Antibody diversity: one enzyme to rule them all

Michel C Nussenzweig & Frederick W Alt

Three reactions diversify antibody genes in human somatic cells of the B lineage: VDJ recombination, somatic hypermutation and class-switch recombination. The discovery of activation-induced cytidine deaminase (AID) has led to the elucidation of a unified molecular mechanism for initiation of the last two reactions and suggests why B cells undergoing these reactions are prone to cancer-associated DNA damage.

Hundreds of billions of different antibodies are produced to fight off invading pathogens. Two general classes of antigen receptor gene diversification processes in B lymphocytes account for this panopoly antigen-independent and antigen-dependent processes. Antigen-independent diversification assembles antibody genes from variable (V), diversity (D) and joining (J) gene segments during B-lymphocyte development in the bone marrow. VDJ recombination is therefore responsible for the initial antibody repertoire. The proteins that mediate VDJ recombination (RAG1 and RAG2) are closely related to transposases, and are thought to have entered the genome of vertebrate ancestors as part of a transposon. Antigen-dependent immunoglobulin gene diversification by somatic hypermutation (SHM) and class-switch recombination (CSR) occurs in mature B cells during immune responses in peripheral lymphoid tissues such as spleen and lymph node. SHM gives rise to point mutations and CSR gives rise to different antibody isotypes by recombination[1,2].

In 1999, Muramatsu *et al.* discovered a cDNA encoding a cytidine deaminase, AID, whose expression is specifically induced in activated B lymphocytes[3]. AID is absolutely required for SHM and CSR[4,5]. Moreover, AID mutation in humans is associated with absence of secondary antibody isotypes and SHM, and produces hyper-IgM syndrome, a disease associated with increased susceptibility to infections[5]. Not only is AID necessary for SHM and CSR, ectopically expressed AID is also sufficient to effect these processes, albeit with reduced rate and spectrum of activity, in diverse cell types[1,2]. Thus, AID is the long sought-after B cell–specific factor required for somatic diversification of antibody genes in mature B cells.

SHM and CSR are entirely different reactions, and so it was unexpected that a single enzyme would be essential for both. SHM (**Fig. 1**) introduces untemplated point mutations in the variable region of the immunoglobulin heavy- and light-chain genes[1]. These mutations are focused primarily over the variable region and occur at a frequency of up to 1 per 10^4 base pairs per cell division, many orders of magnitude higher than spontaneous mutation. Although most somatic mutations have little or no effect on antibody specificity, the few mutations that enhance binding affinity are strongly positively selected and B cells that carry such mutations make a disproportionate contribution to memory responses, thus accounting for the phenomenon of antibody affinity maturation.

CSR is a DNA recombination reaction specific to the immunoglobulin heavy chain (IgH) locus that increases antibody versatility by associating a given IgH variable region with one of several different constant regions, each with unique effector functions[2] (**Fig. 1**). CSR juxtaposes the expressed IgH VDJ variable region exon to a downstream set of constant region exons by a deletional recombination event.

An early clue to how AID might catalyze both reactions came from analysis of the predicted AID protein sequence, which showed that AID is related to APOBEC-1, an mRNA-editing cytosine deaminase[3,4]. Thus, AID might be a cytosine deaminase that edits mRNAs to produce proteins essential for initiating the CSR and SHM reactions. Although there is no direct evidence to support the RNA-editing hypothesis, the finding that *de novo* protein synthesis is required after AID expression during the CSR reaction is consistent with this possibility[6]. This observation, however, has alternative explanations, such as the existence of labile factors that are essential for CSR. Most importantly, an mRNA target for AID has not been identified and AID has not been shown to be capable of using RNA as a substrate *in vitro*.

A second hypothesis, that AID might initiate both SHM and CSR by deaminating cytidine residues in immunoglobulin DNA directly[4,7] is strongly supported by biochemical and genetic experiments. The idea that AID is a DNA cytidine deaminase predicts that this enzyme creates U:G mismatches in DNA and that such mismatches are processed by alternative DNA repair pathways to produce SHM or CSR[7]. Biochemical experiments show that AID binds to and deaminates ssDNA but not dsDNA or RNA *in vitro* and that the ssDNA target for AID is exposed by transcription *in vitro* and in *E. coli*[8–12]. The requirement for transcription *in vitro* and in *E. coli* mirrors the requirement for transcription in SHM and CSR in mice and humans; the ssDNA exposed by transcription may explain this requirement. In addition, the mutations produced in the *in vitro* reactions and in *E. coli* show a preference for WRC motifs (W = A or T; R = A or G)[1] similar to that found in authentic SHM (Ref. 1).

Direct support for the DNA deamination models has been provided by assays that show that AID physically interacts with switch regions in cells stimulated to undergo CSR[13,14]. Finally, biochemical studies show that AID is phosphorylated in activated B cells and thereby associates with replication protein A (RPA), a ssDNA–binding protein. The AID-RPA complex preferentially binds to and deaminates transcribed DNA containing WRC motifs, enhancing AID activity and providing a degree of specificity[14]. The biochemistry strongly suggests that AID is a direct DNA mutator that attacks cytidine residues in ssDNA.

Genetic experiments strengthen the deoxycytidine deamination hypothesis and have begun to flesh out the machinery that takes over once AID damages DNA.These experiments confirm that SHM and CSR are dependent on recognition and processing of U:G mismatches in DNA[7]. U:G mismatches are recognized by either uracyl DNA glycosylase or the mismatch recognition and repair system. During SHM and CSR these enzymes are proposed to process U:G mismatches by alternative repair pathways to produce dsDNA breaks, base substitutions or a distal mutation[1,2]. Deletion

Michael C. Nussenzweig is in the Laboratory of Molecular Immunology and the Howard Hughes Medical Institute, The Rockefeller University, New York, New York 10021, USA.
Frederick W. Alt is at the Howard Hughes Medical Institute, The Children's Hospital, and The Center for Blood Research and Department of Genetics, Harvard Medical School, Boston, Massachusetts 02115, USA.
e-mail: alt@enders.tch.harvard.edu or nussen@mail.rockefeller.edu

of either uracil DNA glycosylase or mismatch repair enzymes (UNG and MSH2) alters the pattern of mutation and interferes with CSR, suggesting that both pathways can effect these reactions downstream of deamination[1,2]. In humans, mutations in UNG result in a hyper-IgM syndrome clinically indistinguishable from AID deficiency[15]. When both UNG and MSH2 are absent, the cell does not recognize the U:G mismatches produced by AID and mutation is limited to C to T changes[16]. Furthermore, in these double mutants, CSR is completely abolished, presumably because there are no additional pathways to process the AID-dependent U:G mismatch[16]. These double-mutant data offer particularly compelling support for the DNA deamination mechanism and for the idea that AID initiates both SHM and CSR by producing U:G mismatches in DNA.

How do the point mutations induced by AID initiate the complex recombination event that is CSR? During CSR, DNA lesions catalyzed by AID are processed to produce double-strand breaks that are recognized by the cellular DNA damage–signaling pathway including ATM, H2AX and 53BP1; repair of the breaks requires components of the nonhomologous end–joining pathway of DNA repair[2,17]. In the absence of these factors, CSR is absent or inefficient[2,17]. In contrast, SHM does not require these factors.

In addition to their role in CSR, ATM, H2AX and 53BP1 are essential for maintaining genomic stability[17,18]. It is believed that these factors mark the site of a DNA lesion for repair and that they may hold broken chromosome ends together. AID is required for accumulation of phosphorylated H2AX and other dsDNA damage signaling factors in foci along the immunoglobulin locus during CSR[19] and their role might be to facilitate accurate synapsis of switch regions by bringing together the dsDNA breaks produced by AID[17,18,20].

B lymphocytes are particularly prone to malignant transformation, and DNA damage incurred during CSR or SHM may account for this cancer susceptibility. Such collateral damage might be the result of mutation of nontarget genes by AID or to abnormal repair of AID-induced DNA damage in physiological targets. Experimental support for bystander gene mutation by AID comes from the finding that in human B cells proto-oncogenes, including *BCL6*, accumulate mutations, and that genes other than antibody genes can be mutated by AID expression[21–23]. In addition, transgenic overexpression of AID in mice leads to T-cell lymphomas and SHM of the *Myc* oncogene[24]. Abnormal repair of physiologically targeted AID-induced lesions is implicated because AID is required for accumulation of chromosome translocations in mice overexpressing IL-6 or injected with mineral oil[25,26].

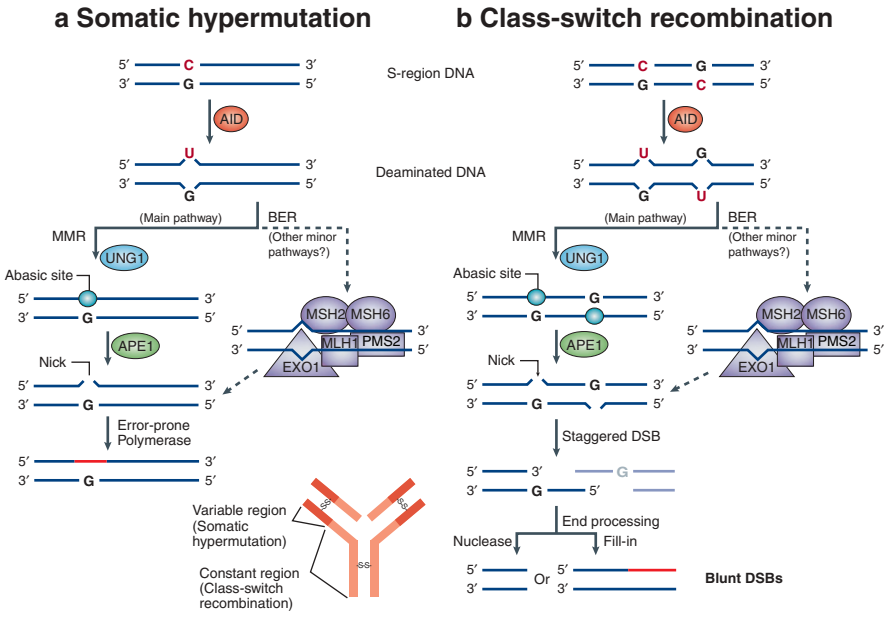

Figure 1 Antibody diversification. Activation-induced cytidine deaminase (AID) deaminates cytidine residues in DNA, converting them to uridine residues. The U:G mismatch can then be processed by either uracil DNA glycosylase (UNG), a component of the base excision repair pathway, or the mismatch-repair machinery (MSH1, MSH6, EXO1, MLH1 and PMS)—resulting in gaps or nicks in DNA. (**a**) During somatic hypermutation the U:G mismatch can simply be replicated to produce a C to T mutation. Alternatively processing the nick by UNG and the mismatch repair machinery can produce an abasic site that will produce a C to A or C or T change; alternatively, a gap can be filled in by error-prone polymerases to produce mutations in nucleotides other than the targeted C. (**b**) During class-switch recombination the nicks induced by the BER pathway are thought to be generated by the following process: UNG removes the AID-introduced deoxyuridine in S-region DNA, creating an abasic site that is processed by the apurine/apyrimidine endonuclease 1 (APE1), which creates the nick. Processing of the staggered ends by unknown exonucleases or by error-prone end-filling reactions can lead to blunt double-stranded breaks that can be ligated to similarly created breaks on downstream S-region DNA to complete class-switch recombination. Figure modified from ref. 2.

The discovery of AID has allowed the formulation of a unifying mechanism for CSR and SHM initiated by programmed DNA damage in ssDNA exposed during transcription. This type of damage is normally restricted to genes that encode antibodies, and its repair leads to diversification by SHM and CSR. The mechanisms that provide for the relative specificity of AID for such genes are under investigation, but appear to include specific transcription-control elements and cofactors such as RPA; this specificity is not absolute and aberrant damage of genes that do not encode immunoglobulins and abnormal repair of AID-induced lesions may account for the propensity of B cells to undergo malignant transformation.

1. Li, Z., Woo, C.J., Iglesias-Ussel, M.D., Ronai, D. & Scharff, M.D. *Genes Dev.* **18**, 1–11 (2004).
2. Chaudhuri, J. & Alt, F.W. *Nat. Rev. Immunol.* **4**, 541–552 (2004).
3. Muramatsu, M. *et al. J. Biol. Chem.* **274**, 18470–18476 (1999).
4. Muramatsu, M. *et al. Cell* **102**, 553–563 (2000).
5. Revy, P. *et al. Cell* **102**, 565–575 (2000).
6. Doi, T., Kinoshita, K., Ikegawa, M., Muramatsu, M. & Honjo, T. *Proc. Natl. Acad. Sci. USA* **100**, 2634–2638 (2003).
7. Petersen-Mahrt, S.K., Harris, R.S. & Neuberger, M.S. AID *Nature* **418**, 99–103 (2002).
8. Bransteitter, R., Pham, P., Scharff, M.D. & Goodman, M.F. *Proc. Natl. Acad. Sci.* **100**, 4102–4107 (2003).
9. Chaudhuri, J. *et al. Nature* **422**, 726–730 (2003).
10. Dickerson, S.K., Market, E., Besmer, E. & Papavasiliou, F.N. *J. Exp. Med.* **197**, 1291–1296 (2003).
11. Ramiro, A.R., Stavropoulos, P., Jankovic, M. & Nussenzweig, M.C. *Nat. Immunol.* **4**, 452–456 (2003).
12. Pham, P., Bransteitter, R., Petruska, J. & Goodman, M.F. *Nature* **424**, 103-7 (2003).
13. Nambu, Y. *et al. Science* **302**, 2137–2140 (2003).
14. Chaudhuri, J., Khuong, C. & Alt, F.W. *Nature* **430**, 992–998 (2004).
15. Imai, K. *et al. Nat. Immunol.* **4**, 1023–1028 (2003).
16. Rada, C., Di Noia, J.M. & Neuberger, M.S. *Mol. Cell* **16**, 163–171 (2004).
17. Fernandez-Capetillo, O., Lee, A., Nussenzweig, M. & Nussenzweig, A. *DNA Repair* **3**, 959–967 (2004).
18. Bassing, C.H. & Alt, F.W. *DNA Repair (Amst)* **3**, 781–796 (2004).
19. Petersen, S. *et al. Nature* **414**, 660–665 (2001).
20. Reina-San-Martin, B. *et al. J. Exp. Med.* **197**, 1767–1778 (2003).
21. Pasqualucci, L. *et al. Nature* **412**, 341–346 (2001).
22. Shen, H.M., Peters, A., Baron, B., Zhu, X. & Storb, U. *Science* **280**, 1750–1752 (1998).
23. Müschen, M. *et al. J. Exp. Med.* **192**, 1833–1840 (2000).
24. Okazaki, I.M. *et al. J. Exp. Med.* **197**, 1173–1181 (2003).
25. Unniraman, S., Zhou, S. & Schatz, D.G. *Nat. Immunol.* **5**, 1117–1123 (2004).
26. Ramiro, A.R. *et al. Cell* **118**, 431–438 (2004).