A Novel Regulator of Activation-Induced Cytidine Deaminase/APOBECs in Immunity and Cancer: Schrödinger’s CATalytic Pocket

Justin J. King and Mani Larijani*

Immunology and Infectious Diseases Program, Division of Biomedical Sciences, Faculty of Medicine, Memorial University of Newfoundland, St. John’s, NL, Canada

Activation-induced cytidine deaminase (AID) and its relative APOBEC3 cytidine deaminases boost immune response by mutating immune or viral genes. Because of their genome-mutating activities, AID/APOBECs are also drivers of tumorigenesis. Due to highly charged surfaces, extensive non-specific protein–protein/nucleic acid interactions, formation of polydisperse oligomers, and general insolubility, structure elucidation of these proteins by X-ray crystallography and NMR has been challenging. Hence, almost all available AID/APOBEC structures are of mutated and/or truncated versions. In 2015, we reported a functional structure for AID using a combined computational–biochemical approach. In so doing, we described a new regulatory mechanism that is a first for human DNA/RNA-editing enzymes. This mechanism involves dynamic closure of the catalytic pocket. Subsequent X-ray and NMR studies confirmed our discovery by showing that other APOBEC3s also close their catalytic pockets. Here, we highlight catalytic pocket closure as an emerging and important regulatory mechanism of AID/APOBEC3s. We focus on three sub-topics: first, we propose that variable pocket closure rates across AID/APOBEC3s underlie differential activity in immunity and cancer and review supporting evidence. Second, we discuss dynamic pocket closure as an ever-present internal regulator, in contrast to other proposed regulatory mechanisms that involve extrinsic binding partners. Third, we compare the merits of classical approaches of X-ray and NMR, with that of emerging computational–biochemical approaches, for structural elucidation specifically for AID/APOBEC3s.

Keywords: lymphoma, antibodies, DNA mutations, leukocytes, enzymes and coenzymes, HIV infections

IMPORTANT AND CHALLENGES OF SOLVING AID/APOBEC STRUCTURES

Activation-induced cytidine deaminase (AID) is a 198 amino acid DNA-editing enzyme that deamminates deoxyxycytidine (dC) to deoxuryridine (dU) in single-stranded DNA (ssDNA) (1–6). It acts on immunoglobulin (Ig) loci initiating hypermutation and recombination events that lead to improved and class-switched antibodies (1, 7–9). However, AID is loosely targeted to Ig loci, and hence, it induces genome-wide mutations and double-strand breaks which can lead to tumors (10–16). In
addition, continued AID expression increases genetic plasticity of tumors thereby accelerating disease progression (17, 18).

Activation-induced cytidine deaminase is a member of the apolipoprotein B RNA-editing catalytic component (AID/APOBEC) family of cytidine deaminases, a Zn-dependent family with 11 members in humans: AID, APOBEC1, APOBEC2, the APOBEC3 sub-branch (A-H, excluding E), and APOBEC4 (9, 19, 20). The APOBEC3 (A3) sub-branch members are anti-retroviral/retroelement restriction factors thereby also playing an immune function (21, 22); however, in the last few years, a role in cancer initiation has also emerged for the A3 sub-branch of the family, in particular, A3A, A3B (23–32), and more recently A3H haplotype I (33).

Given their intimate links to immunity and cancer, much effort has been placed on understanding the molecular structures of AID/APOBECs over the last decade. A major hurdle in this effort has been the isolation and purification of native AID/APOBEC proteins to absolute purity. To different measures for each individual family member, the challenges include cellular genotoxicity, highly charged surfaces mediating extensive non-specific protein–protein, protein–DNA/RNA interactions (9, 34), and polydisperse oligomerization (35, 36). Consequently, in most cases, structure resolution by the traditional methodologies of X-ray crystallography and NMR has necessitated substantial alterations to stabilize protein charge and enhance solubility or crystallization. Therefore, the vast majority (22 of 24) of APOBEC structures solved by X-ray or NMR are of truncated and/or mutated versions, necessary to enhance solubility (Table 1) (37–57). Nevertheless, it has become clear that AID/APOBEC family member enzymes share the core structure of a central β-sheet with 4 or 5 β strands sandwiched between 6 and 7 α-helices, connected by 12–13 flexible loops of variable lengths (Figure 1A). Second, they all have highly charged DNA-binding grooves, necessary to bind negatively charged polynucleotides. The arrangement of core catalytic residues in the catalytic pocket is also conserved, consisting of a Zn-coordinating triad of two cysteines and a histidine, atop a catalytic proton-donor glutamic acid (C87, C90, H56, and E58 in AID) (9, 19, 20, 58).

**SOLVING THE STRUCTURE OF AID AND DISCOVERY OF CATALYTIC POCKET CLOSURE IN AID/APOBECs**

The biochemical properties of AID have been previously described. AID has an exceptionally high affinity (nM range) for binding ssDNA and an unusually slow catalytic rate of one reaction in several minutes (5, 60), ∼2,000 times slower than a typical enzyme (61). We postulated that this catalytic lethargy and high-binding affinity to DNA had evolved to protect genomes from rampant AID activity (9). In direct support of this notion, mutants of AID with higher catalytic rates were shown to mediate higher levels of genome damage in cells (62). Although this body of work led to understanding AID’s behavior, the molecular basis behind these properties remained an enigma. AID is notoriously challenging to isolate to absolute purity and hence its native structure has remained unsolved by X-ray and NMR since its discovery in 1999, despite intense efforts. Thus, we posited that even if AID’s structure were to be solved by traditional methods of X-ray or NMR, it would most likely be of a truncated and/or heavily mutated version. We proposed an alternative methodology to gain insight into the functional and native structure of AID. We utilized eight recently solved structures of AID’s APOBEC relatives as templates to generate thousands of AID predicted model structures followed by identification of the lowest energy clusters (38) (Figure 1A). Concurrently and guided by the computational predictions, we generated a library of 400 AID variants and carried out extensive biochemical characterization of catalytic function and DNA binding to rigorously test key predictions of models. This library included different point mutants for each residue along the length of AID, orthologous AIDs, and chimeras involving regions of other deaminases exchanged into the AID scaffold, or vice-versa. Our rationale for including AID orthologs was that divergent AID from distinctly evolved species may have distinct biochemical properties and characterizing these through a combination of homology structural modeling and functional analysis of mutated and chimeric enzymes would generate structure:function insights. Differences among orthologous AIDs included catalytic rates, substrate preferences, DNA-binding affinities, and thermosensitivity profiles (63–65). Since these differences are typically due to structural features, being reflective of catalytic motifs, surface composition, and overall protein flexibility, respectively, characterizing the basis of these differences among orthologs proved a valuable tool to gain insight into AID’s structure:function relationships. This computational–biochemical approach led to the first relatively detailed 3D maps of AID’s functional structure with special focus on catalytic pocket architecture and ssDNA-binding motifs (9, 58, 63, 64).

The architecture and dynamics of an enzyme’s catalytic pocket are important determinants of its activity. In addition to the core catalytic pocket composed of the aforementioned triad Zinc-coordinating residues and a Glutamic acid, we identified an additional 21 amino acids that are not directly involved in the deamination reaction, but compose the pocket’s physiochemical microenvironment (58). These residues termed secondary catalytic residues form the “walls” and “floors” of the pocket and stabilize dC binding. We noted that the conformations of these secondary catalytic residues exhibited more variability than that of the primary catalytic residues among predictions, because these residues reside on several highly flexible connecting loops without secondary structures of their own, that surround the catalytic pocket (Figure 1B). Because of this placement, we observed that the catalytic pocket of AID appeared to be only marginally stable such that ~75% of conformations exhibit an occluded pocket unable to accommodate dC. Thus, we hypothesized that dynamic catalytic pocket closure is a built-in mechanism that limits AID activity.

There are several lines of indirect but strong evidence for the existence of catalytic pocket closure in AID/APOBEC3s. First, the fact that the majority of AID conformations exist in a state with closed catalytic pockets provides a mechanistic explanation for the relative catalytic lethargy of AID as discussed above (5). Second, it is mathematically compatible with known parameters of AID: AID binds ssDNA sporadically on its surface such that most ssDNA (~95%) neither pass over AID’s catalytic pocket, nor...
position dC for catalytic pocket entry. The proportion of catalytically viable AID:DNA complexes (~5%) multiplied by the ratio of open pockets (25%) yields (1.25%). This correlates closely with our own estimates of active AID complexes based on Michaelis–Menten parameters (5) and with other studies that carried out mathematical modeling of AID’s substrate catalysis (66).

Direct proof for existence and significance of catalytic pocket closure came from two sources: one functional, and the other structure-based. First, we designed a panel of AID variants in which the secondary catalytic loops and surrounding regions were replaced with their equivalents from other APOBECs or orthologous AIDs to alter predicted pocket dynamics such that the pocket would spend either more or less time in the open conformation. We then observed that the proportion of time the pocket was predicted to assume an open conformation correlated exquisitely with catalytic rate differences among said AID variants some of which became up to 100 times more active than wild-type AID because of a catalytic pocket that spends more time in the open conformation. This provided direct functional evidence that pocket closure limits activity. The second proof came from direct observation of closed pockets in several siblings of AID: in APOBEC3A by NMR (40), in APOBEC3B by X-ray crystallography (42, 57), and by NMR (43) (Figures 1C,D).

### DIFFERENTIAL CATALYTIC POCKET STATES MEDIATE VARIABLE BIOLOGICAL ACTIVITIES AMONG AID/APOBECs

From an evolutionary perspective, regulation by catalytic pocket closure provides an effective means to fine-tune variable levels of enzymatic robustness across the AID/APOBEC family, as well as impart varying types of activities among orthologous versions of each family member. This is because the same high degree of movement freedom in the secondary catalytic loops that lead to the fluidity of catalytic pocket dynamics in each AID/APOBEC3...
enzyme also allows for a high level of sequence and length divergence in these loops among individual AID/APOBEC3s, to impart a unique range of open/closed breathing dynamics to the catalytic pocket of each member (Figure 1).

To elaborate, in each APOBEC3, the catalytic pocket “walls” and “floor” are composed of residues contributed by four secondary catalytic loops (Figure 1B). The highest structural variation among the AID/APOBEC family appears in loop 2 (L2), loop 4 (L4), and loop 8 (L8) with respect to sequence homology, length, and compaction relative to the core enzyme structure (Figures 1B,C) (58). L2 contains residues involved in ssDNA-binding, catalytic pocket and dC stabilization, substrate specificity, and 5-mC tolerance (58, 67–69). L4 contains residues critical to catalysis and catalytic pocket occlusion (58). Recently, an allosteric regulatory
role for L4 in A3A and A3G was identified through coordination of a secondary Zn that enhances activity (70). Interestingly, secondary Zn coordination was suggested to fine tune the position of the secondary catalytic residues, thus creating an ideal environment for cytidine deamination (70). Furthermore, secondary Zn coordination was suggested to mediate cooperative dimerization. Lastly, L8 mediates substrate sequence specificity, dC stabilization, and 5-methyl-C (5-mC) tolerance (58, 64, 67, 71). Collectively, the secondary catalytic loops mediate functional differences among the AID/APOBECs and dictate variations in the frequency of open vs. closed catalytic pockets.

Based on these observations, we proposed that differences in secondary catalytic loops mediate variable catalytic pocket breathing dynamics, responsible for different enzymatic robustness among AID/APOBEC enzymes (58). Indeed, in the last year, functional evidence in support of this novel mode of regulation has emerged. First, A3A exhibits open catalytic pockets in more conformations than AID (67 vs. 25.6%), and accordingly it is a more robust enzyme with a faster on/off rate of deaminating DNA (40, 72). Second, A3B-CTD exhibits roughly one third of the pockets in an open conformation compared to A3A (20 vs. 67%, respectively) (Figure 1D) and this also correlates directly with a lower catalytic rate (43). It is intriguing that thus far, catalytic pocket occlusion has been observed in three of the most mutagenic and tumorigenic members of the AID/APOBEC family: AID, A3A, and A3B. This lends credence to the idea that this is an internal protective mechanism to limit genome mutations by these enzymes. As mentioned above, the difference in catalytic activity of purified AID, A3A, A3B-CTD, correlates with the ranking of pocket occlusion (Figure 1D). Though further study is required to clarify the relative contributions of A3A, A3B, and other APOBEC3 branch enzymes (e.g., A3H) in various types of cancers, some emerging evidence indicates that there is a more dominant mutational signature observed from A3A than A3B, at least in a yeast model and in urothelial carcinoma, despite lower levels of A3A expression (28, 73).

In addition to regulation of tumorigenic activity, differences in catalytic pocket dynamics also appear to correlate well with other biological functions of AID/APOBECs. As an example, zebrafish AID has a significantly higher reaction rate than human AID and is also unique among all AID orthologs in that it can deaminate 5-mC in methylated CpG motifs (64). This explains a puzzling previous report that zebrafish AID plays a completely non-immune role. During embryogenesis in zebrafish, AID can mediate promoter demethylation through erasure of gene-silencing CpG methylation marks, thus orchestrating widespread gene expression required for tissue differentiation (74). This is attributable to conformational differences in the aforementioned secondary catalytic loops between human and zebrafish AID, which translate to a higher ratio of open vs. closed catalytic pockets. Consequently, zebrafish AID can accommodate and deaminate 5-mC, as opposed to human AID whose activity on 5-mC is negligible. This enzymatic difference is one factor that enables zebrafish AID to function in genome demethylation during embryonic development, an activity that is completely outside the realm of an immune function (58, 64). Taken together, these lines of evidence are supportive of catalytic pocket occlusion being a key regulator of biological functions of AID/APOBEC3s, including their role in tumorigenic genome damage.

**CATALYTIC POCKET OCCLUSION AS INTERNALLY BUILT-IN REGULATION**

Since the discovery of AID, much effort has been directed at understanding how its activity is regulated, under the supposition that a mutator so threatening must be operating under tight restrictions. To date, almost all efforts have focused on modes of regulation that are extrinsic to the enzyme itself. This has led...
to the identification of over two dozen co-factors proposed to bind AID either directly or indirectly through associations with other proteins or DNA/RNA (75–99). The list of putative binding partners is rather large for a relatively small protein of 198 residues. Hence, one must approach biological relevance with caution for several reasons: first, although some co-factors are modestly enriched at Ig loci, none can account for targeting AID to specific loci. Second, given the relatively small size of AID and the lack of clear conformational protein-binding domains, the number of proposed co-factors seems high. It is rather improbable that a small 198 aa globular enzyme can fold properly to bind ssDNA, deaminate dC, maintain sequence specificity, while still leaving enough non-essential portions free to bind dozens of different co-factors each in a specific and orchestrated fashion. Indeed, a careful analysis of AID’s structure reveals that most of its structure can be ascribed a function directly related to forming the core architecture essential to bind and deaminate a polynucleotide. Furthermore, a portion of the surface is likely unavailable due to forming the oligomerization surfaces, as most AID/APOBECs appear to exist as dimers or tetramers (5, 38, 70, 100). Also, AID has a highly charged surface and a well-known propensity for high affinity non-specific interactions with other proteins (9, 34). Thus, the biological significance of AID binding to many of its putative co-factors is a topic that requires further resolution. Furthermore, the very premise of searching for cofactors to explain regulation of AID targeting may be flawed in that the more AID is studied, the clearer it becomes that its activity is rather not tightly regulated: despite a modest preference for Ig loci which appears to be mediated by unique transcriptional features (101–103), AID mutates endogenous genes and transgenes genome-wide, and can do so in any cell type in which it is naturally or exogenously expressed (10, 104–107).

Like AID, the search for regulatory mechanisms of other APOBEC3s has also focused on extrinsic binding factors, of which several have been identified including various viral proteins, and transcription factors (108–111); the most well-characterized APOBEC3 are virion infectivity factor (Vif) and cytoplasmic ribonucleoside complexes. The Vif protein of HIV binds and targets A3C, A3G, A3F, A3D, and A3H (to varying degrees) for degradation via a ubiquitin-dependent proteasomal pathway (44, 112–120). Thus, when Vif is present, APOBEC3 effectiveness in viral restriction is severely diminished. Second, the activity of the anti-retroviral APOBEC3s is limited by entrapment in high-molecular-mass ribonucleoside complexes (HMM) that may reach megadaltons in size, mediated by non-specific protein/DNA/RNA binding in the cytoplasm, mediated by aforementioned highly charged surfaces (121–127).

In contrast to regulation by extrinsic binding partners, be they protein or nucleic acid, catalytic pocket closure represents a novel intrinsic mode of regulation. This simple mechanism of limiting activity has several attractive features: it is ever-present, biologically reliable, mechanistically simple, and structurally sound. Furthermore, as discussed in the preceding section, its variation is an evolutionary efficient mechanism for diversifying and fine-tuning activity levels of family member enzymes, as catalytic pocket closure rates can be adjusted by minimal amino acid substitutions in secondary catalytic loops. It is also biologically efficient since it does not require any cellular resources, unlike the proposed complex networks of co-factors, which themselves would require regulation in different cells at different stages of differentiation or viral infection, thus amplifying the need for cellular resources.

**IMPORTANCE OF DETERMINING AID/APOBEC3 STRUCTURES THAT ARE NATIVE AND INCLUDE FUNCTIONAL INSIGHTS**

X-ray crystallography and NMR have advanced the AID/APOBEC field with the full or partial structure elucidation of 7 of 11 APOBEC enzymes. Despite these achievements, there are pitfalls in using these traditional methods alone. First, the purification issues discussed above have necessitated working with significantly truncated and/or heavily mutated versions of AID/APOBEC proteins (Table 1). The truncations and mutations are often in functionally critical regions, such as the secondary catalytic loops. Additionally, all double-domain APOBECs whose structure has been characterized (A3B, A3F, and A3G) lack their enzymatically inactive N-terminal half which is implicated in the catalytic activity and dimerization (56, 128, 129). The N-terminal half of A3G's separate structure was recently reported; however, these were also mutated and likewise lack the C-terminal half (55, 56). Second, depending on crystallization or NMR conditions, even the same APOBEC structure determined by different groups can be quite distinct (49, 50). These differences are likely due to differences in solution or crystallization conditions which can bias toward a specific structure or conformation (130).

In contrast, the methodology that we applied to solving AID's structure provides both a functional and native structure (58). By integrating dynamic modeling with the study of a large library of variants to functionally verify key model predictions, the emerging picture integrates the relative abundance of an enzyme's conformations with functional significance (Figure 1E). This approach is particularly advantageous in the case of AID/APOBEC3s because many functional differences among AID/APOBEC3 family members appear to be dictated by subtle differences in breathing dynamics, rather than major architectural differences. It is important to note that despite being a robust methodology for determining functional and native enzyme structures, this approach is not without practical challenges: first, it is laborious and time-intensive since it requires examination of thousands of high confidence models. As modeling efforts progress, there is a continuing need to generate and test a large variant library, often necessitating several mutants of each key residue to rigorously verify its exact position, relative attitude, and role(s). In addition, a sensitive enzyme assay able to detect even small differences in biochemical properties with that of wild-type ought to be in place. Practical difficulties are compounded by the fact that this approach of solving a functional and native structure is often most useful for enzymes that are challenging to purify. Second, modeling efforts depend critically on the availability of solved X-ray or NMR structures to serve as templates, with multiple templates increasing confidence. For instance, at the time of our
efforts on AID, we utilized eight available APOBEC structures as templates. It is important to have numerous templates from different family members, so as to at least partially compensate for aforementioned limitations of each template in terms of encompassing the full range of conformations. Furthermore, template structures ought to be evaluated for their suitability on the basis of extent and location of homologous/identical residues, and model quality itself ought to be rigorously scrutinized, mathematically and functionally using the variant library (58) (Figure 1E). Third, the basic biochemical properties of the enzyme ought to have previously been determined so as to serve as a valuable verification tool for model validity; since we had already determined that AID has an unusually low catalytic rate and high binding affinity for ssDNA, the fact that our structure fully explained both of these properties through the abundance of closed pockets and positively charged surface residues along putative DNA-binding grooves and elsewhere on the surface, provided further confidence. Lastly, definite physical confirmation of findings requires subsequent observation by X-ray and NMR, as in the case of catalytic pocket closure described above. In addition to our observation of catalytic pocket closure being confirmed by direct X-ray and NMR studies as described above (42, 43, 57), other X-ray studies have also confirmed our observation of key catalytic residues as well as important DNA-binding residues of AID: following the publication of AID’s functional structure, the crystal structure of an AID variant was also reported (37). As expected, it included mutations and truncations crucial to solubilize AID for X-ray crystallography (Table 1). Although this structure represents a significant achievement, it necessitated introducing mutations and truncations that removed some of the unique characteristics of AID. For instance, the majority of mutations neutralized the positively charged surface residues lining the DNA-binding groove culminating in a net charge of +4.5, as compared to AID’s native charge of +14 at neutral pH. This high net positive charge of AID is a unique feature amongst AID/as compared to AID’s native charge of +4.5, neutralizations neutralized the positively charged surface residues lining introducing mutations and truncations that removed some of the

However, the structure of L8 is stabilized by L2, which has been shown to modulate its compaction and substrate specificity (69). Additionally, the conformation of α7 relative to the surface of AID is uncertain, although some conformations place it in direct contact with L8 (58). Therefore, although L8 was not directly altered, mutation of L1α1-L2 together with α7 deletion may indirectly perturb its conformation in the AID variant crystal structure.

Using our computational–biochemical approach, we also highlighted two DNA-binding grooves on the surface of AID, for both of which the positively charged R25 residues plays a major role in orienting the negatively charged DNA backbone (58). Recently, DNA-bound crystal structures of mutant A3A and an A3B-CTD chimera were shown to adopt a similar DNA-binding mode, wherein the DNA backbone was bound around the equivalent of R25 in AID (H29 and H212, in mutant A3A and A3B-CTD chimera, respectively) (57).

In this manner, X-ray and NMR structural elucidation of homologous APOBEC3s have provided direct physical support for notable features of AID observed using the computational–biochemical approach. These features include position and identity of catalytic residues, key DNA-contact residues as well as existence of occluded catalytic pockets, a novel regulatory mechanism.

**CONCLUSION**

In summary, we draw the parallel to the Schrödinger’s Cat paradox that the catalytic pockets of AID/APOBEC3s appear to transition between dual states, one of which correlates with activity and the other with catalytic death, each with profound functional consequences. The second parallel between structure determination in the AID/APOBEC3 field and quantum physics is that X-ray crystallography and NMR determination of structures in the AID/APOBEC family have most often necessitated making extensive alterations to structures for and during the very act of observation. In contrast, the computational–biochemical approach used to solve AID’s functional structure relies on unobtrusive observation through prediction. Interventions are strictly reserved for the functional testing phase wherein structure predictions are rigorously scrutinized by conducting enzyme assays on a large library of variants including mutants, orthologs, and chimeras (Figure 1E). As described in the preceding section, it is important to note that this method is nonetheless dependent on the availability of multiple X-ray and NMR structure solutions, both in the beginning as templates and in the end, as independent methods to independently verify the key aspects of the structure.

In the future, as the relative contributions of each individual AID/APOBEC3 enzyme to immunity and cancer in different contexts become clearer, it will be important to test the hypothesis that catalytic pocket breathing differences among the AID/APOBEC3 family members impact their relative contributions, and to understand the extent to which this novel built-in safety switch is intertwined with other regulatory mechanisms, such as perhaps being modulated by aforementioned extrinsic binding partners, oligomerization or post-translational modifications. Although catalytic pocket closure has been described for other
enzymes (131, 132), our discovery of such a functionally critical state in AID/APOBEC3s represents a novel regulatory mechanism for human DNA/RNA-damaging enzymes; hence, it is also important to ascertain how prevalent a regulatory mechanism dynamic catalytic pocket closure is in other DNA/RNA-editing enzymes, or whether it has evolved as a unique regulatory structural feature of the AID/APOBEC3 family.

ACKNOWLEDGMENTS

The authors are grateful to colleague Dr. Michael Grant for discussions which led to the inspiration for the title.

FUNDING

The highlighted work was supported by a Canadian Cancer Society Research Institute (CCSRI) Innovation operating grant and Canadian Institutes of Health Research (CIHR) operating grant (MOP111132) to ML. JK is supported by a CIHR Frederick Banting and Charles Best Canada Doctoral Scholarship.

AUTHOR CONTRIBUTIONS

JK (senior Ph.D. candidate) and ML (PI) both contributed to the writing of this article.

REFERENCES

1. Muramatsu M, Sankaranand VS, Anant S, Sugai M, Kinoshita K, Davidson NO, et al. Specific expression of activation-induced cytidine deaminase (AID), a novel member of the RNA-editing deaminase family in germinal center B cells. J Biol Chem (1999) 274(26):18470–6. doi:10.1074/jbc.274.26.18470
2. Brantsteiter R, Pham P, Scharff MD, Goodman MF. Activation-induced cytidine deaminase deaminates deoxy-cytidine on single-stranded DNA but requires the action of RNase. Proc Natl Acad Sci U S A (2003) 100(7):4102–7. doi:10.1073/pnas.0730835100
3. Dickerson SK, Market E, Besmer E, Papavasiliou FN. AID mediates hypermutation by deaminating single stranded DNA. J Exp Med (2003) 197(10):1291–6. doi:10.1084/jem.20030481
4. Sohail A, Klapac J, Samaranyake M, Ullah A, Bhagwat AS. Human activation-induced cytidine deaminase causes transcription-dependent, strand-biased C to U deaminations. Nucleic Acids Res (2003) 31(12):2990–4. doi:10.1093/nar/gkg646
5. Larjani M, Petrov AP, Kolenchenko O, Berru M, Krylov SN, Martin A. AID associates with single-stranded DNA with high affinity and a long complex half-life in a sequence-independent manner. Mol Cell Biol (2007) 27(1):20–30. doi:10.1128/MCB.00824-06
6. Larjani M, Martin A. Single-stranded DNA structure and positional context of the target cytidine determine the enzymatic efficiency of AID. Mol Cell Biol (2007) 27(23):9038–48. doi:10.1128/MCB.01046-07
7. Muramatsu M, Kinoshita K, Fagarasan S, Yamada S, Shinkai Y, Honjo T. Class switch recombination and hypermutation require activation-induced cytidine deaminase (AID), a potential RNA editing enzyme. Cell (2000) 102(5):553–63. doi:10.1016/S0092-8674(00)00078-7
8. Revy P, Muto T, Levy Y, Feissmann F, Plebani A, Sanal O, et al. Activation-induced cytidine deaminase (AID) deficiency causes the autosomal recessive form of the Hyper-IgM syndrome (HIGM2). Cell (2000) 102(5):565–75. doi:10.1016/S0092-8674(00)00079-7
9. Larjani M, Martin A. The biochemistry of activation-induced deaminase and its physiological functions. Semin Immunol (2012) 24(4):255–63. doi:10.1016/j.smim.2012.03.003
10. Liu M, Duke JL, Richter DJ, Vinaasa CG, Goodnow CC, Kleinstein SH, et al. Two levels of protection for the B cell genome during somatic hypermutation. Nature (2008) 451(7180):841–5. doi:10.1038/nature06547
11. Pasqualucci L, Bhagat G, Jankovic M, Compagno M, Smith P, Muramatsu M, et al. AID is required for germinal center-derived lymphomagenesis. Nat Genet (2008) 40(1):108–12. doi:10.1038/ng.2007.35
12. McCarthy H, Wierda WG, Barron LL, Cromwell CC, Wang J, Coombes KR, et al. High expression of activation-induced cytidine deaminase (AID) and splice variants is a distinctive feature of poor-prognosis chronic lymphocytic leukemia. Blood (2003) 101(12):4903–8. doi:10.1182/blood-2002-09-2906
13. Okazaki IM, Hiai H, Kakazu N, Yamada S, Muramatsu M, Kinoshita K, et al. Constitutive expression of AID leads to tumorigenesis. J Exp Med (2003) 197(9):1173–81. doi:10.1084/jem.20030275
14. Ramiro A, Reina San-Martin B, McBride K, Jankovic M, Barreto V, Nussenzweig A, et al. The role of activation-induced deaminase in antibody diversification and chromosome translocations. Adv Immunol (2007) 94:75–107. doi:10.1016/S0065-2776(06)94003-6
15. Robbiani DF, Bothmer A, Callen E, Reina-San-Martin B, Dorsett Y, Difilippantonio S, et al. AID is required for the chromosomal breaks in c-myc that lead to c-myc/Id4 translocations. Cell (2008) 135(6):1028–38. doi:10.1016/j.cell.2008.09.062
16. Robbiani DF, Bunting S, Feldhahn N, Bothmer A, Camps I, Deroubaix S, et al. AID produces DNA double-strand breaks in non-Ig genes and mature B cell lymphomas with reciprocal chromosome translocations. Mol Cell (2009) 36(4):631–41. doi:10.1016/j.molcel.2009.11.007
17. Klemm L, Duy C, Iacobucci I, Kuchen S, von Levetzow G, Feldhahn N, et al. The B cell mutator AID promotes B lymphoid blast crisis and drug resistance in chronic myeloid leukemia. Cancer Cell (2009) 16(3):232–45. doi:10.1016/j.ccc.2009.07.030
18. Munoz DP, Lee EL, Takayama S, Coppe JP, Heo SJ, Boffelli D, et al. Activation-induced cytidine deaminase (AID) is necessary for the epithelial-mesenchymal transition in mammary epithelial cells. Proc Natl Acad Sci U S A (2013) 110(32):E2797–86. doi:10.1073/pnas.1301021110
19. Conticello SG. The AID/APOBEC family of nucleic acid mutases. Genome Biol (2008) 9(6):229. doi:10.1186/gb-2008-9-6-229
20. Conticello SG, Langlois MA, Yang Z, Neuberger MS. DNA deamination in immunity: AID in the context of its APOBEC relatives. Adv Immunol (2007) 94:37–73. doi:10.1016/S0065-2776(06)94002-4
21. Harris RS. Enhancing immunity to HIV through APOBEC. Nat Biotechnol (2008) 26(10):1089–90. doi:10.1038/nbt1008-1089
22. Stavrakis S, Ross SR. APOBEC3 proteins in viral immunity. J Immunol (2015) 195(10):4565–70. doi:10.4049/jimmunol.1501504
23. Roberts SA, Lawrence MS, Klimczak LJ, Grimm SA, Dorochnow P, Stojanov P, et al. APOBEC cytidine deaminase mutagenesis pattern is widespread in human cancers. Nat Genet (2013) 45(9):970–6. doi:10.1038/ng.2702
24. Roberts SA, Gordenin DA. Clustered and genome-wide transient mutagenesis in human cancers: hypermutation without permanent mutations or loss of fitness. Bioessays (2014). doi:10.1002/bies.201300140
25. Roberts SA, Gordenin DA. Hypermutation in human cancer genomes: footprints and mechanisms. Nat Rev Cancer (2014) 14(12):786–800. doi:10.1038/nrc3816
26. Alexandrov LB, Nik-Zainal S, Wedge DC, Aparicio SA, Behjati S, Biankin AV, et al. Signatures of mutational processes in human cancer. Nature (2013) 500(7463):415–21. doi:10.1038/nature12477
27. Burns MB, Leonard B, Harris RS. APOBEC3B: pathological consequences of an innate immune DNA mutator. Biomolecules (2015) 38(2):102–10. doi:10.4103/2194-6490.1068
28. Chan K, Roberts SA, Klimczak LJ, Sterling JF, Saini N, Malik EP, et al. An APOBEC3A hypermutation signature is distinguishable from the signature of background mutagenesis by APOBEC3B in human cancers. Nat Genet (2015) 47(9):1067–72. doi:10.1038/ng.3378
29. Burns MB, Lackey L, Carpenter MA, Rathore A, Land AM, Leonard B, et al. APOBEC3B is an enzymatic source of mutation in breast cancer. Nature (2013) 494(7437):366–70. doi:10.1038/nature11881
30. Leonard B, Hart SN, Burns MB, Carpenter MA, Temiz NA, Rathore A, et al. APOBEC3B upregulation and genomic mutation patterns in serious
ovarian carcinoma. Cancer Res (2013) 73(24):7222–31. doi:10.1158/0008-5472.CAN-13-1753
31. Knisbacher RA, Gerber D, Levanon EY. DNA editing by APOBECs: a genomic preserver and transformer. Trends Genet (2016) 32(1):16–28. doi:10.1016/j.
tig.2015.01.005
32. Rehband S, Huerer M, Greil R, Geisberger R. AID/APOBEC deaminases and cancer. Oncogene (2015) 2(4):320–33. doi:10.18632/oncogene.155
33. Starrett GJ, Luengas EM, McCann JL, Ebrahimdi H, Temiz NA, Love RP, et al. The DNA cytosine deaminase APOBEC3H hlapotype I likely contributes to breast and lung cancer mutation. Nat Commun (2016) 7:12918. doi:10.1038/ncomms12918
34. Coker HA, Morgan HD, Petersen-Mahrt SK. Genetic and in vitro assays of DNA deamination. Methods Enzymol (2006) 408:156–70. doi:10.1016/s0076-
3227(06)08010-4
35. Chelico L, Sacho EJ, Erie DA, Goodman MF. A model for oligomeric regulation of APOBEC3G cytosine deaminase-dependent restriction of HIV. J Biol Chem (2008) 283(20):13780–91. doi:10.1074/jbc.M801004200
36. Shiyakhtenko LS, Lushnikov AV, Miyagi A, Li M, Harris RS, Lyubchenko YL. Atomic force microscopy studies of APOBECG oligomerization and dynamics. J Struct Biol (2013) 184(2):217–25. doi:10.1016/j.jsb.2013.09.008
37. Pham P, Afif SA, Shimoda M, Maeda K, Sakaguchi N, Pedersen LC, et al. Structural analysis of the activation-induced deoxycytidine deaminase required in immunoglobulin diversification. Biochemistry (2016) 55(21):2944–59. doi:10.1021/acs.biochem.6b00382
38. Prochnow C, Bransteitter R, Klein MG, Goodman MF, Chen XS. The DNA deaminase domain of the HIV-1 restriction factor APOBEC3G. Structure (2015) 23(9):2978–86. doi:10.1016/j.str.2015.04.031
39. King JJ, Manuel CA, Barrett CV, Raber S, Lucas H, Sutter P, et al. Catalytic pocket inaccessibility of activation-induced cytidine deaminase is a safeguard against excessive mutagenic activity. Structure (2015) 23(4):615–27. doi:10.1016/j.str.2015.01.016
40. Trott O, Olson AJ. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multi-threading. J Comput Chem (2010) 31(2):455–61. doi:10.1002/jcc.21334
41. Pham P, Calabrese P, Park SJ, Goodman MF. Analysis of a single-stranded DNA-scanning process in which activation-induced deoxycytidine deaminase (AID) deaminates C to U haphazardly and inefficiently to ensure mutational diversity. J Biol Chem (2011) 286(28):24931–42. doi:10.1074/jbc.
M111.241208
42. Bar-Even A, Noor E, Savir Y, Lieberman D, Avivi-David T, Tawfik DS, et al. The moderately efficient enzyme: evolutionary and physicochemical trends shaping enzyme parameters. Biochemistry (2011) 50(21):4402–10. doi:10.1021/bi2012829
43. Wang M, Yang Z, Rada C, Neuberger MS. AID upmutants isolated using a high-throughput screen highlight the immunity/cancer balance limiting DNA deaminase activity. Nat Struct Mol Biol (2009) 16(7):769–76. doi:10.1038/nsmb.1623
44. Dancyger AM, King JJ, Quinlan MJ, Fifield H, Tucker S, Saunders HL, et al. Differences in the enzymatic efficiency of human and bony fish AID are mediated by a single residue in the C terminus modulating single-stranded DNA binding. FASEB J (2012) 26(4):1517–25. doi:10.1096/fj.11-
198135
45. Abdouni H, King JJ, Suliman M, Quinlan M, Fifield H, Larjani M. Zebrafish AID is capable of deaminating methylated deoxycytidines. Nucleic Acids Res (2013) 41(10):5457–68. doi:10.1093/nar/gkt212
46. Barreto VM, Magor BG. Activation-induced cytidine deaminase structure and functions: a species comparative view. Dev Comp Immunol (2011) 35(9):991–1007. doi:10.1016/j.dci.2011.02.005
47. Mak CH, Pham P, Afif SA, Goodman MF. A mathematical model for scanning and catalysis on single-stranded DNA, illustrated with activation-induced deoxycytidine deaminase. J Biol Chem (2013) 288(11):29786–95. doi:10.1074/jbc.
M112.405565
48. Carpenter MA, Rajagurubandara E, Wijesinghe P, Bhagwat AS. Determinants of sequence-specificity within human AID and APOBEC3G. DNA Repair (Amst) (2010) 9(5):579–87. doi:10.1016/j.dnarep.2010.02.010
49. Carpenter MA, Li M, Rathore A, Lackey L, Law EK, Land AM, et al. Methylosynthetic and normal cytosine deamination by the foreign DNA
tethering of AID to the donor switch region during immunoglobulin class switch recombination. *J Exp Med* (2011) 208(8):1649–60. doi:10.1084/jem.20111118

88. Ueki T, King K, Okawa K, Kobayashi M, Yoshikawa K, Kawai S, Nakaoka H, et al. Histone chaperone Spt6 is required for class switch recombination but not somatic hypermutation. *Proc Natl Acad Sci U S A* (2011) 108(19):7920–5. doi:10.1073/pnas.1104423108

89. Basu U, Chaudhuri J, Alpert C, Dutt S, Ranganath S, Li G, et al. The AID antibody diversification enzyme is regulated by protein kinase A phosphorylation. *Nature* (2005) 438(7067):508–11. doi:10.1038/nature04255

90. Pasqualeucci L, Kitaura Y, Gu H, Dalla-Favera R. PkA-mediated phosphorylation regulates the function of activation-induced deaminase (AID) in B cells. *Proc Natl Acad Sci U S A* (2006) 103(2):395–400. doi:10.1073/pnas.0509969103

91. Demorest ZL, MacDuff DA, Brown WL, Morham SG, Parise LV, Harris RS. The interaction between AID and CIB1 is nonessential for antibody gene diversification by gene conversion or class switch recombination. *PLoS One* (2010) 5(7):e11660. doi:10.1371/journal.pone.0011660

92. MacDuff DA, Neuberger MS, Harris RS. MDM2 can interact with the C-terminus of AID but it is inessential for antibody diversification in DT40 B cells. *Mol Immunol* (2006) 43(8):1099–108. doi:10.1016/j.molimm.2005.07.024

93. Nowak U, Matthews AJ, Zheng S, Chaudhuri J. The splicing regulator PTBP2 interacts with the cytidine deaminase AID and promotes binding of AID to switch-region DNA. *Nat Immunol* (2011) 12(2):160–6. doi:10.1038/ni.1977

94. Hu W, Regum NA, Monda S, Stanile A, Honjo T. Identification of DNA cleavage- and recombination-specific hnRNP cofactors for activation-induced cytidine deaminase. *Proc Natl Acad Sci U S A* (2015) 112(18):5791–6. doi:10.1073/pnas.1506171112

95. Delker RK, Zhou Y, Strikoudis A, Stebbins CE, Papavasiliou FN. Solubility-based genetic screen identifies RING finger protein 126 as an E3 ligase for activation-induced cytidine deaminase. *Proc Natl Acad Sci U S A* (2013) 110(3):1029–34. doi:10.1073/pnas.1214581110

96. Kracker S, Di Virgilio M, Schwartzentruer I, Cuenin C, Forville M, Deau MC, et al. An inherited immunoglobulin class-switch recombination deficiency associated with a defect in the INO80 chromatin remodeling complex. *J Allergy Clin Immunol* (2015) 135(4):579–1. doi:10.1016/j.jaci.2014.08.030

97. Willmann KL, Milosevic S, Pauklin S, Schmitz KM, Rangam S, Simon MT, et al. A role for the RNA pol II-associated PAF complex in AID-induced immune diversification. *J Exp Med* (2012) 209(11):2099–111. doi:10.1084/jem.20112145

98. Xu Z, Fulop Z, Wu G, Pone EJ, Zhang J, Mai T, et al. 14–3–3 adaptepr proteins recruit AID to 5′-AGCT-3′ rich switch regions for class switch recombination. *Nat Struct Mol Biol* (2017) 10(7):1124–35. doi:10.1038/nsmb.3884

99. Kanehiro Y, Todo K, Negishi M, Fukuda G, Gan W, Hikasa T, et al. Activation-induced cytidine deaminase (AID)-dependent somatic hypermutation requires a splice isoform of the serine/arginine-rich (SR) protein SRSF1. *Proc Natl Acad Sci U S A* (2012) 109(4):1216–21. doi:10.1073/pnas.1120368110

100. Bohn MF, Shandilya SM, Silvias TV, Nalivka EA, Kouno T, Kelch BA, et al. The ssDNA mutator APOBEC3A is regulated by cooperative dimerization. *Structure* (2015) 23(5):903–11. doi:10.1016/j.str.2015.03.016

101. Qian J, Wang Q, Dose M, Pruett N, Kieffer-Kwon KR, Resch W, et al. B cell super-enhancers and regulatory clusters recruit AID to tumorigenic could. *Cell* (2014) 159(7):1524–37. doi:10.1016/j.cell.2014.11.013

102. Alikinula J, Schatz DG. Super-enhancer transcription converges on AID. *Cell* (2014) 159(7):1490–2. doi:10.1016/j.cell.2014.12.007

103. Meng FL, Du Z, Federation A, Hu J, Wang Q, Kieffer-Kwon KR, et al. Convergent transcription at intragenic super-enhancers targets AID-initiated genomic instability. *Cell* (2014) 159(7):1538–48. doi:10.1016/j.cell.2014.11.014

104. Martin A, Bardwell PD, Woo CJ, Fan M, Shulman MJ, ScharffMD. Activation-induced cytidine deaminase turns on somatic hypermutation in hybridomas. *Nature* (2002) 415(6873):862–6. doi:10.1038/nature00714

105. Pavri R, Nussenzweig MC. AID targeting in antibody diversity. *Adv Immunol* (2011) 110:1–26. doi:10.1016/B978-0-12-387663-8.00003-3

106. Wang CL, Harper RA, Wahl M. Genome-wide somatic hypermutation. *Proc Natl Acad Sci U S A* (2004) 101(19):7352–6. doi:10.1073/pnas.0402099101
E. coli Vif protein blocks the cytidine deaminase activity of B-cell specific AID in and without protein degradation. J 3G-mediated cytidine deamination by using a single amino acid interaction in residue contacts between X-ray- and NMR-resolved protein structures? Comparison of X-ray and NMR structures: is there a systematic difference in closure of the catalytic site of human CD38 induced by calcium.

Garbuzynskiy SO, Melnik BS, Lobanov MY, Finkelstein AV, Galzitskaya OV. Identification of specific determinants of human APOBEC3F, APOBEC3C, and APOBEC3DE and African green monkey APOBEC3F that interact with HIV-1 Vif. J Virol (2010) 84(4):12599–608. doi:10.1128/JVI.01437-10

Witkopp CJ, Adolph MB, Wu LL, Chelico L, Emerman M. A single nucleotide polymorphism in human APOBEC3C enhances restriction of lentiviruses. PLoS Pathog (2016) 12(10):e1005865. doi:10.1371/journal.ppat.1005865

Wichroski MJ, Robb GB, Rana TM. Human retroviral host restriction factors APOBEC3G and APOBEC3F localize to mRNA processing bodies. PLoS Pathog (2006) 2(5):e111. doi:10.1371/journal.ppat.0020041

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2017 King and Larijani. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction which does not comply with these terms.