



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

Justin J. King and Mani Larjani*

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

OPEN ACCESS

Edited by:

Harry W. Schroeder,
University of Alabama at
Birmingham, USA

Reviewed by:

Linda Chelico,
University of Saskatchewan,
Canada

Silvo G. Conticello,
Istituto Toscano Tumori, Italy

*Correspondence:

Mani Larjani
mlarjani@mun.ca

Specialty section:

This article was submitted to
B Cell Biology,
a section of the journal
Frontiers in Immunology

Received: 07 November 2016

Accepted: 10 March 2017

Published: 06 April 2017

Citation:

King JJ and Larjani M (2017)
A Novel Regulator of Activation-
Induced Cytidine Deaminase/
APOBECs in Immunity and Cancer:
Schrödinger's CATalytic Pocket.
Front. Immunol. 8:351.
doi: 10.3389/fimmu.2017.00351

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

IMPORTANCE AND CHALLENGES OF SOLVING AID/APOBEC STRUCTURES

Activation-induced cytidine deaminase (AID) is a 198 amino acid DNA-editing enzyme that deaminates deoxycytidine (dC) to deoxyuridine (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 (58) (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 distantly 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 a 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

TABLE 1 | All X-ray and NMR solution structures of the APOBEC family.

APOBEC	Experimental method	Truncations	Mutations	PDB ID
Hs-A2	X-ray	Truncated (Δ 1–40)	N/A	2NYT
Mouse-A2	NMR solution	Truncated (Δ 1–45)	N/A	2RPZ
Hs-A3A	NMR solution	N/A	N/A	2M65
Hs-A3A	X-ray	N/A	E72A, C171A	4XXO
Hs-A3A	X-ray	Truncated (Δ 196–199)	E72A	5SWW
Hs-A3B	X-ray	Truncated (Δ 1–186, Δ 242–248)	F200S, W228S, L230K, Y250S, F308K	5CQK
Hs-A3B	NMR solution	Truncated (Δ 1–186)	N/A	2NBQ
Hs-A3B	X-ray	Truncated (Δ 1–186, Δ 205–207, Δ 242–249)	F200S, V205G, L209I, R210G, R212H, Q213K, W228S, L230K, Y250S, E255A, F308K	5TD5
Hs-A3C	X-ray	N/A	N/A	3VOW
Hs-A3F	X-ray	Truncated (Δ 1–184)	Y196D, H247G, C248R, C259A, F302K, W310D, Y314A, Q315A, K355D, K358D, F363D	4IOU
Hs-A3F	X-ray	Truncated (Δ 1–217)	N/A	4J4J
Hs-A3F	X-ray	Truncated (Δ 1–186)	N/A	3WUS
Hs-A3F	X-ray	Truncated (Δ 1–184)	Y196D, H247G, C248R, C259A, F302K, W310D, K355D, K358D, F363D	5HX5
Hs-A3F (Zn-depleted)	X-ray	Truncated (Δ 1–184)	Y196D, H247G, C248R, C259A, F302K, W310D, K355D, K358D, F363D	5HX4
Hs-A3G	NMR solution	Truncated (Δ 1–197)	L234K, C243A, F310K, C321A, C356A	2JYW
Hs-A3G	X-ray	Truncated (Δ 1–196)	N/A	3E1U
Hs-A3G	NMR solution	Truncated (Δ 1–192)	N/A	2KBO
Hs-A3G	NMR solution	Truncated (Δ 1–190)	L234K, C243A, F310K, C321A, C356A	2KEM
Hs-A3G	X-ray	Truncated (Δ 1–194)	L234K, C243A, F310K, C321A, C356A	3IR2
Hs-A3G	X-ray	Truncated (Δ 1–194)	L234K, C243A, F310K, C356A	3V4K
Hs-A3G	X-ray	Truncated (Δ 1–192)	D370A	4ROW
Hs-A3G	NMR	Truncated (Δ 1–11, Δ 78, Δ 143–146, Δ 197–384)	Y13D, R14P, Y22N, L62D, F71L, H72S, W73L, F74V, T101A, A109Q, D110P, P111T, K112H, F126A, C139A, K141A, R142G, M149I, R169G, E170A, L171P, E173Q, N176D, N177G, P179D, K180E, Y181H, Y182S, I183Q, L184A, H186S, I187G, M189R	2MZZ
Primate-A3G	X-ray	Truncated (Δ 139–146, Δ 197–384)	(C139–Q140–K141–R142–D143–G144–P145–H146) replaced with (A–E–A–G) residues	5K83
Hs-AID	X-ray	Truncated (Δ 1–4, Δ 20–22, Δ 184–198)	N7D, R8P, R9H, K10I, L12T, Y13S, Q14N, K16N, V18G, R19I, R25H, E26K, V32E, K34E, R36L	5JJ4

Most APOBEC structures are heavily modified through mutations and/or truncations. Δ denotes the amino acids that were deleted from the structure. In cases where a study has reported several protein databank IDs of highly similar structures, a representative PDB code is listed.

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

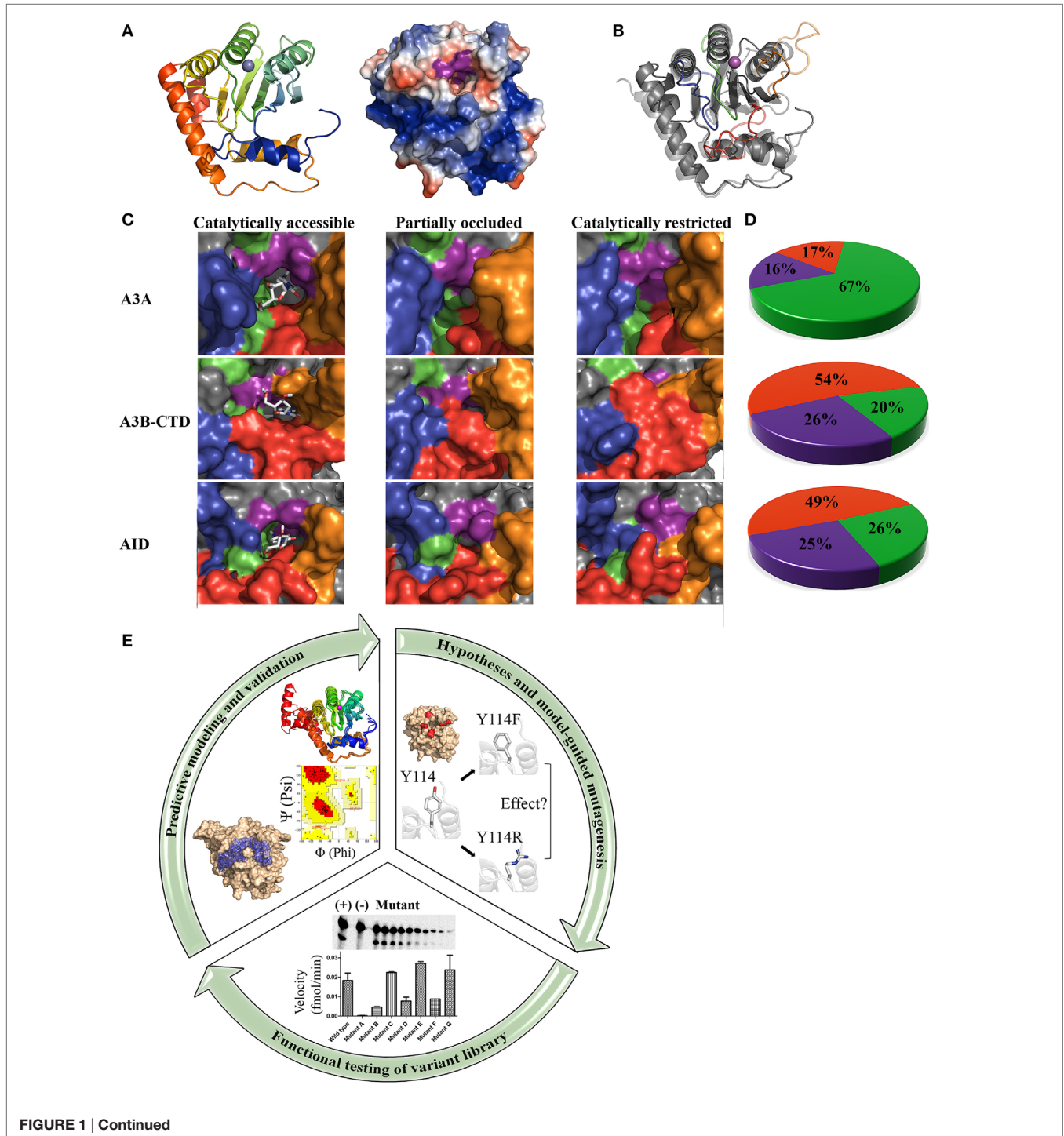


FIGURE 1 | Continued

FIGURE 1 | Continued**Core architecture, catalytic pocket occlusion, and computational/biochemical approaches to solving activation-induced cytidine deaminase (AID)/APOBEC structures.**

(A) Representative ribbon structure (left) and surface topology (right) of AID modeled from APOBEC templates. In the ribbon structure, N- to C-termini progression is shown from blue to red and the gray sphere depicts active site zinc. In the surface topology, positive, negative, and neutral residues have blue, red, and white surfaces, respectively. The Zn-coordinating residues and catalytic glutamic acid surface are colored purple. A distinct feature of AID among the APOBECs is its high positive charge at neutral pH, concentrated along two single-stranded DNA (ssDNA)-binding grooves that pass over the catalytic pocket. **(B)** Ribbon structures of A3A (transparent) and AID (non-transparent) were superimposed. In each protein structure, the secondary catalytic loops 2, 4, 6, and 8 are colored red, orange, green, and blue, respectively. **(C)** Catalytically accessible (left), partially occluded (middle), and catalytically restricted (right) conformations of A3A (top), A3B-CTD (center), and AID (bottom). The surface of secondary catalytic loop 2, 4, 6, and 8 were colored red, orange, green, and blue, respectively, to correspond with the ribbon structure shown in panel B. Catalytically accessible conformations are shown with bound dC in the catalytic pocket. Conformations were deemed catalytically accessible if they bound dC in a deamination-feasible configuration in the catalytic pocket via molecular docking [AutoDock VINA (59)]. In catalytically restricted conformations, the secondary catalytic loops adopt a configuration that block the pocket. **(D)** Proportion of catalytically restricted (red), partially occluded (purple), and catalytically accessible (green) conformations in A3A (top), A3B-CTD (center), and AID (bottom). A3A showed a dramatically higher proportion of catalytically accessible conformations in comparison to A3B-CTD and AID. NMR conformations of A3A (PDB: 2M65), A3B (PDB: 2NBQ), and previously reported structures of AID (58) were used. **(E)** Combinatorial computational/biochemical approach for solution of functional and native enzyme structures. A library of thousands of predicted structures is generated through homology modeling with a range of suitable template structures, generating multiple low energy conformations. The resulting conformational ensemble is then evaluated mathematically (e.g., Ramachandran and other means of evaluating model quality). Models are also checked for concordance with known biochemical properties of the enzyme. Molecular docking can be used to determine the substrate binding regions of the active site and surrounding regions. Concurrently, specific hypotheses are formed based on the highest confidence predicted conformations and their interaction with substrate. To test these hypotheses and to validate the positions and relative attitudes of specific core or surface residues, a large variant library ought to be constructed and tested in functional enzyme assays. This library can include point mutants, multiple mutants, orthologous and chimeric versions of the enzyme. For key residues involved in catalysis regulation, several point mutations spanning the range of synonymous to severe are more informative. Functional evaluation of this variant library ought to be used to confirm the involvement of key residues/motifs in specific biochemical aspects such as substrate binding, catalysis, and structural stability. Collectively, information from functional testing of the variant library is used to refine and validate the predicted enzyme structure and its interactions with substrate, to yield a functional and native structure.

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 promotor 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 ribonuclear complexes. The Vif protein of HIV binds and targets A3C, A3G, A3F, A3D, and A3H (to varying degrees) for degradation *via* a ubiquitin-dependent proteosomal 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 ribonuclear 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/APOBEC3s with known structure (-2, -6, +0.5, -9, and -3.5 of A3A, A3B-CTD, A3C, A3F-CTD, and A3G-CTD, respectively).

With this limitation, this structure nevertheless presents a unique opportunity for a comparison of structure determination methodologies. To this end, we compared the AID variant crystal structure with the computational-biochemical AID conformational ensemble (loop 2, 4, 6, and 8, denoted as loop 1, 3, 5, and 7, respectively, in other publications on APOBEC structures). Overall, the AID structures shared virtually the same tertiary structure and the variant structure confirmed some of the key secondary catalytic residues we posited would stabilize dC in the catalytic pocket (i.e., N51 and Y114). Most of the mutations in the AID variant were localized to L1, L2, and $\alpha 1$, regions, while $\alpha 7$ was deleted. There are also several notable differences: first, L2 adopts a more compact conformation relative to the core structure, likely due to the deletion of three residues in L2. We and others have previously shown that L2 plays a role in catalytic activity and AID:DNA binding of AID (58, 67). Second, L8 adopts a much more extended conformation in the AID variant. It was suggested this extended conformation stabilized larger purine bases upstream of the target cytidine, in contrast to other APOBECs whose shorter L8's preferred pyrimidines upstream.

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 $\alpha 7$ 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 L1- $\alpha 1$ -L2 together with $\alpha 7$ 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.

AUTHOR CONTRIBUTIONS

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

REFERENCES

- 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
- Bransteitter R, Pham P, Scharff MD, Goodman MF. Activation-induced cytidine deaminase deaminates deoxycytidine 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
- 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
- Sohail A, Klapacz 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/gkg464
- 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
- 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):8038–48. doi:10.1128/MCB.01046-07
- 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
- Revy P, Muto T, Levy Y, Geissmann 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)00079-9
- 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.2012.05.003
- Liu M, Duke JL, Richter DJ, Vinuesa 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
- 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
- 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
- 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
- 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
- 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/IgH translocations. *Cell* (2008) 135(6):1028–38. doi:10.1016/j.cell.2008.09.062
- Robbiani DF, Bunting S, Feldhahn N, Bothmer A, Camps J, 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
- 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.ccr.2009.07.030
- 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):E2977–86. doi:10.1073/pnas.1301021110
- Coticello SG. The AID/APOBEC family of nucleic acid mutators. *Genome Biol* (2008) 9(6):229. doi:10.1186/gb-2008-9-6-229
- Coticello 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
- Harris RS. Enhancing immunity to HIV through APOBEC. *Nat Biotechnol* (2008) 26(10):1089–90. doi:10.1038/nbt1008-1089
- Stavrou S, Ross SR. APOBEC3 proteins in viral immunity. *J Immunol* (2015) 195(10):4565–70. doi:10.4049/jimmunol.1501504
- Roberts SA, Lawrence MS, Klimczak LJ, Grimm SA, Fargo D, Stojanov P, et al. An APOBEC cytidine deaminase mutagenesis pattern is widespread in human cancers. *Nat Genet* (2013) 45(9):970–6. doi:10.1038/ng.2702
- Roberts SA, Gordenin DA. Clustered and genome-wide transient mutagenesis in human cancers: hypermutation without permanent mutators or loss of fitness. *Bioessays* (2014). doi:10.1002/bies.201300140
- Roberts SA, Gordenin DA. Hypermutation in human cancer genomes: footprints and mechanisms. *Nat Rev Cancer* (2014) 14(12):786–800. doi:10.1038/nrc3816
- 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
- Burns MB, Leonard B, Harris RS. APOBEC3B: pathological consequences of an innate immune DNA mutator. *Biomed J* (2015) 38(2):102–10. doi:10.4103/2319-4170.148904
- Chan K, Roberts SA, Klimczak LJ, Sterling JF, Saini N, Malc 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
- 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
- Leonard B, Hart SN, Burns MB, Carpenter MA, Temiz NA, Rathore A, et al. APOBEC3B upregulation and genomic mutation patterns in serous

- ovarian carcinoma. *Cancer Res* (2013) 73(24):7222–31. doi:10.1158/0008-5472.CAN-13-1753
31. Knisbacher BA, 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.2015.10.005
 32. Rebhendl S, Huemer M, Greil R, Geisberger R. AID/APOBEC deaminases and cancer. *Oncoscience* (2015) 2(4):320–33. doi:10.18632/oncoscience.155
 33. Starrett GJ, Luengas EM, McCann JL, Ebrahimi D, Temiz NA, Love RP, et al. The DNA cytosine deaminase APOBEC3H haplotype I likely contributes to breast and lung cancer mutagenesis. *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/S0076-6879(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. Shlyakhtenko LS, Lushnikov AY, Miyagi A, Li M, Harris RS, Lyubchenko YL. Atomic force microscopy studies of APOBEC3G 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. *DNA Repair (Amst)* (2016) 43:48–56. doi:10.1016/j.dnarep.2016.05.029
 38. Prochnow C, Bransteitter R, Klein MG, Goodman MF, Chen XS. The APOBEC-2 crystal structure and functional implications for the deaminase AID. *Nature* (2007) 445(7126):447–51. doi:10.1038/nature05492
 39. Krzysiak TC, Jung J, Thompson J, Baker D, Gronenborn AM. APOBEC2 is a monomer in solution: implications for APOBEC3G models. *Biochemistry* (2012) 51(9):2008–17. doi:10.1021/bi300021s
 40. Byeon IJ, Ahn J, Mitra M, Byeon CH, Hercik K, Hritz J, et al. NMR structure of human restriction factor APOBEC3A reveals substrate binding and enzyme specificity. *Nat Commun* (2013) 4:1890. doi:10.1038/ncomms2883
 41. Bohn MF, Shandilya SM, Albin JS, Kouno T, Anderson BD, McDougle RM, et al. Crystal structure of the DNA cytosine deaminase APOBEC3F: the catalytically active and HIV-1 Vif-binding domain. *Structure* (2013) 21(6):1042–50. doi:10.1016/j.str.2013.04.010
 42. Shi K, Carpenter MA, Kurahashi K, Harris RS, Aihara H. Crystal structure of the DNA deaminase APOBEC3B catalytic domain. *J Biol Chem* (2015) 290(47):28120–30. doi:10.1074/jbc.M115.679951
 43. Byeon IJ, Byeon CH, Wu T, Mitra M, Singer D, Levin JG, et al. Nuclear magnetic resonance structure of the APOBEC3B catalytic domain: structural basis for substrate binding and DNA deaminase activity. *Biochemistry* (2016) 55(21):2944–59. doi:10.1021/acs.biochem.6b00382
 44. Kitamura S, Ode H, Nakashima M, Imahashi M, Naganawa Y, Kurosawa T, et al. The APOBEC3C crystal structure and the interface for HIV-1 Vif binding. *Nat Struct Mol Biol* (2012) 19(10):1005–10. doi:10.1038/nsmb.2378
 45. Siu KK, Sultana A, Azimi FC, Lee JE. Structural determinants of HIV-1 Vif susceptibility and DNA binding in APOBEC3F. *Nat Commun* (2013) 4:2593. doi:10.1038/ncomms3593
 46. Nakashima M, Ode H, Kawamura T, Kitamura S, Naganawa Y, Awazu H, et al. Structural insights into HIV-1 Vif-APOBEC3F interaction. *J Virol* (2016) 90(2):1034–47. doi:10.1128/JVI.02369-15
 47. Shaban NM, Shi K, Li M, Aihara H, Harris RS. 1.92 angstrom zinc-free APOBEC3F catalytic domain crystal structure. *J Mol Biol* (2016) 428(11):2307–16. doi:10.1016/j.jmb.2016.04.026
 48. Chen KM, Harjes E, Gross PJ, Fahmy A, Lu Y, Shindo K, et al. Structure of the DNA deaminase domain of the HIV-1 restriction factor APOBEC3G. *Nature* (2008) 452(7183):116–9. doi:10.1038/nature06638
 49. Holden LG, Prochnow C, Chang YP, Bransteitter R, Chelico L, Sen U, et al. Crystal structure of the anti-viral APOBEC3G catalytic domain and functional implications. *Nature* (2008) 456(7218):121–4. doi:10.1038/nature07357
 50. Furukawa A, Nagata T, Matsugami A, Habu Y, Sugiyama R, Hayashi F, et al. Structure and real-time monitoring of the enzymatic reaction of APOBEC3G which is involved in anti-HIV activity. *Nucleic Acids Symp Ser (Oxf)* (2009) (53):87–8. doi:10.1093/nass/nrp044
 51. Harjes E, Gross PJ, Chen KM, Lu Y, Shindo K, Nowarski R, et al. An extended structure of the APOBEC3G catalytic domain suggests a unique holoenzyme model. *J Mol Biol* (2009) 389(5):819–32. doi:10.1016/j.jmb.2009.04.031
 52. Shandilya SM, Nalam MN, Nalivaika EA, Gross PJ, Valesano JC, Shindo K, et al. Crystal structure of the APOBEC3G catalytic domain reveals potential oligomerization interfaces. *Structure* (2010) 18(1):28–38. doi:10.1016/j.str.2009.10.016
 53. Li M, Shandilya SM, Carpenter MA, Rathore A, Brown WL, Perkins AL, et al. First-in-class small molecule inhibitors of the single-strand DNA cytosine deaminase APOBEC3G. *ACS Chem Biol* (2012) 7(3):506–17. doi:10.1021/cb200440y
 54. Lu X, Zhang T, Xu Z, Liu S, Zhao B, Lan W, et al. Crystal structure of DNA cytidine deaminase APOBEC3G catalytic deamination domain suggests a binding mode of full-length enzyme to single-stranded DNA. *J Biol Chem* (2015) 290(7):4010–21. doi:10.1074/jbc.M114.624262
 55. Kouno T, Luengas EM, Shigematsu M, Shandilya SM, Zhang J, Chen L, et al. Structure of the Vif-binding domain of the antiviral enzyme APOBEC3G. *Nat Struct Mol Biol* (2015) 22(6):485–91. doi:10.1038/nsmb.3033
 56. Xiao X, Li SX, Yang H, Chen XS. Crystal structures of APOBEC3G N-domain alone and its complex with DNA. *Nat Commun* (2016) 7:12193. doi:10.1038/ncomms12193
 57. Shi K, Carpenter MA, Banerjee S, Shaban NM, Kurahashi K, Salamango DJ, et al. Structural basis for targeted DNA cytosine deamination and mutagenesis by APOBEC3A and APOBEC3B. *Nat Struct Mol Biol* (2017) 24(2):131–9. doi:10.1038/nsmb.3344
 58. 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
 59. 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
 60. 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
 61. Bar-Even A, Noor E, Savir Y, Liebermeister W, Davidi D, Tawfik DS, et al. The moderately efficient enzyme: evolutionary and physicochemical trends shaping enzyme parameters. *Biochemistry* (2011) 50(21):4402–10. doi:10.1021/bi2002289
 62. 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
 63. 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
 64. 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
 65. 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
 66. 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(41):29786–95. doi:10.1074/jbc.M113.506550
 67. 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
 68. Carpenter MA, Li M, Rathore A, Lackey L, Law EK, Land AM, et al. Methylcytosine and normal cytosine deamination by the foreign DNA

- restriction enzyme APOBEC3A. *J Biol Chem* (2012) 287(41):34801–8. doi:10.1074/jbc.M112.385161
69. Fu Y, Ito F, Zhang G, Fernandez B, Yang H, Chen XS. DNA cytosine and methylcytosine deamination by APOBEC3B: enhancing methylcytosine deamination by engineering APOBEC3B. *Biochem J* (2015) 471(1):25–35. doi:10.1042/BJ20150382
 70. Marx A, Galilee M, Alian A. Zinc enhancement of cytidine deaminase activity highlights a potential allosteric role of loop-3 in regulating APOBEC3 enzymes. *Sci Rep* (2015) 5:18191. doi:10.1038/srep18191
 71. Kohli RM, Abrams SR, Gajula KS, Maul RW, Gearhart PJ, Stivers JT. A portable hot spot recognition loop transfers sequence preferences from APOBEC family members to activation-induced cytidine deaminase. *J Biol Chem* (2009) 284(34):22898–904. doi:10.1074/jbc.M109.025536
 72. Love RP, Xu H, Chelico L. Biochemical analysis of hypermutation by the deoxycytidine deaminase APOBEC3A. *J Biol Chem* (2012) 287(36):30812–22. doi:10.1074/jbc.M112.393181
 73. Faltas BM, Prandi D, Tagawa ST, Molina AM, Nanus DM, Sternberg C, et al. Clonal evolution of chemotherapy-resistant urothelial carcinoma. *Nat Genet* (2016) 48(12):1490–9. doi:10.1038/ng.3692
 74. Rai K, Huggins IJ, James SR, Karpf AR, Jones DA, Cairns BR. DNA demethylation in zebrafish involves the coupling of a deaminase, a glycosylase, and gadd45. *Cell* (2008) 135(7):1201–12. doi:10.1016/j.cell.2008.11.042
 75. Orthwein A, Zahn A, Methot SP, Godin D, Conticello SG, Terada K, et al. Optimal functional levels of activation-induced deaminase specifically require the Hsp40 DnaJ1. *EMBO J* (2012) 31(3):679–91. doi:10.1038/emboj.2011.417
 76. Orthwein A, Patenaude AM, Affar el B, Lamarre A, Young JC, Di Noia JM. Regulation of activation-induced deaminase stability and antibody gene diversification by Hsp90. *J Exp Med* (2010) 207(12):2751–65. doi:10.1084/jem.20101321
 77. McBride KM, Barreto V, Ramiro AR, Stavropoulos P, Nussenzweig MC. Somatic hypermutation is limited by CRM1-dependent nuclear export of activation-induced deaminase. *J Exp Med* (2004) 199(9):1235–44. doi:10.1084/jem.20040373
 78. Han L, Masani S, Yu K. Cutting edge: CTNNB1 is dispensable for Ig class switch recombination. *J Immunol* (2010) 185(3):1379–81. doi:10.4049/jimmunol.1001643
 79. Conticello SG, Ganesh K, Xue K, Lu M, Rada C, Neuberger MS. Interaction between antibody-diversification enzyme AID and spliceosome-associated factor CTNNB1. *Mol Cell* (2008) 31(4):474–84. doi:10.1016/j.molcel.2008.07.009
 80. Uchimura Y, Barton LF, Rada C, Neuberger MS. REG-gamma associates with and modulates the abundance of nuclear activation-induced deaminase. *J Exp Med* (2011) 208(12):2385–91. doi:10.1084/jem.20110856
 81. Ranjit S, Khair L, Linehan EK, Ucher AJ, Chakrabarti M, Schrader CE, et al. AID binds cooperatively with UNG and Msh2-Msh6 to Ig switch regions dependent upon the AID C terminus. *J Immunol* (2011) 187(5):2464–75. doi:10.4049/jimmunol.1101406
 82. Patenaude AM, Orthwein A, Hu Y, Campo VA, Kavli B, Buschiazio A, et al. Active nuclear import and cytoplasmic retention of activation-induced deaminase. *Nat Struct Mol Biol* (2009) 16(5):517–27. doi:10.1038/nsmb.1598
 83. Hasler J, Rada C, Neuberger MS. Cytoplasmic activation-induced cytidine deaminase (AID) exists in stoichiometric complex with translation elongation factor 1alpha (eEF1A). *Proc Natl Acad Sci U S A* (2011) 108(45):18366–71. doi:10.1073/pnas.1106729108
 84. Pavri R, Gazumyan A, Jankovic M, Di Virgilio M, Klein I, Ansarah-Sobrinho C, et al. Activation-induced cytidine deaminase targets DNA at sites of RNA polymerase II stalling by interaction with Spt5. *Cell* (2010) 143(1):122–33. doi:10.1016/j.cell.2010.09.017
 85. Basu U, Meng FL, Keim C, Grinstein V, Pefanis E, Eccleston J, et al. The RNA exosome targets the AID cytidine deaminase to both strands of transcribed duplex DNA substrates. *Cell* (2011) 144(3):353–63. doi:10.1016/j.cell.2011.01.001
 86. Chaudhuri J, Khuong C, Alt FW. Replication protein A interacts with AID to promote deamination of somatic hypermutation targets. *Nature* (2004) 430(7003):992–8. doi:10.1038/nature02821
 87. Jeevan-Raj BP, Robert I, Heyer V, Page A, Wang JH, Cammas F, et al. Epigenetic 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.20110118
 88. Okazaki IM, Okawa K, Kobayashi M, Yoshikawa K, Kawamoto S, Nagaoka 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. Pasqualucci 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, Begum NA, Mondal S, Stanlie 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.1506167112
 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.1214538110
 96. Kracker S, Di Virgilio M, Schwartzentruber J, Cuenin C, Forveille 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):998–1007.e6. doi:10.1016/j.jaci.2014.08.030
 97. Willmann KL, Milosevic S, Pauklin S, Schmitz KM, Rangam G, 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 adaptor proteins recruit AID to 5'-AGCT-3'-rich switch regions for class switch recombination. *Nat Struct Mol Biol* (2010) 17(9):1124–35. doi:10.1038/nsmb.1884
 99. Kanehiro Y, Todo K, Negishi M, Fukuoka J, Gan W, Hikasa T, et al. Activation-induced cytidine deaminase (AID)-dependent somatic hypermutation requires a cytidine 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.1120368109
 100. Bohn MF, Shandilya SM, Silvas TV, Nalivaika 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 tumorigenic activity. *Cell* (2014) 159(7):1524–37. doi:10.1016/j.cell.2014.11.013
 102. Alinikula 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, Scharff MD. Activation-induced cytidine deaminase turns on somatic hypermutation in hybridomas. *Nature* (2002) 415(6873):802–6. doi:10.1038/nature714
 105. Pavri R, Nussenzweig MC. AID targeting in antibody diversity. *Adv Immunol* (2011) 110:1–26. doi:10.1016/B978-0-12-387663-8.00005-3
 106. Wang CL, Harper RA, Wabl M. Genome-wide somatic hypermutation. *Proc Natl Acad Sci U S A* (2004) 101(19):7352–6. doi:10.1073/pnas.0402009101

107. Parsa JY, Basit W, Wang CL, Gommerman JL, Carlyle JR, Martin A. AID mutates a non-immunoglobulin transgene independent of chromosomal position. *Mol Immunol* (2007) 44(4):567–75. doi:10.1016/j.molimm.2006.02.003
108. Maeda K, Almofty SA, Singh SK, Eid MM, Shimoda M, Ikeda T, et al. GANP interacts with APOBEC3G and facilitates its encapsidation into the virions to reduce HIV-1 infectivity. *J Immunol* (2013) 191(12):6030–9. doi:10.4049/jimmunol.1302057
109. Wang X, Ao Z, Chen L, Kobinger G, Peng J, Yao X. The cellular antiviral protein APOBEC3G interacts with HIV-1 reverse transcriptase and inhibits its function during viral replication. *J Virol* (2012) 86(7):3777–86. doi:10.1128/JVI.06594-11
110. Schafer A, Bogerd HP, Cullen BR. Specific packaging of APOBEC3G into HIV-1 virions is mediated by the nucleocapsid domain of the gag polyprotein precursor. *Virology* (2004) 328(2):163–8. doi:10.1016/j.virol.2004.08.006
111. Valera MS, de Armas-Rillo L, Barroso-Gonzalez J, Ziglio S, Batisse J, Dubois N, et al. The HDAC6/APOBEC3G complex regulates HIV-1 infectiveness by inducing Vif autophagic degradation. *Retrovirology* (2015) 12:53. doi:10.1186/s12977-015-0181-5
112. Yu X, Yu Y, Liu B, Luo K, Kong W, Mao P, et al. Induction of APOBEC3G ubiquitination and degradation by an HIV-1 Vif-Cul5-SCF complex. *Science* (2003) 302(5647):1056–60. doi:10.1126/science.1089591
113. Sheehy AM, Gaddis NC, Malim MH. The antiretroviral enzyme APOBEC3G is degraded by the proteasome in response to HIV-1 Vif. *Nat Med* (2003) 9(11):1404–7. doi:10.1038/nm945
114. Santa-Marta M, da Silva FA, Fonseca AM, Goncalves J. HIV-1 Vif can directly inhibit apolipoprotein B mRNA-editing enzyme catalytic polypeptide-like 3G-mediated cytidine deamination by using a single amino acid interaction and without protein degradation. *J Biol Chem* (2005) 280(10):8765–75. doi:10.1074/jbc.M409309200
115. Conticello SG, Harris RS, Neuberger MS. The Vif protein of HIV triggers degradation of the human antiretroviral DNA deaminase APOBEC3G. *Curr Biol* (2003) 13(22):2009–13. doi:10.1016/j.cub.2003.10.034
116. Santa-Marta M, Aires da Silva F, Fonseca AM, Rato S, Goncalves J. HIV-1 Vif protein blocks the cytidine deaminase activity of B-cell specific AID in *E. coli* by a similar mechanism of action. *Mol Immunol* (2007) 44(4):583–90. doi:10.1016/j.molimm.2006.02.005
117. Goncalves J, Santa-Marta M. HIV-1 Vif and APOBEC3G: multiple roads to one goal. *Retrovirology* (2004) 1:28. doi:10.1186/1742-4690-1-28
118. Mehle A, Strack B, Ancuta P, Zhang C, McPike M, Gabuzda D. Vif overcomes the innate antiviral activity of APOBEC3G by promoting its degradation in the ubiquitin-proteasome pathway. *J Biol Chem* (2004) 279(9):7792–8. doi:10.1074/jbc.M313093200
119. Smith JL, Pathak VK. 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(24):12599–608. doi:10.1128/JVI.01437-10
120. Wittkopp CJ, Adolph MB, Wu LI, 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
121. Wichroski MJ, Robb GB, Rana TM. Human retroviral host restriction factors APOBEC3G and APOBEC3F localize to mRNA processing bodies. *PLoS Pathog* (2006) 2(5):e41. doi:10.1371/journal.ppat.0020041
122. Chiu YL, Greene WC. APOBEC3G: an intracellular centurion. *Philos Trans R Soc Lond B Biol Sci* (2009) 364(1517):689–703. doi:10.1098/rstb.2008.0193
123. Chiu YL. Biochemical fractionation and purification of high-molecular-mass APOBEC3G complexes. *Methods Mol Biol* (2011) 718:185–206. doi:10.1007/978-1-61779-018-8_12
124. Gallois-Montbrun S, Holmes RK, Swanson CM, Fernandez-Ocana M, Byers HL, Ward MA, et al. Comparison of cellular ribonucleoprotein complexes associated with the APOBEC3F and APOBEC3G antiviral proteins. *J Virol* (2008) 82(11):5636–42. doi:10.1128/JVI.00287-08
125. Kreisberg JF, Yonemoto W, Greene WC. Endogenous factors enhance HIV infection of tissue naive CD4 T cells by stimulating high molecular mass APOBEC3G complex formation. *J Exp Med* (2006) 203(4):865–70. doi:10.1084/jem.20051856
126. Kozak SL, Marin M, Rose KM, Bystrom C, Kabat D. The anti-HIV-1 editing enzyme APOBEC3G binds HIV-1 RNA and messenger RNAs that shuttle between polysomes and stress granules. *J Biol Chem* (2006) 281(39):29105–19. doi:10.1074/jbc.M601901200
127. Monajemi M, Woodworth CF, Benkaroun J, Grant M, Larjani M. Emerging complexities of APOBEC3G action on immunity and viral fitness during HIV infection and treatment. *Retrovirology* (2012) 9:35. doi:10.1186/1742-4690-9-35
128. Belanger K, Langlois MA. RNA-binding residues in the N-terminus of APOBEC3G influence its DNA sequence specificity and retrovirus restriction efficiency. *Virology* (2015) 483:141–8. doi:10.1016/j.virol.2015.04.019
129. Chelico L, Prochnow C, Erie DA, Chen XS, Goodman MF. Structural model for deoxycytidine deamination mechanisms of the HIV-1 inactivation enzyme APOBEC3G. *J Biol Chem* (2010) 285(21):16195–205. doi:10.1074/jbc.M110.107987
130. Garbuzynskiy SO, Melnik BS, Lobanov MY, Finkelstein AV, Galzitskaya OV. Comparison of X-ray and NMR structures: is there a systematic difference in residue contacts between X-ray- and NMR-resolved protein structures? *Proteins* (2005) 60(1):139–47. doi:10.1002/prot.20491
131. Malinen AM, Turtola M, Parthiban M, Vainonen L, Johnson MS, Belogurov GA. Active site opening and closure control translocation of multisubunit RNA polymerase. *Nucleic Acids Res* (2012) 40(15):7442–51. doi:10.1093/nar/gks383
132. Liu Q, Graeff R, Kriksunov IA, Lam CM, Lee HC, Hao Q. Conformational closure of the catalytic site of human CD38 induced by calcium. *Biochemistry* (2008) 47(52):13966–73. doi:10.1021/bi801642q

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 Larjani. 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 is permitted which does not comply with these terms.