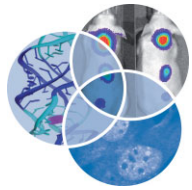
Advanced Review



The multifaceted roles of RNA binding in APOBEC cytidine deaminase functions

Kimberly M. Prohaska, Ryan P. Bennett, Jason D. Salter and Harold C. Smith*

Cytidine deaminases have important roles in the regulation of nucleoside/deoxynucleoside pools for DNA and RNA synthesis. The APOBEC family of cytidine deaminases (named after the first member of the family that was described, *Apo*lipoprotein *B* mRNA *E*diting *C*atalytic Subunit *1*, also known as APOBEC1 or A1) is a fascinating group of mutagenic proteins that use RNA and single-stranded DNA (ssDNA) as substrates for their cytidine or deoxycytidine deaminase activities. APOBEC proteins and base-modification nucleic acid editing have been the subject of numerous publications, reviews, and speculation. These proteins play diverse roles in host cell defense, protecting cells from invading genetic material, enabling the acquired immune response to antigens and changing protein expression at the level of the genetic code in mRNA or DNA. The amazing power these proteins have for interphase cell functions relies on structural and biochemical properties that are beginning to be understood. At the same time, the substrate selectivity of each member in the family and their regulation remains to be elucidated. This review of the APOBEC family will focus on an open question in regulation, namely what role the interactions of these proteins with RNA have in editing substrate recognition or allosteric regulation of DNA mutagenic and host-defense activities.

How to cite this article:
WIREs RNA 2014. doi: 10.1002/wrna.1226

INTRODUCTION

The Apolipoprotein B mRNA editing catalytic subunit (APOBEC) family consists of cytidine deaminases (CDAs) and is a subset of a much larger group of enzymes known to produce modifications of nucleosides/nucleotides and nucleic acids. They are collectively referred to as the RNA and DNA modification enzymes (for a comprehensive listing of enzymes and functions see Ref 1). Cytidine deamination involves a nucleophilic hydrolytic displacement of an amino group at the fourth position of the pyrimidine ring by a hydroxyl group donated from water, which produces a transition of a cytidine or deoxyctyidine to uridine or deoxyuridine. The zinc-dependent catalytic domain (ZDD) of the APOBEC proteins is structurally and functionally similar to that found in CDAs that use free nucleosides and nucleotides as substrates.[2,3]

The intriguing characteristic of APOBEC CDAs is that they identify their cytidine base substrates with remarkable preference within the context of RNA or ssDNA polynucleotides.[4–6] Consequently, not all cytidines or deoxycytidines within a strand of nucleic acid are modified. Moreover, there is cellular regulation at the level of a given population of nucleic acids such that the sites of preferred modification,

*Correspondence to: hsmith@oyageninc.com

OyaGen Inc, Rochester BioVenture Center, Rochester, NY, USA

Conflict of interest: KMP, RPB, and JDS are employees of OyaGen Inc, which is a drug development biotechnology company engaged in the therapeutic targeting of editing enzymes. Dr. Smith is the founder of OyaGen and, as a consultant, serves as CEO and CSO for OyaGen. Dr. Smith is a tenured full professor in the Department of Biochemistry and Biophysics, Center for RNA Biology and Cancer Center at the University of Rochester, School of Medicine and Dentistry in Rochester, NY.

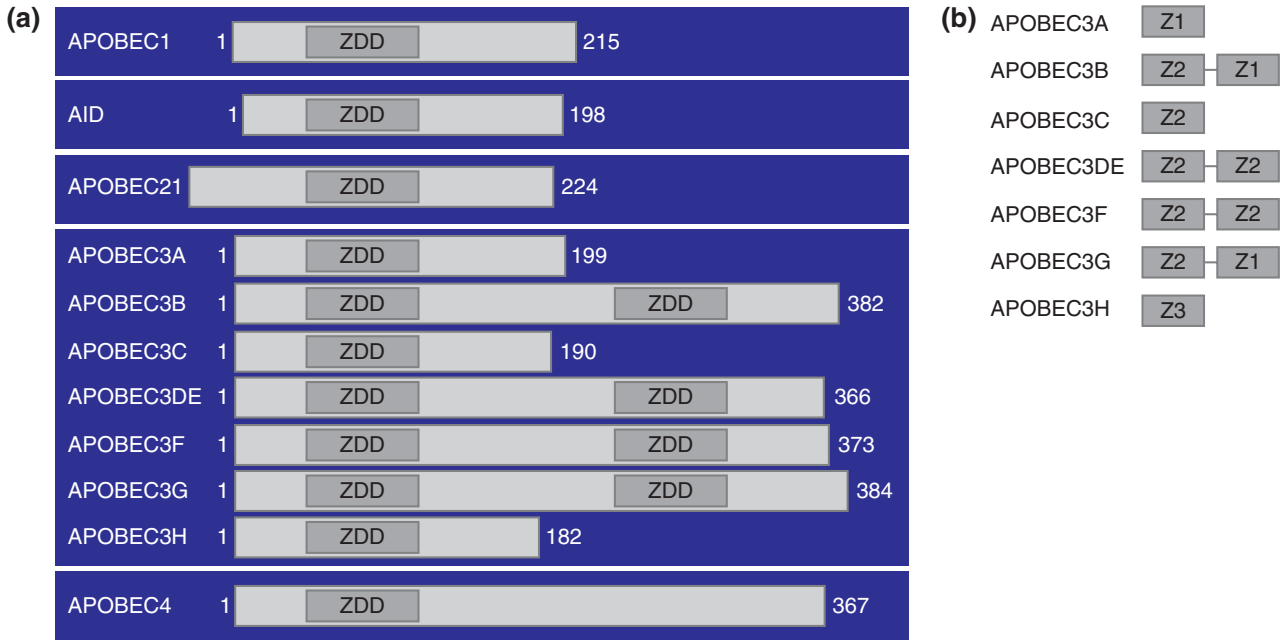


FIGURE 1 | The ZDD of the APOBEC family. (a) The ZDD sequence motif [(H/C)-x-E-$x_{(25–30)}$-P-C-xx-C] is conserved throughout the APOBEC family of cytidine deaminases. Several APOBEC3 members (B, DE, F, and G) have two ZDD motifs in tandem on a single polypeptide. (b) The CDA domains of the A3 family members are subcategorized as Z1, Z2, or Z3 type according to additional conserved sequences within the ZDD motif.

known as 'hot spots', are not always 100% modified within the population (referred to as the editing efficiency). Regulation of sequence targeting and editing efficiency is determined by diverse mechanisms, including developmental, hormonal, and metabolic regulation of APOBEC gene expression and protein post-translational modification, subcellular compartmentalization of APOBEC proteins, APOBEC binding to cofactors and targeted protein degradation.[7–9] Each member of the APOBEC family may be subjected to one or more of these control points, which will determine not only their functionality but also the tissues and the physiological circumstances during which their activity is called into service.

This review considers the APOBEC family from the foundation established by (1) research on A1 as an RNA editing enzyme in lipoprotein metabolism,[10,11] (2) the high-paced rush to discovery following the description of the indispensible role of activation-induced deaminase (AID) as a ssDNA mutation enzyme in immunoglobulin gene diversification[12] and (3) the anti-retroviral, host-defense function of the ssDNA mutational activity of A3G on HIV-1 (human immunodeficiency virus type 1) during reverse transcription.[13] The review examines the major unanswered questions in the field and what is known to differentiate the activities of APOBEC proteins with special attention paid to the role of RNA in these processes.

APOBEC1 (A1) THE RNA EDITING ENZYME

Comprehensive reviews have been written on A1 and the role this enzyme has in editing apolipoprotein B (ApoB) mRNA and the overall role editing has in lipoprotein particle assembly, secretion from liver and the small intestine and lipid metabolism.[14,15] Editing of the ApoB mRNA provided the first example of deaminase-dependent C to U base modification, a discovery that predates any descriptions of the APOBEC gene family.[10,11,16] (Figure 1). The discovery of A1 itself[17] was a difficult task, as this enzyme alone had little to no RNA editing activity and a low affinity for binding to RNA.[18] Biochemical evidence suggested a macromolecular complex, or editosome, was required for site-specific mRNA editing.[19] This hypothesis was controversial until the RNA-binding protein A1CF (A1 complementation factor, also known as ACF) was discovered[20,21] as the auxiliary protein that enabled A1-dependent, site-specific C to U editing. A1CF functions by binding to a conserved RNA motif, the mooring sequence,[22,23] and recruiting A1 to the editing site.[24–26] A1 and A1CF knockout mouse models revealed that A1CF was required for embryogenesis, whereas A1 was not an essential gene,[27,28] suggesting that A1CF is required for other processes. The function of A1 and A1CF can be modulated by the expression of RNA spliced variants[29–32] as well as through metabolic regulation of protein

post-translational modification that determines the subcellular localization of A1 and A1CF, editosome assembly and editing activity.[33–35]

The A1CF mooring sequence is positioned 3′ of the edited C within a tripartite motif that also includes AU-rich enhancer and spacer sequences located immediately 5′ and 3′ of the C, respectively.[14] Using this information, the mRNA encoding the human oncogene neurofibromin (NF1) was identified as an mRNA that was C to U edited.[36] The resulting stop codon truncated NF1, leading to a loss of its Ras regulatory domain. Transcriptome-wide RNA sequencing showed that numerous C to U editing events occurred within the 3′ UTR (untranslated region) of diverse mRNAs, nearly all of which have 3′ proximal mooring sequences.[37] Overexpression of A1 led to promiscuous (multiple sites in the same mRNA) and hyper-editing (multiple mRNAs) that can lead to neoplasia in rodents.[38–40]

A1-dependent RNA editing also has been described as restricting endogenous retroviral elements[41] and HIV.[42] However, independent of editing, A1 binding to AU-rich mooring and 3′ UTR sequences may regulate mRNA stability[37,43] and restrict the mobility of LINE-1 elements.[41] A1 and A1CF binding to mRNAs may regulate protein expression through changes in RNA stability and inhibition of nonsense codon-mediated decay.[28,44–46] How many RNAs are affected by A1 and A1CF and the mechanisms whereby they affect diversity in the transcriptome and proteome remain as open questions.

In addition to its role in RNA biology, A1 also may bind to ssDNA within transcribed regions of genes where it deaminates dC to dU, apparently without the need for an A1CF-like auxiliary protein.[5] When expressed in *Escherichia coli*, A1 is a potent mutagen and has a significantly higher DNA mutagenic activity than other members of the APOBEC family.[5,47] A1 ssDNA editing has been observed on herpes[48] and hepatitis[49] viral genomes and has been proposed to function as a host cell restriction factor. Although editing of mammalian chromosomal DNA by A1 has yet to be described, another member of the APOBEC family, activation induced cytidine deaminase (AID), primarily functions in mutating ssDNA within the transcribed regions of the immunoglobulin locus of B cells.[50] AID ssDNA editing enables antibody diversification through the generation of new antibody variable regions (somatic hypermutation) and heavy chain switching (class switch recombination). Tight control over AID expression and its nuclear import are among the many mechanisms whereby cells subvert AID promiscuous activity at other sites within the genome.[7,51–53] In contrast, A1 and A1CF expression are constitutive and they shuttle from the cytoplasm to the nucleus as part of normal cell physiology.[34,35,54,55] The question of how cells control A1 selection for cellular RNA over ssDNA as substrates and mitigate the oncogenic potential of A1 has never been addressed.

THE APOBEC FAMILY

ssDNA Deaminase Activity

The APOBEC3 (A3) proteins (A3A–A3H) share common features with A1 (Figure 1). The preferred substrate for A3 deaminases is deoxycytidine in ssDNA. It was known for some time that certain T cell lines could restrict HIV infection when the HIV protein viral infectivity factor (Vif) was not expressed (these cell lines were termed nonpermissive).[56,57] Heterokaryons between permissive and nonpermissive cells demonstrated that the latter expressed a viral restriction factor that was overcome by Vif.[58,59] Sheehy et al. utilized subtractive hybridization with two nearly isogenic cell lines, with one being permissive and the other nonpermissive, to identify A3G as the Vif-sensitive anti-HIV factor (originally called CEM15).[13] Permissive cells did not express A3G but introducing its cDNA into these cells was sufficient to render them nonpermissive.[13]

A3G incorporates with nascent viral particles and, upon infection, binds to newly synthesized HIV reverse transcripts (negative strand) and deaminates dC in hotspots (hypermutation).[60–63] This antiviral mechanism also was shown for A3B, A3C (for simian immunodeficiency virus, SIV), A3DE, A3F, and A3H (for HIV).[42,64] Although the C-terminal catalytic domain confers A3G deaminase activity, its antiviral deoxycytidine deaminase activity requires both domains.[16,65,66] APOBEC multimerization has been demonstrated by several techniques, including yeast two-hybrid analysis,[16] small angle X-ray scattering,[67] analytical ultracentrifugation,[68] protein cross-linking[16,65] and co-immunoprecipitation studies with A3G domain deletion constructs.[16,69,70] However, the requirement of multimerization for ssDNA binding and deamination remains unresolved (see section on *Further Reading*).

Although the A3 proteins use ssDNA substrates, they have different selectivities for the *cis*-acting sequences or structural context of the targeted deoxycytidine without an apparent requirement for *trans*-acting factors.[71] For example, A3F targets dC residues that have dT 5′ to the edited dC, while A3G prefers dC residues just 5′ of the editing site.[72] A3G and A3F dC to dU edited nucleotides are clustered within regions of the ssDNA that are distinct from

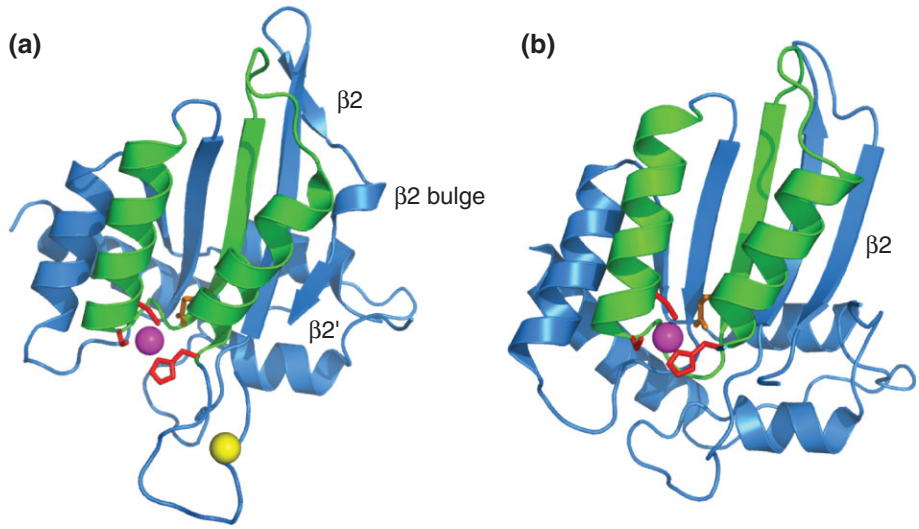


FIGURE 2 | Representative APOBEC crystal structures. Diagram of a representative crystal structure of (a) the C-terminal Z1-type CDA of A3G (PDB 3IR2) at 2.25 Å resolution[73] and (b) the C-terminal Z-2 type CDA of A3F (PDB 4IOU) at 2.75 Å resolution.[74] The $\alpha\beta\alpha$ supersecondary structural element (green) is shown embedded within the core CDA fold, comprising the five-stranded, mixed β-sheet surrounded by six α-helices, typical of the A3 family. Sidechains of the conserved zinc-binding residues and proton-shuttling glutamic acid are illustrated in red and orange, respectively. The catalytic zinc ion is represented as a purple sphere. The additional non-catalytic zinc ion present in the A3G structure is represented by the yellow sphere; it is coordinated by four residues, two from each of adjacent subunits within the crystal and may be an artifact due to crystal packing. The β2 bulge between β2 and β2′ is present in the structure of the Z1-type CDA of the A3G C-terminal CDA (a), but is noticeably absent from the Z2-type CDA of the A3F C-terminal CDA (b).

other APOBEC family members.[5] In addition to ssDNA binding, some A3 proteins are capable of binding to RNA, although they do not utilize it as a substrate for deaminase activity (see section on *Further Reading*).

Common and Distinctive Structures of Deaminases

The APOBEC protein family is part of a larger CDA superfamily that also includes enzymes involved in pyrimidine metabolism that act on free nucleosides, as well as adenosine deaminases that act on tRNA (ADATs) and mRNA (ADARs).[2,3] The distinguishing feature of CDA family members is the evolutionarily conserved ZDD motif [(H/C)xEx(25—30)PCx$_n$C] (Figure 2).[75,76] While cytidine and adenosine deaminases have no global sequence similarities, the conservation of their ZDD motifs suggests a common ancestor. A catalytic zinc ion is coordinated by the three C (or H) residues of the motif and by an activating water molecule that performs a hydrolytic attack on the C4 amine of the substrate cytidine or deoxycytidine. The glutamic acid acts as a proton shuttle during catalysis, while the proline directs the orientation of adjacent conserved residues through its conformational constriction. The catalytic mechanism of deamination is conserved throughout the CDA family and is described in detail for the *E. coli* CDA in Ref 3.

The conserved residues of the ZDD motif are located at adjacent N-terminal ends of two α-helices in an $\alpha\beta\alpha$ supersecondary structural feature that is embedded within the core CDA fold. This fold comprises a five-stranded, mixed β-sheet surrounded by three to six α-helices (Figure 2). CDAs that act on free nucleotides have three to five α-helices, while structures of APOBECs (discussed below) are embellished with a total of six α-helices, a defining feature of the family. In general, subtle structural variations of the CDA domain, including the number and spatial arrangement of helices, the topological order of secondary structural elements, and in particular, the orientation of strand β5, directly impact substrate selection. These structural variations define clades within the CDA family and shed light on the evolutionary relation among family members. The structural and topological variation of the CDA family has been expertly reviewed.[3] Notably, several A3 proteins (B, DE, F, and G) have two ZDD motifs (and presumably, two CDA domains) in tandem on the same polypeptide, a feature unique among all CDA members (Figure 1).

High-resolution structures of APOBEC proteins have not been as forthcoming as those of the free nucleotide CDAs. Nuclear magnetic resonance (NMR) solution structures or X-ray crystal structures of the single CDA domain A2,[77,78] A3A,[79] and A3C,[80] have

been solved. Likewise, crystal structures and NMR solution structures of the C-terminal CDA domain of the dual-deaminase domain A3G[81,82,73] and A3F[74] have been solved. While the topological features and core fold of the CDA family are conserved, there are distinct structural features inherent to the APOBEC family. Regions of positively charged and hydrophobic residues surrounding the zinc-centered active site are common to, but vary in extent among, the structures of the A3 proteins. These patches likely function to neutralize the negatively charged backbone during nucleic acid binding and base stack with nucleic acid substrate, respectively. Several NMR chemical shift perturbations implicated numerous residues in surface grooves adjacent to the active sites of A3G (C-terminal CDA) and A3A for binding a variety of ssDNA substrates.[81,82] However, these models are not consistent with one another, leaving the mode of nucleic acid binding ambiguous. Subtle differences in the length of secondary structural elements and loop regions, deletions/insertions of residues, and specific residues near the active site are likely the primary discriminators for sequence preference, substrate binding affinity and catalytic rate among the A3 family members. For example, differences in the loop between β4 and α4 of the conserved CDA may determine the nucleotide sequence preference surrounding the hot-spot motifs among A3s and other APOBECs[74,83] (see also section on *Further Reading*).

Oligomerization is a hallmark of CDA members that act on free nucleotides and is necessary for catalytic activity. Some CDA proteins, such as the yeast cytidine deaminase (Cdd1), form compact tetramers, burying large amounts of surface area at subunit interfaces.[84] Each subunit's active site requires *trans*-acting (functional groups) elements from adjacent subunits. Other family members, such as the cytosine, guanine, and deoxycytidylate deaminases, form active sites through either domain swapping or subunit *trans*-complementation within dimers.[3] TadA, a tRNA adenosine deaminase, binds the single-stranded anticodon loop of $tRNA^{Arg2}$ and is an obligate dimer for adenosine deamination.

A1 and AID multimerization has been suggested,[70,85,86] however, no structural models have been solved for these proteins. In contrast, purified A2 is monomeric in solution over a broad concentration range as demonstrated by SEC-MALS and heteronuclear NOE analysis by NMR.[77] The single CDA domain proteins, A3A and A3C, are also monomeric. However, the oligomeric state of the dual-deaminase domain A3s is more complicated. Biophysical analyses of purified A3G using analytical ultracentrifugation,[68] SEC-MALS,[87] and atomic force microscopy[88,89] have revealed that while a distribution of oligomeric states (monomers, dimers, tetramers and even higher order oligomers) exist in solution, the dimeric form predominates. Models of dimeric A3G from small angle X-ray scattering[67] and hydrodynamic analyses[68] have revealed elongated molecular envelopes, consistent with an end-to-end quaternary structure in which CDA subunit interactions are minimized. An end-to-end elongated dimeric A3G is consistent with *in vivo* fluorescence-quenching resonance energy transfer (FqRET) experiments that show the C-terminal CDA domain residues 209–336 of A3G are necessary and sufficient for oligomerization.[69] Protein–protein chemical cross-linking of native gel complexes suggested that A3G binds ssDNA as a dimer and must further oligomerize to tetramers before deaminase activity is observed.[16,65] In contrast, A3G mutations that ablate oligomerization were shown to retain catalytic activity.[87,90] The specificities of substrate binding will likely only be fully resolved with a crystal structure of full-length A3G bound to a ssDNA substrate. This is critically important for the dual-deaminase domain A3G, in which the non-catalytic N-terminal CDA is believed to both promote deaminase activity and modulate substrate binding as demonstrated by higher binding affinities of full-length A3G[91,92] compared with its C-terminal-half.[81,82] However, an N-terminal mode of oligomerization has also been suggested[87] that may be bridged by RNA.[93] RNA binding to the N-terminus of A3G can be inhibitory to its deaminase activity[94–96] as will be described below.

The Gene Expansion and Diversification Hypothesis

CDAs acting on free bases, nucleotides and nucelosides involved in pyrimidine salvage pathways have ancient origins and are found in all orders of life. Interestingly, although acting on pyrimidine bases within nucleic acid substrates, the APOBEC family appears to have evolved from an early ancestor of purine metabolism enzymes, namely an ADAT, which binds and edits an adenosine within the single-stranded anticodon loop of tRNA. The implication of this origin is that it may be easier for a nucleic acid binding enzyme to evolve selection from a purine to a pyrimidine than for a CDA acting on free nucleotides to evolve nucleic acid binding activity as is hypothesized in the recent review by Conticello.[97]

The most ancestral of the APOBECs is thought to be AID and its appearance is concurrent with that of vertebrates and the inception of the adaptive immune system. Supporting this contention, homologs of AID

also are found within cartilaginous fish[75,76] and jawless vertebrates,[98] which respectively, have and do not have AID immunoglobulin gene targets. A2[99] and A4,[100] located on different chromosomes than AID, are thought to have originated from retrotranspositional events involving AID,[76] while A1[75] and two A3 family progenitors[16] likely evolved from sequential duplications of the AID gene. The A3 family of proteins is found exclusively in placental mammals and this family has expanded to varying degrees within different lineages.

The expansion of the A3 gene family is thought to have occurred primarily through gene duplication, fusion, and possibly deletion.[94] A3 genes comprise either one or two CDA domains; each CDA domain has been subclassified as Z1, Z2, or Z3 types based on additional conserved residues proximal to the zinc-binding residues of the ZDD motif[75] (Figure 1). The differences among Z-type CDA domains are indicative of their evolutionary relatedness, but also have structural and functional consequences. The dual-deaminase domain A3s are composed of either two tandem Z2 domains or tandem Z1- and Z2-type domains. As was reviewed in Ref 75, mice code for a single A3 gene, while pigs have two A3 genes, the artiodactyls have three, cats have four, horses have six and primates (including humans) have seven. This differential expansion of the A3 gene family in mammals is thought to have occurred under positive selection in response to species-specific pressure from both endogenous and exogenous retroviral elements, including the lentivirus family of retroviruses which infect humans[97] (see below).

THE NON-SUBSTRATE ROLES FOR APOBEC RNA BINDING

Endogenous Retroviral-Like Element RNA

The ability of endogenous retroelements to insert themselves into random locations in the genome can lead to genomic instability and disease.[101] There are three major types of retroelements: (1) long-terminal repeat (LTR-) and non-LTR-based endogenous retroelements, (2) autonomous long interspersed nuclear elements (LINEs), and (3) non-autonomous short interspersed nuclear elements (SINEs). LINEs are autonomous because they code for everything needed to reverse transcribe and reinsert their sequence into another location within the cell's genome. On the other hand, SINEs are non-autonomous because they must hijack the proteins encoded in LINEs for reverse transcription and genomic reinsertion of their sequence.[101]

Interestingly, A3G inhibits SINE retrotransposition (i.e., Alu and hY) by sequestering these RNAs as ribonucleoprotein (RNP) complexes (described below). This deaminase-independent mechanism of A3G stands in contrast to its well-established deaminase-dependent hypermutation mechanism against exogenous retroviruses.[94,95] Furthermore, this mechanism is consistent with the RNA-associated inactivation of catalytic A3G and the localization of A3G within the cytoplasm.[94–96] Moreover, A3G also promotes dissociation of miRNA-targeted mRNA from P bodies, thus allowing for translation of these mRNAs.[102] Collectively, these data support a model whereby the molecular switch controlling A3:RNA interactions affects the oligomeric state and/or localization of A3 proteins, thus influencing both cellular and viral factors.

Various retroelement reporter assays revealed that A3G and A3F inhibited LTR-based endogenous retroviruses (i.e., IAP, Mus-D, and Ty1).[103,104] Hypermutations were detected in these sequences[103,104] as well as a reduced number of reverse transcripts,[103] consistent with the known A3G mechanism on exogenous retroviruses (Table 1). LTR-based retroelements are reverse transcribed in the cytoplasm, where A3G can hypermutate ssDNA and block RT of active retroelements.[103] Conversely, LINEs and *trans*-dependent SINEs are reverse transcribed in the nucleus. Restricted to the cytoplasm, A3G does not have access to the nuclear site of reverse transcription for this class of retroelements.[94,105] However, A3G sequesters these RNAs in the cytoplasm to prevent them from being reverse transcribed in the nucleus (Table 1).

Subcellular localization and deaminase-dependent and -independent activities by other A3s on retroelements also have been characterized. All

TABLE 1 | A3 Localization and Activity Against Endogenous Retroelements

		Endogenous Retroelements		
	Localization	LTR	LINEs	SINEs
A3A	N/C	+	+	+
A3B	N	+ Mut	+ RNA	+ RNA
A3C	N/C	+	+ RNA	+
A3DE	C	ND	+	+
A3F	C	+ Mut	+ RNA	+ RNA
A3G	C	+ Mut	+ RNA	+ RNA
A3H	N/C	+	+	+

Localization: cytoplasmic (C), nuclear (N), or both (N/C). Activity through hypermutation (+ Mut), RNA binding (+ RNA), an as of yet unclear mechanism (+), or not determined (ND).

human A3s inhibit LTR, LINE-1, and/or Alu retrotransposition[105–110] (Table 1). A3B and A3F are able to block LTR-based retroelements through hypermutation and non-LTR retroelements through RNA binding alone[103,108–110] (Table 1). A3A may directly inhibit nuclear reverse transcription of both LTR and non-LTR retroelements, but hypermutation was not detected in LTR retroelements despite the requirement for an intact active site.[106] A3A and A3C do not induce hypermutation in their restriction of LINE1 retrotransposons,[105] rather, A3C requires dimerization and RNA binding to ORF1-p of LINE1, suggesting a mechanism of RNA sequestration similar to that of A3G and a block to reverse transcription in the nucleus[111] (Table 1). A3H is located in both the nucleus and cytoplasm and is active against LINEs and SINEs but its weak interaction with Alu RNA suggests that it does not sequester RNA.[110] A3DE antiviral activity seems to be related to A3G and A3F but has weaker effects on non-LTR retroelements (LTR retroelements have not been tested to date)[64,109,110] (Table 1).

The expansion of A3s in primates correlates well with the decreased presence of active retroelements in humans.[101] There are seven A3s in humans and the only currently active retroelements in humans are non-LTR-based retroelements (i.e., LINEs and SINEs). Mice have only one A3 protein and their genomes contain active LTR-based and non-LTR-based retroelements.[101,103] Moreover, mice carry 50–60 times more active LINE retroelements in their genomes than humans and the proportion of LINE causing disease is 35% greater in mice compared with humans.[101] As with A1, A2 and AID, up-regulation of nuclear A3A, A3B, and A3C deaminases may become genotoxic and pose a risk of inducing cancer.[112,113] An open question in the field is what determines A3 expression levels and regulates restriction of their activities to different mobile genetic elements?

RNA as an Allosteric Modulator of ssDNA Deaminase Activity

Some of the A3 proteins have differential cellular localization dependent on cell cycle stage.[9] A3A, A3C, and A3H each become uniformly localized throughout the cell during telophase; A3B translocates to the nucleus after mitosis, and A3DE, A3F, and A3G remain cytoplasmic throughout the cell cycle and do not have access to host DNA [108,114,115] (Table 1).

In the cytoplasm, A3G is associated with RNA and proteins as large (megaDalton-sized) RNP complexes. RNA binding is mediated through the N-terminal half of A3G as shown in Figure 3 (see section on *Further Reading*). In this RNA-bound form, A3G is enzymatically inactive, although this can be reversed *in vitro* by reversion to a lower molecular weight complex through RNase treatment.[116] Aggregation of A3G also has been observed *in vivo* where low molecular weight A3G in resting nonproliferating CD4 T cells formed RNP complexes upon cytokine and mitogen activation of these cells.[116,117] Interest in this finding stemmed from the knowledge that resting T cells are resistant to HIV infection, whereas proliferating T cells are not. This phenomenon is not cell-type specific, as freshly isolated monocytes are resistant to HIV infection and express the enzymatically active form of A3G, while differentiation into macrophages with cytokines causes the formation of A3G–RNP and allows for active HIV infection and replication.[118]

Tandem affinity purification and mass spectrometry (MS) identified more than 90 proteins that are associated with A3G–RNP complexes. A large proportion of these proteins, like A3G, are RNA-binding proteins and they are bridged to A3G via association with a common RNA. A3G–RNP complexes are largely found in P bodies and stress granules, which are cytoplasmic complexes involved in RNA function, metabolism and stability.[94,119] A3G–RNA complexes are unable to bind or deaminate ssDNA substrates *in vitro*, as RNA has been shown to displace ssDNA from A3G.[96] These data suggest that additional cellular proteins do not inhibit A3G–RNP activity but, rather, RNA binding to A3G alone is inhibitory. It has been suggested that once A3G assembles as a cellular RNP, it cannot be packaged into virions and, therefore, does not contribute to anti-lentiviral activity.[120]

During studies to identify the preferred substrate for AID, it was shown that this protein was also inhibited by RNA.[121] This suggests a possible role for RNA in regulating AID activity. Studies in both mouse and *Xenopus laevis* embryos have demonstrated expression of AID prior to the appearance of mature B cells, indicating a role for this protein in embryogenesis.[122,123] Considering the tumorigenic potential of AID, it is tempting to speculate that AID activity in these tissues is regulated by an inhibitory RNA(s), although this has not been studied. In contrast to A3G, AID functions to initiate demethylation of methyl-cytosines (meC) through its deaminase activity.[124] This can result in two outcomes: (1) deamination of meC that is not repaired will lead to a meC → T transition or (2) base excision repair will remove the T residue, which will be replaced with an unmethylated C.[124] AID is highly expressed in oocytes and primordial germ cells (PGCs) that undergo

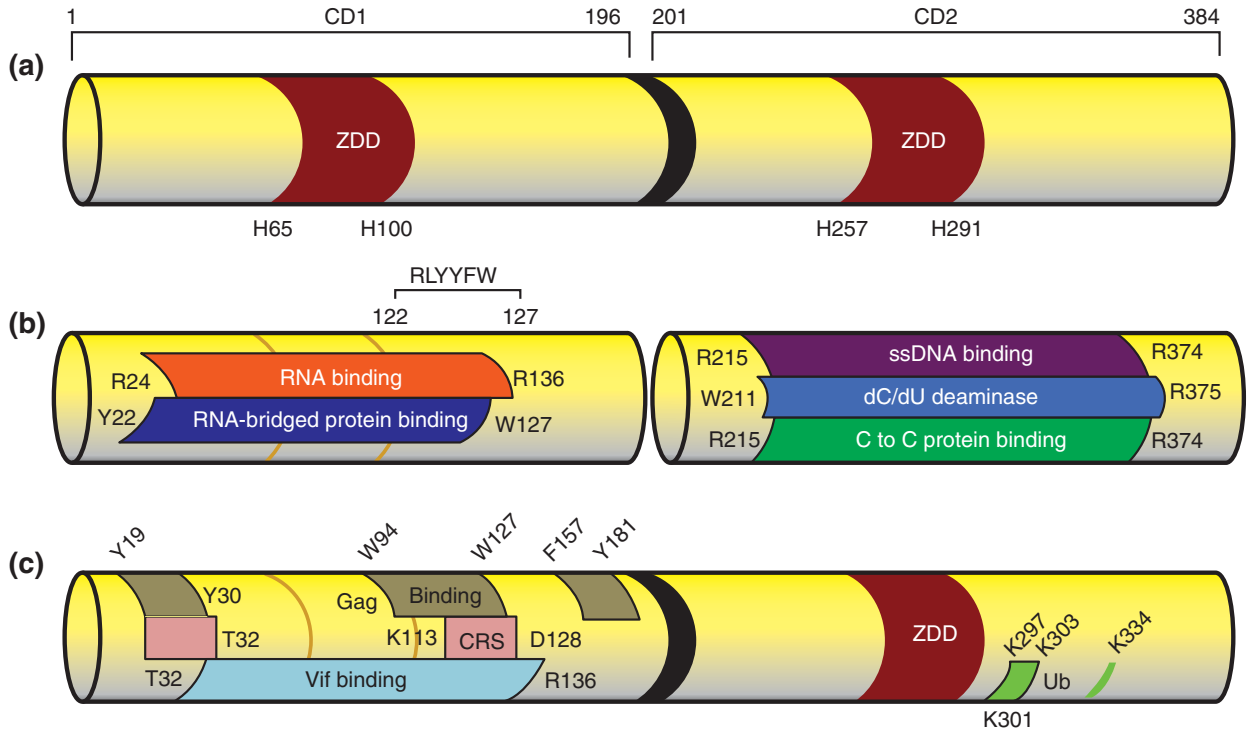


FIGURE 3 | Important functional domains of A3G. (a) Schematic of full-length A3G protein highlighting the regions comprising the two ZDDs (N-terminal ZDD spans residues 65–100, C-terminal ZDD spans residues 257–291). (b) Separation of A3Gs N- and C-termini to highlight the characteristics and functions of each half of the protein. The N-terminal half of A3G is required for both RNA binding (residues 24–136) and RNA bridged A3G multimerization (residues 22–127). Residues 122–127 have been suggested to mediate N-terminal dimerization of A3G in a head-to-head fashion. The C-terminal portion of A3G is required for its ssDNA binding and deaminase activity (residues 215–374 and 211–375, respectively). The C-terminus is also responsible for A3G tail-to-tail dimerization (residues 215–374). (c) Representation of the A3G domains required for interaction with viral proteins and cellular localization. The cytoplasmic retention signal (CRS) of A3G is located in the N-terminal half of the protein, requiring residues 19–32 and 113–128. A3G incorporation with virions requires an interaction with the HIV Gag protein, which is facilitated by multiple regions of the N-terminal domain (residues 19–30, 94–127, 157–181). In addition to Gag binding, the N-terminal domain of A3G has been shown to interact with HIV Vif (residues 32–136). Upon binding to Vif, A3G is ubiquitinated (Ub) at C-terminal lysine residues, including K297, K301, K303, and K334.

epigenetic programming (in which DNA methylation plays a significant role), where they are colocalized within a cluster of pluripotency genes (reviewed in Ref 125). In PGCs lacking AID, there was a significant increase in meC levels, leading to a plausible function for AID in germline methylation erasure.[125] In addition to a role in development, AID's activity on meC residues has also been implicated in tumorigenesis. It had been demonstrated that ectopic expression of AID in mice could lead to cancer.[126] Furthermore, many dC mutations within oncogenes and tumor suppressor genes in human tissue have been discovered within the context of the preferred target sequence for AID-induced meC deamination, AG^{me}CG,[124] suggesting a role for AID in cancer regulation.[71] In an assay to determine the prevalence of AID meC deamination in colorectal cancer, it was found that 75% of all CpG mutations occurred in the context of this preferred AID target sequence.[124] This is a 25-fold enrichment over the expected frequency of random CpG mutations giving rise to a stop codon, suggesting misregulation of AID activity. Although not as thoroughly studied, A2 also functions early in development, having a role in bilateral axis orientation during embryogenesis.[127] As demethylation is dependent on AID deaminase activity, it is possible that RNA binding to AID also regulates this process. However, this is most likely not the case for A2 as its preferred substrate is RNA.[128] Since both AID and A2 have been shown to be tumorigenic, the study of RNA inhibition of AID and RNA editing by A2 represent an interesting new focus.

RNAs Assisting Lentivirus Packaging of APOBEC

While RNA-mediated A3G complexes are inactive for editing, binding of A3G to specific RNAs is important

for its incorporation into viral particles. Pulse-chase experiments showed that shortly after A3G was translated, there was a 30–90 min window before it was recruited into RNP complexes.[120] This newly synthesized A3G was preferentially encapsidated into virions.[120] Although A3G has the ability to bind RNAs nonselectively to form RNP complexes (described above), several studies have shown that A3G (and A3F) can select for specific cellular RNAs, such as 7SL and viral genomic RNA, that become incorporated with A3G in viral particles.[129,130] The N-terminal domain of A3G and A3G–viral RNA interactions facilitate the association of A3G with the nucleocapsid (NC) region of HIV-1 Gag, a requirement for A3G viral packaging[131] (Figure 3). A3G–RNA complexes in viral particles are catalytically inactive until viral genomic RNA is removed by RNaseH subsequent to reverse transcription.[120]

RNA binding also contributes deaminase-independent antiviral activities linked to A3G.[132] These activities include inhibition of $tRNA^{lys3}$ priming of first strand synthesis,[133] strand transfer activity,[134] reverse transcript elongation,[135] and inhibition of double stranded proviral DNA integration.[134,136] A3C also is incorporated into HIV particles and although it does not associate with 7SL RNA or HIV NC, it does interact with the matrix (MA) region of HIV Gag in what may be a 5.8S RNA-dependent manner.[137,138] Overall, selective RNA interactions of A3 family members are crucial for their recruitment into virions and antiviral activity.

SUMMARY

APOBEC is a family of DNA mutagenic enzymes that display either ssDNA or RNA editing activity. The discoveries of AID and A3G ssDNA deaminases answered long-standing questions of the mechanisms for immunoglobulin gene diversification and lentiviral host cell restriction, respectively. The fact that many of the APOBEC proteins have retained their ability to bind RNA while only using ssDNA as substrate for deamination suggests a biological function that we do not fully appreciate. At the moment, we understand that A3 proteins bind to retroviral RNAs to inhibit their expression and function. We also know that RNAs of a wide range of sequence bind to A3G, A3F, and AID, resulting in inhibition of ssDNA deaminase function and, hence, regulation APOBEC biological activities. The biological advantage of binding (and regulation by) RNA is not clear. This is most evident for A3 proteins, which bind HIV genomic RNA for incorporation into nascent viral particles (a part of the antiviral mechanism) and yet, upon binding RNA in cells or in viral particles, their antiviral deaminase activity is inactivated. Why do A3s not edit viral RNA and instead wait until after infection for reverse transcriptase to make viral cDNA?

There remain many open questions regarding APOBEC structure that must be addressed before we will be able to understand how RNA and ssDNA are bound, the basis for substrate selectivity and the mechanism whereby RNA binding inhibits ssDNA deaminase activity. Progress has been made in crystallizing some APOBEC members, however, crystallization of A3G requires removal of the N-terminal RNA-binding domain. While structural information of the C-terminal half has provided some information, the fact remains that the C-terminal half of A3G requires the N-terminal half for full catalytic activity. In addition to RNA binding, Vif and Gag only bind to the N-terminal half. Given this, it is likely that the structure of the N-terminal half of A3G will have distinctive features that differentiate it from that of the C-terminal half. Biophysical approaches are critical for assessing how A3 enzymes bind, but do not edit, RNA and will likely shed light on the catalytic requirements for ssDNA substrate recognition and deamination.

We are also learning that the DNA mutation activity of APOBEC enzymes is stringently controlled by cells through the formation of higher order complexes of protein and RNA, as well as through subcellular compartmentalization and turnover at the level of their mRNAs and protein. To address these mechanisms and importantly, to know the oncogenic potential of APOBEC proteins, we will need to understand if there is an RNA sequence or secondary structure specificity in APOBEC binding to RNA that determines functionalities such as viral particle inclusion, P body sequestration and inhibition of deaminase activity. Also significant are the unresolved questions concerning *cis*- and *trans*-acting factors influencing the selection of ssDNA sequences for deamination and the mechanism(s) for processivity of editing and hyper-editing by AID and A3 enzymes. Answers to these questions will undoubtedly go beyond our current understanding of lax nearest-neighbor nucleotide preferences and involve higher order enzyme complexes.

If AID was the first ssDNA-editing enzyme to evolve, did it evolve from a primordial RNA deaminase whose active site had lax substrate selectivity to acquire the ability to mutate ssDNA for adaptive immunity? It is not clear how this could happen while retaining the ability to bind RNA but losing the ability to edit RNA. This question is particularly

puzzling given that A1 presumably arose from a gene duplication event and has both RNA and ssDNA deaminase activity. What was the selection pressure that drove the diversification of the APOBEC gene family? The co-evolution of the A3 gene expansion with the emergence of endogenous and exogenous retroviral elements is a plausible and intriguing hypothesis. RNA binding by A3 proteins functions to sequester retroviral elements and their deaminase activity suppresses reverse transcription, which may have selected for these proteins with both RNA and ssDNA binding capacities. It would strengthen this hypothesis if we knew whether the amino acid residues that bind to RNA and ssDNA are the same or different. In conclusion, although early work focused on APOBECs in RNA editing and currently, much of the research focuses on ssDNA editing, RNA binding to APOBEC remains a highly relevant and exciting area for future research on the catalytic mechanism, evolutionary biology, and current-day biomedical functional significance of this family of enzymes.

ACKNOWLEDGMENTS

The authors apologize that some important contributions to the field may have not been cited owing to limitations of space. Data from the authors' laboratory that have been described in this manuscript were obtained with the support of Public Health Services Grants NIDDK DK43739, NIAID AI54369, AI095007, AI58789, NINDS NS067671, an Air Force Office of Scientific Research Grant F49620, and a Bill and Melinda Gates Foundation Grand Challenges in Exploration grant awarded to HCS. Additional support was from a Public Health Services Grant NIAID AI076085 awarded to Joseph E. Wedekind.

REFERENCES

1. Machnicka MA, Milanowska K, Osman Oglou O, Purta E, Kurkowska M, Olchowik A, Januszewski W, Kalinowski S, Dunin-Horkawicz S, Rother KM, et al. MODOMICS: a database of RNA modification pathways--2013 update. *Nucleic Acids Res* 2013, 41:D262–D267.
2. Wedekind JE, Dance GS, Sowden MP, Smith HC. Messenger RNA editing in mammals: new members of the APOBEC family seeking roles in the family business. *Trends Genet* 2003, 19:207–216.
3. MacElrevey C, Wedekind JE. Chapter 16: Chemistry, Phylogeny, and Three-Dimensional Structure of the APOBEC Protein Family. In: Smith HC, ed. *RNA and DNA Editing: Molecular Mechanisms and Their Integration into Biological Systems*. Hoboken, NJ: John Wiley & Sons, Inc; 2008, 371–421.
4. Smith HC, Wedekind JE, Kefang X, Sowden MP. Mammalian C to U editing. *Top Curr Biol* 2005, 12:365–400.
5. Harris RS, Petersen-Mahrt SK, Neuberger MS. RNA editing enzyme APOBEC1 and some of its homologs can act as DNA mutators. *Mol Cell* 2002, 10:1247–1253.
6. Honjo T, Muramatsu M, Fagarasan S. AID: how does it aid antibody diversity? *Immunity* 2004, 20:659–668.
7. Hamilton CE, Papavasiliou FN, Rosenberg BR. Diverse functions for DNA and RNA editing in the immune system. *RNA Biol* 2010, 7:220–228.
8. Jaszczur M, Bertram JG, Pham P, Scharff MD, Goodman MF. AID and Apobec3G haphazard deamination and mutational diversity. *Cell Mol Life Sci* 2013, 70:3089–3108.
9. Smith HC. *The APOBEC1 Paradigm for Mammalian Cytidine Deaminases that Edit DNA and RNA*. Landes BioScience: Austin, TX; 2009.
10. Chen SH, Habib G, Yang CY, Gu ZW, Lee BR, Weng SA, Silberman SR, Cai SJ, Deslypere JP, Rosseneu M, et al. Apolipoprotein B-48 is the product of a messenger RNA with an organ-specific in-frame stop codon. *Science* 1987, 238:363–366.
11. Powell LM, Wallis SC, Pease RJ, Edwards YH, Knott TJ, Scott J. A novel form of tissue-specific RNA processing produces apolipoprotein- B48 in intestine. *Cell* 1987, 50:831–840.
12. Muramatsu MSV, Anant S, Sugai M, Kinoshita K, Davidson NO, Honjo T. 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, 274:18470–18476.
13. Sheehy AM, Gaddis NC, Choi JD, Malim MH. Isolation of a human gene that inhibits HIV-1 infection and is suppressed by the viral Vif protein. *Nature* 2002, 418:646–650.
14. Smith HC. *Deaminase-dependent and deaminase-independent functions of APOBEC1 and APOBEC1 complementation factor in the context of the*

APOBEC family. Norfolk: Caister Academic Press; 2013.

15. Blanc V, Davidson NO. APOBEC-1-mediated RNA editing. *WIREs: Syst Biol Med* 2010, 2:594–602.
16. Jarmuz A, Chester A, Bayliss J, Gisbourne J, Dunham I, Scott J, Navaratnam N. An anthropoid-specific locus of orphan C to U RNA-editing enzymes on chromosome 22. *Genomics* 2002, 79:285–296.
17. Teng B, Burant CF, Davidson NO. Molecular cloning of an apolipoprotein B messenger RNA editing protein. *Science* 1816–1819, 1993:260.
18. Anant S, MacGinnitie AJ, Davidson NO. apobec-1, the catalytic subunit of the mammalian apolipoprotein B mRNA editing enzyme, is a novel RNA-binding protein. *J Biol Chem* 1995, 270:14762–14767.
19. Smith HC, Kuo SR, Backus JW, Harris SG, Sparks CE, Sparks JD. In vitro apolipoprotein B mRNA editing: identification of a 27S editing complex. *Proc Natl Acad Sci U S A* 1991, 88:1489–1493.
20. Mehta A, Kinter MT, Sherman NE, Driscoll DM. Molecular cloning of apobec-1 complementation factor, a novel RNA- binding protein involved in the editing of apolipoprotein B mRNA. *Mol Cell Biol* 1846–1854, 2000:20.
21. Lellek H, Kirsten R, Diehl I, Apostel F, Buck F, Greeve J. Purification and molecular cloning of a novel essential component of the apolipoprotein B mRNA editing enzyme-complex. *J Biol Chem* 2000, 275:19848–19856.
22. Shah RR, Knott TJ, Legros JE, Navaratnam N, Greeve JC, Scott J. Sequence requirements for the editing of apolipoprotein B mRNA. *J Biol Chem* 1991, 266:16301–16304.
23. Backus JW, Schock D, Smith HC. Only cytidines 5' of the apolipoprotein B mRNA mooring sequence are edited. *Biochim Biophys Acta* 1994, 1219:1–14.
24. Blanc V, Henderson JO, Kennedy S, Davidson NO. Mutagenesis of apobec-1 complementation factor reveals distinct domains that modulate RNA binding, protein-protein interaction with apobec-1, and complementation of C to U RNA-editing activity. *J Biol Chem* 2001, 276:46386–46393.
25. Mehta A, Driscoll DM. Identification of domains in APOBEC-1 complementation factor required for RNA binding and apolipoprotein B mRNA editing. *RNA* 2002, 8:69–82.
26. Sowden MP, Ballatori N, de Mesy Jensen KL, Hamilton Reed L, Smith HC. The editosome for cytidine to uridine mRNA editing has a native complexity of 27S: identification of intracellular domains containing active and inactive editing factors. *J Cell Sci* 2002, 115:1027–1039.
27. Hirano K, Young SG, Farese RV Jr, Ng J, Sande E, Warburton C, Powell-Braxton LM, Davidson NO. Targeted disruption of the mouse apobec-1 gene abolishes apolipoprotein B mRNA editing and eliminates apolipoprotein B48. *J Biol Chem* 1996, 271:9887–9890.
28. Blanc V, Henderson JO, Newberry EP, Kennedy S, Luo J, Davidson NO. Targeted deletion of the murine apobec-1 complementation factor (acf) gene results in embryonic lethality. *Mol Cell Biol* 2005, 25:7260–7269.
29. Nakamuta M, Oka K, Krushkal J, Kobayashi K, Yamamoto M, Li WH, Chan L. Alternative mRNA splicing and differential promoter utilization determine tissue-specific expression of the apolipoprotein B mRNA-editing protein (Apobec1) gene in mice. Structure and evolution of Apobec1 and related nucleoside/nucleotide deaminases. *J Biol Chem* 1995, 270:13042–13056.
30. Lee RM, Hirano K, Anant S, Baunoch D, Davidson NO. An alternatively spliced form of apobec-1 messenger RNA is overexpressed in human colon cancer. *Gastroenterology* 1998, 115:1096–1103.
31. Dance GSC, Sowden MP, Cartegni L, Cooper E, Krainer AR, Smith HC. Two proteins essential for apolipoprotein B mRNA editing are expressed from a single gene through alternative splicing. *J Biol Chem* 2002, 277:12703–12709.
32. Sowden MP, Lehmann DM, Lin X, Smith CO, Smith HC. Identification of novel alternative splice variants of APOBEC-1 complementation factor with different capacities to support ApoB mRNA editing. *J Biol Chem* 2004, 278:197–206.
33. Lehmann DM, Galloway CA, Macelrevey C, Sowden MP, Wedekind JE, Smith HC. Functional characterization of APOBEC-1 complementation factor phosphorylation sites. *Biochim Biophys Acta* 2007, 1773:408–418.
34. Lehmann DM, Galloway CA, Sowden MP, Smith HC. Metabolic regulation of apoB mRNA editing is associated with phosphorylation of APOBEC-1 complementation factor. *Nucleic Acids Res* 2006, 34:3299–3308.
35. Galloway CA, Ashton J, Sparks JD, Mooney RA, Smith HC. Metabolic regulation of APOBEC-1 complementation factor trafficking in mouse models of obesity and its positive correlation with the expression of ApoB protein in hepatocytes. *Biochim Biophys Acta* 1802, 2010:976–985.
36. Skuse GR, Cappione AJ, Sowden M, Metheny LJ, Smith HC. The neurofibromatosis type I messenger RNA undergoes base-modification RNA editing. *Nucleic Acids Res* 1996, 24:478–485.
37. Rosenberg BR, Hamilton CE, Mwangi MM, Dewell S, Papavasiliou FN. Transcriptome-wide sequencing reveals numerous APOBEC1 mRNA-editing targets in transcript 3' UTRs. *Nat Struct Mol Biol* 2011, 18:230–236.
38. Sowden M, Hamm JK, Smith HC. Overexpression of APOBEC-1 results in mooring sequence-dependent

promiscuous RNA editing. *J Biol Chem* 1996, 271:3011–3017.

39. Yamanaka S, Balestra M, Ferrell L, Fan J, Arnold KS, Taylor S, Taylor JM, Innerarity TL. Apolipoprotein B mRNA editing protein induces hepatocellular carcinoma and dysplasia in transgenic animals. *Proc Natl Acad Sci U S A* 1995, 92:8483–8487.
40. Yamanaka S, Poksay KS, Arnold KS, Innerarity TL. A novel translational repressor mRNA is edited extensively in livers containing tumors caused by the transgene expression of the apoB mRNA- editing enzyme. *Genes Dev* 1997, 11:321–333.
41. Ikeda T, Abd El Galil KH, Tokunaga K, Maeda K, Sata T, Sakaguchi N, Heidmann T, Koito A. Intrinsic restriction activity by apolipoprotein B mRNA editing enzyme APOBEC1 against the mobility of autonomous retrotransposons. *Nucleic Acids Res* 2011, 39:5538–5554.
42. Bishop KN, Holmes RK, Sheehy AM, Malim MH. APOBEC-mediated editing of viral RNA. *Science* 2004, 305:645.
43. Anant S, Davidson NO. An AU-rich sequence element (UUUN[A/U]U) downstream of the edited C in apolipoprotein B mRNA is a high-affinity binding site for Apobec-1: binding of Apobec-1 to this motif in the 3' untranslated region of c-myc increases mRNA stability. *Mol Cell Biol* 1982–1992, 2000:20.
44. Blanc V, Sessa KJ, Kennedy S, Luo J, Davidson NO. Apobec-1 complementation factor modulates liver regeneration by post-transcriptional regulation of interleukin-6 mRNA stability. *J Biol Chem* 2010, 285:19184–19192.
45. Xie Y, Blanc V, Kerr TA, Kennedy S, Luo J, Newberry EP, Davidson NO. Decreased expression of cholesterol 7α-hydroxylase and altered bile acid metabolism in Apobec-1−/− mice lead to increased gallstone susceptibility. *J Biol Chem* 2009, 284:16860–16871.
46. Blanc V, Henderson JO, Newberry RD, Xie Y, Cho SJ, Newberry EP, Kennedy S, Rubin DC, Wang HL, Luo J, et al. Deletion of the AU-rich RNA binding protein Apobec-1 reduces intestinal tumor burden in Apc(min) mice. *Cancer Res* 2007, 67:8565–8573.
47. Petersen-Mahrt SK, Neuberger MS. In vitro deamination of cytosine to uracil in single-stranded DNA by apolipoprotein B editing complex catalytic subunit 1 (APOBEC1). *J Biol Chem* 2003, 278:19583–19586.
48. Gee P, Ando Y, Kitayama H, Yamamoto SP, Kanemura Y, Ebina H, Kawaguchi Y, Koyanagi Y. APOBEC1-mediated editing and attenuation of herpes simplex virus 1 DNA indicate that neurons have an antiviral role during herpes simplex encephalitis. *J Virol* 2011, 85:9726–9736.
49. Gonzalez MC, Suspene R, Henry M, Guetard D, Wain-Hobson S, Vartanian JP. Human APOBEC1 cytidine deaminase edits HBV DNA. *Retrovirology* 2009, 6:96.
50. Muramatsu MKK, 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, 102:553–563.
51. McCarthy HWW, Barron LL, Cromwell CC, Wang J, Coombes KR, Rangel R, Elenitoba-Johnson KS, Keating MJ, Abruzzo LV. High expression of activation-induced cytidine deaminase (AID) and splice variants is a distinctive feature of poor-prognosis chronic lymphocytic leukemia. *Blood* 2003, 101:4903–4908.
52. Patenaude AM, Di Noia JM. The mechanisms regulating the subcellular localization of AID. *Nucleus* 2010, 1:325–331.
53. Aoufouchi S, Faili A, Zober C, D'Orlando O, Weller S, Weill JC, Reynaud C-A. Proteasomal degradation restricts the nuclear lifespan of AID. *J Exp Med* 2008, 205:1357–1368.
54. Yang Y, Smith HC. Multiple protein domains determine the cell type-specific nuclear distribution of the catalytic subunit required for apolipoprotein B mRNA editing. *Proc Natl Acad Sci U S A* 1997, 94:13075–13080.
55. Blanc V, Kennedy SM, Davidson NO. A novel nuclear localization signal in the auxiliary domain of apobec-1 complementation factor (ACF) regulates nucleo-cytoplasmic import and shuttling. *J Biol Chem* 2003, 278:41198–41204.
56. Gabuzda DH, Lawrence K, Langhoff E, Terwillinger E, Dorfman T, Haseltine WA, Sodroski J. Role of vif in replication of human immunodeficiency virus type 1 in CD4+ T lymphocytes. *J Virol* 1992, 66:6489–6495.
57. von Schwedler U, Song J, Aiken C, Trono D. Vif is crucial for human immunodeficiency virus type 1 proviral DNA synthesis in infected cells. *J Virol* 1993, 67:4945–4955.
58. Madani NaK D. An endogenous inhibitor of human immunodeficiency virus in human lymphocytes is overcome by the viral Vif protein. *J Virol* 1998, 72:10251–10255.
59. Simon JH, Gaddis NC, Foucier RA, Malim MH. Evidence for a newly discovered cellular anti-HIV-1 phenotype. *Nat Med* 1998, 4:1397–1400.
60. Lecossier D, Bouchennet F, Clavel F, Hance AJ. Hypermutation of HIV-1 DNA in the absence of the Vif protein. *Science* 2003, 300:1112.
61. Zhang H, Yang B, Pomerantz RJ, Zhang C, Arunachalam SC, Gao L. The cytidine deaminase CEM15 induces hypermutation in newly synthesized HIV-1 DNA. *Nature* 2003, 424:94–98.
62. Mangeat B, Turelli P, Caron G, Friedli M, Perrin L, Trono D. Broad antiretriviral defence by human APOBEC3G through lethan editing of nascent reverse transcripts. *Nature* 2003, 424:99–103.

63. Harris RS, Bishop KN, Sheehy AM, Craig HM, Petersen-Mahrt SK, Watt IN, Neuberger MS, Malim MH. DNA deamination mediates innate immunity to retroviral infection. *Cell* 2003, 113:803–809.
64. Hultquist JF, Lengyel JA, Refsland EW, Larue RS, Lackey L, Brown WL, Harris RS. Human and Rhesus APOBEC3D, APOBEC3F, APOBEC3G, and APOBEC3H demonstrate a conserved capacity to restrict Vif-deficient HIV-1. *J Virol* 2011, 85:11220–11234.
65. McDougall WM, Okany C, Smith HC. Deaminase activity on single-stranded DNA (ssDNA) occurs in vitro when APOBEC3G cytidine deaminase forms homotetramers and higher-order complexes. *J Biol Chem* 2011, 286:30655–30661.
66. Navarro F, Bollman B, Chen H, Konig R, Yu Q, Chiles K, Landau NR. Complementary function of the two catalytic domains of APOBEC3G. *Virology* 2005, 333:374–386.
67. Wedekind JE, Gillilan R, Janda A, Krucinska J, Salter JD, Bennett RP, Raina J, Smith HC. Nanostructures of APOBEC3G support a hierarchical assembly model of high molecular mass ribonucleoprotein particles from dimeric subunits. *J Biol Chem* 2006, 281:38122–38126.
68. Salter JD, Krucinska J, Raina J, Smith HC, Wedekind JE. A hydrodynamic analysis of APOBEC3G reveals a monomer-dimer-tetramer self-association that has implications for anti-HIV function. *Biochemistry* 2009, 48:10685–10687.
69. Bennett RP, Salter JD, Liu X, Wedekind JE, Smith HC. APOBEC3G subunits self-associate via the C-terminal deaminase domain. *J Biol Chem* 2008, 283:33329–33336.
70. Lau PP, Zhu H-J, Baldini HA, Charnsangavej C, Chan L. Dimeric structure of a human apo B mRNA editing protein and cloning and chromosomal localization of its gene. *Proc Natl Acad Sci U S A* 1994, 91:8522–8526.
71. Beale RC, Petersen-Mahrt SK, Watt IN, Harris RS, Rada C, Neuberger MS. Comparison of the differential context-dependence of DNA deamination by APOBEC enzymes: correlation with mutation spectra in vivo. *J Mol Biol* 2004, 26:585–596.
72. Langlois M-A, Beale RC, Conticello SG, Neuberger MS. Mutational comparison of the single-domained APOBEC3C and double-domained APOBEC3F/G anti-retroviral cytidine deaminases provides insight into their DNA target site specificities. *Nucleic Acids Res* 2005, 33:1913–1923.
73. Shandilya SM, Nalam MN, Nalivaika EA, Gross PJ, Valesano JC, Shindo K, Li M, Munson M, Royer WE, Harjes E, et al. Crystal structure of the APOBEC3G catalytic domain reveals potential oligomerization interfaces. *Structure* 2010, 18:28–38.
74. Bohn MF, Shandilya SM, Albin JS, Kouno T, Anderson BD, McDougle RM, Carpenter MA, Rathore A, Evans L, Davis AN, et al. Crystal structure of the DNA cytosine deaminase APOBEC3F: the catalytically active and HIV-1 Vif-binding domain. *Structure* 2013, 21:1042–1050.
75. LaRue RS, Andresdottir V, Blanchard Y, Conticello SG, Derse D, Emerman M, Greene WC, Jonsson SR, Landau NR, Lochelt M, et al. Guidelines for naming nonprimate APOBEC3 genes and proteins. *J Virol* 2009, 83:494–497.
76. Conticello SG. The AID/APOBEC family of nucleic acid mutators. *Genome Biol* 2008, 9:229.
77. Krzysiak TC, Jung J, Thompson J, Baker D, Gronenborn AM. APOBEC2 is a monomer in solution: implications for APOBEC3G models. *Biochemistry* 2012, 51:2008–2017.
78. Prochnow C, Bransteitter R, Klein MG, Goodman MF, Chen XS. The APOBEC-2 crystal structure and functional implications for the deaminase AID. *Nature* 2007, 445:447–451.
79. Byeon IJ, Ahn J, Mitra M, Byeon CH, Hercik K, Hritz J, Charlton LM, Levin JG, Gronenborn AM. NMR structure of human restriction factor APOBEC3A reveals substrate binding and enzyme specificity. *Nat Commun* 1890, 2013:4.
80. Kitamura S, Ode H, Nakashima M, Imahashi M, Naganawa Y, Kurosawa T, Yokomaku Y, Yamane T, Watanabe N, Suzuki A, et al. The APOBEC3C crystal structure and the interface for HIV-1 Vif binding. *Nat Struct Mol Biol* 2012, 19:1005–1010.
81. Chen KM, Harjes E, Gross PJ, Fahmy A, Lu Y, Shindo K, Harris RS, Matsuo H. Structure of the DNA deaminase domain of the HIV-1 restriction factor APOBEC3G. *Nature* 2008, 452:116–119.
82. Furukawa A, Nagata T, Matsugami A, Habu Y, Sugiyama R, Hayashi F, Kobayashi N, Yokoyama S, Takaku H, Katahira M. Structure, interaction and real-time monitoring of the enzymatic reaction of wild-type APOBEC3G. *EMBO J* 2009, 28:440–451.
83. 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, 284:22898–22904.
84. Xie K, Sowden MP, Dance GS, Torelli AT, Smith HC, Wedekind JE. The structure of a yeast RNA-editing deaminase provides insight into the fold and function of activation-induced deaminase and APOBEC-1. *Proc Natl Acad Sci U S A* 2004, 101:8114–8119.
85. Dickerson SK, Market E, Besmer E, Papavasiliou FN. AID mediates hypermutation by deaminating single stranded DNA. *J Exp Med* 2003, 197:1291–1296.
86. Ta VT, Nagaoka H, Catalan N, Durandy A, Fischer A, Imai K, Nonoyama S, Tashiro J, Ikegawa M, Ito S, et al. AID mutant analyses indicate requirement

for class-switch-specific cofactors. *Nat Immunol* 2003, 4:843–848.

87. Chelico L, Prochnow C, Erie DA, Chen XS, Goodman MF. A structural model for deoxycytidine deamination mechanisms of the HIV-1 inactivation enzyme APOBEC3G. *J Biol Chem* 2010, 285:16195–16205.
88. 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, 283:13780–13791.
89. 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, 184:217–225.
90. Feng Y, Chelico L. Intensity of deoxycytidine deamination of HIV-1 proviral DNA by the retroviral restriction factor APOBEC3G is mediated by the noncatalytic domain. *J Biol Chem* 2011, 286:11415–11426.
91. Chelico L, Pham P, Calabrese P, Goodman MF. APOBEC3G DNA deaminase acts processively 3' --> 5' on single-stranded DNA. *Nat Struct Mol Biol* 2006, 13:392–399.
92. Iwatani Y, Takeuchi H, Strebel K, Levin JG. Biochemical activities of highly purified, catalytically active human APOBEC3G: correlation with antiviral effect. *J Virol* 2006, 80:5992–6002.
93. Huthoff H, Autore F, Gallois-Montbrun S, Fraternali F, Malim MH. RNA-dependent oligomerization of APOBEC3G is required for restriction of HIV-1. *PLoS Pathog* 2009, 5:e1000330.
94. Chiu YL, Witkowska HE, Hall SC, Santiago M, Soros VB, Esnault C, Heidmann T, Greene WC. High-molecular-mass APOBEC3G complexes restrict Alu retrotransposition. *Proc Natl Acad Sci U S A* 2006, 103:15588–15593.
95. Hulme AE, Bogerd HP, Cullen BR, Moran JV. Selective inhibition of Alu retrotransposition by APOBEC3G. *Gene* 2007, 390:199–205.
96. McDougall WM, Smith HC. Direct evidence that RNA inhibits APOBEC3G ssDNA cytidine deaminase activity. *Biochem Biophys Res Commun* 2011, 412:612–617.
97. Conticello SG. Creative deaminases, self-inflicted damage, and genome evolution. *Ann N Y Acad Sci* 2012, 1267:79–85.
98. Rogozin IB, Iyer LM, Liang L, Glazko GV, Liston VG, Pavlov YI, Aravind L, Pancer Z. Evolution and diversification of lamprey antigen receptors: evidence for involvement of an AID-APOBEC family cytosine deaminase. *Nat Immunol* 2007, 8:647–656.
99. Liao W, Hong SH, Chan BH, Rudolph FB, Clark SC, Chan L. APOBEC-2, a cardiac- and skeletal muscle-specific member of the cytidine deaminase supergene family. *Biochem Biophys Res Commun* 1999, 260:398–404.
100. Rogozin IB, Basu MK, Jordan IK, Pavlov YI, Koonin EV. APOBEC4, a new member of the AID/APOBEC family of polynucleotide (deoxy)cytidine deaminases predicted by computational analysis. *Cell Cycle* 2005, 4:1281–1285.
101. Schumann GG. APOBEC3 proteins: major players in intracellular defence against LINE-1-mediated retrotransposition. *Biochem Soc Trans* 2007, 35:637–642.
102. Huang J, Liang Z, Yang B, Tian H, Ma J, Zhang H. Derepression of microRNA-mediated protein translation inhibition by apolipoprotein B mRNA-editing enzyme catalytic polypeptide-like 3G (APOBEC3G) and its family members. *J Biol Chem* 2007, 282:33632–33640.
103. Esnault C, Millet J, Schwartz O, Heidmann T. Dual inhibitory effects of APOBEC family proteins on retrotransposition of mammalian endogenous retroviruses. *Nucleic Acids Res* 2006, 34:1522–1531.
104. Schumacher AJ, Nissley DV, Harris RS. APOBEC3G hypermutates genomic DNA and inhibits Ty1 retrotransposition in yeast. *Proc Natl Acad Sci U S A* 2005, 102:9854–9859.
105. Kinomoto M, Kanno T, Shimura M, Ishizaka Y, Kojima A, Kurata T, Sata T, Tokunaga K. All APOBEC3 family proteins differentially inhibit LINE-1 retrotransposition. *Nucleic Acids Res* 2007, 35:2955–2964.
106. Bogerd HP, Wiegand HL, Doehle BP, Lueders KK, Cullen BR. APOBEC3A and APOBEC3B are potent inhibitors of LTR-retrotransposon function in human cells. *Nucleic Acids Res* 2006, 34:89–95.
107. Chen H, Lilley CE, Yu Q, Lee DV, Chou J, Narvaiza I, Landau NR, Weitzman MD. APOBEC3A is a potent inhibitor of adeno-associated virus and retrotransposons. *Curr Biol* 2006, 16:480–485.
108. Muckenfuss H, Hamdorf M, Held U, Perkovic M, Lower J, Cichutek K, Flory E, Schumann GG, Munk C. APOBEC3 proteins inhibit human LINE-1 retrotransposition. *J Biol Chem* 2006, 281:22161–22172.
109. Stenglein MD, Harris RS. APOBEC3B and APOBEC3F inhibit L1 retrotransposition by a DNA deamination-independent mechanism. *J Biol Chem* 2006, 281:16837–16841.
110. Tan L, Sarkis PT, Wang T, Tian C, Yu XF. Sole copy of Z2-type human cytidine deaminase APOBEC3H has inhibitory activity against retrotransposons and HIV-1. *FASEB J* 2009, 23:279–287.
111. Horn AV, Klawitter S, Held U, Berger A, Jaguva Vasudevan AA, Bock A, Hofmann H, Hanschmann KM, Trosemeier JH, Flory E, et al. Human LINE-1 restriction by APOBEC3C is deaminase independent and mediated by an ORF1p interaction that affects LINE reverse transcriptase activity. *Nucleic Acids Res* 2013, 42:396–416.
112. Landry S, Narvaiza I, Linfesty DC, Weitzman MD. APOBEC3A can activate the DNA damage

response and cause cell-cycle arrest. *EMBO Rep* 2011, 12:444–450.

113. Bonvin M, Achermann F, Greeve I, Stroka D, Keogh A, Inderbitzin D, Candinas D, Sommer P, Wain-Hobson S, Vartanian JP, et al. Interferon-inducible expression of APOBEC3 editing enzymes in human hepatocytes and inhibition of hepatitis B virus replication. *Hepatology* 2006, 43:1364–1374.
114. Bennett RP, Presnyak V, Wedekind JE, Smith HC. Nuclear Exclusion of the HIV-1 host defense factor APOBEC3G requires a novel cytoplasmic retention signal and is not dependent on RNA binding. *J Biol Chem* 2008, 283:7320–7327.
115. Lackey L, Law EK, Brown WL, Harris RS. Subcellular localization of the APOBEC3 proteins during mitosis and implications for genomic DNA deamination. *Cell Cycle* 2013, 12:762–772.
116. Chiu YL, Soros VB, Kreisberg JF, Stopak K, Yonemoto W, Greene WC. Cellular APOBEC3G restricts HIV-1 infection in resting CD4+ T cells. *Nature* 2005, 435:108–114.
117. Stopak KS, Chiu YL, Kropp J, Grant RM, Greene WC. Distinct patterns of cytokine regulation of APOBEC3G expression and activity in primary lymphocytes, macrophages, and dendritic cells. *J Biol Chem* 2007, 282:3539–3546.
118. Sonza S, Maerz A, Deacon N, Meanger J, Mills J, Crowe S. Human immunodeficiency virus type 1 replication is blocked prior to reverse transcription and integration in freshly isolated peripheral blood monocytes. *J Virol* 1996, 70:3863–3869.
119. 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, 281:29105–29119.
120. Soros VB, Yonemoto W, Greene WC. Newly synthesized APOBEC3G is incorporated into HIV virions, inhibited by HIV RNA, and subsequently activated by RNase H. *PLoS Pathog* 2007, 3:e15.
121. 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, 100:4102–4107.
122. Popp C, Dean W, Feng S, Cokus SJ, Andrews S, Pellegrini M, Jacobsen SE, Reik W. Genome-wide erasure of DNA methylation in mouse primordial germ cells is affected by AID deficiency. *Nature* 2010, 463:1101–1105.
123. Marr S, Morales H, Bottaro A, Cooper M, Flajnik M, Robert J. Localization and differential expression of activation-induced cytidine deaminase in the amphibian Xenopus upon antigen stimulation and during early development. *J Immunol* 2007, 179:6783–6789.
124. Morgan HD, Dean W, Coker HA, Reik W, Petersen-Mahrt SK. Activation-induced cytidine deaminase deaminates 5-methylcytosine in DNA and is expressed in pluripotent tissues: implications for epigenetic reprogramming. *J Biol Chem* 2004, 279:52353–52360.
125. Fritz EL, Papavasiliou FN. Cytidine deaminases: AIDing DNA demethylation? *Genes Dev* 2010, 24:2107–2114.
126. Okazaki IMHH, Kakazu N, Yamada S, Muramatsu M, Kinoshita K, Honjo T. Constitutive expression of AID leads to tumorigenesis. *J Exp Med* 2003, 197:1173–1181.
127. Vonica A, Rosa A, Arduini BL, Brivanlou AH. APOBEC2, a selective inhibitor of TGF-β signaling, regulates left-right axis specification during early embryogenesis. *Dev Biol* 2011, 350:13–23.
128. Okuyama S, Marusawa H, Matsumoto T, Ueda Y, Matsumoto Y, Endo Y, Takai A, Chiba T. Excessive activity of apolipoprotein B mRNA editing enzyme catalytic polypeptide 2 (APOBEC2) contributes to liver and lung tumorigenesis. *Int J Cancer* 2012, 130:1294–1301.
129. Svarovskaia ES, Xu H, Mbisa JL, Barr R, Gorelick RJ, Ono A, Freed EO, Hu WS, Pathak VK. Human APOBEC3G is incorporated into HIV-1 virions through interactions with viral and nonviral RNAs. *J Biol Chem* 2004, 279:35822–35828.
130. Wang TTC, Zhang W, Sarkis PT, Yu XF. Interaction with 7SL RNA but not with HIV-1 genomic RNA or P bodies is required for APOBEC3F virion packaging. *J Mol Biol* 2008, 375:1098–1112.
131. Alce TM, Popik W. APOBEC3G is incorporated into virus-like particles by a direct interaction with HIV-1 Gag nucleocapsid protein. *J Biol Chem* 2004, 279:34083–34086.
132. Smith HC. APOBEC3G: a double agent in defense. *Trends Biochem Sci* 2011, 36:239–244.
133. Guo F, Cen S, Niu M, Yang Y, Gorelick RJ, Kleiman L. The interaction of APOBEC3G with human immunodeficiency virus type 1 nucleocapsid inhibits tRNA3Lys annealing to viral RNA. *J Virol* 2007, 81:11322–11331.
134. Mbisa JBR, Thomas J, Vandergraaff N, Dorweiler I, Svarovskaia E, Brown W, Mansky L, Gorelick R, Harris R, Engelman R, et al. HIV-1 cDNAs produced in the presence of APOBEC3G exhibit defects in plus-strand DNA transfer and integration. *J Virol* 2007, 81:7099–7110.
135. Bishop KN, Verma M, Kim EY, Wolinsky SM, Malim MH. APOBEC3G inhibits elongation of HIV-1 reverse transcripts. *PLoS Pathog* 2008, 4:e1000231.
136. Luo K, Wang T, Liu B, Tian C, Xiao Z, Kappes J, Yu XF. Cytidine deaminases APOBEC3G and APOBEC3F interact with human immunodeficiency

virus type 1 integrase and inhibit proviral DNA formation. *J Virol* 2007, 81:7238–7248.

137. Wang T, Zhang W, Tian C, Liu B, Yu Y, Ding L, Spearman P, Yu XF. Distinct viral determinants for the packaging of human cytidine deaminases APOBEC3G and APOBEC3C. *Virology* 2008, 377:71–79.

138. Stauch BHH, Perkovic M, Weisel M, Kopietz F, Cichutek K, Münk C, Schneider G. Model structure of APOBEC3C reveals a binding pocket modulating ribonucleic acid interaction required for encapsidation. *Proc Natl Acad Sci U S A* 2009, 106:12079–12084.

FURTHER READING

APOBEC Structure

Autore F, Bergeron JR, Malim MH, Fraternali F, Huthoff H. Rationalisation of the differences between APOBEC3G structures from crystallography and NMR studies by molecular dynamics simulations. *PLoS One* 2010, 5:e11515.

Harjes E, Gross PJ, Chen KM, Lu Y, Shindo K, Nowarski R, Gross JD, Kotler M, Harris RS, Matsuo H. An extended structure of the APOBEC3G catalytic domain suggests a unique holoenzyme model. *J Mol Biol* 2009, 389:819–832.

Holden LG, Prochnow C, Chang YP, Bransteitter R, Chelico L, Sen U, Stevens RC, Goodman MF, Chen XS. Crystal structure o the anti-viral APOBEC3G catalytic domain and functional implications. *Nature* 2008, 456:121–124.

Ichikawa HT, Sowden MP, Torelli A, Bachl J, Huang P, Dance GSC, Marr SH, Robert J, Wedekind JE, Smith HC, et al. Structural phylogenetic analysis of activation-induced deaminase function. *J Immun* 2006, 177:355–361.

APOBEC Protein- and RNA-Interactions and Catalytic Activity

Arakawa H, Hauschild J, Buerstedde JM. Requirement of the activation-induced deaminase (AID) gene for immunoglobulin gene conversion. *Science* 2002, 295:1301–1306.

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, 285:16195–16205.

Feng Y, Chelico L. Intensity of deoxycytidine deamination of HIV-1 proviral DNA by the retroviral restriction factor APOBEC3G is mediated by the noncatalytic domain. *J Biol Chem* 2011, 286:11415–11426.

Friew YN, Boyko V, Hu WS, Pathak VK. Intracellular interactions between APOBEC3G, RNA, and HIV-1 Gag: APOBEC3G multimerization is dependent on its association with RNA. *Retrovirology* 2009, 6:4690–4696.

Lavens D, Peelman F, Van der Heyden J, Uyttendaele I, Catteeuw D, Verhee A, Van Schoubroeck B, Kurth J, Hallenberger S, Clayton R, et al. Definition of the interacting interfaces of APOBEC3G and HIV-1 Vif using MAPPIT mutagenesis analysis. *Nucleic Acids Res* 2010, 38:1902–1912.

Li X, Ma J, Zhang Q, Zhou J, Yin X, Zhai C, You X, Yu L, Guo F, Zhao L, et al. Functional analysis of the two cytidine deaminase domains in APOBEC3G. *Virology* 2011, 414:130–136.

Shlyakhtenko LS, Lushnikov AY, Li M, Lackey L, Harris RS, Lyubchenko YL. Atomic force microscopy studies provide direct evidence for dimerization of the HIV restriction factor APOBEC3G. *J Biol Chem* 2011, 286:3387–3395.

Uyttendaele I, Lavens D, Catteeuw D, Lemmens I, Bovijn C, Tavernier J, Peelman F. Random mutagenesis MAPPIT analysis identifies binding sites for Vif and Gag in both cytidine deaminase domains of APOBEC3G. *PLoS One* 2012, 7:e44143.

APOBEC Subcellular Distribution

Ito S, Nagaoka H, Shinkura R, Begum N, Muramatsu M, Nakata M, Honjo T. Activation-induced cytidine deaminase shuttles between nucleus and cytoplasm like apolipoprotein B mRNA editing catalytic polypeptide 1. *Proc Natl Acad Sci U S A* 2004, 101:1975–1980.

Stenglein MD, Matsua H, Harris RS. Two regions within the amino-terminal half of APOBEC3G cooperate to determine cytoplasmic localization. *J Virol* 2008, 82:9591–9599.