7494–7506 *Nucleic Acids Research, 2017, Vol. 45, No. 12* *Published online 30 May 2017*
doi: 10.1093/nar/gkx362

Structural determinants of APOBEC3B non-catalytic domain for molecular assembly and catalytic regulation

Xiao Xiao[1,2,†], Hanjing Yang[2,†], Vagan Arutiunian[2], Yao Fang[2,3,4], Guillaume Besse[2,5], Cherie Morimoto[2], Brett Zirkle[1,2] and Xiaojiang S. Chen[1,2,6,7,*]

[1]Genetic, Molecular and Cellular Biology Program, Keck School of Medicine, University of Southern California, Los Angeles, CA 90089, USA, [2]Molecular and Computational Biology, Departments of Biological Sciences and Chemistry, University of Southern California, Los Angeles, CA 90089, USA, [3]Department of Clinical Microbiology and Immunology of Southwest Hospital, Third Military Medical University, Chongqing 400038, China, [4]161 Hospital, Wuhan 430012, China, [5]Polytech' Clermont-Ferrand, Université Blaise Pascal, Clermont-Ferrand, France, [6]Center of Excellence in NanoBiophysics, University of Southern California, Los Angeles, CA 90089, USA and [7]Norris Comprehensive Cancer Center, University of Southern California, Los Angeles, CA 90089, USA

Received October 11, 2016; Revised April 04, 2017; Editorial Decision April 19, 2017; Accepted May 27, 2017

ABSTRACT

The catalytic activity of human cytidine deaminase APOBEC3B (A3B) has been correlated with kataegic mutational patterns within multiple cancer types. The molecular basis of how the N-terminal non-catalytic CD1 regulates the catalytic activity and consequently, biological function of A3B remains relatively unknown. Here, we report the crystal structure of a soluble human A3B-CD1 variant and delineate several structural elements of CD1 involved in molecular assembly, nucleic acid interactions and catalytic regulation of A3B. We show that (i) A3B expressed in human cells exists in hypoactive high-molecular-weight (HMW) complexes, which can be activated without apparent dissociation into low-molecular-weight (LMW) species after RNase A treatment. (ii) Multiple surface hydrophobic residues of CD1 mediate the HMW complex assembly and affect the catalytic activity, including one tryptophan residue W127 that likely acts through regulating nucleic acid binding. (iii) One of the highly positively charged surfaces on CD1 is involved in RNA-dependent attenuation of A3B catalysis. (iv) Surface hydrophobic residues of CD1 are involved in heterogeneous nuclear ribonucleoproteins (hnRNPs) binding to A3B. The structural and biochemical insights described here suggest that unique structural features on CD1 regulate the molecular assembly and catalytic activity of A3B through distinct mechanisms.

INTRODUCTION

The generation of genome instability plays critical roles in both the initiation and progression of neoplastic disease and is therefore one of the defining hallmarks of cancer (1). The acquisition of excessive somatic mutations creates cancer heterogeneity that promotes cancer development and evolution. In recent years, kataegis patterning of hypermutation clusters dominated by C-to-T transitions has been identified in many types of cancers (2–6). Several studies have revealed that mutational patterns within kataegis are correlated with the upregulated expression of a cytidine deaminase APOBEC3B (A3B) (4–7). A3B catalyzes the cytidine deamination with a highly preferred TC dinucleotide hotspot, consistent with a subtype of mutational signatures found in bladder, cervical, head and neck, lung, breast, ovarian cancers and myeloma (2,4–9). Overexpression of A3B causes excessive mutations that can lead to cancer and the upregulated expression level of A3B is also correlated with the increased tendency of *TP53* mutations in tumors and poor prognosis in cancer patients (5,8,10–15). Given the relationship between tumor pathogenesis and the potential role of A3B as an enzymatic source of genome instability, significant insights can reside in understanding the basis of A3B's biological and enzymatic function.

APOBEC is a family of cytidine deaminases consisting AID, APOBEC1 (A1), APOBEC2 (A2), APOBEC3s (A3s, A3A to A3H) and APOBEC4 (A4). They mostly catalyze C-to-U transitions and result in C-to-T mutations on single-stranded DNA (ssDNA), in addition A1 and A3A can edit RNA (16,17). A3B belongs to subfamily of A3s that were originally identified for their role as host restriction factors

*To whom correspondence should be addressed. Tel: +1 213 740 5487; Fax: +1 213 740 4340; Email: xiaojiac@usc.edu
†These authors contributed equally to this work as first authors.

against retroviruses and DNA viruses, as well as endogenous retroelements (18–24). Within the human A3 subfamily, there are seven members, consisting of either one (A3A, A3C, and A3H) or two (A3B, A3D, A3F, A3G) homologous zinc-coordinating cytidine deaminase (CD) domains. All the single-domain A3s and the C-terminal deaminase domains (CD2 or CTD) of double-domain A3s are catalytically active, with preference of TC or CC dinucleotide on ssDNA substrates. In contrast, the N-terminal domain (CD1 or NTD) of double-domain A3s, is catalytically inactive and presumably involved in the interactions with nucleic acids and protein oligomerization (25,26). While the catalytic A3 domains have been extensively studied, CD1 is much less characterized mostly due to their poor solubility and complexity of interactions with nucleic acids.

So far, the only well-studied CD1 domain is from A3G, an HIV-1 restriction factor (27–29). Previous studies showed that A3G-CD1 not only greatly enhances the substrate binding, processivity, and deamination activity *in vitro* (30–34), but also mediates the RNA-dependent oligomerization of A3G (35–39). The oligomerization of A3G is shown to be required for A3G to incorporate into HIV virion and inhibit viral replication (34,40–42). In the cytoplasm of activated CD4+ T cells, A3G resides in RNA-depended high-molecular-weight (HMW, also termed high-molecular-mass or HMM) complexes as an inactive form, which can be converted into low-molecular weight (LMW, also termed low-molecular-mass or LMM) complex and activated by RNase A treatment (43). This phenomenon has also been observed in A3G purified from Sf9 cells that displayed similar conversion between inactive HMW aggregates and active LMW form (30). Mutational analysis has resolved two tryptophan residues on A3G-CD1, W94 and W127, are involved in HMW complex assembly (39). Recently the crystal structure of a primate A3G-CD1 (rA3G-CD1) revealed that the interactions within the dimer interface are critical for nucleic acid binding and oligomerization (27). Studies on the CD1 of A3F further support the observation that CD1 enhances the catalytic activities and mediates HMW assembly of A3 double-domain proteins (44), however, the HMW complexes of A3F are more resistant to RNase A (37,45). The formation of HMW species has emerged to be a significant property of double-domain A3 biology and function. It has been suggested that this unique property is shared with A3B as molecular brightness analysis has shown that A3B possesses a propensity to form concentration-dependent oligomers similar to other double-domain A3s (46). Of a further note, A3B interactions with hnRNP K and other RNPs have been observed (47). While the crystal and NMR structures of A3B-CD2 and its complex with ssDNA have revealed the important residues and loops of CD2 for its catalytic activity (48–50), it is unclear how A3B-CD1 can greatly enhance the catalytic activity of A3B-CD2 in the full-length A3B (51,52) and a complete atomic-level understanding of A3B remains elusive.

In this study, we investigate the biochemical and structural characteristics of A3B-CD1 for regulating HMW complex assembly and catalytic activity of A3B. Here, we present the crystal structure of a soluble human A3B-CD1 variant (A3B-CD1m) at 1.9 Å resolution. Surprisingly, the pursuit of a monomeric form of A3B-CD1 has led to the identification of four tyrosine (4Y) residues involved in the A3B-HMW complex assembly in HEK293T cells. Mutating these tyrosine residues generates a LMW mutant of A3B that no longer forms high-order complexes in HEK293T cells and displays a reduced catalytic activity. Addition of one tryptophan (W127) mutation to the 4Y mutant abolishes ssDNA binding and further impairs its catalytic activity, even though this W127 mutation alone does not elicit a detectable change on the activity. Moreover, structure-guided mutagenesis study has identified a role for the positive charge patch around loop-2, loop-4 and β5 of CD1 in the RNA-dependent regulation of catalytic activity. In addition, co-immunoprecipitation assay suggests that the combination of 4Y and W127 mutations affect preferential binding of multiple hnRNPs to A3B. Taken together, this study reveals several critical elements of A3B-CD1 that regulate the molecular assembly and catalytic behavior of A3B, an enzyme that is emerging to play an important role in cancer development and cancer evolution.

MATERIALS AND METHODS

Protein expression and purification

A codon-optimized A3B-CD1m (Y13D/Y28S/Y83D/W127S/Y162D/Y191H) was cloned into pMAL-c5X (NEB) with an N-terminal MBP tag and a PreScission cleavage site. *Escherichia coli* cells transformed with the plasmid were grown in LB media at 37 °C until OD_{600} reached 0.4. The media with cells were then transferred into 16 °C shaker and induced with 0.2 mM IPTG when OD_{600} reached 0.6 for overnight induction. The harvested bacteria cell pellets were resuspended in buffer A (25 mM HEPES pH 7.5, 250 mM NaCl, 20 mM $MgCl_2$, 1 mM DTT) with 0.1 mg ml^{-1} RNase A (Qiagen) and lysed by Microfluidizer. After centrifugation, the supernatant of cell lysates was incubated with amylose resin (NEB), washed with buffer A with 500 mM, 1 M and 500 mM NaCl gradient, and eluted with buffer B (50 mM HEPES pH 7.5, 500 mM NaCl, 40 mM maltose, 0.5 mM TCEP). MBP-A3B-CD1m was applied to Superdex 75 gel filtration column (GE Healthcare) equilibrated with buffer B. Monomeric fractions of MBP-A3B-CD1m were pooled, concentrated, and cleaved by PreScission protease in buffer C (50 mM HEPES pH 7.5, 250 mM NaCl, 0.5 M arginine, 0.5 mM TCEP), then purified again using Superdex 75 column. The cleaved monomeric A3B-CD1m was collected for crystallization. For the EMSA assay, MBP-fusion A3B-CD1m and mutants were purified by the same methods without PreScission cleavage.

Protein crystallization, data collection, structure determination and refinement

Purified A3B-CD1m was concentrated to 6 mg ml^{-1} for crystallization screening. Crystals were obtained by sitting drop vapor-diffusion method in the condition including 0.01 M iron (III) chloride hexahydrate, 0.1 M sodium citrate dehydrate pH 5.6, 10% (v/v) Jeffamine M-600 from an aged screening block (Hampton Research). Diffraction data was collected from Advanced Light Source BL-821 and Advanced Photon Source 19-ID/23-ID. Data sets were

Table 1. Crystallographic data collection and refinement statistics

	A3B-CD1m
Data collection	
Space group	*P*21 21 21
Cell dimensions	
a, *b*, *c* (Å)	60.7, 60.8, 111.5
α, β, γ (Å)	90, 90, 90
Resolution (Å)	50–1.90 (1.97–1.90)[a]
R_{sym} or R_{merge}	6.6 (67.7)
I / σI	29.2 (5.1)
Completeness (%)	99.8 (100)
Redundancy	11.3 (10.2)
Molecules per ASU	2
Refinement	
Resolution (Å)	30.73–1.90
No. reflections	32 957
R_{work}/R_{free}	18.6/20.9
No. atoms	3368
Protein	3132
Ligand/ion	2
Water	234
B-factors	36.39
Protein	35.93
Ligand/ion	24.57
Water	42.67
r.m.s. deviations	
Bond lengths (Å)	0.007
Bond angles (°)	0.790

Structure was determined from a single crystal.
[a]Highest-resolution shell is shown in parentheses

indexed, integrated and scaled using HKL2000 program package. The structure of A3B-CD1m was determined by molecular replacement method by molrep (CCP4 suite) using rA3G-CD1 (PDB: 5K81) as the template. The structure was then refined by PHENIX and manually checked in COOT. The statistics for diffraction data and structural determination/refinement is shown in Table 1.

Construction of A3B and mutants for human cell-based assays

A3B-CD1, A3B-CD2, and FL A3B E255A without intron were cloned into a pcDNA 3.1(+) vector (Invitrogen) with an N-terminal FLAG tag. Our original cloning of an N-terminal FLAG tagged wild-type A3B coding sequence into pcDNA 3.1(+) vector was not successful, likely due to the genotoxicity of A3B to *E. coli*. Another trial of cloning A3B retaining the entire C-terminus intron 7 (53) was successful in *E. coli*, but the intron 7 was not spliced out efficiently during expression in HEK293T cells in our system, resulting an A3B carrying extra 30 amino acid residues at the C-terminus due to the un-spliced intron 7 and the subsequent frameshift. To resolve this problem, we tested a 235 bp intron (Supplementary Figure S1A) from adenovirus type 2 RNase gene between exon 1 and exon 2 (54,55) into pcDNA 3.1(+)-FLAG-A3B E255A construct before codon E255A. E255 is the critical residue for A3B catalytic activity. If the introduced intron sequence retains during A3B expression in 293T cells, there is an in-frame stop codon generated within the intron, produces a non-catalytic 33 kDa truncated product. Western blot against the N-terminal FLAG confirms that this A3B construct has the efficient splicing, and generates the expression product of the same size as the no-intron A3B E255A control construct (Supplementary Figure S1B). The codon of E255A mutation was then corrected back to E255 to generate the wild-type A3B. Derivatives were then generated by site-directed mutagenesis PCR.

Cell culture, transfection and western blot

HEK293T cells (ATCC) were maintained in DMEM medium (10013CV, Corning), supplemented with 10% FBS, 100 U ml^{-1} penicillin and 100 μg ml^{-1} streptomycin. MDA-MB231 cells (ATCC) were maintained in DMEM low-glucose medium (11885084, Gibco), supplemented with 10% FBS, 100 U ml^{-1} penicillin and 100 μg ml^{-1} streptomycin. Transfections were done by using X-tremeGENE 9 DNA Transfection Reagent (Roche) and following manufacturer's recommendation. For western blot, cell lysate samples were analyzed by SDS-PAGE, transferred onto PVDF membrane (EMD Millipore), and blotted with anti-FLAG M2 mAb (F3165, Sigma, 1:3000), anti-alpha tubulin mAb (GT114, GeneTex, 1:5000), anti-GAPDH mAb (GTX627408, GeneTex, 1:5000), anti-hnRNP K (ab134060, Abcam, 1:5000), anti-PTBP1 (1:8000), anti-hnRNP Q (R5653, Sigma, 1:1000), anti-hnRNP U (A300-A689A, Bethyl, 1:1000), anti-A3B (5210-87-13, NIH AIDS Reagent Program from Dr Reuben Harris, 1:20) (56).

FPLC analysis for native complex formation

The FPLC analysis of native complex formation protocol was modified from a previous report (43). At 72 h post-transfection, A3B-transfected 293T cells in 150 mm^2 dishes were harvested, washed with PBS, and lysed in lysis buffer (50 mM HEPES pH 7.5, 125 mM NaCl, 0.6% NP-40 alternative, 1 mM DTT, final total volume is 1 ml after mixing with cells) with 1 × Halt protease and phosphatase inhibitor (Thermo Fisher) for 10 min. After centrifugation and removing the surface lipid fraction, the clear supernatant fraction was loaded onto Superdex 200 10/300 GL column equilibrated with 50 mM HEPES pH 7.5, 125 mM NaCl, 0.1% NP-40 alternative, 1 mM DTT, 10% glycerol. Fractions were subjected to western blot. For RNase A treatment, the clear supernatant after lysis were incubated with 100 μg ml^{-1} RNase A (Qiagen) on ice for 2 h before loading onto Superdex 200 10/300 GL column. For analysis of endogenous A3B in MDA-MB231 cells, one dish of 150 mm^2 cells was harvested, lysed, and analyzed by FPLC with the same method as above. After FPLC, 450 μl of each fraction was incubated with nine times of volumes of pre-cooled ethanol in −80°C for overnight precipitation. The precipitated protein was resuspended in 30 μl lysis buffer for western analysis.

In vitro deamination assay with HEK293T cell lysates

At 48 h post-transfection, A3B-transfected 293T cells in 12-well plates were harvested, washed with PBS, and the whole cell lysates were prepared using M-PER protein extraction reagent (Thermo Fisher) with 1 × Halt protease and phosphatase inhibitor (Thermo Fisher). After centrifugation, the clear supernatant fraction was used for deamination. Prior to deamination reaction, the expression level

of each A3B sample was quantified by western blot and the total protein concentration was quantified by BCA protein assay (Pierce), and normalized with the whole cell lysate transfected with the empty pcDNA 3.1(+) vector.

Deamination assay was modified from previous reports (51,57). 5′-FAM labeled 50 nt ssDNA (ATTATT ATTATT CAAATT TATTTA TTTATT TATGGT GTTTGG TGTGGT TG, 300 nM) was incubated with 8 μl lysate in deamination buffer (25 mM HEPES, pH 7.2, 50 mM NaCl, 1 mM DTT and 0.1% Triton X-100) with a final volume of 20 μl. RNase A of a final concentration 10 ng/μl (Qiagen) was added together with the substrate and reaction buffer as indicated. The reactions with or without RNase A were done concurrently. The deamination reaction was incubated at 37°C for 1 h terminated by heat inactivation at 90°C for 5 min, and followed by UDG treatment (2.5 U, New England Biolabs) at 37°C for 1 h. Hydrolysis of abasic sites was carried out in the presence of 0.1 M NaOH at 90°C for 10 min, then the samples were mixed with an equal volume of 2× loading dye (25 mM EDTA and 95% formamide) and analyzed by 20% denaturing PAGE. Biorad PharosFX scanner was used to visualize the images and Quantity One 1D analysis software (Bio-Rad Laboratories) was used for quantification. The averaged values (of three independent transfections) with S.E.M. are shown in the scattering plots and bar graphs.

Electrophoresis mobility shift assay (EMSA)

Purified MBP-fused A3B-CD1m and mutants were incubated with 75 nM of 5′-FAM labeled 30 nt ssDNA (ATTTAT ATTATT TATCCC TATTTA TATTTA) and 25 nM of 5′-FAM labeled 50 nt RNA (AUUAUU AUUAUU UAUCCC UAUUUA UAUUUA UUGUUA UUGUUA UUGUUA UA) on ice in the condition of 60 mM Tris pH 8.0, 100 mM NaCl (10 μl reaction volume) for 10 min. The reaction mixture was then mixed with 2 μl of 80% glycerol and analyzed by native PAGE.

Co-immunoprecipitation assay

At 60 h post-transfections, A3B-transfected 293T cells were harvested, washed with PBS, and lysed in lysis buffer (50 mM HEPES pH 7.5, 150 mM NaCl, 0.6% NP-40 alternative, 1 mM EDTA, 1 mM DTT, final total volume was 1 ml after mixing with cells) with 1 × Halt protease and phosphatase inhibitor (Thermo Fisher) for 10 min. Cell lysates were incubated with anti-FLAG M2 agarose (Sigma) beads with and without 100 μg ml^{-1} RNase A (Qiagen) at 4 °C for 3 h. Beads were then washed four times with 1 ml wash buffer (50 mM HEPES pH 7.5, 150 mM NaCl, 0.6% NP-40 alternative, 1 mM EDTA, 1 mM DTT) before eluting with FLAG elution buffer (wash buffer supplemented with 250 μg ml^{-1} of 3xFLAG peptide (APExBIO)). The eluted samples were analyzed by western blot.

Structural modeling and analysis

Structure of wild-type A3B-CD1 was modeled by Phyre2 one-to-one threading method (58), using the structure of A3B-CD1m as template. Surface electrostatic potential of A3B-CD1m and other A3s was calculated by APBS (59).

RESULTS

Crystal structure of a monomeric variant A3B-CD1m

To pursue structural studies on A3B-CD1, the insolubility and aggregation issues of the wild type A3B-CD1 domain (1–191) expressed in *E. coli* needed to be resolved. With the goal of producing soluble and homogeneous A3B-CD1 constructs, we generated an A3B-CD1 homologous model based on the rA3G-CD1 crystal structure (27) and the sequence alignments of several APOBEC homologs including A3B-CD1 and rA3G-CD1 (Supplementary Figure S2). We engineered several variants of A3B-CD1 carrying specific mutations based on the previously determined rA3G-CD1 dimerization interface residues as well as solvent exposed surface hydrophobic residues. After systematically testing a series of mutations, we obtained a mutant carrying six mutations Y13D /Y28S /Y83D /W127S /Y162D /Y191H that had reduced A3B-CD1 aggregation and markedly improved solubility. Four of these six residues are surface tyrosine mutations (Y13D /Y28S /Y83D /Y162D), one (W127S) is from previously determined rA3G-CD1 dimerization interface, and one (Y191H) is the C-terminus hydrophobic residue mutation. This A3B-CD1 variant, referred to as A3B-CD1m, was readily purified in monomeric form when fused to an N-terminal MBP tag (Supplementary Figure S3). After cleaving off MBP, the tag-free A3B-CD1m was crystallized, and the crystal structure to 1.9 Å resolution was solved (Figure 1, Table 1).

The crystal structure of A3B-CD1m has the typical, conserved APOBEC CD fold that is composed of five core β-sheets, surrounded by six α-helices and connecting loops (Figure 1A). The structure superimposes well with previously determined APOBEC structures, especially at the Zn-coordinated center and the core five stranded β-sheet. A representative structural superimposition between A3B-CD1 and a previously published A3B-CD2 is shown in Figure 1B (49). A zinc atom is coordinated with one histidine (H66) and two cysteines (C97, C100) (Figure 1C and D), the three-dimensional conformation of which resembles the conserved structural configuration that is previously observed in other catalytic APOBEC proteins (33,48,49,60–69) and non-catalytic rA3G-CD1 (27). Like some other Z2-type domain structures including A3C, rA3G-CD1 and A3F-CD2 (27,62,63,69), the length of loop-3 in A3B-CD1m is relatively short, resulting in a closed formation of the Zn-coordinated center. This is distinct from Z1-type domains including A3A and A3G-CD2 (33,61,66) (Figure 1E). The mutated residues Y13, Y28, Y83, W127, Y162 are scattered on N-terminus, loop-1, loop-4, loop-7 and loop-10, respectively (Figure 1G). One unique feature of A3B-CD1m is the discontinuous α4, forming a small 3_{10} helix α4′ at the beginning of the helix (Figure 1B and F), and compared to other APOBEC structures, this beginning part of α4′ in A3B-CD1m is shifted away from the interior core of the Zn-coordinated center (Figure 1F). As a result, this α4′ at beginning of α4 and the adjacent loop-7 residues show different conformations from that of rA3G-CD1 (Figure 1F).

The plot of surface charge features show that A3B-CD1m clearly has a much more positively charged molecular sur-

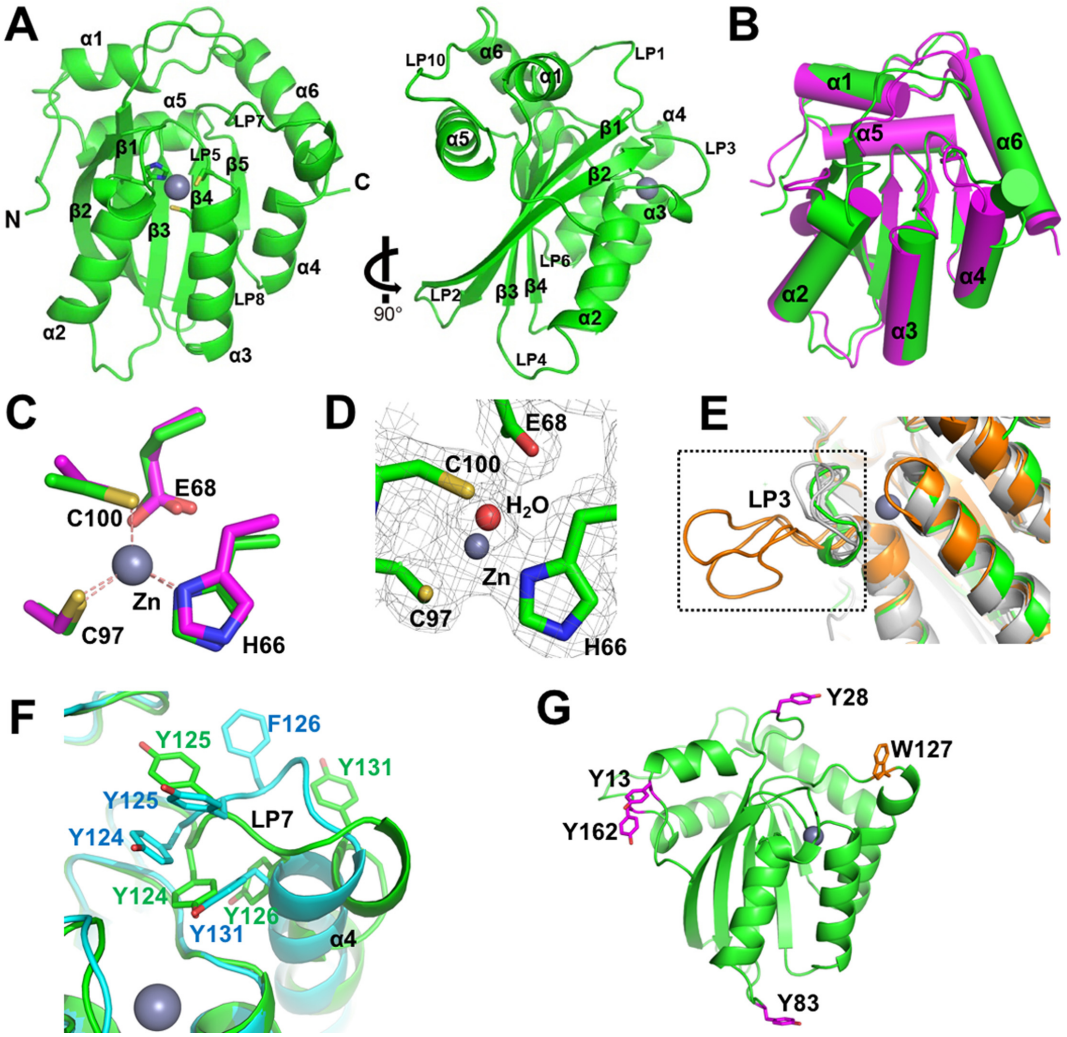

Figure 1. Crystal structure of A3B-CD1m. (**A**) Two views of the 1.9 Å monomer structure of A3B-CD1m, related by 90° rotation. Secondary structure nomenclature is as described previously (84). (**B** and **C**) Superimposition of A3B-CD1m (green) with the crystal structure of A3B-CD2 (49) (magenta) in secondary structures (B) and zinc-coordinated active center (C). The beginning of α4 of A3B-CD1 breaks into a small 3_{10} helix α4′ (B). (**D**) Electron density map of A3B-CD1m around active center, contouring at 1.0σ level. (**E**) Structural comparison of loop-3 (LP3) between A3B-CD1m (green), Z1 (A3A (66), A3G-CD2 (33), orange) and Z2 (A3F-CD2c (63), A3C (62), rA3G-CD1 (27), white) domains. Dashed box indicates loop-3. (**F**) Structural comparison of loop-7 (LP7) and α4 between A3B-CD1m (green) and rA3G-CD1 (cyan). The α4′ at the beginning of α4 on A3B-CD1m is labeled. (**G**) Model of A3B-CD1 WT based on A3B-CD1m structure. Four surface tyrosine residues (magenta) and W127 (orange) mutated in A3B-CD1m are shown as magenta sticks.

face than A3B-CD2 (compare Figure 2A top and bottom), but less than rA3G-CD1 (compare Figure 2A top and middle). The charged residues on A3B-CD1m surface are clustered to form two major positive charge patches: one (patch 1) is around the pseudo active center, including β2, loop-3, loop-1, α2, α3 and α4 (patch 1 in Figure 2A and B). The other one (patch 2) is on the opposing side that starts from loop-2, and extends through β sheets to α6 (patch 2 in Figure 2A, C). In summary, the resolved crystallographic information demonstrates that while the CD1 domain exhibits the conserved APOBEC architecture, unique atomic-level details are present that may give new insights into the fundamental mechanisms of A3B function.

A3B forms hypoactive HMW complexes and is activated by RNase A without apparent conversion to LMW

Double-domain APOBEC oligomerization and formation of HMW complexes in human cells have been extensively studied with A3G and to a lesser extent, A3F (30,32,33,35,37,39,43,45,46,70–74). All four double-domain A3s are shown to behave as large oligomeric forms in living cells based on brightness characterization to the EGFP-labeled APOBEC proteins (46). So far, the available evidence suggests that the CD1 domains of A3G and A3F are mainly responsible for the HMW complexes assembly and high-order oligomerization (32,35,37,39). The monomeric nature of A3B-CD1m prompted us to examine the HMW complex formation of the full-length A3B in human cells, and whether or not the mutations that generate

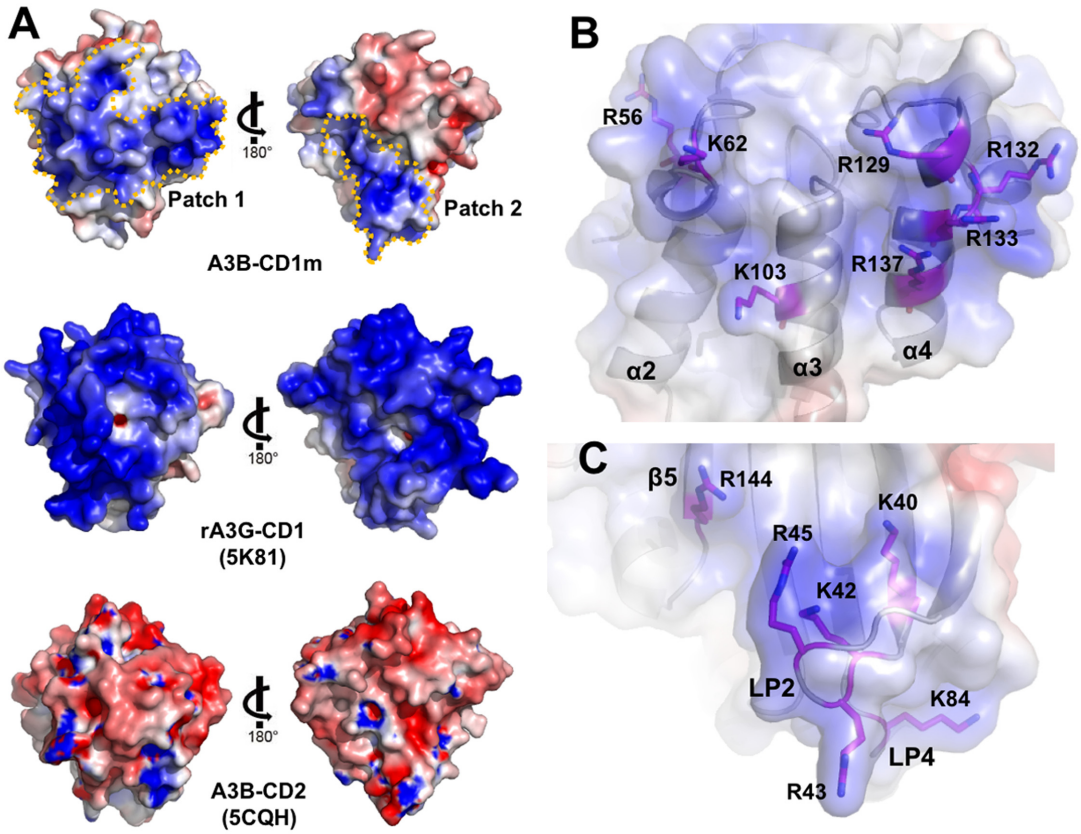

Figure 2. Surface electrostatic potential of A3B-CD1m colored according to calculated electrostatic potential of accessible surface area from $-5kT/e$ (red) to $5kT/e$ (blue). (**A**) Two major positive charge patches (patch 1, 2, blue) of A3B-CD1m are encircled with yellow dashed lines. This is different from rA3G-CD1 (27) and A3B-CD2 (49). The left panel represents the same orientation as the Figure 1A. (B, C) Positive charge residues within patch 1 (**B**) and patch 2 (**C**) are shown in magenta sticks.

the monomeric A3B-CD1m would impact the HMW complex assembly of the full-length A3B.

We used a pcDNA 3.1 (+) mammalian expression vector with an N-terminal FLAG tag to express the wild-type A3B in HEK293T cells. Initial attempts of cloning the wild-type A3B gene were unsuccessful, likely due to the genotoxicity of A3B to *E. coli* (52,53). To overcome this obstacle, we interrupted the wild-type A3B coding sequence by introducing a heterologous intron sequence originated from adenovirus (54,55) (see Materials and Methods for details). This heterologous intron was placed in between the two codons A254 and E255 of A3B and un-spliced transcripts would result in loss of enzymatic activity of A3B. Subsequent cloning of this construct in *E. coli* was successful and the correct splicing of the intron sequence in HEK293T cells was confirmed by western blot (Supplementary Figure S1B).

To examine HMW complex formation of the expressed A3B, the soluble fractions of HEK293T cell lysates were prepared and fractionated by Superdex 200 analytical size-exclusion chromatography (Figure 3, Supplementary Figure S4). The distribution of A3B in the eluted fractions was detected by western blot using monoclonal antibody against the FLAG epitope fused to A3B. Endogenous α-tubulin was used as an internal size marker, as it assembles into ∼100 kDa heterodimers (75) (Figure 3). As expected, the majority of the wild-type A3B expressed in human cells eluted out in early fractions indicating that A3B predominantly exists in HMW forms (Figure 3A, Supplementary Figure S4A and B). Consistently, the endogenous A3B from MDA-MB231 breast cancer cells that are previously reported with up-regulated A3B expression (5,76), also eluted as HMW species (Figure 3B). Surprisingly, after RNase A treatment, most of A3B remains as HMW species (Figure 3A, Supplementary Figure S4A). In contrast, the HMW species of A3G has been dissociated into LMW species under the same experimental condition (Figure 3A), consistent with prior reports (43).

We examined the deaminase activity of the full-length A3B in the whole cell lysates of the transfected HEK293T cells. The results showed that relatively low deaminase activity of A3B was present, but the activity was drastically activated by RNase A treatment of the lysates (Supplementary Figure S5A). This RNase A-dependent enhancement of deaminase activity is much less evident in the CD2 construct alone, and no detectable activity is seen with either CD1 or E255A inactive mutant (Supplementary Figure S5A). In fact, the activation of the deaminase activity by RNase A treatment was also observed in the elution fractions that contained A3B HMW species (Figure 3C). These results suggest that the deamination activity of A3B is attenuated in an RNA-dependent manner, which is consistent

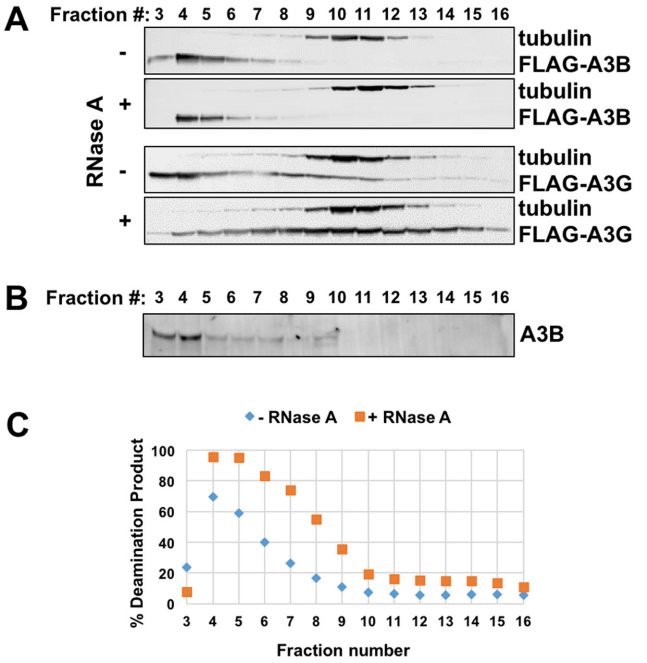

Figure 3. Analysis of the oligomeric status of the wild-type A3B. (**A**) Western blot of FPLC fractions of HEK293T cell lysate expressing A3B and A3G under no RNase A and with RNase A conditions. α-tubulin is an endogenous control. The fraction shift of A3B under with RNase A condition is due to the slightly variation of FPLC, as shown in Supplementary Figure S5A. (**B**) Western blot of FPLC fractions from MDA-MB231 cells lysate, showing the endogenous A3B. (**C**) The deamination activity of A3B FPLC fractions from A.

with previous observations in other APOBECs (30,77–79). However, the activation of deaminase activity by RNase A treatment of the HMW species of A3B is not dependent on the conversion from the HMW complex to LMW species (Figure 3A and C, and Supplementary Figure S4A).

Surface tyrosine residues and W127 on CD1 affect the assembly of A3B HMW complexes

To examine whether the residues mutated in A3B-CD1m affect A3B HMW complex assembly, we generated three A3B mutants with mutations on CD1: a single mutant W127A on loop 7, a quadruple-tyrosine mutant Y13D/Y28A/Y83D/Y162D ('4Y') on the exposed surface, and a combination mutant W127A + Y13D/Y28A/Y83D/Y162D ('W+4Y'). The C-terminal mutation Y191H was not included as it is located at the CD1–CD2 linker rather than on the CD1 surface. The theoretical molecular weight of A3B is ∼46 kDa, thus the soluble A3B from the cell lysates that appears in the fractions eluted later than the tubulin marker (∼100 kDa) likely belongs to dimeric or monomeric forms (or LMW species). The results showed that, unlike the wild-type A3B that eluted in the early HMW fractions, the W+4Y mutant of A3B eluted in the LMW fractions, indicating disruption of HMW complex formation by W+4Y mutations (Figure 4). The W127A single mutant, however, formed complexes with a wide range of molecular weights from HMW to LMW (Figure 4), suggesting that it may partially impair

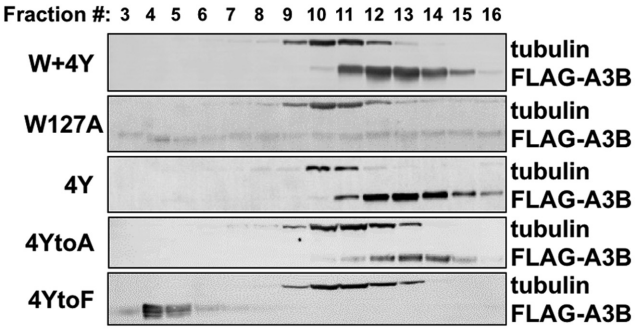

Figure 4. Analysis of the oligomeric status of the A3B mutants. W+4Y refers to W127A + Y13D/Y28A/Y83D/Y162D mutant, 4Y refers to Y13D /Y28A/Y83D/Y162D mutant, 4YtoA refers to Y13A/Y28A/Y83A/Y162A mutant, 4YtoF refers to Y13F/Y28F/Y83F/Y162F mutant.

the formation of the HMW complexes. Interestingly, the mutant with 4Y mutation alone behaved similarly to the W+4Y mutant in terms of disrupting the HMW complex formation (Figure 4). Further investigation showed that a quadruple-alanine mutant Y13A/Y28A/Y83A/Y162A ('4YtoA') behaved similarly to '4Y' and 'W+4Y' mutants, whereas a quadruple-phenylalanine mutant Y13F/Y28F/Y83F/Y162F ('4YtoF') restores the formation of HMW species (Figure 4). These results indicate that the aromatic sidechains of these four surface tyrosine residues Y13, Y28, Y83, and Y162 on CD1 are critical for A3B HMW complex assembly in HEK293T cells, and W127 also affects HMW complex properties, even though W127 mutation alone did not completely dissociate the HMW complexes.

Surface tyrosine residues and W127 on CD1 are important for A3B deaminase activity

We next explored whether the residues mutated in A3B-CD1m affect the catalytic activity of the full-length A3B. We performed deamination assays with A3B expressed 293T cell lysate under conditions with or without RNase A treatment. The amount of 4Y and W+4Y mutants were normalized to the level of the wild-type A3B, as their original expression levels were significantly higher than the wild-type A3B. The results showed that under both conditions, the four single tyrosine mutants, Y13D ('Y_1'), Y28A ('Y_2'), Y83D ('Y_3'), Y162D ('Y_4'), displayed the activities similar to or slightly reduced than that of the wild type (Figure 5A–C). However, the catalytic activity of the combined 4Y mutant is greatly reduced (Figure 5A–D), indicating that these four tyrosine residues on CD1 collectively are involved in the regulation of the full-length A3B catalytic activity. The activity of W+4Y mutant is even more severely impaired than the 4Y mutant, with ∼80% reduction of activity compared to the wild-type A3B (Figure 5A–D).

This reduction of activity is likely due to the loss of ssDNA substrate binding. The ssDNA binding to the CD1 domain of W+4Y mutant (equivalent to A3B-CD1m) is greatly impaired compared with that of the CD1 domain of 4Y mutant (Figure 5F), suggesting that W127 is critical for ssDNA binding in the context of 4Y mutant. The CD1 domains of both W+4Y and 4Y mutants showed some weak

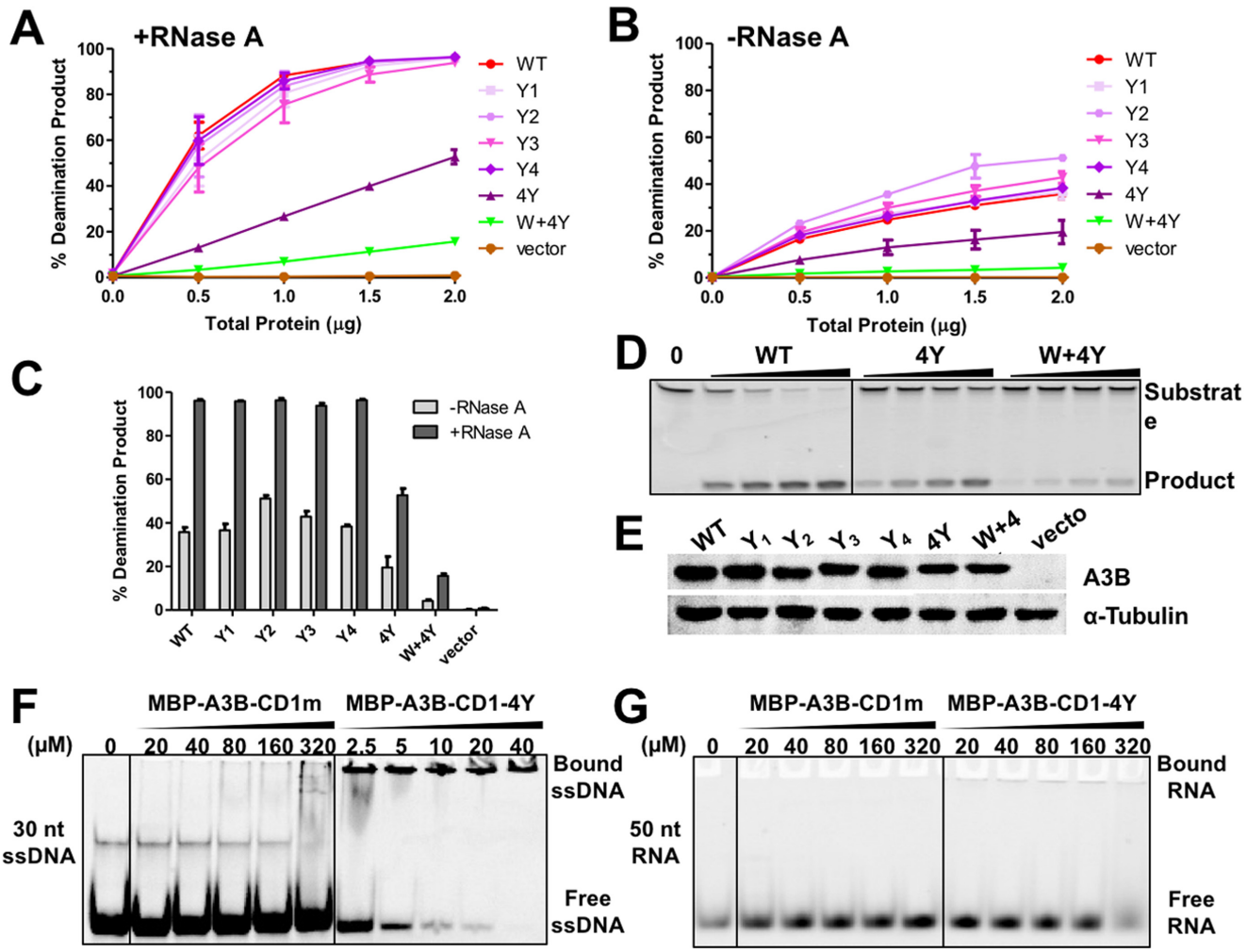

Figure 5. Deamination assay of HEK293T cell lysate expressing the wild-type A3B and mutants. (A, B) Quantified deamination results under conditions with RNase A (**A**) and without RNase A (**B**). (**C**) The percentage of the product with cell lysate of 2 μg total protein is also shown in bar graphs for comparison. (**D**) The represented results of deamination assay for wild-type A3B, 4Y and W+4Y mutants under the condition with RNase A. (**E**) The amount of A3B and mutants in the 293T cells lysate are normalized to the similar levels, and confirmed by western blot. (F, G) EMSA assay of MBP-A3B-CD1m and MBP-A3B-CD1-4Y mutant with 30 nt ssDNA (F) and 50nt RNA (**G**).

RNA binding, with the RNA binding to the CD1 domain of 4Y is slightly stronger (Figure 5G). Because the wild-type CD1 is not soluble, we could not include it in this study.

To further investigate the synergistic/additive effect of mutating W127 and each of the four tyrosine residues on CD1, we tested the activity of a set of mutants carrying W127A alone, or W127A in combination with each single tyrosine mutation respectively ('W+Y_1', 'W+Y_2', 'W+Y_3', 'W+Y_4' (Supplementary Figure S6). Among this group, the expression levels of W127A and W+Y_4 are the lowest, thus we normalized other mutants to the level of W127A and W+Y_4 (Supplementary Figure S6D). Interestingly, unlike W+4Y mutant that severely impaired the deaminase activity, W127A alone did not impair the deamination activity, and W127A in combination with any of the individual four tyrosine mutation only has minor but consistent reduction of deaminase activity under RNase A treated condition (Supplementary Figure S6). These results suggest that W127 and the four surface tyrosine residues on CD1 affect the catalytic activity of the full-length A3B in an additive or synergistic manner.

RNA-mediated attenuation of A3B catalytic activity is dependent on a site-specific, positively charged region on CD1

RNase A treatment of the wild-type A3B and the 4Y/W+4Y mutants exhibited an increase of deaminase activity under RNase A treatment, indicating that this class of mutants display similar levels of RNA-dependent activity attenuation as the wild type. In other words, the four surface tyrosine residues or W127 related mutants likely still interact with RNA in a similar manner that attenuates the deaminase activity of the wild-type A3B. The relatively more positively charged surface of A3B-CD1m in comparison to A3B-CD2 (Figure 2) suggests a possibility that these positive charge residues on CD1 may interact with RNA to affect the deaminase activity of the full-length A3B. Based on the structure of A3B-CD1m, we de-

signed two mutants, each harboring mutations of several positively charged residues located at the center of the positively charged patch 1 or patch 2 of CD1 (Figure 2B and C). Specifically, for patch 1, we introduced four mutations K62E/K103E/R132Q/R133E that reverse the charge of α2–α4 (designated as patch 1 mutant) and for patch 2, we simply replaced the loop-2 with the equivalent of A3B-CD2 (42KRGR45→DNGT) in combination with R144E on β5 (designated as patch 2 mutant).

We tested the deamination activities of these two patch mutants and compared with the wild-type A3B under the conditions with or without RNase A treatment (Figure 6A–E). In the presence of RNase A, both patch mutants showed a similar level of deamination activity as that of the wild type (Figure 6A and C), indicating that they have similar catalytic activity when there is no attenuation by RNA. In the absence of RNase A, however, a much reduced RNA-dependent activity attenuation was observed with the patch 2 mutant (Figure 6A–D). For example, with 2 μg cell lysate, the activity of the wild type and both mutants deaminated almost 100% substrate. In contrast, in the absence of RNase A, ∼70% of substrate was deaminated by patch 2 mutant, whereas only ∼30% of substrate was deaminated by patch 1 mutant and the wild-type A3B (Figure 6C). These results indicate that patch 2 plays a major role in mediating the RNA-dependent attenuation. We also examined whether any individual positive charge residue in patch 2 is sufficient to mediate this effect. Single mutations of all positive charge residues in patch 2 were tested (Supplementary Figure S7A-D). While most single residue mutants, namely K42A, K43A, K45A, K84A and R144A showed minor attenuation effect (Supplementary Figure S7A–C), none of them had a dramatic effect comparable to that of the combined patch 2 mutant (Figure 6C). Therefore, we concluded that these multiple positive charge residues within loop-2, loop-4 and β5 surface area on patch 2 collectively contribute to the RNA-dependent attenuation of A3B catalytic activity.

Surface hydrophobic residues of CD1 affect interactions between A3B and multiple hnRNPs

Previously A3B and other APOBECs have been reported to interact with multiple cellular hnRNPs (45,47,74,80). Here we asked whether four surface tyrosine residues, W127, and patch 2 of CD1 affect A3B–hnRNP interactions. We conducted pull-down of FLAG-tagged wild-type A3B and several A3B mutants expressed in 293T cells with anti-FLAG beads and examined the associated hnRNPs by western blot (Figure 7). Samples with vector alone and a FLAG-tagged A3G were also included as controls. The experiments were conducted under both with RNase A and without RNase A conditions. The results show that in the absence of RNase A, hnRNP Q, hnRNP U, as well as the previously reported hnRNP K and PTBP1 (hnRNP I), preferentially bind to the wild-type A3B. In contrast, the interactions of these hnRNPs with A3G are near background levels. Mutants carrying 4Y, 4YtoA, or patch 2 mutations retained the association with these hnRNPs similar to the wild-type A3B (Figure 7B). Remarkably, W+4Y mutant shows no detectable interactions with hnRNP K or PTBP1, and only very weak interactions with Q and U despite its high expression level among all tested constructs (Figure 7B). In the presence of RNase A, the association of A3B with hnRNP K and PTBP1 were much reduced, while apparently unaffected with hnRNP Q and hnRNP U (Figure 7B). Taken together, A3B interacts with multiple hnRNPs and the combination of four surface tyrosine residues and W127 mutations on CD1 significantly impaired A3B–hnRNP interactions.

DISCUSSION

The A3 family of cytidine deaminases has evolved to restrict viral infection and endogenous retrotransposition through pathways including deamination-induced hypermutation. However, this is a double-edged sword, as misregulation of deamination may result in excessive mutations that can lead to host genome instability. The non-catalytic CD1 domain of A3s apart from the catalytically active CD2 domain becomes an integral part of the biological activity of the double domain A3s. For A3B, the catalytic activity of CD2 alone is relatively weaker and less genotoxic, yet the overall activity of the full-length double-domain A3B can be enhanced or attenuated depending on the states of CD1. Understanding the role of A3B-CD1 in regulating the catalytic activity in the context of the full-length A3B would help to understand A3B function in cells and misregulation in cancer genomes.

In this study, we report the crystal structure of a human A3B-CD1 variant, and identified several structural elements of A3B-CD1 that play important roles in HMW complex formation and RNA mediated attenuation of A3B. The four surface tyrosine substitutions that greatly improved the solubility of A3B-CD1 are critical for A3B HMW complex assembly in human cells (Figure 4) and are also important for A3B deaminase activity (Figure 5). These four surface tyrosine residues are located at N-terminus (Y13), loop-1 (Y28), loop-4 (Y83), loop-10 (Y162). They are not equivalent to previously identified residues involved in A3G HMW complex formation (39) and oligomerization (27). Indeed, different from A3G HMW complexes which dissociate after RNase A treatment (43), A3B HMW complexes apparently do not become LMW species under the same RNase A treatment condition (Figure 4). This observation is similar to previous reports regarding RNase resistant A3F-HMW complex (37,45). Given the fact that A3BCD1 and A3FCD1 share 82% sequence identity, it is possible that the structural elements involved in HMW formation are conserved between these two CD1s. Results from 4YtoA and 4YtoF mutants (Figure 4) suggest that hydrophobic interactions are critical for A3B-HMW complex assembly in HEK293T cells. In addition to four surface tyrosine residues, W127 and the positive charge patch 2 also contribute to HMW formation. In each case, mutations cause partial dissociation of HMW species (Figure 4). It has been reported that W127 of A3G-CD1 and A3F-CD1 are involved in oligomerization of A3G and A3F (27,32,39,44). Therefore, a similar role may be played by W127 of A3B-CD1. The patch 2 mutant of A3BCD1 shows alleviated RNA-dependent attenuation of deaminase activity (Figure 6) suggesting that RNAs may be involved in A3B-HMW assembly. Furthermore, A3B-CD2 has been re-

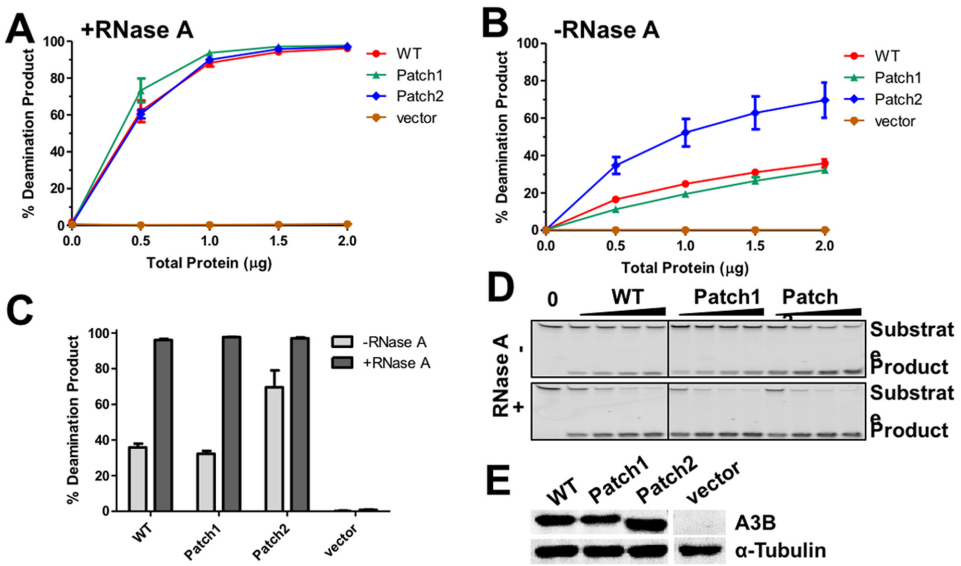

Figure 6. Deamination assay of HEK293T cell lysate expressing patch 1 and 2 mutants. (A, B) Quantified deamination results under conditions with RNase A (**A**) and without RNase A (**B**). (**C**) The percentage of the product with cell lysate of 2 μg total protein is also shown in bar graphs for comparison. (**D**) The represented results of deamination assay for wild-type A3B, 4Y and W+4Y mutants under the condition with RNase A and without RNase A. (**E**) The expression of A3B and mutants in the 293T cells lysate are at similar levels, confirmed by western blot.

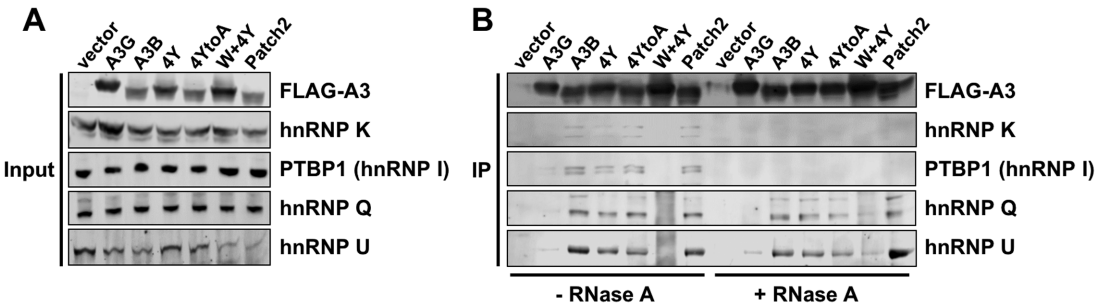

Figure 7. Co-immunoprecipitation of wild-type A3B and mutants with hnRNPs.

ported to dimerize on its own (52), thus it may play an additional role in the HMW formation of the full-length A3B.

Structural analysis of A3B CD1 resolved two positively charged patches on A3B-CD1m (Figure 2) that may bind nucleic acids through electrostatic interactions. These interactions may have dual roles in regulating the catalytic activity of A3B. On one hand, the interactions with ssDNA substrates facilitate the deaminase activity. On the other hand, the interactions with RNA molecules can attenuate the deaminase activity. Deamination assays of mutants that neutralized the positively charged regions on CD1 under the conditions with and without RNase A, allowed us to distinguish different interactions with ssDNA and RNA. This led to the identification of a positive charge patch around loop-2, loop-4, and β5 (on patch 2 surface in Figure 2A) that plays an important role in the RNA-dependent attenuation of deaminase activity, demonstrated by that the catalytic activity of an A3B containing mutation that reduces the surface charge within this region exhibited comparable catalytic activity in absence of RNase A when compared to treatment with RNase A (Figure 6); a phenomena not observed for the wild type A3B or a mutant aimed at reducing the positively charged residues of α2-α4 (on patch 1 surface in Figure 2A). Of note, the catalytic activity of A3B-CD2 alone was also slightly increased after RNase A treatment (Supplementary Figure S5A), suggesting that RNA appears to interact CD2 domain as well to compete with the ssDNA substrate. Similar competitive regulatory role of RNA has been reported in the study of A3G (81). It is possible that RNAs may be stabilized by interacting with patch 2 regions of CD1 domain and extend to CD2 domain, competing with the ssDNA substrate. In the presence of RNase A, the activity of both positive surface charge reduction mutants is similar to that of the wild type (Figure 6) also indicating that neither one of the two patch mutants disrupted ssDNA substrate binding for deamination. Consequently, our data also suggest that a different region of A3B-CD1 other than the two positively charged regions may participate in the ssDNA substrate binding. Alternatively, the α2–α4 and loop-2/loop-4/β5 of CD1 may be involved in ssDNA binding redundantly, thus mutating either patch is not sufficient to abolish ssDNA substrate binding through A3B-CD1. The nature of RNAs involved in RNA-dependent attenuation of A3B is unclear. Extraction of RNA following IP of the

size exclusion chromatography fractions or pooled fraction may provide valuable insights.

The four surface tyrosine substitutions on A3B-CD1m also partially impair the catalytic activity of A3B (Figure 5A–E), suggesting that HMW complex assembly facilitate A3B catalytic activity. Introducing W127A in the context of four surface tyrosine substitutions further impairs the catalytic activity significantly (Figure 5A–E), likely due to the loss of ssDNA substrate binding (Figure 5F). Interestingly, W127A mutation alone does not affect the activity of A3B (Supplementary Figure S6). One possibility is that the HMW assembly of A3B provides additional ssDNA binding sites from other copies of A3B or associated proteins, thus facilitates A3B catalytic activity. In summary, enzymatic activity is likely regulated by dimeric/tetrameric structure of A3B or by a higher order protein complex formation, which is unclear in the present study and requires further investigation.

A major difference from A3G is that A3B is dominantly localized in nucleus (5,82), which is composed of unique subcellular structures and soluble components including hnRNPs (83). Previous studies and our data show that A3B interacts with multiple hnRNPs (47). Interestingly, combination of four tyrosine mutation and W127A severely impaired A3B interaction with multiple hnRNPs, therefore, these surface hydrophobic residues on CD1 are critical in mediating A3B and hnRNPs interactions. This mutant would serve as a valuable control in further investigation of the endogenous A3B protein complex. It is also worth noting that hnRNPs have distinct RNA binding motifs (83), therefore, these hnRNP binding RNAs may act as potential regulatory elements of A3B activity through hnRNP and A3B interactions.

In summary, we obtained a soluble form of A3B-CD1 through systematic and iterative mutational studies, and determined its high-resolution structure to 1.9 Å resolution. Biophysical and biochemical analysis combined with structure-guided mutagenesis of CD1 in the context of the full-length A3B protein illuminated that CD1 plays a key role in regulating catalytic activity through mechanisms potentially associated with the formation of HMW complexes of A3B in mammalian cells and RNA mediated attenuation. With the knowledge and resolution gained during the pursuit of A3B-CD1 structural biology, we have uncovered several key amino acid residues on CD1 involved in these regulatory processes.

ACCESSION NUMBER

PDB ID: 5TKM.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We thank BL-8.2.1 of Advanced Light Source at Berkeley, 19ID and 23ID of Advanced Photon Source at Argonne National Laboratory and the staffs there for assisting data collection, the core laboratory in the Center of Excellence in NanoBiophysics at University of Southern California (USC) for assisting biophysics characterization, Dr Y. Ying at University of California Los Angeles for providing pcDNA 3.1 vector and cell-based assays advices. Dr B. Xie, Dr M.F. Press, J. Guo, Dr W. Li, for their generous gift of breast cancer cell lines.

FUNDING

National Institutes of Health (NIH) [R01GM087986 to X.S.C.].
Conflict of interest statement. None declared.

REFERENCES

1. Hanahan,D. and Weinberg,R.A. (2011) Hallmarks of cancer: the next generation. *Cell*, **144**, 646–674.
2. Nik-Zainal,S., Alexandrov,L.B., Wedge,D.C., Van Loo,P., Greenman,C.D., Raine,K., Jones,D., Hinton,J., Marshall,J., Stebbings,L.A. *et al.* (2012) Mutational processes molding the genomes of 21 breast cancers. *Cell*, **149**, 979–993.
3. Roberts,S.A., Sterling,J., Thompson,C., Harris,S., Mav,D., Shah,R., Klimczak,L.J., Kryukov,G.V., Malc,E., Mieczkowski,P.A. *et al.* (2012) Clustered mutations in yeast and in human cancers can arise from damaged long single-strand DNA regions. *Mol. Cell*, **46**, 424–435.
4. Burns,M.B., Temiz,N.A. and Harris,R.S. (2013) Evidence for APOBEC3B mutagenesis in multiple human cancers. *Nat. Genet.*, **45**, 977–983.
5. Burns,M.B., Lackey,L., Carpenter,M.A., Rathore,A., Land,A.M., Leonard,B., Refsland,E.W., Kotandeniya,D., Tretyakova,N., Nikas,J.B. *et al.* (2013) APOBEC3B is an enzymatic source of mutation in breast cancer. *Nature*, **494**, 366–370.
6. Roberts,S.A., Lawrence,M.S., Klimczak,L.J., Grimm,S.A., Fargo,D., Stojanov,P., Kiezun,A., Kryukov,G.V., Carter,S.L., Saksena,G. *et al.* (2013) An APOBEC cytidine deaminase mutagenesis pattern is widespread in human cancers. *Nat. Genet.*, **45**, 970–976.
7. Leonard,B., Hart,S.N., Burns,M.B., Carpenter,M.A., Temiz,N.A., Rathore,A., Vogel,R.I., Nikas,J.B., Law,E.K., Brown,W.L. *et al.* (2013) APOBEC3B upregulation and genomic mutation patterns in serous ovarian carcinoma. *Cancer Res.*, **73**, 7222–7231.
8. Walker,B.A., Wardell,C.P., Murison,A., Boyle,E.M., Begum,D.B., Dahir,N.M., Proszek,P.Z., Melchor,L., Pawlyn,C., Kaiser,M.F. *et al.* (2015) APOBEC family mutational signatures are associated with poor prognosis translocations in multiple myeloma. *Nat. Commun.*, **6**, 6997.
9. Nik-Zainal,S., Wedge,D.C., Alexandrov,L.B., Petljak,M., Butler,A.P., Bolli,N., Davies,H.R., Knappskog,S., Martin,S., Papaemmanuil,E. *et al.* (2014) Association of a germline copy number polymorphism of APOBEC3A and APOBEC3B with burden of putative APOBEC-dependent mutations in breast cancer. *Nat. Genet.*, **46**, 487–491.
10. Sieuwerts,A.M., Willis,S., Burns,M.B., Look,M.P., Meijer-Van Gelder,M.E., Schlicker,A., Heideman,M.R., Jacobs,H., Wessels,L., Leyland-Jones,B. *et al.* (2014) Elevated APOBEC3B correlates with poor outcomes for estrogen-receptor-positive breast cancers. *Hormones Cancer*, **5**, 405–413.
11. Cescon,D.W., Haibe-Kains,B. and Mak,T.W. (2015) APOBEC3B expression in breast cancer reflects cellular proliferation, while a deletion polymorphism is associated with immune activation. *Proc. Natl. Acad. Sci. U.S.A.*, **112**, 2841–2846.
12. Xu,L., Chang,Y., An,H., Zhu,Y., Yang,Y. and Xu,J. (2015) High APOBEC3B expression is a predictor of recurrence in patients with low-risk clear cell renal cell carcinoma. *Urologic Oncol.*, **33**, 340.
13. Zhang,J., Wei,W., Jin,H.C., Ying,R.C., Zhu,A.K. and Zhang,F.J. (2015) The roles of APOBEC3B in gastric cancer. *Int. J. Clin. Exp. Pathol.*, **8**, 5089–5096.
14. Yan,S., He,F., Gao,B., Wu,H., Li,M., Huang,L., Liang,J., Wu,Q. and Li,Y. (2016) Increased APOBEC3B predicts worse outcomes in lung cancer: a comprehensive retrospective study. *J. Cancer*, **7**, 618–625.

15. Hoopes,J.I., Cortez,L.M., Mertz,T.M., Malc,E.P., Mieczkowski,P.A. and Roberts,S.A. (2016) APOBEC3A and APOBEC3B preferentially deaminate the lagging strand template during DNA replication. *Cell Rep.*, **14**, 1273–1282.
16. Prohaska,K.M., Bennett,R.P., Salter,J.D. and Smith,H.C. (2014) The multifaceted roles of RNA binding in APOBEC cytidine deaminase functions. *Wiley Interdiscipl. Rev. RNA*, **5**, 493–508.
17. Salter,J.D., Bennett,R.P. and Smith,H.C. (2016) The APOBEC protein family: united by structure, divergent in function. *Trends Biochem. Sci.*, **41**, 578–594.
18. Harris,R.S. and Liddament,M.T. (2004) Retroviral restriction by APOBEC proteins. *Nat. Rev. Immunol.*, **4**, 868–877.
19. Chiu,Y.L. and Greene,W.C. (2008) The APOBEC3 cytidine deaminases: an innate defensive network opposing exogenous retroviruses and endogenous retroelements. *Annu. Rev. Immunol.*, **26**, 317–353.
20. Malim,M.H. (2009) APOBEC proteins and intrinsic resistance to HIV-1 infection. *Philos. Trans. Roy. Soc. London B, Biol. Sci.*, **364**, 675–687.
21. Aydin,H., Taylor,M.W. and Lee,J.E. (2014) Structure-guided analysis of the human APOBEC3-HIV restrictome. *Structure*, **22**, 668–684.
22. Verhalen,B., Starrett,G.J., Harris,R.S. and Jiang,M. (2016) Functional upregulation of the DNA cytosine deaminase APOBEC3B by polyomaviruses. *J. Virol.*, **90**, 6379–6386.
23. Wang,K., Li,Y., Dai,C., Wang,K., Yu,J., Tan,Y., Zhang,W. and Yu,X.F. (2013) Characterization of the relationship between APOBEC3B deletion and ACE Alu insertion. *PLoS One*, **8**, e64809.
24. Wissing,S., Montano,M., Garcia-Perez,J.L., Moran,J.V. and Greene,W.C. (2011) Endogenous APOBEC3B restricts LINE-1 retrotransposition in transformed cells and human embryonic stem cells. *J. Biol. Chem.*, **286**, 36427–36437.
25. Feng,Y., Baig,T.T., Love,R.P. and Chelico,L. (2014) Suppression of APOBEC3-mediated restriction of HIV-1 by Vif. *Front. Microbiol.*, **5**, 450.
26. Bransteitter,R., Prochnow,C. and Chen,X.S. (2009) The current structural and functional understanding of APOBEC deaminases. *Cell. Mol. Life Sci.: CMLS*, **66**, 3137–3147.
27. Xiao,X., Li,S.X., Yang,H. and Chen,X.S. (2016) Crystal structures of APOBEC3G N-domain alone and its complex with DNA. *Nat. Commun.*, **7**, 12193.
28. Kouno,T., Luengas,E.M., Shigematsu,M., Shandilya,S.M., Zhang,J., Chen,L., Hara,M., Schiffer,C.A., Harris,R.S. and Matsuo,H. (2015) Structure of the Vif-binding domain of the antiviral enzyme APOBEC3G. *Nat. Struct. Mol. Biol.*, **22**, 485–491.
29. Sheehy,A.M., Gaddis,N.C., Choi,J.D. and Malim,M.H. (2002) Isolation of a human gene that inhibits HIV-1 infection and is suppressed by the viral Vif protein. *Nature*, **418**, 646–650.
30. Chelico,L., Pham,P., Calabrese,P. and Goodman,M.F. (2006) APOBEC3G DNA deaminase acts processively 3′ –>5′ on single-stranded DNA. *Nat. Struct. Mol. Biol.*, **13**, 392–399.
31. Chelico,L., Sacho,E.J., Erie,D.A. and Goodman,M.F. (2008) A model for oligomeric regulation of APOBEC3G cytosine deaminase-dependent restriction of HIV. *J. Biol. Chem.*, **283**, 13780–13791.
32. Chelico,L., Prochnow,C., Erie,D.A., Chen,X.S. and Goodman,M.F. (2010) Structural model for deoxycytidine deamination mechanisms of the HIV-1 inactivation enzyme APOBEC3G. *J. Biol. Chem.*, **285**, 16195–16205.
33. Holden,L.G., Prochnow,C., Chang,Y.P., Bransteitter,R., Chelico,L., Sen,U., Stevens,R.C., Goodman,M.F. and Chen,X.S. (2008) Crystal structure of the anti-viral APOBEC3G catalytic domain and functional implications. *Nature*, **456**, 121–124.
34. Iwatani,Y., Takeuchi,H., Strebel,K. and Levin,J.G. (2006) Biochemical activities of highly purified, catalytically active human APOBEC3G: correlation with antiviral effect. *J. Virol.*, **80**, 5992–6002.
35. Huthoff,H., Autore,F., Gallois-Montbrun,S., Fraternali,F. and Malim,M.H. (2009) RNA-dependent oligomerization of APOBEC3G is required for restriction of HIV-1. *PLoS Pathogens*, **5**, e1000330.
36. Opi,S., Takeuchi,H., Kao,S., Khan,M.A., Miyagi,E., Goila-Gaur,R., Iwatani,Y., Levin,J.G. and Strebel,K. (2006) Monomeric APOBEC3G is catalytically active and has antiviral activity. *J. Virol.*, **80**, 4673–4682.
37. Wang,X., Dolan,P.T., Dang,Y. and Zheng,Y.H. (2007) Biochemical differentiation of APOBEC3F and APOBEC3G proteins associated with HIV-1 life cycle. *J. Biol. Chem.*, **282**, 1585–1594.
38. Bulliard,Y., Turelli,P., Rohrig,U.F., Zoete,V., Mangeat,B., Michielin,O. and Trono,D. (2009) Functional analysis and structural modeling of human APOBEC3G reveal the role of evolutionarily conserved elements in the inhibition of human immunodeficiency virus type 1 infection and Alu transposition. *J. Virol.*, **83**, 12611–12621.
39. Belanger,K., Savoie,M., Rosales Gerpe,M.C., Couture,J.F. and Langlois,M.A. (2013) Binding of RNA by APOBEC3G controls deamination-independent restriction of retroviruses. *Nucleic Acids Res.*, **41**, 7438–7452.
40. Navarro,F., Bollman,B., Chen,H., Konig,R., Yu,Q., Chiles,K. and Landau,N.R. (2005) Complementary function of the two catalytic domains of APOBEC3G. *Virology*, **333**, 374–386.
41. Svarovskaia,E.S., Xu,H., Mbisa,J.L., Barr,R., Gorelick,R.J., Ono,A., Freed,E.O., Hu,W.S. and Pathak,V.K. (2004) Human apolipoprotein B mRNA-editing enzyme-catalytic polypeptide-like 3G (APOBEC3G) is incorporated into HIV-1 virions through interactions with viral and nonviral RNAs. *J. Biol. Chem.*, **279**, 35822–35828.
42. Wang,T., Tian,C., Zhang,W., Luo,K., Sarkis,P.T., Yu,L., Liu,B., Yu,Y. and Yu,X.F. (2007) 7SL RNA mediates virion packaging of the antiviral cytidine deaminase APOBEC3G. *J. Virol.*, **81**, 13112–13124.
43. Chiu,Y.L., Soros,V.B., Kreisberg,J.F., Stopak,K., Yonemoto,W. and Greene,W.C. (2005) Cellular APOBEC3G restricts HIV-1 infection in resting CD4+ T cells. *Nature*, **435**, 108–114.
44. Chen,Q., Xiao,X., Wolfe,A. and Chen,X.S. (2016) The in vitro biochemical characterization of an HIV-1 restriction factor APOBEC3F: importance of loop 7 on both CD1 and CD2 for DNA binding and deamination. *J. Mol. Biol.*, **428**, 2661–2670.
45. Gallois-Montbrun,S., Holmes,R.K., Swanson,C.M., Fernandez-Ocana,M., Byers,H.L., Ward,M.A. and Malim,M.H. (2008) Comparison of cellular ribonucleoprotein complexes associated with the APOBEC3F and APOBEC3G antiviral proteins. *J. Virol.*, **82**, 5636–5642.
46. Li,J., Chen,Y., Li,M., Carpenter,M.A., McDougle,R.M., Luengas,E.M., Macdonald,P.J., Harris,R.S. and Mueller,J.D. (2014) APOBEC3 multimerization correlates with HIV-1 packaging and restriction activity in living cells. *J. Mol. Biol.*, **426**, 1296–1307.
47. Zhang,W., Zhang,X., Tian,C., Wang,T., Sarkis,P.T., Fang,Y., Zheng,S., Yu,X.F. and Xu,R. (2008) Cytidine deaminase APOBEC3B interacts with heterogeneous nuclear ribonucleoprotein K and suppresses hepatitis B virus expression. *Cell. Microbiol.*, **10**, 112–121.
48. Byeon,I.J., Byeon,C.H., Wu,T., Mitra,M., Singer,D., Levin,J.G. and Gronenborn,A.M. (2016) Nuclear magnetic resonance structure of the APOBEC3B catalytic domain: structural basis for substrate binding and DNA deaminase activity. *Biochemistry*, **55**, 2944–2959.
49. Shi,K., Carpenter,M.A., Kurahashi,K., Harris,R.S. and Aihara,H. (2015) Crystal structure of the DNA deaminase APOBEC3B catalytic domain. *J. Biol. Chem.*, **290**, 28120–28130.
50. Shi,K., Carpenter,M.A., Banerjee,S., Shaban,N.M., Kurahashi,K., Salamango,D.J., McCann,J.L., Starrett,G.J., Duffy,J.V., Demir,O. *et al.* (2017) Structural basis for targeted DNA cytosine deamination and mutagenesis by APOBEC3A and APOBEC3B. *Nat. Struct. Mol. Biol.*, **24**, 131–139.
51. Fu,Y., Ito,F., Zhang,G., Fernandez,B., Yang,H. and Chen,X.S. (2015) DNA cytosine and methylcytosine deamination by APOBEC3B: enhancing methylcytosine deamination by engineering APOBEC3B. *Biochem. J.*, **471**, 25–35.
52. Siriwardena,S.U., Guruge,T.A. and Bhagwat,A.S. (2015) Characterization of the catalytic domain of human APOBEC3B and the critical structural role for a conserved methionine. *J. Mol. Biol.*, **427**, 3042–3055.
53. Caval,V., Suspene,R., Shapira,M., Vartanian,J.P. and Wain-Hobson,S. (2014) A prevalent cancer susceptibility APOBEC3A hybrid allele bearing APOBEC3B 3′UTR enhances chromosomal DNA damage. *Nat. Commun.*, **5**, 5129.
54. Konarska,M.M., Grabowski,P.J., Padgett,R.A. and Sharp,P.A. (1985) Characterization of the branch site in lariat RNAs produced by splicing of mRNA precursors. *Nature*, **313**, 552–557.

55. Black,D.L. (1991) Does steric interference between splice sites block the splicing of a short c-src neuron-specific exon in non-neuronal cells? *Genes Dev.*, **5**, 389–402.
56. Leonard,B., McCann,J.L., Starrett,G.J., Kosyakovsky,L., Luengas,E.M., Molan,A.M., Burns,M.B., McDougle,R.M., Parker,P.J., Brown,W.L. *et al.* (2015) The PKC/NF-kappaB signaling pathway induces APOBEC3B expression in multiple human cancers. *Cancer Res.*, **75**, 4538–4547.
57. Gu,J., Chen,Q., Xiao,X., Ito,F., Wolfe,A. and Chen,X.S. (2016) Biochemical characterization of APOBEC3H variants: implications for their HIV-1 restriction activity and mC modification. *J. Mol. Biol.*, **428**, 4626–4638.
58. Kelley,L.A., Mezulis,S., Yates,C.M., Wass,M.N. and Sternberg,M.J. (2015) The Phyre2 web portal for protein modeling, prediction and analysis. *Nat. Protoc.*, **10**, 845–858.
59. Baker,N.A., Sept,D., Joseph,S., Holst,M.J. and McCammon,J.A. (2001) Electrostatics of nanosystems: application to microtubules and the ribosome. *Proc. Natl. Acad. Sci. U.S.A.*, **98**, 10037–10041.
60. Chen,K.M., Harjes,E., Gross,P.J., Fahmy,A., Lu,Y., Shindo,K., Harris,R.S. and Matsuo,H. (2008) Structure of the DNA deaminase domain of the HIV-1 restriction factor APOBEC3G. *Nature*, **452**, 116–119.
61. Shandilya,S.M., Nalam,M.N., Nalivaika,E.A., Gross,P.J., Valesano,J.C., Shindo,K., Li,M., Munson,M., Royer,W.E., Harjes,E. *et al.* (2010) Crystal structure of the APOBEC3G catalytic domain reveals potential oligomerization interfaces. *Structure*, **18**, 28–38.
62. Kitamura,S., Ode,H., Nakashima,M., Imahashi,M., Naganawa,Y., Kurosawa,T., Yokomaku,Y., Yamane,T., Watanabe,N., Suzuki,A. *et al.* (2012) The APOBEC3C crystal structure and the interface for HIV-1 Vif binding. *Nat. Struct. Mol. Biol.*, **19**, 1005–1010.
63. Siu,K.K., Sultana,A., Azimi,F.C. and Lee,J.E. (2013) Structural determinants of HIV-1 Vif susceptibility and DNA binding in APOBEC3F. *Nat. Commun.*, **4**, 2593.
64. Bohn,M.F., Shandilya,S.M., Albin,J.S., Kouno,T., Anderson,B.D., McDougle,R.M., Carpenter,M.A., Rathore,A., Evans,L., Davis,A.N. *et al.* (2013) Crystal structure of the DNA cytosine deaminase APOBEC3F: the catalytically active and HIV-1 Vif-binding domain. *Structure*, **21**, 1042–1050.
65. Byeon,I.J., Ahn,J., Mitra,M., Byeon,C.H., Hercik,K., Hritz,J., Charlton,L.M., Levin,J.G. and Gronenborn,A.M. (2013) NMR structure of human restriction factor APOBEC3A reveals substrate binding and enzyme specificity. *Nat. Commun.*, **4**, 1890.
66. Bohn,M.F., Shandilya,S.M., Silvas,T.V., Nalivaika,E.A., Kouno,T., Kelch,B.A., Ryder,S.P., Kurt-Yilmaz,N., Somasundaran,M. and Schiffer,C.A. (2015) The ssDNA Mutator APOBEC3A is regulated by cooperative dimerization. *Structure*, **23**, 903–911.
67. Harjes,E., Gross,P.J., Chen,K.M., Lu,Y., Shindo,K., Nowarski,R., Gross,J.D., Kotler,M., Harris,R.S. and Matsuo,H. (2009) An extended structure of the APOBEC3G catalytic domain suggests a unique holoenzyme model. *J. Mol. Biol.*, **389**, 819–832.
68. Pham,P., Afif,S.A., Shimoda,M., Maeda,K., Sakaguchi,N., Pedersen,L.C. and Goodman,M.F. (2016) Structural analysis of the activation-induced deoxycytidine deaminase required in immunoglobulin diversification. *DNA Repair (Amst.)*, **43**, 48–56.
69. Nakashima,M., Ode,H., Kawamura,T., Kitamura,S., Naganawa,Y., Awazu,H., Tsuzuki,S., Matsuoka,K., Nemoto,M., Hachiya,A. *et al.* (2015) Structural Insights into HIV-1 Vif-APOBEC3F Interaction. *J. Virol.*
70. Ara,A., Love,R.P. and Chelico,L. (2014) Different mutagenic potential of HIV-1 restriction factors APOBEC3G and APOBEC3F is determined by distinct single-stranded DNA scanning mechanisms. *PLoS Pathogens*, **10**, e1004024.
71. Wedekind,J.E., Gillilan,R., Janda,A., Krucinska,J., Salter,J.D., Bennett,R.P., Raina,J. and Smith,H.C. (2006) Nanostructures of APOBEC3G support a hierarchical assembly model of high molecular mass ribonucleoprotein particles from dimeric subunits. *J. Biol. Chem.*, **281**, 38122–38126.
72. Bennett,R.P., Salter,J.D., Liu,X., Wedekind,J.E. and Smith,H.C. (2008) APOBEC3G subunits self-associate via the C-terminal deaminase domain. *J. Biol. Chem.*, **283**, 33329–33336.
73. Salter,J.D., Krucinska,J., Raina,J., Smith,H.C. and Wedekind,J.E. (2009) A hydrodynamic analysis of APOBEC3G reveals a monomer-dimer-tetramer self-association that has implications for anti-HIV function. *Biochemistry*, **48**, 10685–10687.
74. Gallois-Montbrun,S., Kramer,B., Swanson,C.M., Byers,H., Lynham,S., Ward,M. and Malim,M.H. (2007) Antiviral protein APOBEC3G localizes to ribonucleoprotein complexes found in P bodies and stress granules. *J. Virol.*, **81**, 2165–2178.
75. Nogales,E., Wolf,S.G. and Downing,K.H. (1998) Structure of the alpha beta tubulin dimer by electron crystallography. *Nature*, **391**, 199–203.
76. Cescon,D.W. and Haibe-Kains,B. (2016) DNA replication stress: a source of APOBEC3B expression in breast cancer. *Genome Biol.*, **17**, 202.
77. Mitra,M., Singer,D., Mano,Y., Hritz,J., Nam,G., Gorelick,R.J., Byeon,I.J., Gronenborn,A.M., Iwatani,Y. and Levin,J.G. (2015) Sequence and structural determinants of human APOBEC3H deaminase and anti-HIV-1 activities. *Retrovirology*, **12**, 3.
78. Bransteitter,R., Pham,P., Scharff,M.D. and Goodman,M.F. (2003) Activation-induced cytidine deaminase deaminates deoxycytidine on single-stranded DNA but requires the action of RNase. *Proc. Natl. Acad. Sci. U.S.A.*, **100**, 4102–4107.
79. McDougall,W.M. and Smith,H.C. (2011) Direct evidence that RNA inhibits APOBEC3G ssDNA cytidine deaminase activity. *Biochem. Biophys. Res. Commun.*, **412**, 612–617.
80. Mondal,S., Begum,N.A., Hu,W. and Honjo,T. (2016) Functional requirements of AID's higher order structures and their interaction with RNA-binding proteins. *Proc. Natl. Acad. Sci. U.S.A.*, **113**, E1545–E1554.
81. Polevoda,B., McDougall,W.M., Tun,B.N., Cheung,M., Salter,J.D., Friedman,A.E. and Smith,H.C. (2015) RNA binding to APOBEC3G induces the disassembly of functional deaminase complexes by displacing single-stranded DNA substrates. *Nucleic Acids Res.*, **43**, 9434–9445.
82. Caval,V., Bouzidi,M.S., Suspene,R., Laude,H., Dumargne,M.C., Bashamboo,A., Krey,T., Vartanian,J.P. and Wain-Hobson,S. (2015) Molecular basis of the attenuated phenotype of human APOBEC3B DNA mutator enzyme. *Nucleic Acids Res.*, **43**, 9340–9349.
83. Dreyfuss,G., Kim,V.N. and Kataoka,N. (2002) Messenger-RNA-binding proteins and the messages they carry. *Nat. Rev. Mol. Cell Biol.*, **3**, 195–205.
84. Prochnow,C., Bransteitter,R., Klein,M.G., Goodman,M.F. and Chen,X.S. (2007) The APOBEC-2 crystal structure and functional implications for the deaminase AID. *Nature*, **445**, 447–451.