Chapter 1

Introduction

F. Nina Papavasiliou[1] *and Janet Stavnezer*[2]

[1]*The Rockefeller University*
New York, NY 10021, USA
E-mail: papavasiliou@rockefeller.edu

[2]*University of Massachusetts Medical School*
Worcester, MA 01655, USA
E-mail: Janet.Stavnezer@umassmed.edu

Biological information is coded in the base sequence of DNA and/or RNA. It follows that the fidelity of this information is meticulously preserved during its replication, transcription and maintenance, particularly in higher organisms; alterations at the level of genome or transcriptome can have dramatic downstream functional implications. Unintended sequence changes often lead to deleterious consequences. However, despite the tremendous pressure to guard against such effects, many biological systems have developed targeted mechanisms which alter DNA or RNA sequences and their corresponding information content. Though such sequence modification pathways have diverse roles throughout biology, many provide important host defense functions in innate and adaptive immunity.

Programmed sequence alterations that change genomic DNA or the genetic meaning of a genomically-encoded transcript have been termed "editing" (Grosjean *et al.*, 2004). To date, the only known enzymatic activities involved in polynucleotide editing catalyze either the deamination of cytidine to uridine in tRNA, mRNA or DNA (Conticello *et al.*, 2007) or the deamination of adenosine to inosine in tRNA or

DNA DEAMINATION AND THE IMMUNE SYSTEM - AID in Health and Disease
© Imperial College Press
http://www.worldscibooks.com/lifesci/p729.html

mRNA (Hamilton *et al.*, 2010). This book focuses on the central role of activation-induced cytidine deaminase (AID) in establishing the genetic diversity required for an effective humoral adaptive immune response by editing DNA at immunoglobulin (Ig) loci.

1.1 Discovery of AID

The technical catalyst for the discovery of AID was the generation, in 1996, of the B lymphocyte cell line, CH12F3 (Nakamura *et al.*, 1996). Derived from the CH12.LX lymphoma cell line, CH12F3 was selected to undergo class switch recombination (CSR) at a high frequency and exclusively to the isotype IgA upon stimulation with IL-4, TGFβ and CD40L. Theorizing that a specific recombinase was responsible for CSR, Muramatsu and Honjo (Muramatsu *et al.*, 1999) applied a PCR-based subtraction method to screen genes upregulated upon stimulation of CH12F3 cells for class-switching. Among the four novel genes discovered, AID proved to be especially interesting because of: its germinal center B cell restriction; its homology to the APOBEC family of RNA cytidine deaminases; and its *in vitro* deaminase activity, unique from that of APOBEC–1. Furthermore, AID-deficient mice were unable to undergo CSR and, surprisingly, also somatic hypermutation (SHM; Muramatsu *et al.*, 2000), underscoring the central role of the protein in Ig diversification reactions. Concurrently, through the use of standard human genetics, the Durandy laboratory independently identified AID as the gene responsible for HIGM in a subset of patients with an autosomal form of hyper-IgM syndrome (henceforth named type II, or HIGM2) who had a severe CSR defect and lacked somatically-mutated immunoglobulin genes (Revy *et al.*, 2000).

Initial sequence comparison of AID revealed that it possessed a cytidine deaminase domain with homology to the only well-characterized RNA deaminase at the time, which was a protein termed Apolipoprotein-B mRNA editing catalytic polypeptide–1 (APOBEC–1; Muramatsu *et al.*, 1999). APOBEC–1 was known to specifically edit the mRNA of apolipoprotein-B, converting a glutamine to a stop codon and thereby generating two distinct apoB isoforms with different functions

DNA DEAMINATION AND THE IMMUNE SYSTEM - AID in Health and Disease
© Imperial College Press
http://www.worldscibooks.com/lifesci/p729.html

(Wedekind *et al.*, 2003). Based on this homology, it was initially hypothesized that AID also edited a single mRNA that functioned in both SHM and CSR. Though this was an entirely reasonable proposition at the time, a wealth of additional evidence over the years has clearly tipped the balance toward the notion that AID directly edits DNA at the Ig locus.

1.2 Current Model of AID Function

Experimental support for the hypothesis that AID is a DNA deaminase was first provided by showing that ectopic expression of AID results in mutation of the *E. coli* genome (Petersen-Mahrt *et al.*, 2002). As it is unlikely that AID would edit the same mRNA in prokaryotic and eukaryotic cells to generate a novel DNA mutator, the simplest interpretation of these data is that AID is a *bona fide* DNA mutator, and as such, the first member of a family of polynucleotide deaminases that act on DNA. Other studies demonstrated that ectopic expression of AID was able to mutate the genome of a number of mammalian cell types (plasma cells, HEK–293T cells, NIH 3T3 cells; Martin *et al.*, 2002; Yoshikawa *et al.*, 2002), and also of yeast (Mayorov *et al.*, 2005; Poltoratsky *et al.*, 2004).

Further experimental support for the notion that AID is a DNA mutator emerged from studies focused on an important intermediate – U:G mismatches. Specifically, the Neuberger laboratory conducted experiments to study the role of uracil DNA glycosylase (UNG) in SHM and CSR. UNG is the major glycosylase that removes uracil from DNA in the context of base excision repair. Thus, if AID does indeed catalyze the formation of uracil in DNA, UNG would be expected to be central to the resolution of U:G mismatches. Indeed, Rada and colleagues (Rada *et al.*, 2002) found that UNG-deficient animals mutated their Ig locus at rates identical to those of their wild-type littermates. However, the spectra of mutations that accumulated at Ig sequences were very different. Specifically, mutations at G and C were strongly biased towards C to T and G to A events, as a result of direct replication of such mismatches, and the mutations observed at A and T bases were similar

DNA DEAMINATION AND THE IMMUNE SYSTEM - AID in Health and Disease
© Imperial College Press
http://www.worldscibooks.com/lifesci/p729.html

for both *ung*$^{-/-}$ and wild-type mice. In addition to its importance for SHM, Rada and colleagues found that UNG was central to CSR. In the absence of UNG, CSR is nearly abrogated (Rada *et al.*, 2004). The notion that UNG is central to SHM and CSR was further bolstered by experiments in the Durandy laboratory, which found that a subset of HIGM patients carried mutations in their UNG genes (Imai *et al.*, 2003).

Finally, a number of groups have studied the biochemical activities of purified AID from several expression systems (activated B cells, baculovirus-infected insect cells, recombinant expression in bacteria; Bransteitter *et al.*, 2003; Chaudhuri *et al.*, 2003; Dickerson *et al.*, 2003) and shown that AID exhibits activity only on single-stranded DNA substrates (naked DNA; Bransteitter *et al.*, 2003; Dickerson *et al.*, 2003), transcribed double-stranded DNA (Besmer *et al.*, 2006; Chaudhuri *et al.*, 2003) or transcribed DNA complexed with nucleosomes (Shen *et al.*, 2009), but not on non-transcribed double-stranded DNA. In addition, the local sequence preference of AID *in vitro*, namely that the WRC motif is an AID activity hotspot (Pham *et al.*, 2003), coincides well with SHM hotspots observed *in vivo*: changing the AID coding sequence results in concomitant changes in DNA hotspot motif preferences (Wang *et al.* 2010).

1.3 Open Questions

AID has been shown to be an active DNA mutator in a number of settings. It follows that in the cell AID must be meticulously regulated, and this is indeed the case: AID is regulated transcriptionally, post-transcriptionally and post-translationally (Ramiro and Di Noia discuss AID regulation in Chapter 7).

An additional mode of regulation appears to be the targeting of the molecule to the Ig locus, and locus-specific elements important for SHM/CSR targeting have been described (and are discussed by Dunnick and Fugmann in Chapter 3). Chromatin modifications important for SHM/CSR targeting have also been described, as well as the presence of local sequence features (Gearhart and Kenter, Chapter 2).

DNA DEAMINATION AND THE IMMUNE SYSTEM - AID in Health and Disease
© Imperial College Press
http://www.worldscibooks.com/lifesci/p729.html

It is important to distinguish between AID targeting to the Ig loci and repair targeting that eventually results in SHM/CSR. AID could be directly targeted to the Ig loci via a factor or factors that tether it to each locus (discussed by Reina-San-Martin and Chaudhuri in Chapter 4). A plausible non-mutually exclusive alternative is that faulty resolution of uracil lesions is mostly unique to the Ig loci. Krijger *et al.* will discuss error-prone and error-free lesion resolution in SHM (Chapter 6); CSR lesion resolution is the topic for Yu and Lieber (Chapter 5). Experiments in mice (Rada *et al.*, 2004) and zebrafish (Rai *et al.*, 2008) have demonstrated a genetic association between AID and UNG as well as AID and thymine DNA glycosylase (TDG): i.e. whereas UNG is clearly downstream of AID in a pathway that results in error-prone repair at the Ig locus, a genetic association of TDG with AID appears to be required for active CpG demethylation in zebrafish, a process which is error-free. Though there is no evidence thus far that AID directly interacts with either of those molecules, it is intriguing that the AID:UNG pathway results in error-prone repair, in contrast to the AID:TDG pathway for which the associated outcome is thought to be faithful repair of uracil lesions. The intriguing possibility therefore remains that AID itself selects between error-prone and error-free repair, possibly through the acquisition of cell- or locus-specific interaction partners, or cell- or locus-specific post-translational modifications.

Finally, lack of AID clearly causes immune deficiency. Conversely, AID overexpression (or ectopic expression) is also thought to be causal for a number of cancers of many different tissues. It is not hard to imagine that off-target action of AID in the germinal center can cause mutations and eventually even cancer-causing translocations (such as the well-studied *IgH:c-myc*). Indeed, mutations due to AID activity have been documented at a number of oncogenes and tumor-suppressors isolated from germinal center cells (e.g. *p53, c-myc, pim1, bcl-6*: Liu *et al.*, 2008; Pasqualucci *et al.*, 2001; Shen *et al.*, 1998). The link between AID expression and occurrence of B cell lymphoma is strong, and is discussed by Willmann *et al.* (Chapter 8), as well as by Diaz *et al.* (Chapter 9). Curiously, AID has also been detected in a large number of different tumors (e.g. breast, prostate, bone marrow-derived lymphomas, hepatocellular carcinomas, gastric cancers; Chiba and

DNA DEAMINATION AND THE IMMUNE SYSTEM - AID in Health and Disease
© Imperial College Press
http://www.worldscibooks.com/lifesci/p729.html

Marusawa, 2009), where it appears to be not only central to oncogenic transformation (Pasqualucci *et al.*, 2008), but also to tumor relapse after initial therapy (Chiba and Marusawa, 2009; Feldhahn *et al.*, 2007; Mullighan *et al.*, 2008).

1.4 A Unifying Model for AID Function

How can all these observations be reconciled to our current knowledge of AID and its role in the immune system? A clue toward a "unifying model" for AID function is perhaps provided by the hypothesis, first put forth by Petersen-Mahrt (Morgan *et al.*, 2004) that AID can act as a *de facto* active demethylase, i.e. it can deaminate mCpGs yielding TpG, which can be faithfully repaired back to CpG through the action of TDG. A flood of recent experimental evidence supports this notion: AID (with APOBEC−2 and TDG) appears to be at the core of an active demethylation system in zebrafish (Rai *et al.*, 2008); ectopic expression of AID can lead to specific reprogramming toward pluripotency in stem cells (Bhutani *et al.*, 2009); and finally, the primordial germ cell genome in AID-deficient animals appears to be hypermethylated (Popp *et al.*, 2010). Though these recent experiments are by no means flawless, they do provide support for the interesting hypothesis that AID might have evolved to demethylate CpGs in an error-free fashion for purposes of epigenetic reprogramming.

Taken together, the data discussed above suggest that the finding that AID deaminates genes other than Ig genes is not merely due to dysregulation and a by-product of its "real" function. This novel function for AID raises some interesting issues. One is that, though AID can deaminate methyl-dC *in vitro* (Morgan *et al.*, 2004) it has higher activity on dC (Bransteitter *et al.*, 2003; Larijani *et al.*, 2005). Another is that AID activity appears to be directed exclusively toward single-stranded DNA (Bransteitter *et al.*, 2003; Dickerson *et al.*, 2003), whereas methylated DNA would presumably be double-stranded, prior to transcription. However, this difficulty might not be insurmountable as many non-coding gene regions have been shown to be transcribed. In addition, promoter-proximal regions that are dense in CpG motifs, often

DNA DEAMINATION AND THE IMMUNE SYSTEM - AID in Health and Disease
© Imperial College Press
http://www.worldscibooks.com/lifesci/p729.html

contain non-canonical DNA structures which expose single-stranded regions (Ball *et al.*, 2009; Tsai *et al.*, 2009; Wittig *et al.*, 1991).

But is there a need for reprogramming toward pluripotency in healthy, functional, adult cells in which AID is expressed? One could argue that germinal center B cells could be just such cells. Germinal centers are hotbeds of epigenetic reprogramming, leading to the differentiation of activated B cells into memory cells and terminal differentiation into long-lived plasma cells. Plasma cells express a very different genetic profile from naïve, germinal center or memory B cells (Bhattacharya *et al.*, 2007; Klein *et al.*, 2003). Perhaps AID-mediated CpG demethylation at non-Ig loci contributes to B cell memory and plasma cell differentiation.

Conversely, there is tremendous pressure toward reprogramming (and eventually pluripotency) in cancers. If we assume that one role for AID is epigenetic reprogramming, and that reprogramming is central to tumorigenesis, then there might be a common signaling pathway that would lead to AID expression under conditions of transformation. One intriguing possibility is that this signal is the switch from mitochondrial oxidative phosphorylation to aerobic glycolysis, also referred to as "the Warburg effect" (Vander Heiden *et al.*, 2009), which allows rapidly dividing cells (such as germinal center B cells as well as all cancer cells) to generate the energy and building blocks needed for growth. Aerobic glycolysis is a universal requirement for cancer cells, and hypoxia (which functionally mimics the Warburg effect) is known to rapidly induce AID expression in cultured B cells (Kim *et al.*, 2006). Of course, each cancerous tissue would have different requirements for reprogramming, which would explain the curious correspondence between the types of AID-promoted translocations seen in different tumors (e.g. *IgH:myc* for B cells [Ramiro *et al.*, 2004]; ETS:TMPRSS2 for prostate cells [Lin *et al.*, 2009]) and the types of stimuli required for these tumors to reprogram toward pluripotency and progress toward malignancy. The expectation then would be that AID targets different genes for demethylation in different settings, and that the selection of translocation partners and their amplification in such settings is lineage-specific.

DNA DEAMINATION AND THE IMMUNE SYSTEM - AID in Health and Disease
© Imperial College Press
http://www.worldscibooks.com/lifesci/p729.html

This year has marked a decade of research into AID and its role in generating antibody diversity. Though we have still a lot to learn in that context, recent work would place AID in a much wider milieu with regard to its contributions to health and disease, and we are looking forward to an integrated understanding of AID function.

1.5 Acknowledgements

We gratefully acknowledge NIH support to the Stavnezer lab (RO1 AI23283) and to the Papavasiliou lab (R01 CA098495).

1.6 References

1. Ball M.P., Li J.B., Gao Y. *et al.* (2009). Targeted and genome-scale strategies reveal gene-body methylation signatures in human cells. *Nat Biotechnol* **27**: 361–368.
2. Besmer E., Market E., Papavasiliou F.N. (2006). The transcription elongation complex directs activation-induced cytidine deaminase-mediated DNA deamination. *Mol Cell Biol* **26**: 4378–4385.
3. Bhattacharya D., Cheah M.T., Franco C.B. *et al.* (2007). Transcriptional profiling of antigen-dependent murine B cell differentiation and memory formation. *J Immunol* **179**: 6808–6819.
4. Bhutani N., Brady J.J., Damian M. *et al.* (2009). Reprogramming towards pluripotency requires AID-dependent DNA demethylation. *Nature* **463**: 1042–1047.
5. Bransteitter R., Pham P., Scharff M.D. *et al.* (2003). Activation-induced cytidine deaminase deaminates deoxycytidine on single-stranded DNA but requires the action of RNase. *Proc Natl Acad Sci USA* **100**: 4102–4107.
6. Chaudhuri J., Tian M., Khuong C. *et al.* (2003). Transcription-targeted DNA deamination by the AID antibody diversification enzyme. *Nature* **422**: 726–730.
7. Chiba T., Marusawa H. (2009). A novel mechanism for inflammation-associated carcinogenesis; an important role of activation-induced cytidine deaminase (AID) in mutation induction. *J Mol Med* **83**: 1023–1027.
8. Conticello S.G., Langlois M.-A., Yang Z. *et al.* (2007). DNA deamination in immunity: AID in the context of its APOBEC relatives. *Adv Immunol* **94**: 37–73.
9. Dickerson S.K., Market E., Besmer E. *et al.* (2003). AID mediates hypermutation by deaminating single stranded DNA. *J Exp Med* **197**: 1291–1296.
10. Feldhahn N., Henke N., Melchior K. *et al.* (2007). Activation-induced cytidine deaminase acts as a mutator in BCR-ABL1-transformed acute lymphoblastic leukemia cells. *J Exp Med* **204**: 1157–1166.

DNA DEAMINATION AND THE IMMUNE SYSTEM - AID in Health and Disease
© Imperial College Press
http://www.worldscibooks.com/lifesci/p729.html

11. Grosjean H., de Crecy-Lagard V. *et al.* (2004). Aminoacylation of the anticodon stem by a tRNA-synthetase paralog: relic of an ancient code? *Trends Biochem Sci* **29**: 519–522.
12. Hamilton C., Papavasiliou F., Rosenberg B. (2010). Diverse functions for DNA and RNA editing in the immune system. *RNA Biol* **7**: 1–10.
13. Imai K., Slupphaug G., Lee W.I. *et al.* (2003). Human uracil-DNA glycosylase deficiency associated with profoundly impaired immunoglobulin class-switch recombination. *Nat Immunol* **4**: 1023–1028.
14. Kim J.W., Tchernyshyov I., Semenza G.L. *et al.* (2006). HIF-1-mediated expression of pyruvate dehydrogenase kinase: a metabolic switch required for cellular adaptation to hypoxia. *Cell Metab* **3**: 177–185.
15. Klein U., Tu Y., Stolovitzky G.A. *et al.* (2003). Transcriptional analysis of the B cell germinal center reaction. *Proc Natl Acad Sci USA* **100**: 2639–2644.
16. Larijani M., Frieder D., Sonbuchner T.M. *et al.* (2005). Methylation protects cytidines from AID-mediated deamination. *Mol Immunol* **42**: 599–604.
17. Lin C., Yang L., Tanasa B. *et al.* (2009). Nuclear receptor-induced chromosomal proximity and DNA breaks underlie specific translocations in cancer. *Cell* **139**: 1069–1083.
18. Liu M., Duke J.L., Richter D.J. *et al.* (2008). Two levels of protection for the B cell genome during somatic hypermutation. *Nature* **451**: 841–845.
19. Martin A., Bardwell P.D., Woo C.J. *et al.* (2002). Activation-induced cytidine deaminase turns on somatic hypermutation in hybridomas. *Nature* **415**: 802–806.
20. Mayorov V.I., Rogozin I.B., Adkison L.R. *et al.* (2005). Expression of human AID in yeast induces mutations in context similar to the context of somatic hypermutation at G-C pairs in immunoglobulin genes. *BMC Immunol* **6**: 10.
21. Morgan H.D., Dean W., Coker H.A. *et al.* (2004). Activation-induced cytidine deaminase deaminates 5-methylcytosine in DNA and is expressed in pluripotent tissues: implications for epigenetic reprogramming. *J Biol Chem* **279**: 52353–52360.
22. Mullighan C.G., Phillips L.A., Su X. *et al.* (2008). Genomic analysis of the clonal origins of relapsed acute lymphoblastic leukemia. *Science* **322**: 1377–1380.
23. Muramatsu M., Kinoshita K., Fagarasan S. *et al.* (2000). Class switch recombination and hypermutation require activation-induced cytidine deaminase (AID), a potential RNA editing enzyme. *Cell* **102**: 553–563.
24. Muramatsu M., Sankaranand V.S., Anant S. *et al.* (1999). Specific expression of activation-induced cytidine deaminase (AID), a novel member of the RNA-editing deaminase family in germinal center B cells. *J Biol Chem* **274**: 18470–18476.
25. Nakamura M., Kondo S., Sugai M. *et al.* (1996). High frequency class switching of an IgM+ B lymphoma clone CH12F3 to IgA+ cells. *Int Immunol* **8**: 193–201.
26. Pasqualucci L., Bhagat G., Jankovic M. *et al.* (2008). AID is required for germinal center-derived lymphomagenesis. *Nat Genet* **40**: 108–112.
27. Pasqualucci L., Neumeister P., Goossens T. *et al.* (2001). Hypermutation of multiple proto-oncogenes in B-cell diffuse large-cell lymphomas. *Nature* **412**: 341–346.

DNA DEAMINATION AND THE IMMUNE SYSTEM - AID in Health and Disease
© Imperial College Press
http://www.worldscibooks.com/lifesci/p729.html

28. Petersen-Mahrt S.K., Harris R.S., Neuberger M.S. (2002). AID mutates E. coli suggesting a DNA deamination mechanism for antibody diversification. *Nature* **418**: 99–103.
29. Pham P., Bransteitter R., Petruska J. *et al.* (2003). Processive AID-catalysed cytosine deamination on single-stranded DNA simulates somatic hypermutation. *Nature* **424**: 103–107.
30. Poltoratsky V.P., Wilson S.H., Kunkel T.A. *et al.* (2004). Recombinogenic phenotype of human activation-induced cytosine deaminase. *J Immunol* **172**: 4308–4313.
31. Popp C., Dean W., Feng S. *et al.* (2010). Genome-wide erasure of DNA methylation in mouse primordial germ cells is affected by AID deficiency. *Nature* **463**: 1101–1105.
32. Rada C., Di Noia J.M., Neuberger, M.S. (2004). Mismatch recognition and uracil excision provide complementary paths to both Ig switching and the A/T-focused phase of somatic mutation. *Mol Cell* **16**: 163–171.
33. Rada C., Williams G.T., Nilsen H. *et al.* (2002). Immunoglobulin isotype switching is inhibited and somatic hypermutation perturbed in UNG-deficient mice. *Curr Biol* **12**: 1748–1755.
34. Rai K., Huggins I.J., James S.R. *et al.* (2008). DNA demethylation in zebrafish involves the coupling of a deaminase, a glycosylase, and gadd45. *Cell* **135**: 1201–1212.
35. Ramiro A.R., Jankovic M., Eisenreich T. *et al.* (2004). AID is required for c-myc/IgH chromosome translocations in vivo. *Cell* **118**: 431–438.
36. Revy P., Muto T., Levy Y. *et al.* (2000). Activation-induced cytidine deaminase (AID) deficiency causes the autosomal recessive form of the Hyper-IgM syndrome (HIGM2). *Cell* **102**: 565–575.
37. Shen H.M., Peters A., Baron B. *et al.* (1998). Mutation of BCL-6 gene in normal B cells by the process of somatic hypermutation of Ig genes. *Science* **280**: 1750–1752.
38. Shen H.M., Poirier M.G., Allen M.J. *et al.* (2009). The activation-induced cytidine deaminase (AID) efficiently targets DNA in nucleosomes but only during transcription. *J Exp Med* **206**: 1057–1071.
39. Tsai A.G., Engelhart A.E., Hatmal M.M. *et al.* (2009). Conformational variants of duplex DNA correlated with cytosine-rich chromosomal fragile sites. *J Biol Chem* **284**: 7157–7164.
40. Vander Heiden M.G., Cantley L.C., Thompson C.B. (2009). Understanding the Warburg effect: the metabolic requirements of cell proliferation. *Science* **324**: 1029–1033.
41. Wang M., Rada C., Neuberger M.S. (2010) Altering the spectrum of immunoglobulin V gene somatic hypermutation by modifying the active site of AID. *J Exp Med* **207**: 141–153.
42. Wedekind J.E., Dance G.S.C., Sowden M.P. *et al.* (2003). Messenger RNA editing in mammals: new members of the APOBEC family seeking roles in the family business. *Trends Genet* **19**: 207–216.

DNA DEAMINATION AND THE IMMUNE SYSTEM - AID in Health and Disease
© Imperial College Press
http://www.worldscibooks.com/lifesci/p729.html

43. Wittig B., Dorbic T., Rich A. (1991). Transcription is associated with Z-DNA formation in metabolically active permeabilized mammalian cell nuclei. *Proc Natl Acad Sci USA* **88**: 2259–2263.
44. Yoshikawa K., Okazaki I.M., Eto T. *et al.* (2002). AID enzyme-induced hypermutation in an actively transcribed gene in fibroblasts. *Science* **296**: 2033–2036.

DNA DEAMINATION AND THE IMMUNE SYSTEM - AID in Health and Disease
© Imperial College Press
http://www.worldscibooks.com/lifesci/p729.html