Published in final edited form as:

Adv Cancer Res. 2012 ; 113: 167–190. doi:10.1016/B978-0-12-394280-7.00005-1.

Activation Induced Deaminase in Antibody Diversification and Chromosome Translocation

Anna Gazumyan^{1,2}, Anne Bothmer¹, Isaac A. Klein¹, Michel C Nussenzweig^{1,2}, and Kevin M. McBride³

¹Laboratory of Molecular Immunology, The Rockefeller University, New York, New York, USA

²Howard Hughes Medical Institute, The Rockefeller University, New York, New York, USA

³Department of Molecular Carcinogenesis, University of Texas MD Anderson Cancer Center, Smithville, Texas 78957, USA

Abstract

DNA damage, rearrangement and mutation of the human genome are the basis of carcinogenesis and thought to be avoided at all costs. An exception is the adaptive immune system where lymphocytes utilize programmed DNA damage to effect antigen receptor diversification. Both B and T lymphocytes diversify their antigen receptors through RAG1/2 mediated recombination, but B cells undergo two additional processes – somatic hypermutation (SHM) and class switch recombination (CSR), both initiated by Activation Induced Deaminase (AID). AID deaminates cytidines in DNA resulting in U:G mismatches that are processed into point mutations in SHM or double strand breaks in CSR. Although AID activity is focused at Immunoglobulin (*Ig*) gene loci, it also targets a wide array of non-*Ig* genes including oncogenes associated with lymphomas. Here we review the molecular basis of AID regulation, targeting, and initiation of CSR and SHM, as well as AID's role in generating chromosome translocations that contribute to lymphomagenesis.

Introduction

Adaptive immunity is an exquisitely specific immune response that vertebrates have evolved to recognize and remember specific pathogens. A key event is the somatic assembly of unique immune receptors; antibodies from immunoglobulin (*Ig*) genes in B lymphocytes and the T cell receptor in T lymphocytes. During development each lymphocyte is capable of generating a unique receptor specific to one of a vast array of possible antigens. For example, in humans B cells generate a repertoire of receptors that exceed trillions of unique antibodies (Abbas et al., 2010). This diversity is generated by a system of programmed DNA damage that recombines and alters the coding sequence of lymphocyte immune receptors.

Initial antibody assembly occurs in the bone marrow. There, B lymphocytes express RAG1/2 to catalyze V(D)J recombination, a site-specific recombination reaction that juxtaposes variable (V), diversity (D), and joining (J) gene segments to one another

(Fugmann et al., 2000; Jung et al., 2006). T lymphocytes assemble the T-cell receptor in a similar RAG1/2-mediated manner. However, in contrast to T cells, B cells can further diversify *Ig* genes through somatic hypermutation (SHM) and class switch recombination (CSR) after antigen encounter (Figure 1). SHM alters antibody affinity by introducing nucleotide changes in the antigen binding variable region of antibodies. B cells producing antibodies with improved antigen affinity are positively selected during the process of affinity maturation. CSR is a region-specific recombination reaction that replaces one antibody-constant region with another, thereby altering antibody effector function while leaving the variable region and its antigen binding specificity intact (Di Noia and Neuberger, 2007; Peled et al., 2008; Rajewsky, 1996; Stavnezer et al., 2008; Teng and Papavasiliou, 2007). While CSR and SHM are very different reactions, both are initiated by AID (Muramatsu et al., 2000; Revy et al., 2000), which introduces uracil:guanine (U:G) mismatches in transcribed DNA (Bransteitter et al., 2003; Chaudhuri et al., 2003; Dickerson et al., 2003; Petersen-Mahrt et al., 2002; Ramiro et al., 2003). These U:G mismatches are fixed to a mutation in the case of SHM or processed to double stranded DNA breaks (DSB), which serve as obligate intermediates in the recombination reaction during CSR (Di Noia and Neuberger, 2007; Stavnezer, et al., 2008).

Chromosome translocations

If not properly repaired, physiological DSBs that arise during CSR may pose a threat to genome integrity. For example, they can be substrates for chromosome rearrangements such as deletions and translocations and can lead to malignant transformation (Gostissa et al., 2009; Nussenzweig and Nussenzweig, 2010; Tsai and Lieber, 2010; Zhang et al., 2010). While deletions may occur by joining breaks on one chromosome in *cis*, chromosome translocations involve the joining of paired DSBs on different chromosomes. Translocations may arise in either reciprocal or non-reciprocal conformations. In the latter, formation of genetically unstable dicentric or acentric chromosomes with or without the loss of chromosome segments may result. Conversely, a reciprocal translocation involves the exchange of telomeric and centromeric chromosome portions, resulting in the formation of two stable hybrid chromosomes with complete sequence retention (Figure 2). These balanced events can then be stably propagated during cell division. Most hematopoietic malignancies harbor clonal reciprocal translocations (Kuppers, 2005). The translocations are frequently a consequence of genetically programmed DSB generation during lymphocyte development. And, since these events may induce oncogenic transformation through several mechanisms, they are thought to be etiologic in many cases. For example, a highly active promoter or cis-regulatory element can be juxtaposed to a proto-oncogene thereby deregulating its expression. An “infamous” example is the *c-myc/IgH* translocation, a hallmark of Burkitt's lymphoma, which places *IgH* regulatory elements upstream of the *c-myc* proto-oncogene. A chromosome translocation may bring together disparate coding sequences to form a chimeric fusion protein. For example, the BCR-ABL fusion, found in chronic myeloid leukemia results in constitutively active ABL kinase (Kuppers, 2005; Potter, 2003).

Unlike V(D)J recombination, which introduces DSBs at specific recombination signal sequences (RSSs), CSR is a regional, imprecise recombination event. Aberrant DSBs during

CSR can occur in non-Ig loci and participate in chromosome translocations (Robbiani et al., 2008). The promiscuity of this B cell-specific process is reflected in the relative prevalence of B cell lymphomas in the population; more than 90% of human lymphomas in the western world arise from B cells as opposed to T cells (Kuppers, 2005). The majority of these originate from mature B cells or post-germinal center B cell compartments where AID expression is normally induced. Therefore, AID and aberrant switching events may be substantial contributors to the molecular etiology of B cell lymphomas.

Activation Induced Deaminase

A seminal discovery in understanding the molecular mechanism of CSR and SHM was the identification of AID by Honjo and colleagues (Muramatsu et al., 1999). The finding that AID deficiency abolished CSR and SHM in mice and humans confirmed its essential role in both processes (Muramatsu, et al., 2000; Revy, et al., 2000). Defects in the AID gene (*AICDA*), located at 12p13, are the cause of autosomal recessive hyper-IgM syndrome type 2 (HIGM2) in humans (Revy, et al., 2000). Initial sequence analysis of AID revealed a deaminase homologous to the mRNA editing cytidine deaminase, APOBEC1 (Muramatsu, et al., 1999); however, attempts to show deaminase activity on RNA have failed and no RNA substrate has been identified. In contrast, *in vitro* analysis has revealed that AID can directly deaminate single stranded DNA (Bransteitter, et al., 2003; Chaudhuri, et al., 2003; Dickerson, et al., 2003; Petersen-Mahrt, et al., 2002; Pham et al., 2003; Ramiro, et al., 2003; Sohail et al., 2003). Indeed, the preponderance of biochemical, cell biology and genetic evidence supports a model in which AID deaminates DNA to initiate CSR and SHM (Di Noia and Neuberger, 2007; Petersen-Mahrt, et al., 2002). During SHM, mutations may be generated by replication over U:G mismatches. Alternatively, lesions may be processed by uracil DNA glycosylase (UNG), or the mismatch repair proteins MSH2/MSH6 and translesion polymerases involved in error prone DNA synthesis. In CSR U:G mismatches in donor and acceptor switch regions are processed to DSBs that are joined by both the classical and alternative non-homologous end joining pathways (NHEJ) (Delker et al., 2009; Di Noia and Neuberger, 2007; Maul and Gearhart, 2010; Maul et al.; Peled, et al., 2008; Stavnezer, et al., 2008).

Regulation of AID

AID has significant potential for inducing genomic damage and tumorigenesis across cell types. Mice carrying a ubiquitously expressed AID transgene driven by the actin promoter develop tumors in a wide variety of tissues, but curiously not in B cells (Okazaki et al., 2003; Rucci et al., 2006). Development of AID dependent B cell cancers requires additional inactivation of p53 (Ramiro et al., 2006; Robbiani et al., 2009). Normally AID expression is cell type restricted - mainly to germinal-center or activated B cells. *In vitro*, cytokines such as IL-4 and TGF- β in combination with B cell mitogens such as CD40 ligation or LPS induce AID (Dedeoglu et al., 2004; Muramatsu, et al., 1999; Zhou et al., 2003). A highly conserved regulatory region within the AID gene is thought to be responsible for this restriction (Crouch et al., 2007) by enforcing a requirement for PAX5 and E47 transcription factor binding that is balanced by inhibitors of differentiation (Id) proteins Id2 and Id3 (Gonda et al., 2003; Sayegh et al., 2003; Tran et al., 2010). A number of cytokine response

elements including NFkB, STAT6 and Smad3/4, also reside there and confer response to signaling by CD40L, IL-4 and TGFbeta (Tran, et al., 2010).

The regulation of AID expression plays an important role in controlling its activity since AID levels are rate-limiting for both SHM and CSR. AID exhibits haploinsufficiency or hyperactivity when overexpressed by virtue of a retrovirus or transgene (McBride et al., 2008; Robbiani, et al., 2009; Sernandez et al., 2008; Takizawa et al., 2008). Importantly, when overexpressed AID not only increases the rates of physiological SHM and CSR, but also generates widespread genomic instability and associated rearrangements (Robbiani, et al., 2009). Therefore, AID levels are strictly controlled to maintain genomic stability and to support proper antibody development, which is achieved by multiple levels of regulation.

Once transcribed, the stability and half-life of AID mRNA is controlled by miR-155 and miR-181b binding to sites in the 3' untranslated region (de Yebenes et al., 2008; Dorsett et al., 2008; Teng et al., 2008). Deficiency of miR-155 or mutation of the miR-155 binding site in AID mRNA results in increased AID levels and a marked increase in *c-myc/IgH* translocations. This indicates that miR-155 could act as a tumor suppressor by directly reducing AID levels (Dorsett, et al., 2008). In the cytoplasm, AID protein levels are stabilized by HSP90 (heat shock protein 90) interaction, which prevents ubiquitylation and degradation. Inhibition of HSP90 results in decreased AID levels, decreased CSR rates and fewer off-target mutations (Orthwein et al., 2010).

AID, like many proteins with nuclear function, is regulated on the level of subcellular localization. Although AID is actively transported to the nucleus, (Patenaude et al., 2009) its steady-state localization is predominantly cytoplasmic (Rada et al., 2002). This is accomplished by a combination of cytoplasmic retention, (Patenaude, et al., 2009), degradation of nuclear AID (Aoufouchi et al., 2008) and active nuclear export by virtue of a carboxyl terminal nuclear export signal (NES) recognized by the CRM1 shuttling factor (Brar et al., 2004; Ito et al., 2004; McBride et al., 2004).

Protein modification by phosphorylation is one of the most common means to alter activity. Compared to the previously mentioned mechanisms, phosphorylation has the potential to rapidly and dynamically modulate activity. In switching B cells, mass spectrometry analysis of AID has revealed phosphorylation on serine 3, serine 38, threonine 140 and tyrosine 184 (S3, S38, T140, Y184 respectively) (S3, S38, T140, Y184) (Basu et al., 2005; Gazumyan et al., 2011; McBride, et al., 2008). With the exception of Y184, these phosphorylation sites have been confirmed by AID anti-phospho antibodies (Gazumyan, et al., 2011; McBride et al., 2006; McBride, et al., 2008). Mutation of S38 or T140 to alanine does not impact AID catalytic activity, but do interfere with CSR and SHM, indicating that they are positive regulators (Basu, et al., 2005; Chatterji et al., 2007; Gazumyan, et al., 2011; McBride, et al., 2006; Pasqualucci et al., 2006). While S38 is equally important for CSR and SHM, T140 phosphorylation preferentially affects SHM suggesting post translational modification may contribute to the choice between SHM and CSR (McBride, et al., 2008). In contrast to S38 and T140, phosphorylation of S3 acts as a negative regulator. AID carrying an S3A mutation displays increased CSR and *c-myc/IgH* translocation frequencies, demonstrating that the oncogenic potential of AID can be regulated by phosphorylation (Gazumyan, et al., 2011).

The signaling pathways that control AID phosphorylation are unknown; however, evidence suggests PKA phosphorylates AID at S38: it is within a PKA consensus site, it is phosphorylated by PKA *in vitro*, and inhibitors of PKA decrease CSR (Basu, et al., 2005; Pasqualucci, et al., 2006). Although the overall levels of S38 phosphorylation are low (5% total), chromatin associated AID is highly phosphorylated (McBride, et al., 2006). Consistent with this finding, PKA is specifically recruited to switch regions (Vuong et al., 2009).

There is unique control of each AID phosphorylation site, as AID-T140 can be phosphorylated by PKC, but not PKA, and protein phosphatase 2A (PP2A) preferentially controls AID-S3 levels (Gazumyan, et al., 2011; McBride, et al., 2008). It appears that multiple signaling pathways converge on AID to control activity through phosphorylation. The complete story of AID regulation by phosphorylation is likely to be even more complicated; little is known about the function of Y184 phosphorylation and additional phosphorylation sites at T27, S41 and S43 have been detected in AID purified from Sf9 insects cells (Pham et al., 2008). Taken together, the evidence suggests that increasing AID levels or activity increases SHM and CSR efficiency at the expense of genomic integrity. Therefore, the summation of AID regulatory mechanisms may strike an important equilibrium between the rate of antibody evolution and oncogenesis.

AID is targeted by transcription

Transcription is an absolute requirement for CSR and SHM, and their rates directly relate to the rate of transcription (Bachl et al., 2001). This connection between transcription and SHM and CSR was appreciated long before the discovery of AID. Early studies of mutation patterns in B cells revealed that SHM was largely confined to the transcribed 1-2 kb region 3' to the V(D)J or switch region promoter. Promoter regions and distal downstream regions were largely devoid of mutations (Both et al., 1990; Lebecque and Gearhart, 1990; Rothenfluh et al., 1993; Steele et al., 1992; Weber et al., 1991). The finding that *Ig* transgenes undergo SHM at levels that correlated with transcription rates provided additional evidence for this connection (Bachl, et al., 2001). Similar relationships between transcription and recombination of switch regions in CSR were also found (Chaudhuri and Alt, 2004; Guikema et al., 2008; Stavnezer, et al., 2008). Replacement of the V(D)J region with unrelated sequences did not significantly alter SHM rates, nor did replacing the V promoter with promoters from non-*Ig* (B-globin or B29) loci (Betz et al., 1994; Tumas-Brundage and Manser, 1997). This suggests that transcription, and not specific sequence elements per se, were essential to targeting SHM. A conclusive link between SHM and transcription was shown by Storb and colleagues. They demonstrated that a normally unmutated exon underwent SHM when repositioned to a promoter proximal location, demonstrating position within a transcription unit dictated SHM (Peters and Storb, 1996). These experiments together with the spatial positioning of mutations from the promoter suggested that the active Pol II complex is playing an essential role in AID targeting *in vivo*. Based on their study Peters and Storb proposed a model linking somatic mutation to transcription. According to that model a “mutator factor” present only in mutating B cells loads on to the transcription initiation complex assembled at the promoter and remains associated with the elongating complex (Peters and Storb, 1996).

The mechanistic role of transcription in SHM and CSR became evident after AID was discovered to be a ssDNA deaminase. *In vitro* assays with purified AID demonstrated no activity on double stranded DNA, but showed activity against single stranded DNA and transcribed double stranded DNA (Bransteitter, et al., 2003; Chaudhuri, et al., 2003; Dickerson, et al., 2003; Petersen-Mahrt, et al., 2002; Pham, et al., 2003; Ramiro, et al., 2003; Sohail, et al., 2003). This suggested that transcription was creating the ssDNA substrate. The precise structure of ssDNA that is accessible to AID is not known but transcription bubbles, supercoiled DNA in the wake of the polymerase and R-loops (RNA-DNA hybrid structures that can displace the non transcribed DNA strand) are suggested to be targets (Besmer et al., 2006; Lumsden et al., 2004; Yu et al., 2003).

Consistent with this, DNA transcribed by *E. coli*, T7 and mammalian RNA Pol II all become AID targets both *in vitro* and *in vivo*. Interestingly, there is an inherent bias for AID to target the non-template transcribed DNA strand in *E. coli* but not in mammalian cells. This suggests the existence of a specific mechanism targeting AID to template DNA. A number of groups have suggested that RNA-DNA hybrid R loops may influence AID access (Basu et al., 2011; Bransteitter, et al., 2003; Pham, et al., 2003; Ramiro, et al., 2003). The likely role of transcription therefore, is to create the ssDNA substrate for AID. However, most transcribed genes do not undergo SHM in B cells (Liu et al., 2008). The finding that AID physically associates with the Pol II complex in B cells suggested that there were additional factors that influence AID targeting (Nambu et al., 2003; Pavri et al., 2010).

AID targeting by Spt5 and stalled Pol II

Studies of Pol II, AID and Spt5 distribution at the *Ig* locus suggested a relationship between Pol II density, elongation stalling and AID activity. ChIP analysis of Pol II distribution demonstrated a spike in density through the core of *Ig* switch regions in activated B cells (Rajagopal et al., 2009). Such density suggested Pol II accumulation due to elongation stalling. Consistent with this, nuclear run-on analysis showed a concomitant accumulation of transcripts within the transcribed switch regions (Rajagopal, et al., 2009). A high Pol II density within the highly repetitive switch regions seemed to be an intrinsic feature of the locus, occurring even in the absence of AID (Rajagopal, et al., 2009; Wang et al., 2009). Impressively, Pol II density correlated with the rate and distribution of mutation through the switch region. A stalled Pol II complex would be predicted to expose ssDNA AID substrate created by transcription for prolonged periods. The observed stalling at the *Ig* switch regions provided a logical mechanism by which Pol II-associated AID would have preferential access to sites of elongation stalling. Consistent with this concept, stalling has been reported at other AID target sites including *Igk* and *c-myc* (Bentley and Groudine, 1986; Raschke et al., 1999).

The study by Pavri et al. of the Pol II stalling factor Spt5 provided a direct mechanistic link between AID and stalled Pol II. In an unbiased loss of function screen, Suppressor of Ty5 homolog (Spt5) was found to be an essential factor for CSR (Pavri, et al., 2010). Spt5 is a Pol II interacting elongation factor that associates with stalled Pol II and ssDNA (Gilmour, 2009; Lis, 2007; Rahl et al., 2010). Spt5 partners together with Spt4 to form the DRB sensitivity inducing factor (DSIF) complex (Swanson et al., 1991; Wada et al., 1998). DSIF

can interact with the negative elongation factor (NELF) complex to promote the stalling of Pol II (Yamaguchi et al., 1999). This stall effect of DSIF and NELF is relieved by P-TEFb, a kinase that phosphorylates Pol II and Spt5 (Kim and Sharp, 2001; Yamada et al., 2006). Co-immunoprecipitation of Spt5 interacting factors revealed an association with AID in activated B cells, while binding studies with recombinant proteins demonstrated a direct interaction (Pavri, et al., 2010). Since both Spt5 and AID associate with the Pol II complex (Nambu, et al., 2003; Wada, et al., 1998) the contribution of each protein to this interaction was examined. siRNA mediated depletion of Spt5 decreased the amount of Pol II co-immunoprecipitated by AID. In contrast, Pol II depletion did not alter AID-Spt5 interaction suggesting that Spt5 serves as a direct adaptor to recruit AID to Pol II (Pavri, et al., 2010).

AID induced mutations are known to be widely distributed among transcribed genes in B cells (Liu, et al., 2008). Genome-wide AID ChIP-Seq analysis confirmed the widespread association of AID with thousands of promoter proximal transcribed areas. Strikingly, Pol II occupancy rather than overall transcription rate correlated with AID load (Yamane et al., 2011). AID was specifically enriched in Pol II dense promoter proximal regions, which are known sites of Pol II stalling (Yamane, et al., 2011). There was also a strong correlation between Spt5 and AID recruitment in these regions, consistent with Spt5-AID association at Pol II stalling sites (Pavri, et al., 2010). Importantly, a high density of AID, Spt5 and Pol II were found in the area of known switch region hypermutation (Pavri, et al., 2010; Yamane, et al., 2011). Thus, a Spt5 mediated targeting mechanism orchestrated by Pol II stalling could explain the targeting enigma of AID and the widespread promiscuity of AID mutations.

The exact mechanism by which AID-Spt5 accesses ssDNA is unknown, however the recently solved crystal structure of *Pyrococcus furiosus* (Pfu) Spt 4/5 in complex with the RNA polymerase clamp domain gives some insight. RNA polymerases (RNAP) are found in all organisms and structural studies reveal universal conservation of architecture and active site mechanism throughout evolution (Cramer et al., 2008; Klein et al., 2011a). The only RNAP associated factor universally conserved through all evolution is Spt5 (Cramer, et al., 2008; Martinez-Rucobo et al., 2011). The Spt4/5-Pol II structure reveals that Spt5 is in close proximity to the upstream edge of the transcription bubble (Figure 3) (Klein, et al., 2011a; Martinez-Rucobo, et al., 2011). Based on this model AID interaction with Spt5 in the Pol II complex would result in close spatial proximity to ssDNA at the transcription bubble and the newly formed RNA-DNA hybrid (Figure 3). In combination, AID interaction with Spt5 place AID near ssDNA substrates in a Pol II elongation complex, while stalling of Pol II would result in a prolonged ssDNA substrate exposure (Figure 4).

Template strand targeting by Exosomes

AID targets both the template and non-template strand of transcribed genes *in vivo*, however in *E. coli* or *in vitro* the non-template strand is preferentially targeted (Bransteitter, et al., 2003; Chaudhuri, et al., 2003; Dickerson, et al., 2003; Petersen-Mahrt, et al., 2002; Pham, et al., 2003; Ramiro, et al., 2003; Sohail, et al., 2003). Transcription readily exposed ssDNA on the non-template strand but the mechanism for template strand targeting was unclear. However, RNA-DNA hybrids such as those formed on transcribed template strands are

inhibitory to AID (Bransteitter, et al., 2003). The recent discovery by Alt and colleagues that the exosome complex associates with AID and provides access to the template strand *in vitro* provides an explanation (Basu, et al., 2011). Components of the exosome were identified by mass spectrometry from Ramos cell fractions that enhanced deamination activity of purified AID on *in vitro* transcribed dsDNA. They were confirmed to interact *in vivo* and siRNA knockdown resulted in diminished SHM however most importantly, purified exosome resulted in the appearance of mutations on the template strand (Basu, et al., 2011). The evolutionally conserved exosome complexes are 3' to 5' exoribonucleases involved in RNA processing, turnover and degradation of unstable stalled transcripts (Houseley et al., 2006). The exosome is only known to engage free 3' ends of RNA that would be buried and inaccessible in an elongating Pol II complex. However stalled polymerase may represent a special situation since they are known to undergo backtracking resulting in an exposed 3' RNA substrate that exosomes may access (Adelman et al., 2005). Furthermore there is a known association of Spt5 and exosome components (Andrulis et al., 2002). Therefore, AID associated Spt5/Pol II stalled complexes may be special targets for exosome action to remove nascent RNA transcripts and expose the template strand (Pavri, et al., 2010). The stalling model of AID targeting would therefore provide a mechanism for not only AID access but for exposure of the template strand via exosome processing (Figure 4).

AID Genome wide damage

Following the discovery that *Ig* variable regions undergo high levels of hypermutation (McKean et al., 1984; Weigert et al., 1970) analysis of the surrounding promoter and constant regions revealed few mutations (Both, et al., 1990; Lebecque and Gearhart, 1990; Rothenfluh, et al., 1993; Steele, et al., 1992; Weber, et al., 1991). It was therefore concluded that hypermutation was confined to the *Ig* variable region. This assumption was extended to other areas of the genome as it seemed logical that a dangerous process like hypermutation would be confined. Over the next decades reports accumulated that a few non-*Ig* genes did undergo SHM in germinal center B cells (*Bcl-6*, *IgA*, *IgB*, *FAS*, *c-myc*, *miR-142*) (Gordon et al., 2003; Muschen et al., 2000; Pasqualucci et al., 1998; Pasqualucci et al., 2001; Pavri, et al., 2010; Robbiani, et al., 2009; Shen et al., 1998; Yamane, et al., 2011). Furthermore, it was demonstration that a retroviral indicator gene with no *Ig* elements underwent hypermutation at multiple locations (Wang et al., 2004) challenging the notion that hypermutation was confined to *Ig* loci.

A mutational analysis of ~120 highly transcribed genes in germinal center B cells showed that more than 25% accumulated mutations in an AID dependent manner (Liu, et al., 2008). However, most of these genes mutated at rates much lower than the *Ig* variable region. Strikingly, almost 50% of assayed genes accumulated mutations in UNG/MSH2 double knock out B cells, also in an AID dependent manner (Liu, et al., 2008). This suggested that low level AID targeting was widespread in the genome and that high-fidelity repair served to suppress AID induced mutations (Liu and Schatz, 2009). The extent of potential AID targets genome-wide was revealed by AID-CHIP seq analysis that showed AID loading on more than 5000 genes (Yamane, et al., 2011).

In addition to mutations, AID lesions can also be processed into DSBs that serve as substrates for CSR and chromosome translocations. Recurrent translocations are frequently found in lymphomas and leukemias and, reciprocal chromosomal translocations between oncogenes and the *IgH* regulatory elements are hallmark cytogenetic abnormalities in many mature B cell lymphomas. The first evidence that AID induced DSBs act as translocation substrates came from studies examining *c-myc/IgH* translocations, a hallmark of Burkitt's lymphoma. AID was shown to be essential for the appearance of *c-myc/IgH* chromosome translocations *in vivo* (Ramiro et al., 2004). Further characterization demonstrated that AID was creating the DSB necessary for *c-myc/IgH* translocations at both the *c-myc* and *IgH* loci (Robbiani, et al., 2008). Consistent with widespread AID hypermutation, it was soon shown that AID is capable of inducing DSB at several non-Ig loci, thereby catalyzing their rearrangement and triggering oncogenesis (Robbiani, et al., 2009). Like SHM, AID induced translocations are suppressed by internal surveillance and protection pathways. Stress-induced checkpoint control proteins (ATM, Nbs1, p19-Arf and p53) suppress translocations, the loss of any results in greatly enhanced *c-myc/IgH* translocation frequencies (Ramiro, et al., 2006).

Until recently, the genome-wide incidence of AID induced DSBs had not been measured. Only specific loci, such as *c-myc*, have been conclusively shown to be targeted by AID for DSBs (Robbiani, et al., 2008; Robbiani, et al., 2009). And, while ChIP-Seq studies revealed thousands of potential AID targets (Staszewski et al., 2011; Yamane, et al., 2011), and selective sequencing of specific loci suggested widespread AID activity (Liu, et al., 2008), the unbiased measurement of AID DSB activity had yet to be achieved. The recently developed technique of genome-wide translocation capture sequencing facilitated the unbiased discovery of AID targets, thereby revealing the loci that suffer AID-mediated DSBs and confirming the current model of AID targeting (Chiarle et al., 2011; Klein et al., 2011b). The two groups that simultaneously developed this method both engineered I-SceI meganuclease sites at the *IgH* or the *c-myc* locus. Defined DSBs were induced by retroviral introduction of the I-SceI nuclease in activated B cells in the presence or absence of AID. After short-term culture, junctions between the I-SceI induced DSB and other endogenous DSB sites were identified by PCR amplification and deep sequencing (Oliveira et al., 2011). The capture strategy allowed unbiased amplification of DNA surrounding the I-SceI site but selected against intact I-SceI sites. Thereby events that disrupted the I-SceI site, such as translocations would be identified. The power of this methodology is the capture of single cell events in short-term culture. This allows for the unbiased analysis of rare translocation partners, providing insight into the mechanisms of genomic rearrangement at the earliest stages of oncogenic transformation. By defining recurrent AID-dependent rearrangement hotspots these studies revealed the DSB targets of AID (Chiarle, et al., 2011; Klein, et al., 2011b).

Using this method nearly 200,000 independent junctions were identified. Whole genome rearrangement maps reveal that the pluralities of rearrangement partners are local, within several kb of the engineered I-SceI break site at both the *IgH* and *c-myc* locus. Farther afield, double strand breaks still showed a strong preference towards intrachromosomal rearrangements, suggesting the presence of chromosome territories (Chiarle, et al., 2011;

Klein, et al., 2011b). This suggests that both chromosome positioning and proximal DNA damage response play a role in DSB partner ligation. The tendency toward local translocation partners could have the functional affect of abating more deleterious rearrangements to distant parts of the genome (Chiarle, et al., 2011; Klein, et al., 2011b).

When interchromosomal translocations did occur they did so across the genome, most frequently to genic regions near active transcription start sites. The enrichment of rearrangements in transcription start sites, which are sites of ssDNA exposure, implicates ssDNA as a significant source of genomic instability even in the absence of AID. Inducing expression of AID revealed hundreds of AID-dependent rearrangement hotspots that frequently rearranged to both the *c-myc* and *IgH* locus. When the regions bearing AID-dependent rearrangements were examined for shared genomic features several interesting trends emerged. Consistent with an ssDNA substrate, AID dependent hotspots were clustered in the transcription start sites of transcribed genes. However, transcription itself was neither rate-limiting nor sufficient for AID targeting; thousands of highly transcribed genes remained undamaged by AID. (Chiarle, et al., 2011; Klein, et al., 2011b). In contrast, stalled Pol II and the presence of Spt5 correlated well with AID translocation hotspots (Klein, et al., 2011b; Pavri, et al., 2010; Yamane, et al., 2011). Furthermore, analysis of mutations rates at AID hotspots showed a strong correlation suggesting that the rate of SHM and DSB creation at non-*Ig* are proportional (Klein, et al., 2011b). Hence the genome wide mapping of AID induced DSB is consistent with an AID-Spt5-Pol II targeting complex at stalled Pol II sites (Figure 4). Finally, many loci targeted for DSBs by AID are found to participate in clonal translocations in mature B cell lymphomas. This indicates that AID is an important source of genomic instability in these tumors to generate chromosomal abnormalities that drive lymphomagenesis.

CSR specific cofactors

The model of AID targeting by Pol II (Fig 4) explains several key observations of SHM and CSR including the transcription requirement and promoter proximal occupancy. However the spread of mutations and DSB over several kb suggests that AID action is coincident along the elongating Pol II. Therefore together with targeting, the stability of this interaction may be important to maintain AID association across the switch region. Multiple mechanisms may impact this including AID phosphorylation and interactions with partner proteins. For example the splicing factor Ptbp2 and scaffold proteins of the 14-3-3 family are critical for CSR and necessary for AID switch region association (Nowak et al., 2011; Xu et al., 2010). AID phosphorylation regulates AID activity without affecting catalytic function, suggesting it may also impact targeting or stability with Pol II. AID phosphorylation recruits the single stranded DNA binding protein RPA to the switch regions where it augments CSR (Basu, et al., 2005; Vuong, et al., 2009; Yamane, et al., 2011). Hence, AID phosphorylation could stabilize partner interactions and therefore contribute to prolonged AID presence at the switch regions. Future studies of the basis of AID interaction with partners and regulation via post-translational modification may provide a more complete view.

AID influences CSR outcome

The carboxyl terminus of AID, the same domain that encodes an NES, is required for CSR but not SHM. Interestingly, this domain is a nexus for multiple functions. Deletion of this region results in: i) increased nuclear AID; ii) decreased protein stability; iii) increased off target DNA damage with loss of CSR and hyper-IGM syndrome in humans (Barreto et al., 2003; Geisberger et al., 2009; McBride, et al., 2004; Shinkura et al., 2004; Ta et al., 2003). That AID lacking the C-terminal region is able to target and mutate the switch regions but does not support CSR (Barreto, et al., 2003; Ta, et al., 2003) suggests that AID interacts with class-switch specific cofactors (Doi et al., 2009). Interestingly, mutant AID lacking the C-terminal region acts as a dominant negative molecule in heterozygous individuals resulting in severely compromised CSR efficiency. Examination of the switch region junctions in the few cells that accomplished CSR revealed a bias toward junctions bearing microhomology (Kracker et al., 2010). This suggested that domains of AID including the C-terminal region impacted the repair of created lesions, possibly by interacting with CSR specific factors. Consistent with this, the C-terminus of AID is involved in cooperative binding of AID, UNG and mismatch repair proteins MSH2-MSH6 to the switch regions (Ranjit et al., 2011). Other AID interacting factors including KAP1 and Spt6 have been found to be critical for CSR but not SHM (Jeevan-Raj et al., 2011; Okazaki et al., 2011). Altogether this suggests that AID itself plays an important role in dictating the downstream processing of the U:G mismatches it creates, likely by a physically interacting with CSR specific factors.

Conclusions

DNA damage DNA DSB breaks and chromosome translocations are events fundamental to the genesis of cancer. The adaptive immune system induces these genetic insults as a byproduct of diversifying immune receptors. Even though both T and B cells undergo RAG1/2 V(D)J recombination lymphomas are overwhelmingly of B cell origin. The DSB and mutations created by AID during SHM and CSR *Ig* diversification are major contributors to this prevalence. Recent studies have revealed a surprisingly wide array of AID target genes including oncogenes frequently involved in lymphoid cancer (Chiarle, et al., 2011; Klein, et al., 2011b; Yamane, et al., 2011). AID targeting by Spt5 to stalled Pol II gives major mechanistic insight into the basis of genome wide damage. However, AID action is highly focused at the *Ig* loci compared to elsewhere in the genome. This differential of activity could involve preferential targeting, Pol II stalling, and the stabilization of the AID complex to the *Ig* region. AID post-translational modifications and interacting factors are likely players and future studies addressing these should lead to a more comprehensive picture of AID targeting. During CSR, AID induced U:G mismatches in the switch regions triggers DSBs and DNA recombination. However, U:G mismatches also occur through spontaneously or radiation induced deamination of cytidines. These occur not infrequently in the genome, but are normally faithfully repaired (Krokan et al., 2002). Whether a mechanism governs preferential processing of AID lesions to DSB and mutations is still unclear. However, AID lacking the C-terminal target switch regions does not support CSR suggesting that specific domains of AID influence the resolution of lesions. Overall, the wide targeting of AID poses a significant threat to genome stability. Therefore mechanisms

that regulate AID targeting, activity and U:G mismatch processing represent important means to suppress pro-oncogenic damage.

Acknowledgments

We would like to thank Alex Scheinker for his assistance in rendering figures 3 and 4 .

References

- Abbas, AK.; Lichtman, AH.; Pillai, S. Cellular and Molecular Immunology. Elsevier; Philadelphia: 2010.
- Adelman K, Marr MT, Werner J, Saunders A, Ni Z, Andrulis ED, Lis JT. Mol Cell. 2005; 17:103–12. [PubMed: 15629721]
- Andrulis ED, Werner J, Nazarian A, Erdjument-Bromage H, Tempst P, Lis JT. Nature. 2002; 420:837–41. [PubMed: 12490954]
- Aoufouchi S, Faili A, Zober C, D'Orlando O, Weller S, Weill JC, Reynaud CA. J Exp Med. 2008; 205:1357–68. [PubMed: 18474627]
- Bachl J, Carlson C, Gray-Schopfer V, Dessing M, Olsson C. J Immunol. 2001; 166:5051–7. [PubMed: 11290786]
- Barreto V, Reina-San-Martin B, Ramiro AR, McBride KM, Nussenzweig MC. Mol Cell. 2003; 12:501–8. [PubMed: 14536088]
- Basu U, Chaudhuri J, Alpert C, Dutt S, Ranganath S, Li G, Schrum JP, Manis JP, Alt FW. Nature. 2005; 438:508–11. [PubMed: 16251902]
- Basu U, Meng FL, Keim C, Grinstein V, Pefanis E, Eccleston J, Zhang T, Myers D, Wasserman CR, Wesemann DR, Januszyk K, Gregory RI, Deng H, Lima CD, Alt FW. Cell. 2011; 144:353–63. [PubMed: 21255825]
- Bentley DL, Groudine M. Nature. 1986; 321:702–6. [PubMed: 3520340]
- Besmer E, Market E, Papavasiliou FN. Mol Cell Biol. 2006; 26:4378–85. [PubMed: 16705187]
- Betz AG, Milstein C, Gonzalez-Fernandez A, Pannell R, Larson T, Neuberger MS. Cell. 1994; 77:239–48. [PubMed: 8168132]
- Both GW, Taylor L, Pollard JW, Steele EJ. Mol Cell Biol. 1990; 10:5187–96. [PubMed: 2118991]
- Branstetter R, Pham P, Scharff MD, Goodman MF. Proc Natl Acad Sci U S A. 2003; 100:4102–7. [PubMed: 12651944]
- Brar SS, Watson M, Diaz M. J Biol Chem. 2004; 279:26395–401. [PubMed: 15087440]
- Chatterji M, Unniraman S, McBride KM, Schatz DG. J Immunol. 2007; 179:5274–80. [PubMed: 17911613]
- Chaudhuri J, Alt FW. Nat Rev Immunol. 2004; 4:541–52. [PubMed: 15229473]
- Chaudhuri J, Tian M, Khuong C, Chua K, Pinaud E, Alt FW. Nature. 2003; 422:726–30. [PubMed: 12692563]
- Chiarle R, Zhang Y, Frock RL, Lewis SM, Molinie B, Ho YJ, Myers DR, Choi VW, Compagno M, Malkin DJ, Neuberger D, Monti S, Giallourakis CC, Gostissa M, Alt FW. Cell. 2011; 147:107–19. [PubMed: 21962511]
- Cramer P, Armache KJ, Baumli S, Benkert S, Brueckner F, Buchen C, Damsma GE, Dengl S, Geiger SR, Jasiak AJ, Jawhari A, Jennebach S, Kamenski T, Kettenberger H, Kuhn CD, Lehmann E, Leike K, Sydow JF, Vannini A. Annu Rev Biophys. 2008; 37:337–52. [PubMed: 18573085]
- Crouch EE, Li Z, Takizawa M, Fichtner-Feigl S, Gourzi P, Montano C, Feigenbaum L, Wilson P, Janz S, Papavasiliou FN, Casellas R. J Exp Med. 2007; 204:1145–56. [PubMed: 17452520]
- de Yébenes VG, Belver L, Pisano DG, Gonzalez S, Villasante A, Croce C, He L, Ramiro AR. J Exp Med. 2008; 205:2199–206. [PubMed: 18762567]
- Dedeoglu F, Horwitz B, Chaudhuri J, Alt FW, Geha RS. Int Immunol. 2004; 16:395–404. [PubMed: 14978013]
- Delker RK, Fugmann SD, Papavasiliou FN. Nat Immunol. 2009; 10:1147–53. [PubMed: 19841648]

- Di Noia J, Neuberger M. *Annu Rev Biochem.* 2007; 76:1–22. [PubMed: 17328676]
- Dickerson SK, Market E, Besmer E, Papavasiliou FN. *J Exp Med.* 2003; 197:1291–6. [PubMed: 12756266]
- Doi T, Kato L, Ito S, Shinkura R, Wei M, Nagaoka H, Wang J, Honjo T. *Proc Natl Acad Sci U S A.* 2009; 106:2758–63. [PubMed: 19202055]
- Dorsett Y, McBride KM, Jankovic M, Gazumyan A, Thai TH, Robbani DF, Di Virgilio M, San-Martin BR, Heidkamp G, Schwickert TA, Eisenreich T, Rajewsky K, Nussenzweig MC. *Immunity.* 2008; 28:630–8. [PubMed: 18455451]
- Fugmann SD, Lee AI, Shockett PE, Villey IJ, Schatz DG. *Annu Rev Immunol.* 2000; 18:495–527. [PubMed: 10837067]
- Gazumyan A, Timachova K, Yuen G, Siden E, Di Virgilio M, Woo EM, Chait BT, Reina San-Martin B, Nussenzweig MC, McBride KM. *Mol Cell Biol.* 2011; 31:442–9. [PubMed: 21135131]
- Geisberger R, Rada C, Neuberger MS. *Proc Natl Acad Sci U S A.* 2009; 106:6736–41. [PubMed: 19351893]
- Gilmour DS. *Chromosoma.* 2009; 118:1–10. [PubMed: 18830703]
- Gonda H, Sugai M, Nambu Y, Katakai T, Agata Y, Mori KJ, Yokota Y, Shimizu A. *J Exp Med.* 2003; 198:1427–37. [PubMed: 14581609]
- Gordon MS, Kanegai CM, Doerr JR, Wall R. *Proc Natl Acad Sci U S A.* 2003; 100:4126–31. [PubMed: 12651942]
- Gostissa M, Yan CT, Bianco JM, Cogne M, Pinaud E, Alt FW. *Nature.* 2009; 462:803–7. [PubMed: 20010689]
- Guikema JE, Schrader CE, Leus NG, Ucher A, Linehan EK, Werling U, Edelmann W, Stavnezer J. *J Immunol.* 2008; 181:8450–9. [PubMed: 19050263]
- Houseley J, LaCava J, Tollervey D. *Nat Rev Mol Cell Biol.* 2006; 7:529–39. [PubMed: 16829983]
- Ito S, Nagaoka H, Shinkura R, Begum N, Muramatsu M, Nakata M, Honjo T. *Proc Natl Acad Sci U S A.* 2004; 101:1975–80. [PubMed: 14769937]
- Jeevan-Raj BP, Robert I, Heyer V, Page A, Wang JH, Cammas F, Alt FW, Losson R, Reina-San-Martin B. *J Exp Med.* 2011; 208:1649–60. [PubMed: 21746811]
- Jung D, Giallourakis C, Mostoslavsky R, Alt FW. *Annu Rev Immunol.* 2006; 24:541–70. [PubMed: 16551259]
- Kim JB, Sharp PA. *J Biol Chem.* 2001; 276:12317–23. [PubMed: 11145967]
- Klein BJ, Bose D, Baker KJ, Yusoff ZM, Zhang X, Murakami KS. *Proc Natl Acad Sci U S A.* 2011a; 108:546–50. [PubMed: 21187417]
- Klein IA, Resch W, Jankovic M, Oliveira T, Yamane A, Nakahashi H, Di Virgilio M, Bothmer A, Nussenzweig A, Robbani DF, Casellas R, Nussenzweig MC. *Cell.* 2011b; 147:95–106. [PubMed: 21962510]
- Kracker S, Imai K, Gardes P, Ochs HD, Fischer A, Durandy AH. *Proc Natl Acad Sci U S A.* 2010; 107:22225–30. [PubMed: 21135220]
- Krokan HE, Drablos F, Slupphaug G. *Oncogene.* 2002; 21:8935–48. [PubMed: 12483510]
- Kuppers R. *Nat Rev Cancer.* 2005; 5:251–62. [PubMed: 15803153]
- Lebecque SG, Gearhart PJ. *J Exp Med.* 1990; 172:1717–27. [PubMed: 2258702]
- Lis JT. *Nature.* 2007; 450:198–202. [PubMed: 17994086]
- Liu M, Duke JL, Richter DJ, Vinuesa CG, Goodnow CC, Kleinstein SH, Schatz DG. *Nature.* 2008; 451:841–845. [PubMed: 18273020]
- Liu M, Schatz DG. *Trends Immunol.* 2009; 30:173–81. [PubMed: 19303358]
- Lumsden JM, McCarty T, Petiniot LK, Shen R, Barlow C, Wynn TA, Morse HC 3rd, Gearhart PJ, Wynshaw-Boris A, Max EE, Hodes RJ. *J Exp Med.* 2004; 200:1111–21. [PubMed: 15504820]
- Martinez-Rucobo FW, Sainsbury S, Cheung AC, Cramer P. *EMBO J.* 2011; 30:1302–10. [PubMed: 21386817]
- Maul RW, Gearhart PJ. *Adv Immunol.* 2010; 105:159–91. [PubMed: 20510733]
- Maul RW, Saribasak H, Martomo SA, McClure RL, Yang W, Vaisman A, Gramlich HS, Schatz DG, Woodgate R, Wilson DM 3rd, Gearhart PJ. *Nat Immunol.* 2011; 12:70–6. [PubMed: 21151102]

- McBride KM, Barreto V, Ramiro AR, Stavropoulos P, Nussenzweig MC. J Exp Med. 2004; 199:1235–44. [PubMed: 15117971]
- McBride KM, Gazumyan A, Woo EM, Barreto VM, Robbani DF, Chait BT, Nussenzweig MC. Proc Natl Acad Sci U S A. 2006; 103:8798–803. [PubMed: 16723391]
- McBride KM, Gazumyan A, Woo EM, Schwickert TA, Chait BT, Nussenzweig MC. J Exp Med. 2008; 205:2585–94. [PubMed: 18838546]
- McKean D, Huppi K, Bell M, Staudt L, Gerhard W, Weigert M. Proc Natl Acad Sci U S A. 1984; 81:3180–4. [PubMed: 6203114]
- Muramatsu M, Kinoshita K, Fagarasan S, Yamada S, Shinkai Y, Honjo T. Cell. 2000; 102:553–63. [PubMed: 11007474]
- Muramatsu M, Sankaranand VS, Anant S, Sugai M, Kinoshita K, Davidson NO, Honjo T. J Biol Chem. 1999; 274:18470–6. [PubMed: 10373455]
- Muschen M, Re D, Jungnickel B, Diehl V, Rajewsky K, Kuppers R. J Exp Med. 2000; 192:1833–40. [PubMed: 11120779]
- Nambu Y, Sugai M, Gonda H, Lee CG, Katakai T, Agata Y, Yokota Y, Shimizu A. Science. 2003; 302:2137–40. [PubMed: 14684824]
- Nowak U, Matthews AJ, Zheng S, Chaudhuri J. Nat Immunol. 2011; 12:160–6. [PubMed: 21186367]
- Nussenzweig A, Nussenzweig MC. Cell. 2010; 141:27–38. [PubMed: 20371343]
- Okazaki IM, Hiai H, Kakazu N, Yamada S, Muramatsu M, Kinoshita K, Honjo T. J Exp Med. 2003; 197:1173–81. [PubMed: 12732658]
- Okazaki IM, Okawa K, Kobayashi M, Yoshikawa K, Kawamoto S, Nagaoka H, Shinkura R, Kitawaki Y, Taniguchi H, Natsume T, Iemura S, Honjo T. Proc Natl Acad Sci U S A. 2011; 108:7920–5. [PubMed: 21518874]
- Oliveira T, Resch W, Jankovic M, Casellas R, Nussenzweig MC, Klein IA. J Immunol Methods. 2011
- Orthwein A, Patenaude AM, Affar el B, Lamarre A, Young JC, Di Noia JM. J Exp Med. 2010; 207:2751–65. [PubMed: 21041454]
- Pasqualucci L, Kitaura Y, Gu H, Dalla-Favera R. Proc Natl Acad Sci U S A. 2006; 103:395–400. [PubMed: 16387847]
- Pasqualucci L, Migliazza A, Fracchiolla N, William C, Neri A, Baldini L, Chaganti RS, Klein U, Kuppers R, Rajewsky K, Dalla-Favera R. Proc Natl Acad Sci U S A. 1998; 95:11816–21. [PubMed: 9751748]
- Pasqualucci L, Neumeister P, Goossens T, Nanjangud G, Chaganti RS, Kuppers R, Dalla-Favera R. Nature. 2001; 412:341–6. [PubMed: 11460166]
- Patenaude AM, Orthwein A, Hu Y, Campo VA, Kavli B, Buschiazzi A, Di Noia JM. Nat Struct Mol Biol. 2009; 16:517–27. [PubMed: 19412186]
- Pavri R, Gazumyan A, Jankovic M, Di Virgilio M, Klein I, Ansarah-Sobrinho C, Resch W, Yamane A, San-Martin BR, Barreto V, Nieland TJ, Root DE, Casellas R, Nussenzweig MC. Cell. 2010; 143:122–33. [PubMed: 20887897]
- Peled JU, Kuang FL, Iglesias-Ussel MD, Roa S, Kalis SL, Goodman MF, Scharff MD. Annu Rev Immunol. 2008; 26:481–511. [PubMed: 18304001]
- Peters A, Storb U. Immunity. 1996; 4:57–65. [PubMed: 8574852]
- Petersen-Mahrt SK, Harris RS, Neuberger MS. Nature. 2002; 418:99–103. [PubMed: 12097915]
- Pham P, Bransteitter R, Petruska J, Goodman MF. Nature. 2003; 424:103–7. [PubMed: 12819663]
- Pham P, Smolka MB, Calabrese P, Landolph A, Zhang K, Zhou H, Goodman MF. J Biol Chem. 2008; 283:17428–39. [PubMed: 18417471]
- Potter M. Immunol Rev. 2003; 194:177–95. [PubMed: 12846815]
- Rada C, Jarvis JM, Milstein C. Proc Natl Acad Sci U S A. 2002; 99:7003–8. [PubMed: 12011459]
- Rahl PB, Lin CY, Seila AC, Flynn RA, McCuine S, Burge CB, Sharp PA, Young RA. Cell. 2010; 141:432–45. [PubMed: 20434984]
- Rajagopal D, Maul RW, Ghosh A, Chakraborty T, Khamlichi AA, Sen R, Gearhart PJ. J Exp Med. 2009; 206:1237–44. [PubMed: 19433618]
- Rajewsky K. Nature. 1996; 381:751–8. [PubMed: 8657279]

- Ramiro AR, Jankovic M, Callen E, Difilippantonio S, Chen HT, McBride KM, Eisenreich TR, Chen J, Dickins RA, Lowe SW, Nussenzweig A, Nussenzweig MC. *Nature*. 2006; 440:105–9. [PubMed: 16400328]
- Ramiro AR, Jankovic M, Eisenreich T, Difilippantonio S, Chen-Kiang S, Muramatsu M, Honjo T, Nussenzweig A, Nussenzweig MC. *Cell*. 2004; 118:431–8. [PubMed: 15315756]
- Ramiro AR, Stavropoulos P, Jankovic M, Nussenzweig MC. *Nat Immunol*. 2003; 4:452–6. [PubMed: 12692548]
- Ranjit S, Khair L, Linehan EK, Ucher AJ, Chakrabarti M, Schrader CE, Stavnezer J. *J Immunol*. 2011; 187:2464–75. [PubMed: 21804017]
- Raschke EE, Albert T, Eick D. *J Immunol*. 1999; 163:4375–82. [PubMed: 10510378]
- Revy P, Muto T, Levy Y, Geissmann F, Plebani A, Sanal O, Catalan N, Forveille M, Dufourcq-Labelouse R, Gennery A, Tezcan I, Ersoy F, Kayserili H, Ugazio AG, Brousse N, Muramatsu M, Notarangelo LD, Kinoshita K, Honjo T, Fischer A, Durandy A. *Cell*. 2000; 102:565–75. [PubMed: 11007475]
- Robbiani DF, Bothmer A, Callen E, Reina-San-Martin B, Dorsett Y, Difilippantonio S, Bolland DJ, Chen HT, Corcoran AE, Nussenzweig A, Nussenzweig MC. *Cell*. 2008; 135:1028–38. [PubMed: 19070574]
- Robbiani DF, Bunting S, Feldhahn N, Bothmer A, Camps J, Deroubaix S, McBride KM, Klein IA, Stone G, Eisenreich TR, Ried T, Nussenzweig A, Nussenzweig MC. *Mol Cell*. 2009; 36:631–41. [PubMed: 19941823]
- Rothenfluh HS, Taylor L, Bothwell AL, Both GW, Steele EJ. *Eur J Immunol*. 1993; 23:2152–9. [PubMed: 8370398]
- Rucci F, Cattaneo L, Marrella V, Sacco MG, Sobacchi C, Lucchini F, Nicola S, Della Bella S, Villa ML, Imberti L, Gentili F, Montagna C, Tiveron C, Tatangelo L, Facchetti F, Vezzoni P, Villa A. *Gene*. 2006; 377:150–8. [PubMed: 16787714]
- Sayegh CE, Quong MW, Agata Y, Murre C. *Nat Immunol*. 2003; 4:586–93. [PubMed: 12717431]
- Sernandez IV, de Yebenes VG, Dorsett Y, Ramiro AR. *PLoS One*. 2008; 3:e3927. [PubMed: 19079594]
- Shen HM, Peters A, Baron B, Zhu X, Storb U. *Science*. 1998; 280:1750–2. [PubMed: 9624052]
- Shinkura R, Ito S, Begum NA, Nagaoka H, Muramatsu M, Kinoshita K, Sakakibara Y, Hijikata H, Honjo T. *Nat Immunol*. 2004; 5:707–12. [PubMed: 15195091]
- Sohail A, Klapacz J, Samaranayake M, Ullah A, Bhagwat AS. *Nucleic Acids Res*. 2003; 31:2990–4. [PubMed: 12799424]
- Staszewski O, Baker RE, Ucher AJ, Martier R, Stavnezer J, Guikema JE. *Mol Cell*. 2011; 41:232–42. [PubMed: 21255732]
- Stavnezer J, Guikema JE, Schrader CE. *Annu Rev Immunol*. 2008; 26:261–92. [PubMed: 18370922]
- Steele EJ, Rothenfluh HS, Both GW. *Immunol Cell Biol*. 1992; 70(Pt 2):129–44. [PubMed: 1398773]
- Swanson MS, Malone EA, Winston F. *Mol Cell Biol*. 1991; 11:4286. [PubMed: 2072920]
- Ta VT, Nagaoka H, Catalan N, Durandy A, Fischer A, Imai K, Nonoyama S, Tashiro J, Ikegawa M, Ito S, Kinoshita K, Muramatsu M, Honjo T. *Nat Immunol*. 2003; 4:843–8. [PubMed: 12910268]
- Takizawa M, Tolarova H, Li Z, Dubois W, Lim S, Callen E, Franco S, Mosaico M, Feigenbaum L, Alt FW, Nussenzweig A, Potter M, Casellas R. *J Exp Med*. 2008
- Teng G, Hakimpour P, Landgraf P, Rice A, Tuschl T, Casellas R, Papavasiliou FN. *Immunity*. 2008; 28:621–9. [PubMed: 18450484]
- Teng G, Papavasiliou FN. *Annu Rev Genet*. 2007; 41:107–20. [PubMed: 17576170]
- Tran TH, Nakata M, Suzuki K, Begum NA, Shinkura R, Fagarasan S, Honjo T, Nagaoka H. *Nat Immunol*. 2010; 11:148–54. [PubMed: 19966806]
- Tsai AG, Lieber MR. *BMC Genomics*. 2010; 11 Suppl 1:S1. [PubMed: 20158866]
- Tumas-Brundage K, Manser T. *J Exp Med*. 1997; 185:239–50. [PubMed: 9016873]
- Vuong BQ, Lee M, Kabir S, Irimia C, Macchiarulo S, McKnight GS, Chaudhuri J. *Nat Immunol*. 2009; 10:420–6. [PubMed: 19234474]

- Wada T, Takagi T, Yamaguchi Y, Ferdous A, Imai T, Hirose S, Sugimoto S, Yano K, Hartzog GA, Winston F, Buratowski S, Handa H. *Genes Dev.* 1998; 12:343–56. [PubMed: 9450929]
- Wang CL, Harper RA, Wabl M. *Proc Natl Acad Sci U S A.* 2004; 101:7352–6. [PubMed: 15123833]
- Wang L, Wuerffel R, Feldman S, Khamlichi AA, Kenter AL. *J Exp Med.* 2009; 206:1817–30. [PubMed: 19596805]
- Weber JS, Berry J, Manser T, Claflin JL. *J Immunol.* 1991; 146:3652–5. [PubMed: 1902859]
- Weigert MG, Cesari IM, Yonkovich SJ, Cohn M. *Nature.* 1970; 228:1045–7. [PubMed: 5483159]
- Xu Z, Fulop Z, Wu G, Pone EJ, Zhang J, Mai T, Thomas LM, Al-Qahtani A, White CA, Park SR, Steinacker P, Li Z, Yates J 3rd, Herron B, Otto M, Zan H, Fu H, Casali P. *Nat Struct Mol Biol.* 2010; 17:1124–35. [PubMed: 20729863]
- Yamada T, Yamaguchi Y, Inukai N, Okamoto S, Mura T, Handa H. *Mol Cell.* 2006; 21:227–37. [PubMed: 16427012]
- Yamaguchi Y, Takagi T, Wada T, Yano K, Furuya A, Sugimoto S, Hasegawa J, Handa H. *Cell.* 1999; 97:41–51. [PubMed: 10199401]
- Yamane A, Resch W, Kuo N, Kuchen S, Li Z, Sun HW, Robbani DF, McBride K, Nussenzweig MC, Casellas R. *Nat Immunol.* 2011; 12:62–9. [PubMed: 21113164]
- Yu K, Chedin F, Hsieh CL, Wilson TE, Lieber MR. *Nat Immunol.* 2003; 4:442–51. [PubMed: 12679812]
- Zhang Y, Gostissa M, Hildebrand DG, Becker MS, Boboila C, Chiarle R, Lewis S, Alt FW. *Adv Immunol.* 2010; 106:93–133. [PubMed: 20728025]
- Zhou C, Saxon A, Zhang K. *J Immunol.* 2003; 170:1887–93. [PubMed: 12574355]

FIG 1. Schematic of AID dependent CSR and SHM

Schematic representation of the IgH locus and rearrangement that takes place during CSR. Constant region exons depicted by solid rectangles, switch regions by solid ovals, promoters by black boxes, enhancer and 3' regulatory regions by white cylinders. Areas that accumulate DSB are denoted by discontinuous lines and mutations area denoted by black circles.

FIG 2. Depiction of reciprocal translocations

Schematic representation of normal chromosome structures or following a reciprocal translocation such as the *c-myc/IgH*. PCR primers (arrows) can be used to specifically amplify translocations for detection in Southern assays.

FIG 3. Model of AID within a Spt4/5 RNA polymerase complex

Rendering based on the crystal structure of a complex of Spt4 (red)/Spt5 (purple) with RNA clamp domain (orange) (Martinez-Rucobo, et al.). Melted DNA is loaded into the cleft to trigger RNA (red ribbon) synthesis on template DNA (blue ribbon). Spt5 domain locks nucleic acids and contacts the nontemplate DNA (green ribbon) in the cleft, preventing collapse of the transcription bubble. The interaction with Spt5 brings AID (pink) into the proximity with the both template and non-template ssDNA.

FIG 4. A model of AID transcription mediated targeting by stalled Pol II/Spt5 elongation complex

Stalling model for AID access based on refinement of a proposed association with Pol II (Peters and Storb, 1996).

(A) Elongation factors including Spt5/Spt4 (DSIF) (purple/red), and other factors (not shown) bound to RNA Pol II (grey) create the stalled state in the promoter proximal region near the transcription start site (TSS).

(B) AID, via interaction with Spt5 and association with stalled polymerase, gains access to single stranded non-template DNA.

(C) Exosome degrades 3' end of the nascent RNA exposed by backtracking of stalled polymerase and thus exposes the template DNA (blue) to AID.