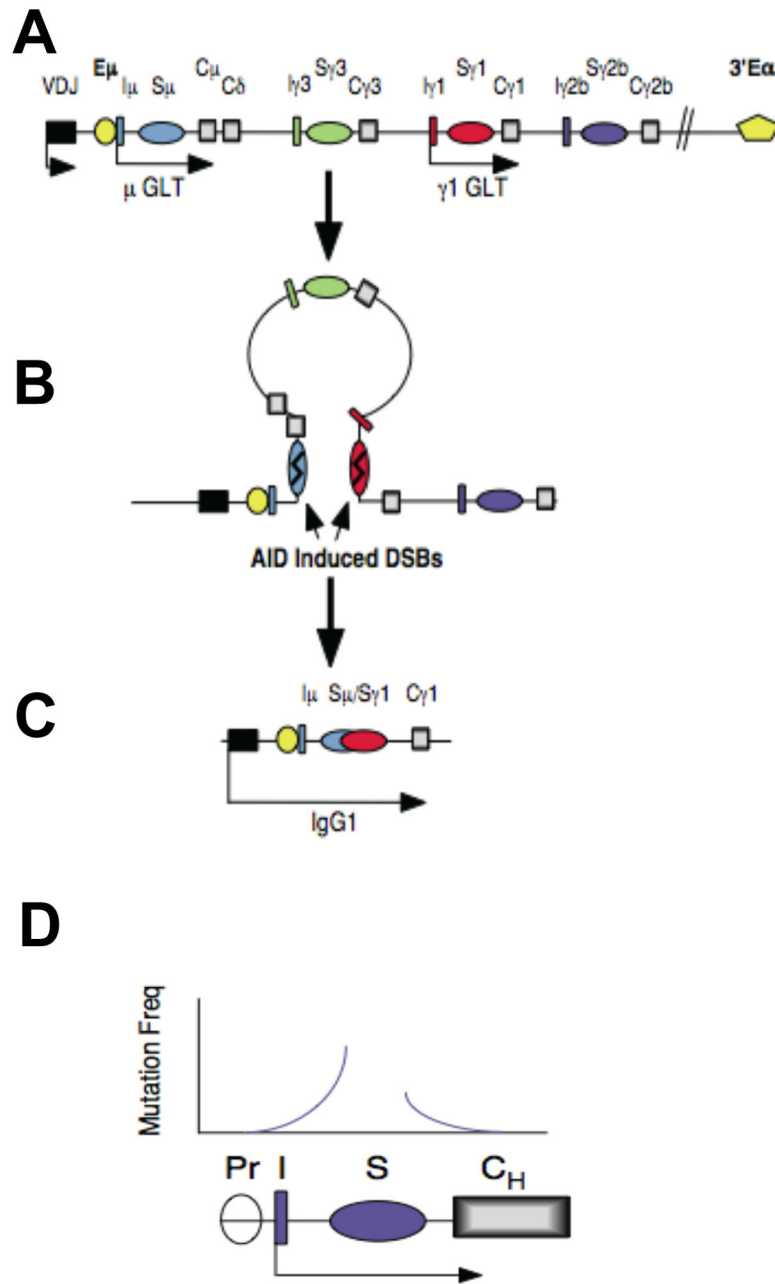


Figure 1. The looping-out and deletion model of Ig switch recombination

A) A partial schematic map of *Igh* locus before CSR is shown not to scale. A productive V(D)J rearrangement has occurred allowing expression of the γ and μ IgH chains. Intact S_{μ} and $S_{\gamma 1}$ are separated by approximately 70 kb. Stimulation of B cells with antigen or mitogen induces germline transcription through the $I_{\gamma 1}$ - $S_{\gamma 1}$ - $C_{\gamma 1}$ region prior to recombination. **B)** The S_{μ} and $S_{\gamma 1}$ regions are aligned causing the intervening genomic DNA to form a loop. **C)** A reciprocal crossover between S_{μ} and $S_{\gamma 1}$ results in the formation of a new hybrid transcriptional unit containing the original VDJ exons contiguous with $C_{\gamma 1}$ and the formation of hybrid $S_{\mu/S_{\gamma 1}}$ molecules. **D)** A schematic diagram depicts a generic I-

S-C_H region downstream of the GLT promoter (Pr). Above the diagram a summary of mutation frequencies 5' and 3' of the S region is shown.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript