Published in final edited form as:

Methods. 2016 September 1; 107: 10–22. doi:10.1016/j.ymeth.2016.03.006.

Structural and Functional Assessment of APOBEC3G Macromolecular Complexes

Bogdan Polevoda[a], **William M. McDougall**[a,e], **Ryan P. Bennett**[d], **Jason D. Salter**[d], and **Harold C. Smith**[a,b,c,d,f]

[a]Department of Biochemistry and Biophysics, University of Rochester, School of Medicine and Dentistry, 601 Elmwood Avenue, Rochester, NY 14642 USA

[b]Wilmot Cancer Institute, 601 Elmwood Avenue, Rochester, NY 14642 USA

[c]Center for RNA Biology, 601 Elmwood Avenue, Rochester, NY 14642 USA

[d]OyaGen, Inc, BioVenture Center, 77 Ridgeland Road Rochester, NY 14623 USA

[f]Center for AIDS Research, 601 Elmwood Avenue, Rochester, NY 14642 USA

Abstract

There are eleven members in the human APOBEC family of proteins that are evolutionarily related through their zinc-dependent cytidine deaminase domains. The human APOBEC gene clusters arose on chromosome 6 and 22 through gene duplication and divergence to where current day APOBEC proteins are functionally diverse and broadly expressed in tissues. APOBEC serve enzymatic and non enzymatic functions in cells. In both cases, formation of higher-order structures driven by APOBEC protein-protein interactions and binding to RNA and/or single stranded DNA are integral to their function. In some circumstances, these interactions are regulatory and modulate APOBEC activities. We are just beginning to understand how macromolecular interactions drive processes such as APOBEC subcellular compartmentalization, formation of holoenzyme complexes, gene targeting, foreign DNA restriction, anti-retroviral activity, formation of ribonucleoprotein particles and APOBEC degradation. Protein-protein and protein-nucleic acid cross-linking methods coupled with mass spectrometry, electrophoretic mobility shift assays, glycerol gradient sedimentation, fluorescence anisotropy and APOBEC deaminase assays are enabling mapping of interacting surfaces that are essential for these functions. The goal of this methods review is through example of our research on APOBEC3G, describe the application of cross-linking methods to characterize and quantify macromolecular interactions and their functional implications. Given the homology in structure and function, it is proposed that these methods will be generally applicable to the discovery process for other APOBEC and RNA and DNA editing and modifying proteins.

[f]*Corresponding author*. harold.smith@rochester.edu, *Phone*: (585) 275-4267.

[e]Current address: Microbiology and Physiological Systems, University of Massachusetts Medical School, 368 Plantation Street, AS8-2055, Worcester MA 01605

1. Introduction

The APOBEC family of proteins plays significant roles in human health and disease [1]. Editing of apolipoprotein B mRNA by APOBEC1 (A1) was the first mammalian C to U editing enzyme activity discovered and hence the namesake of the family. Subsequently ten additional family members were discovered. APOBECs are zinc-binding proteins that have either one or two evolutionarily conserved zinc-dependent deaminase domains (ZDD) [2–4]. APOBECs as they are expressed in mammalian cells either have little or no deaminase activity (APOBEC 2 and 4) [5, 6], only edit RNA (A1) [7–9], only deaminate single-stranded (ssDNA) (Activation Induced Deaminase (AID), A3B, 3D, 3F, 3G and 3H) [10–16] or deaminate both RNA and ssDNA (APOBEC3A) [17, 18]. Identification of these proteins has led to new understandings of how APOBECs are involved in genomic evolution and genetic stability [19–22], cancer [9, 23–28], control of retrotransposition [13, 29–35], class switch recombination and somatic hypermutation of the IgG locus for acquired immunity [36–39] and anti-retroviral activity [11, 15, 22, 27, 37, 40, 41]. Many of the functions of APOBECs rely on protein-protein and protein-nucleic acids interactions and in some instances post-translational modifications, to regulate their subcellular compartmentalization, substrate specificity and biological activity (reviewed in [1, 42]). Consequently the current and future research emphasis will be structural characterization of the APOBEC proteins and the development of methods for identifying and functionally characterizing macromolecular assemblies containing APOBEC proteins. In this review we focus on cross-linking methods developed with APOBEC3G (A3G) to evaluate protein-protein and proteinnucleic acids interactions in the macromolecular assemblies required for enzyme activity. For additional methods developed for APOBEC1 and RNA binding cofactor A1CF and yeast cytidine deaminase CDD1 the reader is referred to a prior review in Methods [43].

2. Methods

2.1. Recombinant APOBEC expression and purification

Expression and purification of APOBEC/AID proteins in bulk has been very difficult using bacterial expression systems as these proteins are genotoxic [12, 44, 45], are frequently insoluble as full length proteins [46] and with the possible exceptions of single deaminase domain-containing proteins such as AID [47], APOBEC2 [48], A3A [49], A3C [50] and yeast CDD1 [51], required deletion or mutagenesis in order to be made soluble. Even N- or C-terminal truncations of APOBECs [52–57] may require mutations to improve their solubility. *In vitro* translation of APOBECs has been successful in producing analytical quantities of soluble APOBEC [58] whereas larger quantities of soluble A3A and A3G have been purified from HEK293T cells [59, 60]. Bulk quantities of native and full length A3G has been achievable using the Baculovirus expression in Sf9 insect cells and we recommend this as the system of choice [61].

Regardless of the method of expression, recombinant APOBECs typically can be purified from cell extracts using a single affinity chromatography step with a Glutathione, Myc or 6xHis tag that has typically been placed at the C-terminus [46, 50, 52–57]. For each application, the position of the affinity tag should be evaluated to optimize expression and to

ensure no interference with functional interactions. Cleavable affinity tags have been used in some of the applications cited above.

APOBEC proteins bind to nucleic acids and regardless of the source of their expression, will copurify with RNA unless measures are taken to remove it [42] (Figure 1). RNase A and DNase I (in presence of $MgCl_2$ and $CaCl_2$) digestion of the cellular extract at 1M NaCl will reduce the amount of the bound nucleic acid. A3G protein structure determined by circular dichroism was unaffected by 1M NaCl or even 1M urea [61]. The ratio of 260/280 nm absorption of RNase treated preparations still may be high 0.9–1.1 but without RNase digestion, the ratio can be 1.6 or higher and proteins will be highly aggregated, that will be apparent using a variety of physical techniques [60, 62– 64] (Figure 1). The following protocol is recommended for the purification of enzymatically active A3G from cells and will yield protein suitable for structural studies and protein-protein or protein nucleic acid binding studies [65–68].

2.1.1 A3G expression in Sf9 cells—The protocol relies on frozen Sf9 cells (−80 °C) that had expressed APOBEC3G. There are numerous institutional core facilities and commercial services available for recombinant protein expression. These may provide custom synthesis of APOBEC cDNAs with a variety of tags and appropriate expression vectors for prokaryotic and eukaryotic cells. cDNA encoding human A3G with a C-terminal 6xHis tag (synthesized by GenScript, Piscataway, NJ or Blue Sky BioServices, Worchester, MA) was cloned into the Baculovirus vector pFastBac using the Bac-to-Bac system from Invitrogen™ Thermo Fisher Scientific (Waltham, MA). Virus production, Sf9 insect cell transduction and four liter cultures of Sf9 cells expressing A3G-4xHis was optimized by GenScript (Piscataway, NJ), Blue Sky BioServices (Worchester, MA) or ImmunoDx (Woburn, MA).

2.1.2 Cell extract preparation and purification of A3G by nickel affinity chromatography—All protocols should be performed at room temperature unless otherwise stated. Two to four grams of frozen cell paste should be resuspended in 30 ml of 1X A3G buffer (25 mM HEPES pH=7.2, 1.0 M NaCl, 5 mM $MgCl_2$, 3.5 mM β-mercaptoethanol, 5 mM imidazole, 0.1 mM $CaCl_2$) containing a combination of protease inhibitors (Complete Mini EDTA-free, cat # 11836170001, Roche, IN). The solution will become viscous as cells lyse occurs and the histones release from high molecular weight DNA. Bring the solution to 0.4% Triton X-100 (cat # 16-10407 BioRad, CA) and add 2.5 mg each of DNase I (cat # DN25, Sigma Aldrich, St. Louis, MO) and RNase A (cat # R5503, Sigma-Aldrich) and incubate the lysate at 37 °C for 30 minutes with occasional inversion of the tube.

At the end of the incubation, the viscosity in the solution should no longer be apparent. Add ultrapure urea to a final concentration of 1M and incubate the extract at room temperature for 10 minutes. A cleared supernatant should be prepared by centrifuging the extract in a swinging bucket rotor at 16,000 × g for 10 minutes at 7 °C to pellet cellular debris. At the conclusion of the centrifugation, draw off the supernatant being careful to avoid disturbing the pellet or drawing up flocculent material that may float above the pellet. Add the cleared supernatant to 0.5 to 1.2 ml of packed nickel affinity (Ni-NTA) resin (cat # 30210, Ni-NTA,

Qiagen, CA) prewashed with 1X A3G buffer and mix by tumbling for two hours at 7 °C. The cleared supernatant-NiNTA mixture should be placed into a small column with a glass frit (cat # 731-1550, BioRad, CA) and allowed to drain completely before beginning the resin wash steps.

Initially wash the inside of the column walls and resin with 1.0 ml of 1X A3G buffer contain 1M urea followed sequentially by: 5 ml of 1X A3G containing 0.5M urea, 10 ml of 1X A3G, 10 ml of 1X A3G that contains 0.5M NaCl (1X A3G-0.5), 5 ml of 1X A3G-0.5 containing 70 mM imidazole and 3 ml of 1X A3G-0.5 containing 100 mM imidazole (final concentration). Once the column has stopped dripping, begin eluting the column with 1X A3G-0.5 containing 350 mM imidazole while collecting 1 ml fractions. The recovery of A3G using the described protocol in 350 mM containing fractions is evident by Coomassie blue staining of SDS-PAGE (Figure 2), with minor amounts eluting with the 100 mM imidazole wash and significantly less than peak A3G recovery in a subsequent 500 mM imidazole elution.

2.2. Validating the functional activity of A3G

Purified APOBEC proteins should be assessed and quantified for their functional characteristics as a laboratory benchmark for comparing batch-to-batch variability. In addition to the physical assessment of nucleic acid associated with APOBEC and SDS-PAGE evaluation of APOBEC purity (described above), size exclusion chromatography [50, 54, 69, 70], circular dichroism [61, 70] equilibrium sedimentation [68] and velocity sedimentation with western blotting [65, 71] and mass spectrometry [72] have been useful. APOBEC preparations frequently contain a mixture of monomers to higher molecular mass homo oligomers and hence the signal from dynamic light scattering almost always will be polydisperse and not necessarily correlate with the functionality of the purified APOBEC protein. The specific activity of the purified APOBEC is the best criteria for well folded proteins as it reflects both the ability to bind to nucleic acid substrates and the presence of a functional cytidine deaminase catalytic domain. Second to enzymatic activity is the functional criteria of binding to nucleic acids and the quantification of binding affinities.

Over time in storage, aggregation of purified APOBEC will take place that in severe cases can be centrifuged out of solution. Once aggregation begins it is not readily reversed. The resultant homodimers and higher order multimers will be resistant to disruption by 6M urea or 6M guanidine HCl and will appear as dimers and multimers on SDS-PAGE following boiling in SDS and DTT-containing gel loading buffer. It is therefore recommended that quantities of APOBEC be prepared that are appropriate for the experiments planned over 1–4 weeks and that thawed aliquots of protein are cleared by centrifugation and protein concentration determined prior to use.

2.2.1. Quantifying cytidine/deoxycytidine deaminase activity—Methods for quantifying recombinant APOBEC deaminase activity on short strands of RNA or ssDNA substrates have been extensively described in the literature cited above. Briefly, quantifying deaminase activity of APOBEC and mutants has been achieved buffer by: (1) primer extension sequencing method or RNA or ssDNA from *in vitro* assays consisting of

recombinant protein, nucleic acid substrate [73, 74] (Figure 3C), (2) uracil DNA glycosylase-dependent ssDNA cleavage [75–77] or (3) PCR and restriction enzyme cleavage [78]. These assays are simple to set up and are quantitative. Deaminase activity has also been quantified based on the ability of wild type or mutant APOBEC transformed into *E. coli* to induce DNA mutations that will lead to antibiotic resistance [79]. The bacterial assay has been extensively used but also has been criticized as not reflecting activity on biologically relevant nucleic acid substrates and occasionally mutant APOBEC that otherwise are not active in mammalian cells, have activity in the antibiotic sensitivity reversion assay [43].

2.2.2. Determining APOBEC binding to nucleic acids—An intrinsic property of the many of the members of the APOBEC family is the ability to bind to RNA and ssDNA [42]. Methods for quantifying nucleic acid binding to APOBEC proteins that frequently are encountered in the literature are native gel electrophoretic mobility shift assays (EMSA) [65, 66, 75, 80, 81], ultraviolet (UV) light-induced cross-linking [67, 82–84] sedimentation analysis [65], fluorescence anisotropy [52, 67], chemical shifts changes in APOBEC nuclear magnetic resonance profiles due to nucleic acid binding [49, 57, 70, 85, 86] and small angle X-ray scattering [61].

Many laboratories have relied on EMSA in their analysis of APOBEC binding to nucleic acid as the equipment for electrophoresis and gel image scanning and the reagents are readily available. Providing that protein and nucleic acid input to the binding reactions is carefully controlled, EMSA can be semi-quantitative and binding isotherms can be determined for a titration of protein relative to a fixed input nucleic acid [65] (Figure 3D).

Due to the tendency of APOBECs to aggregate, some or all of the protein will either not readily enter native gels or migrate as polydisperse bands. Scans of native gels for fluorescently tagged APOBECs or western blots probed with antibodies for APOBECs typically reveal signal in the wells of gels or smears. In contrast, APOBECs bound to nucleic acid will enter native gels as defined complexes with discrete mobility (bands) whose abundance and mobility are directly dependent on the ratio of APOBEC to nucleic acid used in the binding reaction (Figure 3A). For this reason, nucleic acid binding studies that are evaluated by EMSA should be carried out with labeled nucleic acid and unlabeled APOBECs rather than by western blotting and probing for APOBECs.

Fluorescence anisotropy (FA) has advantages over the aforementioned methods for quantifying APOBEC binding to nucleic acid in that it involves significantly less protein and nucleic acid, does not involve gels or lengthy set up following the reaction and is easily scalable for kinetic studies where multiple samples and replicates are required. The disadvantages of FA are that it lacks the visual images associated with gel-based methods that will reveal whether different sized complexes have been assembled and requires access to a fluorometer and specialized cuvettes. The theory is shown in Figure 4 and involves excitation of the sample with polarized light and measuring changes from a low anisotropic fluorescence emission due to the high molecular tumbling rate of a small nucleic acid in solution to a higher amount of anisotropic fluorescence emission due to the relatively slower rate of tumbling when the nucleic acid forms a larger complex by binding to protein.

A typical application for FA would be to evaluate the binding affinity of APOBEC for ssDNA or RNA (Figure 5A and C) as well as the dissociation of APOBEC from nucleic acids (Figure 5B and D) or the relative affinity of APOBEC for different sequences and lengths of nucleic acids (Figure 6A–D) [67]. For these experiments, Alexa Fluor® 647 labeled ssDNA or RNA can be combined at a fixed concentration (0–5 nM) with a range of A3G concentrations (1 to 200 nM) in buffer (40 mM Tris pH 7.2, 50 mM NaCl, 10 mM $MgCl_2$, 1 mM DTT, 0.1 % Triton X-100, 2 % glycerol) using a final reaction volume of 20 μl. Each assembly condition should be performed in triplicate and each experiment should be performed in triplicate. The assembly reaction should be incubated at 37 °C for 20 minutes; though shorter and longer times could be evaluated. We use a Fluormax-4 Spectrofluorometer (HORIBA Scientific, Edison, NJ) for the anisotropy measurements (ex = 647 nM, em = 670 nM) with 5 nm band passes. The change in anisotropy can be calculated by subtracting the average anisotropy of the protein-free labeled ssDNA or RNA from the average anisotropy of the A3G:nucleic acid assembly reactions. The K_d can be calculated through non-linear regression by fitting the data to equation 1, where A3G:NA is the fraction of A3G bound to nucleic acid (NA) and *Bmax* is a protein capacity to bind substrate.

$$A3G{:}NA = B\max\left(\frac{[A3G]}{Kd+[A3G]}\right) \quad \text{eq. 1}$$

Competition analyses shown in Figure 5 and 6 employed Alexa Fluor® 647 labeled ssDNA or RNA incubated with A3G and unlabeled ssDNA or RNA of the same sequence as competitor. The change in the average anisotropy was plotted against the unlabeled competitor nucleic acid concentration. The data can be fit through non-linear regression analysis using equation 2, and the concentration necessary for 50% of the FA to be decreased (half the complexes with labeled nucleic acid have been dissociated) or IC50 can be determined.

$$A3G{:}NA = y\mathrm{Max}\left(\frac{1-([\mathrm{Comp}])}{[\mathrm{Comp}]+IC50}\right) \quad \text{eq. 2}$$

Ymax is a maximum value of the curve and [*Comp*] is the concentration of the competitor.

3. Higher order complexes containing A3G

APOBECs form homo oligomers with nucleic acids [51, 55, 60, 61, 63, 65, 66, 68, 84, 87–92] and with other proteins [93–102] or are bridged to other proteins through RNA [62, 103–109]. Proteomic approaches exploited in these studies to identify and quantify interactions included Forester Resonance Energy Transfer (FRET), Atomic Force Microscopy (AFM) and mass spectrometry (MS) protein sequence analysis or western blotting for proteins that co-immunoprecipitated or co-sedimented with APOBECs. Protein-protein or protein-nucleic acid cross-linking is a powerful means by which the physical proximity of APOBEC with other proteins or nucleic acids can be determined. The following methods will describe cross-linking of APOBEC homo oligomers within distinct native complexes of APOBEC

and ssDNA substrate for mass spectrometry determination of peptides that may bind to ssDNA.

3.1. Three dimensional method for determining A3G homo oligomerization

3.1.1. First dimension, resolving native complexes of A3G bound to ssDNA— The first step is to resolve APOBEC bound to nucleic acid into discrete complexes. Radiolabeled or AlexaFluor-labeled substrates should be incubated at varying molar ratios of A3G to ssDNA or RNA in deaminase buffer (40 mM Tris pH 7.2, 50 mM NaCl, 10 mM MgCl2, 1 mM DTT, 0.1 % Triton X-100, 2 % glycerol) for 20 minutes at 37 °C. Typically 2 to 6 pMol of labeled nucleic acid can be combined with varying amounts of A3G from 0 to 10 μg in 10 μl final volume assembly reactions. The resulting complexes should be resolved on a 5% native gel. The composition of the native gel is given in Table 1.

After the native gel polymerizes (2 hr), it should be chilled to 7 °C before applying the samples. The assembly reaction should be combined with 2 μl 5X loading buffer (50% Glycerol in 0.5X TBE, without tracking dyes) and held on ice or frozen at −80 °C until loaded on the gel. Precise loading of sample into the wells and using low sample volumes (7 to 10 μl) are essential for good resolution as the gel does not have a stacking system. Run the samples at constant 75V for 15 minutes and then at 150V for 75 to 120 minutes. The duration of electrophoresis necessary for resolving complexes may need to be optimized for each application. The native gel complexes can be visualized after electrophoresis using a Typhoon™ Trio imager (GE Healthcare, Piscataway, NJ) and excitation at 633 nm and measuring fluorescence emission at 670 nm (Figure 7A) [65]. Higher acrylamide concentrations with longer electrophoresis times may increase homogeneity in the migration of native complexes (tighter bands) but will not allow bigger complexes to be resolved.

3.1.2. The second dimension, resolving cross-linked proteins—Since the assembly reactions are of defined composition, the APOBEC stoichiometry in each complex will be known from the size of the complex. The goal of the second dimension is to determine the size of cross-linked complexes based on their electrophoretic mobility on denaturing SDS-PAGE. This will require stabilizing the protein complexes by cross-linking.

Lanes of interest from the first dimension native gels should be identified and cut out as a gel strip using a printout of the native gel scan as a guide. To cross-link APOBEC complexes in the gel, transfer the gel strip to a plastic weigh boat and incubate it with 3.0 ml of 30 mM of DTBP (dimethyl 3,3′-dithiobispropionimidate, cat #20665 Thermo Fisher Scientific) in 0.2 M Triethanolamine (TEA) (cat # T-1377, Sigma-Aldrich), pH= 8.0 to 8.5 for 30 minutes with occasional mixing (or buffer alone as a control for no cross-linking). DTBP is a 16 Å cross-linking reagent that will form covalent bonds spontaneously through appropriately spaced amine groups on proteins using reactive imino esters at each end of the compound. This reagent has a centrally located, cleavable disulfide so that cross-links can be reversed for certain applications. The yield of cross-linked complexes may be optimized by selecting from several commercially available cross-linking reagents with different linker lengths and different reactive groups available through Thermo Fisher Scientific.

After cross-linking, the gel strip should be rinsed with 0.2 M TEA and then incubated in 200 to 500 μl of 2X SDS sample loading buffer (125 mM Tris pH=6.8, 4 % SDS, 20% glycerol, 0.016 % (w/v) bromophenol blue, without reducing agent) at room temperature for 10 minutes with occasional mixing. The treated gel strip should be layered horizontally onto the stacking gel (4.5 % acrylamide) of a 10.5 % denaturing SDS-PAGE. To calibrate the size of cross linked complexes resolved on the second dimension, squares of gel into which prestained molecular mass markers have been polymerized should be placed on the second dimension gel at either end of the native gel strip (to form their own lanes). Molten agarose (at ~45 °C) in 0.5X SDS gel running buffer (0.5% agarose in 12.5 mM Tris pH= 8.3, 0.096 mM glycine, 0.05% SDS) should be used to seal the gel strip and molecular mass marker squares to the top of the second dimension gel by positioning the plates at a 45 degree pitch to allow the agarose to pool while solidifying. Run the gel at 25 mA at room temperature until the molecular mass equivalent to monomeric APOBEC migrates close to bottom of the gel (Figure 7B and C).

3.1.3. The third dimension, immunoblot detection of A3G in cross-linked complexes—

At the end of the second dimensional gel electrophoresis, transfer the proteins by western blot to nitrocellulose or nylon membrane. Block the membrane and probe it with the antibody specific for A3G (cat # 9906, rabbit polyclonal antibody, AIDS Research and Reference Reagent Program) (or other relevant primary antibody) and a secondary antibody using appropriate membrane blocking reagents (0.5 % nonfat milk or bovine serum albumin). The proteins on the membranes can be detected with Lumi-Light Western Blotting Substrate reagent (Roche, cat # 12015196001) or any other type of detection chemicals (Figure 7B and C) [65]. The position of monomeric proteins of interest can be determined from the no cross-linking control or by polymerizing the protein of interest in a gel slice and sealing it next the gel strip on the second dimension gel. To address the question of other proteins that may be associated with APOBEC, the blots can be stripped, re-blocked and re-probed with an antibody selective for the protein of interest or its unique epitope tag. In the particular example in Figure 7, the kinetics of assembling C1 and C3 complexes revealed that while A3G homo dimeric complexes (C1) initially bind to ssDNA (Figure 3B and C), dC to dU deamination was associated with A3G homo tetrameric complexes (C3) (Figure 3C, D and E).

3.2. UV-induced cross-linking of A3G to nucleic acids

Photo cross-linking of proteins to nucleic acids is a common procedure for identifying protein:nucleic acid interactions and binding sites both *in vitro* and *in vivo*. When certain nucleotides in nucleic acid are in close proximity to amino acid residue(s) of interacting proteins, high energy UV light will activate reactive hydrogen or hydroxyl groups to induce a covalent bond that forms the cross link between macromolecules. Amino acids that can form photo-adducts with pyrimidines are cysteine, methionine, serine, arginine, lysine, histidine, phenylalanine, tryptophan and tyrosine. Pyrimidines, especially uridine, are typically 10-times more efficient for cross-linking to amino acids as compared to purines. In addition, halogenated analogs of uridine, such as bromodeoxyuridine (BrdU) for ssDNA and RNA or thiouridine (ThioU) for RNA are more photo activatable by UV light for cross-

linking because substitution of a small methyl group in uridine with a large moiety increases its susceptibility for free radical formation induced by UV.

For cross-linking, purified protein should be mixed with single stranded DNA (ssDNA) or RNA to assemble complexes and exposed to 254 nm UV light on ice for 10–30 min in quartz cuvettes one centimeter from the light source. If cross-linking of nucleic acid at preferred positions is required then modified nucleotides at the corresponding sites could be introduced, for example BrdU for ssDNA or ThioU for RNA. In addition to higher selectivity, BrdU cross linking requires less energy, therefore a medium wavelength (302 nm) UV-light should be used. Similarly, ThioU-modified RNA can be cross linked with longer wavelength (354 nm) UV-light.

In order to cross-link nucleic acid to protein several factors should be considered. First, the reaction volume should be minimal since in most cases the efficiency of cross linking is only 5–10% and if cross-linked products will be separated on SDS-PAGE then large loading volumes could become an issue (although reactions could be concentrated by precipitation with TCA or acetone). Second, whenever possible avoid using Tris or other amine-containing buffer reagents in cross-linking assays as they will compete for free amine cross linking (however it may be useful to add Tris or glycine to end a conjugation process and stop the reaction). Third, the protein:nucleic acid molar ratio in binding reaction should be optimized to saturate the protein with nucleic acid. Fourth, if MS is your downstream application, carry out cross-linking reactions in quartz cuvettes to minimize any undesirable chemistry from plastic tubes or containers that may interfere with LC-MS separation, detection and product identification.

3.2.1 A3G cross-linking to BrdU-modified substrate ssDNA—To cross-link A3G to substrate DNA, assemble protein with DNA on ice in 1:4 ratio by combining 100 pMol of Ni-NTA affinity purified protein and 400 pMol of 25 nt BrdU-modified substrate ssDNA (5’TTATTBrdUUUAAGGATTTATTTATTTA-3’) (Integrated DNA Technologies, Coralville, IA) in 1XPBS buffer with 0.1% NP-40, in a total volume 20 μl. Prepare a control reaction containing A3G without DNA. Incubate the reactions for 20 minutes at 37 °C then cool them on ice and transfer into 0.1–0.5 ml quartz cuvettes. Keeping cuvettes on ice irradiate the reactions in CL-1000 UV-cross linker (UVP, Upland, CA) with medium wavelength 302 nm UV-bulb for 20 min (optimal UV-treatment time may vary and should be tested in a range between 5 to 30 min). Collect the reactions into 1.5 ml Protein LoBind Eppendorf tubes (Eppendorf, Hamburg, Germany, cat # 022431081), add 20 μl of 2XSDS (containing reducing agent) gel loading buffer and denature the samples at 95 °C for 10 min followed by 2 min cooling to room temperature. Clear any insoluble material by centrifugation before loading the samples onto 12% SDS-PAGE gel. If the downstream application includes MS, stain the gel with MS-compactible dye, for example Coomassie G-250-type SimpleBlue SafeStain (Thermo Fisher Scientific) and use a glass container when treating the gel, otherwise stain the gel with standard Coomassie R-250 dye (methanol:acetic acid solubilized). Due to covalent modification, A3G cross-linked protein typically appears as a light intensity band on the gel with mobility 5–6 kDa slower than monomeric uncross linked A3G (Figure 8A). Destain the gel to low background before taking a picture (saving an image). If no visible cross-linked protein band is observed, then optimize the reaction

parameters to enhance cross linking by increasing the time of exposure to UV light and vary the protein:DNA ratio in the assembly reaction. Optimization of these conditions should be validated using fluorescently labeled DNA and EMSA.

3.2.2 A3G cross-linking to RNA—Cross-linking of A3G to RNA should be performed by a similar procedure as described for ssDNA, except that to prevent RNA from degradation 0.5 μl/per reaction of Protector RNase inhibitor (Roche Diagnostics, Indianapolis, IN, cat # 03335402001) should be included. Short RNA oligonucleotides can be obtained from Integrated DNA Technologies and longer RNAs are typically prepared using *in vitro* transcription kits mMESSAGE mMACHINE (T3 or T7, depending on a vector in which the corresponding DNA insert is subcloned) and purifying using MEGAclear columns (all from Ambion/Thermo Fisher Scientific, cat #s AM1348, AM1344 and AM1908, respectively). Cross-linking of unmodified RNA should be carried out using 254 nm UV-light and or with 354 nm UV light for RNA containing ThioU. If necessary, RNA cross linking reaction parameters should be optimized with Alexa Fluor dye-labeled RNA as describe for ssDNA cross-linking in the preceding section. If the downstream application will be MS, the gel should be stained with Coomassie G-250-type SimpleBlue SafeStain (Thermo Fisher Scientific) as described for A3G cross-linking to ssDNA. Otherwise stain the gel with standard Coomassie R-250 dye (methanol:acetic acid solubilized). Due to covalent modification, A3G cross-linked protein typically may not appear as a discrete shifted band following PAGE due to the heterogeneity that may be associated with APOBEC-RNA complex formation. A3G-RNA cross-linked complexes appear as material that slightly enters the resolving gel of 10–12% SDS PAGE (Figure 8B).

3.3. Mass spectrometry identification of A3G peptides interacting with ssDNA

3.3.1 Preparation of A3G cross-linked samples for mass spectrometry—To prepare the protein samples for MS the gel bands of interest should be obtained by accurately excising minimal sized slices of the gel containing A3G:DNA cross-linked complexes or A3G monomeric protein band and transferring them into LoBind tubes. The gel slices should be washed several times with Mili-Q-water and then with DNase I digestion buffer (20 mM Tris pH=7.5, 2.5 mM $MgCl_2$ and 0.5 mM $CaCl_2$). Cut the gel slices into about 1 mm pieces and add enough DNase I buffer to cover the gel pieces and digest with 10 μg of DNase I (Sigma-Aldrich, cat # DN25) and 300 units of micrococcal nuclease (Thermo Fisher Scientific, cat # ENO181) for 4–6 hours at 37 °C. After digestion, pulse-spin the tubes, remove the liquid and wash the gel slices three times with 1 ml of 25 mM ammonium bicarbonate (AmBc, Sigma-Aldrich, cat # 09830) with 5 min gentle vortexing each time. Pulse-spin the samples and remove all of the residual liquid and dehydrate the gel slices with two changes of solution containing 50% acetonitrile (ACN, LC-MS quality, J.T. Baker, cat # 9835-03) and 25 mM AmBc (10 min vortexing each time). Dry out the gel slices in a speed vacuum centrifuge (typically 15–20 min) and rehydrate them in cold AmBc in a volume just sufficient to cover the gel pieces. To remove any residual amount of nucleases and digested oligonucleotides in the samples, perform two more rounds of dehydration-rehydration followed by a treatment with 10 mM DTT in 25 mM AmBc at 55 °C for 1 hour and then with 55 mM iodoacetamide (Sigma-Aldrich, I-6125) in 25 mM AmBc at room temperature in the dark for 45 min. Wash the gel fragments with 25 mM

AmBc, dehydrate with 50% ACN/25 mM AmBc solution and vacuum dry. Digest the gel fragments overnight at 37 °C with Trypsin Gold (0.5 μg/μl; Promega, Madison, WI, cat # V5280) in 25 mM AmBc at 20:1 ratio (A3G:trypsin).

The next morning collect the supernatants into clean 1.5 ml LoBind tubes and extract the gel fragment digests twice with 50 μl of 50% ACN/5% TFA (trifluoroacetic acid, Thermo Scientific, cat #28901) while vortexing, for 20 min each time. Pulse-spin the tubes, sonicate for 5 min in Branson 1800 sonicator (Branson Ultrasonics, Danbury, CT), combine all three supernatants for each sample and dry the samples in a vacuum-speed centrifuge (typically 15–20 min). Dissolve the samples in 10 μl of LC-MS water (Burdick & Jackson, Morris Plains, NJ, cat # LC365) containing 0.1% of formic acid (MS quality; Thermo Fisher Scientific, cat # 28905). Importantly, when using high-accuracy high-resolution MS instrument, for example Q Exactive XL it is critical to ensure that MS samples have no contamination of any type, therefore clean the peptide samples with Pierce C18 spin columns (Thermo Fisher Scientific, cat. # 89870) as suggested by manufacturer and wear fresh gloves for all sample handling procedures.

3.3.2 Mass spectrometry analysis of A3G cross-linked samples—Tryptic peptide samples can be analyzed by MS using any high accuracy MS instrument, including ion-trap-type, MALDI or superior FT-MS that are manufactured by several companies (Bruker Daltonics, Thermo Scientific, Agilent, AB Sciex, Shimadzu or Waters). It may be beneficial to run samples on two different types of MS instruments, for example ion-trap and MALDI to determine which of the instruments provide better protein/peptide coverage or to obtain complementary results to increase data confidence. Detailed discussion on MS instrument choice is beyond the scope of this article. We have had good success using LTQ Orbitrap XL MS and Q Exactive XL MS (both from Thermo Scientific; located at the University of Rochester Medical Center Mass Spectrometry Core Resource Laboratory).

Briefly, in LTQ Orbitrap XL MS, a nano-HPLC system (Easy-nLC II, Thermo Scientific) is coupled to the electrospray ionization source and each sample is loaded by autosampler. We use a C_{18} reverse phase column, 5 μm particle size, 200Å pore diameter, 10cm in length, internal diameter 75 μm that should be equilibrated to initial run conditions prior to loading the sample. Solvent A is typically 0.1% formic acid in LC-MS grade water, and Solvent B is 0.1% formic acid in ACN. Peptides are normally eluted with the following chromatographic profile: 0% B for 2 min, ramping to 40% B over 13 min then to 70% B over 1 min, remaining at 70% B for 3 min, and finally returning to initial run conditions; flow rate 350 nl/min.

MS data are usually collected as RAW files and converted to MGF format using Proteome Discoverer (Thermo) or other software. Resultant MGF files should be searched via MASCOT (MatrixScience, Boston, MA), or imported and searched in Proteome Discoverer or other software, for example ProteinScape (Bruker Daltonics). In addition, identification of peptides that are cross-linked to the BrdU modification typically show the appropriate ratio of peptide masses that are due to $Br^{79/81}$ isotopes in the parent peak.

3.3.3 Approaches for identification of A3G cross-linked peptides—Several different approaches can be used to identify nucleic acid cross-linked peptides, primarily depending on MS instrument that is used in experiments and the knowledge or prediction of nucleotide modification (chemistry). With ion-trap type MS instrument, LTQ Obitrap XL, we have applied two different approaches. First, a deductive or comparative approach, wherein control A3G (no cross-linking) and cross-linked samples were analyzed by MS and MASCOT identified peptides in both samples are compared for A3G peptide coverage (Figure 9). Peptides modified with cross-linked nucleotides are expected to have different mass/charge values as compared to unmodified A3G control peptides of the same sequence. Therefore, the peptides that are not detected in MASCOT searches with parameters set to identify native (unmodified) A3G tryptic peptides, and thus missing in the cross-linked sample, are deduced to be involved in nucleic acid interactions. By comparing MS coverage of A3G peptides from cross-linked samples to the MS coverage of uncross-linked A3G samples, candidate peptides with covalently bound nucleotides could be easily identified (Figure 10).

By using such comparative MS approach we identified several candidate A3G ssDNA-interacting peptides [67]. For example, in control uncross-linked A3G sample, the MASCOT search for the tryptic peptide with monoisotopic mass of 865.3961 matched the predicted monoisotopic mass of IYDDQGR peptide (865.3930, aa 314–320). This peptide was not identified among peptides from the 25 nt BrdU-modified ssDNA cross linked sample, suggesting a different mass/charge value due to a covalent addition of nucleoside(s). Similarly, we observed a peptide with monoisotopic mass of 1761.9892 that matched YYILLHIMLGELR tryptic peptide (1761.9851, aa 181–194) in control A3G sample but again could not identify the peptide with this mass in A3G:DNA cross linked sample. These two A3G peptides together with a peptide aa 345–374 were reproducibly found as involved in A3G-DNA interaction in three independent experiments (Figure. 10). Peptide coverage typically observed in A3G control sample was reproducibly in a range of 76–77%. Interestingly, MS data on 25 nt ApoB RNA cross-linkiing to A3G obtained by using comparative MS approach showed an overlap in peptides cross-linked to ssDNA that suggested the presence of competitive binding sites where both RNA and DNA can bind to A3G in the C-terminus and extending into the N-terminus of A3G (Fig. 10) [67]. On the other hand, there were distinct peptides that only RNA cross-linked to N-terminal peptides (Fig. 10) [67].

An alternative or direct approach that we found helpful in the identification of A3G:RNA cross-linked peptides used an affinity purification procedure to separate cross-linked peptides before loading the samples onto MS. Biotin-conjugated DNA or RNA oligonucleotides of appropriate length sufficient to enable efficient binding of A3G were obtained from Integrated DNA Technology. Biotin is typically conjugated to 5’-end of the oligo via C6-linker. The cross-linking reaction should be performed as described in Section 3.2.1, however in this case it is more reasonable to keep the protein:nucleic acid molar ratio close to 1:1 and not to oversaturate the reactions with nucleic acid. After UV cross-linking the mixture should be digested overnight at 37 °C with Trypsin Gold (0.5 μg/μl) at 40:1 ratio (A3G:trypsin) and a fresh aliquot of trypsin added to the mixture the next morning for 2

hours in order to ensure a complete digestion. Trypsin Gold did not show nuclease activity during overnight incubation with A3G at 10:1 ratio as indicated by the recovery of intact oligonucleotides of expected length and resolved on 5% denaturing PAGE.

Purification of cross-linked peptides can be accomplished using NeutrAvidin (NA)-agarose (Thermo Fisher Scientific, cat # 29201). Before adding NA-resin to a mixture of tryptic peptides we usually dilute the latter with 1XPBS to 0.5–1.0 ml total volume so that the protein concentration does not exceed 1 mg/ml. Then add 50 μl of 1XPBS-pre-washed NA-resin in 2 ml Bio-Spin disposable chromatography columns (BioRad, Hercules, CA, cat # 7326008). These conditions are sufficient to pull-down cross-linked peptides from the mixture, assuming that out of initial 100 μgs of A3G only 10–20% is cross linked and that the NA-resin binding affinity is >3 mg protein/ml. Samples should be incubated at room temperature for 30 min or at 7 °C for 60 min.

After binding, the resin should be washed 3 times with 5–10 volumes of 1XPBS before eluting cross-linked peptides 3 times with 100 μl of 8M Guanidine·HCl pH=1.5 directly into the tubes containing 10 μl of 1M Trizma base (Sigma-Aldrich, unadjusted pH) to neutralize acidic pH. The resultant three fractions should be combined in one 1.5 ml LoBind tube and the peptides precipitated with 1.2 ml of cold acetone at −20 °C for 2 hrs. The precipitate can be recovered by microfuge centrifugation at 14,000×g and the pellet should be air dried for 30 min. The dried pellet should be dissolved in 50 μl of 10 mM Tris pH=7.5, 2 mM EDTA, and in the case of RNA-protein cross linking, 10 μgs of RNase A and RNase T1 added and the sample incubated for 1 hr at 37 °C to digest covalently attached RNA oligonucleotides.

Before loading the samples on high-accuracy high-resolution MS instruments nucleotides, salt, RNases and other contaminations should be removed from the samples with Pierce C18 spin columns (Thermo Fisher Scientific). However, at this step we prefer to use HPLC separation with C18 reverse phase column since in addition to cleaning the samples and better removing all contaminants including RNases, it will be possible to determine the number of peaks (peptides) in the samples by monitoring absorbance at A_{214} for peptide bond and A_{260} for cross-linked nucleotides. Each A_{214}/A_{260} peak could be analyzed separately on MS instrument to produce better quality data. It is important to keep in mind that large and more hydrophylic peptides, especially when they are modified by nucleotide attachment, may elute very early in a typical 0–70% ACN HPLC gradient. We normally use 0–40% ACN gradient and a low elution rate (<0.5 ml/min) to avoid this problem. Usually each HPLC peak that shows both A_{214}/A_{260} absorbances contains a peptide cross-linked to nucleic acid. The direct approach for identification of A3G cross-linked peptides enables selective analysis of peptides of interest, however requires more reagents, expensive equipment and more equipment time.

The third approach for identification of nucleic acid cross-linked peptides is also a direct approach however it requires a different MS instrument. Ion-trap MS instruments mentioned so far, including Orbitrap and Q Exactive are utilizing collision-induced dissociation (CID) technique to induce fragmentation of molecular ions (peptides) in the gas phase. Peptides are accelerated by electrical potential to high kinetic energy then allowed to collide with gas (usually helium) and some of that kinetic energy is converted into internal molecular energy

that results in peptide bond breakage and fragmentation into smaller fragments (peptides). These fragment ions can then be analyzed by a tandem MS, however peptides can be fragmented at different positions that makes their identification more complicated.

In contrast, Electron Transfer Dissociation (ETD) and similar electro-capture dissociation (ECD) MS techniques induce fragmentation of cations, proteins or peptides, by transferring electrons to them. During ETD/ECD, proteins or peptides are cleaved randomly along the peptide backbone while side chains and modifications such as phosphorylation are left intact. Therefore, ETD/ECD-type MS instruments afford advantages for the fragmentation of long peptides or even small proteins (top-down proteomics) and, importantly, are effective for the detection of protein/peptide modifications. These MS instruments maybe better suited for the direct identification of cross-linked peptides as well as modified individual amino acid residues. The only drawback of this procedure is that it requires a significant amount of highly pure protein for analysis.

ETD/ECD MS technology so far has not been used to study A3G interactions with nucleic acids primarily due to a fact that A3G protein is too large (46 kDa) to obtain a high quality ETD fragmentation data. Therefore, to make this approach successful it may require an additional step to construct a functional A3G protein with internal protease cleavable site between N-terminal and C-terminal domains that will enable MS analysis of both A3G parts separately but with much higher level of confidence. On the other hand, it should be possible to use this approach to study nucleic acid interactions with one CDD-containing APOBECs, for example A3C.

4. Summary

The following discussion of methods while largely based on our experience with A3G are likely to be generalizable because of the significant similarity in amino acid sequence and conserved structure among the APOBEC family members. We have advised the reader that we anticipate optimization of the protocols may be necessary when proteins other than A3G are evaluated. An important take home message from this review is to validate the purity and functional integrity of the protein preparation before embarking on studies of their protein-protein and protein-nucleic acid interactions. To this end we have described the approaches and metrics we use to establish a lab standard for comparing batch-to-batch preparation of recombinant A3G protein. We recommend that validation of purified protein should be based on both biophysical criteria and functional end-point quantification.

This review is focused on methods for preparation of functional APOBEC proteins that can be used to characterize their binding to nucleic acids and deaminase activity. While RNA binds to APOBEC1 as an editing substrate, RNA can bind to APOBEC proteins, such as A3G, and not serve as a substrate for deamination. Rather, RNA competitively and noncompetitively regulates ssDNA substrate binding to A3G and deaminase activity. We anticipate that this and other roles for RNA binding to APOBEC will be discovered and that the methods described here will enable future mechanistic studies. However, with the exceptions of APOBEC1 and to a more limited extent AID, very little is known concerning the interactions of APOBEC with other cellular proteins. Most of what has been reported for

other APOBEC proteins is related to their homo oligomerization. In the case of A3G, the three dimensional PAGE analysis revealed that homo oligomerization is critical for holoenzyme formation and function [61–63]. Glycerol gradient sedimentation has been a valuable analytical tool in the discovery of the APOBEC editosome, the A3G holoenzyme as well as inactivation of A3G antiviral activity through ribonucleoprotein complex formation. Under explored is the functional significance of RNA-bridged protein-protein interactions that may contribute to subcellular localization of APOBEC and their functional interactions with nucleic acid substrates [62, 103–109].

We hasten to add that APOBEC and RNA editing and DNA modification as a field is rapidly developing and that new instrumentation and technology for MS and proteomics are rapidly emerging. Therefore the reader should take what we describe here as fundamental principles and first steps toward the analysis of editing enzymes but be prepared to optimize and develop new protocols that will address relevant emergent questions.

Acknowledgments

The authors wish to acknowledge the numerous contributions of investigators in the field of RNA and DNA editing and modification and especially the trainees in Dr. Smith's laboratory who embrace the spirit of inventors and discovery. The work in Dr. Smith's laboratory was supported by PHS grants DK043739, AI054369, AI095007, AI058789, and GM110568 (HCS, PI) and NS067671 (awarded to JE Wedekind). W.M. McD was supported during his graduate thesis research in part NIH T32 training grants GM068411 and AI049815 and an Elon Huntington Hooker Graduate Research Fellowship. Mass spectrometry was performed by the University of Rochester Medical Center Mass Spectrometry Resource Laboratory.

Abbreviations

APOBEC	Apolipoprotein B Editing Catalytic Subunit-like
CD	cytidine deaminase domain
RNP	ribonucleoprotein particle
ZDD	zinc-dependent deaminase domain
EMSA	fluorescence anisotropy, mass spectroscopy
PAGE	velocity sedimentation

References

1. Smith HC, et al. Functions and regulation of the APOBEC family of proteins. Semin Cell Dev Biol. 2012; 23(3):258–268. [PubMed: 22001110]
2. MacElrevey, C.; Wedekind, JE. Chapter 16: Chemistry, Phylogeny, and Three-Dimensional Structure of the APOBEC Protein Family. In: Smith, HC., editor. RNA and DNA Editing: Molecular Mechanisms and Their Integration into Biological Systems. John Wiley & Sons Inc; 2008. p. 371-421.
3. LaRue RS, et al. Guidelines for naming nonprimate APOBEC3 genes and proteins. Journal of Virology. 2009; 83(2):494–497. [PubMed: 18987154]
4. Wedekind JE, et al. Messenger RNA editing in mammals: new members of the APOBEC family seeking roles in the family business. Trends in genetics : TIG. 2003; 19(4):207–216. [PubMed: 12683974]

5. Sato Y, et al. Deficiency in APOBEC2 leads to a shift in muscle fiber type, diminished body mass, and myopathy. J Biol Chem. 2010; 285(10):7111–7118. [PubMed: 20022958]
6. Almeida RR, et al. Modulating APOBEC expression enhances DNA vaccine immunogenicity. Immunol Cell Biol. 2015
7. Rosenberg BR, et al. Transcriptome-wide sequencing reveals numerous APOBEC1 mRNA-editing targets in transcript 3' UTRs. Nat Struct Mol Biol. 2011; 18(2):230–236. [PubMed: 21258325]
8. Smith HC, Wedekind JE, Kefang X, Sowden MP. Mammalian C to U Editing. Topics in Current Bioogy. 2005; 12:365–400.
9. Skuse GR, et al. The neurofibromatosis type I messenger RNA undergoes base-modification RNA editing. Nucleic Acids Research. 1996; 24(3):478–485. [PubMed: 8602361]
10. Dickerson SK, et al. AID mediates hypermutation by deaminating single stranded DNA. J Exp Med. 2003; 197(10):1291–1296. [PubMed: 12756266]
11. 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(6):803–809. [PubMed: 12809610]
12. 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(5):1247–1253. [PubMed: 12453430]
13. Stenglein MD, Harris RS. APOBEC3B and APOBEC3F inhibit L1 retrotransposition by a DNA deamination-independent mechanism. The Journal of Biological chemistry. 2006; 281(25):16837–16841. [PubMed: 16648136]
14. Zhen A, et al. Reduced APOBEC3H variant anti-viral activities are associated with altered RNA binding activities. PLoS One. 2012; 7(7):e38771. [PubMed: 22859935]
15. Smith HC. APOBEC3G: a double agent in defense. Trends Biochem Sci. 2011; 36(5):239–244. [PubMed: 21239176]
16. Hultquist JF, et al. Human and Rhesus APOBEC3D, APOBEC3F, APOBEC3G, and APOBEC3H Demonstrate a Conserved Capacity to Restrict Vif-deficient HIV-1. Journal of Virology. 2011
17. Sharma S, et al. APOBEC3A cytidine deaminase induces RNA editing in monocytes and macrophages. Nat Commun. 2015; 6:6881. [PubMed: 25898173]
18. Chan K, et al. An APOBEC3A hypermutation signature is distinguishable from the signature of background mutagenesis by APOBEC3B in human cancers. Nat Genet. 2015; 47(9):1067–1072. [PubMed: 26258849]
19. Conticello SG. The AID/APOBEC family of nucleic acid mutators. Genome Biol. 2008; 9(6):229. [PubMed: 18598372]
20. Munk C, Willemsen A, Bravo IG. An ancient history of gene duplications, fusions and losses in the evolution of APOBEC3 mutators in mammals. BMC Evol Biol. 2012; 12:71. [PubMed: 22640020]
21. Wang Q, et al. Epigenetic targeting of activation-induced cytidine deaminase. Proc Natl Acad Sci U S A. 2014; 111(52):18667–18672. [PubMed: 25512519]
22. Franchini DM, Petersen-Mahrt SK. AID and APOBEC deaminases: balancing DNA damage in epigenetics and immunity. Epigenomics. 2014; 6(4):427–443. [PubMed: 25333851]
23. Burns MB, Temiz NA, Harris RS. Evidence for APOBEC3B mutagenesis in multiple human cancers. Nat Genet. 2013; 45(9):977–983. [PubMed: 23852168]
24. Sowden M, Hamm JK, Smith HC. Overexpression of APOBEC-1 results in mooring sequence-dependent promiscuous RNA editing. The Journal of Biological chemistry. 1996; 271(6):3011–3017. [PubMed: 8621694]
25. 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 USA. 1995; 92:8483–8487. [PubMed: 7667315]
26. Yamanaka S, et al. 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(3):321–333. [PubMed: 9030685]
27. Rebhandl S, et al. AID/APOBEC deaminases and cancer. Oncoscience. 2015; 2(4):320–333. [PubMed: 26097867]

28. Robbiani DF, Nussenzweig MC. Chromosome translocation, B cell lymphoma, and activation-induced cytidine deaminase. Annu Rev Pathol. 2013; 8:79–103. [PubMed: 22974238]
29. Chiu YL, et al. High-molecular-mass APOBEC3G complexes restrict Alu retrotransposition. Proc Natl Acad Sci U S A. 2006; 103(42):15588–15593. [PubMed: 17030807]
30. Horn AV, 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 research. 2013
31. Huang J, et al. Derepression of microRNA-mediated protein translation inhibition by apolipoprotein B mRNA-editing enzyme catalytic polypeptide-like 3G (APOBEC3G) and its family members. The Journal of Biological chemistry. 2007; 282(46):33632–33640. [PubMed: 17848567]
32. Hulme AE, et al. Selective inhibition of Alu retrotransposition by APOBEC3G. Gene. 2007; 390(1–2):199–205. [PubMed: 17079095]
33. Kinomoto M, et al. All APOBEC3 family proteins differentially inhibit LINE-1 retrotransposition. Nucleic Acids Research. 2007; 35(9):2955–2964. [PubMed: 17439959]
34. Muckenfuss H, et al. APOBEC3 proteins inhibit human LINE-1 retrotransposition. Journal of Biological Chemistry. 2006; 281(31):22161–22172. [PubMed: 16735504]
35. Schumann GG. APOBEC3 proteins: major players in intracellular defence against LINE-1-mediated retrotransposition. Biochem Soc Trans. 2007; 35(Pt 3):637–642. [PubMed: 17511669]
36. Honjo T, Muramatsu M, Fagarasan S. AID: how does it aid antibody diversity? Immunity. 2004; 20(6):659–668. [PubMed: 15189732]
37. Longerich S, et al. AID in somatic hypermutation and class switch recombination. Curr Opin Immunol. 2006; 18(2):164–174. [PubMed: 16464563]
38. Chiarle R, et al. Genome-wide translocation sequencing reveals mechanisms of chromosome breaks and rearrangements in B cells. Cell. 2011; 147(1):107–119. [PubMed: 21962511]
39. Klein IA, et al. Translocation-capture sequencing reveals the extent and nature of chromosomal rearrangements in B lymphocytes. Cell. 2011; 147(1):95–106. [PubMed: 21962510]
40. Koito A, Ikeda T. Intrinsic immunity against retrotransposons by APOBEC cytidine deaminases. Front Microbiol. 2013; 4:28. [PubMed: 23431045]
41. Esnault C, et al. Dual inhibitory effects of APOBEC family proteins on retrotransposition of mammalian endogenous retroviruses. Nucleic Acids Research. 2006; 34(5):1522–1531. [PubMed: 16537839]
42. Prohaska KM, et al. The multifaceted roles of RNA binding in APOBEC cytidine deaminase functions. Wiley Interdiscip Rev RNA. 2014; 5(4):493–508. [PubMed: 24664896]
43. Smith HC. Measuring editing activity and identifying cytidine-to-uridine mRNA editing factors in cells and biochemical isolates. Methods in enzymology. 2007; 424:389–416. [PubMed: 17662851]
44. Beale RC, et al. Comparison of the differential context-dependence of DNA deamination by APOBEC enzymes: correlation with mutation spectra in vivo. J Mol Biol. 2004; 337(3):585–596. [PubMed: 15019779]
45. Petersen-Mahrt SK, Harris RS, Neuberger MS. AID mutates E. coli suggesting a DNA deamination mechanism for antibody diversification. Nature. 2002; 418:99–103. [PubMed: 12097915]
46. Prochnow C, et al. The APOBEC-2 crystal structure and functional implications for the deaminase AID. Nature. 2007; 445(7126):447–451. [PubMed: 17187054]
47. Brar SS, et al. Activation-induced deaminase, AID, is catalytically active as a monomer on single-stranded DNA. DNA Repair (Amst). 2008; 7(1):77–87. [PubMed: 17889624]
48. Krzysiak TC, et al. APOBEC2 is a monomer in solution: implications for APOBEC3G models. Biochemistry. 2012; 51(9):2008–2017. [PubMed: 22339232]
49. Mitra M, et al. Structural determinants of human APOBEC3A enzymatic and nucleic acid binding properties. Nucleic Acids Res. 2014; 42(2):1095–1110. [PubMed: 24163103]
50. Kitamura S, et al. The APOBEC3C crystal structure and the interface for HIV-1 Vif binding. Nat Struct Mol Biol. 2012; 19(10):1005–1010. [PubMed: 23001005]

51. Xie K, et al. 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(21):8114–8119. [PubMed: 15148397]
52. Bohn MF, et al. The ssDNA Mutator APOBEC3A Is Regulated by Cooperative Dimerization. Structure. 2015; 23(5):903–911. [PubMed: 25914058]
53. Kouno T, et al. Structure of the Vif-binding domain of the antiviral enzyme APOBEC3G. Nat Struct Mol Biol. 2015; 22(6):485–491. [PubMed: 25984970]
54. Chen KM, et al. Extensive mutagenesis experiments corroborate a structural model for the DNA deaminase domain of APOBEC3G. FEBS Lett. 2007; 581(24):4761–4766. [PubMed: 17869248]
55. Shandilya SM, et al. Crystal structure of the APOBEC3G catalytic domain reveals potential oligomerization interfaces. Structure. 2010; 18(1):28–38. [PubMed: 20152150]
56. Harjes E, et al. An extended structure of the APOBEC3G catalytic domain suggests a unique holoenzyme model. J Mol Biol. 2009; 389(5):819–832. [PubMed: 19389408]
57. Furukawa A, et al. Structure, interaction and real-time monitoring of the enzymatic reaction of wild-type APOBEC3G. EMBO J. 2009; 28(4):440–451. [PubMed: 19153609]
58. Bulliard Y, et al. Structure-function analyses point to a polynucleotide-accommodating groove essential for APOBEC3A restriction activities. J Virol. 2011; 85(4):1765–1776. [PubMed: 21123384]
59. Shlyakhtenko LS, et al. Interaction of APOBEC3A with DNA assessed by atomic force microscopy. PLoS One. 2014; 9(6):e99354. [PubMed: 24905100]
60. Shlyakhtenko LS, et al. Atomic force microscopy studies provide direct evidence for dimerization of the HIV restriction factor APOBEC3G. J Biol Chem. 2011; 286(5):3387–3395. [PubMed: 21123176]
61. Wedekind JE, et al. Nanostructures of APOBEC3G support a hierarchical assembly model of high molecular mass ribonucleoprotein particles from dimeric subunits. J Biol Chem. 2006; 281(50): 38122–38126. [PubMed: 17079235]
62. Friew YN, et al. Intracellular interactions between APOBEC3G, RNA, and HIV-1 Gag: APOBEC3G multimerization is dependent on its association with RNA. Retrovirology. 2009; 6:56. [PubMed: 19497112]
63. Shlyakhtenko LS, et al. Atomic force microscopy studies of APOBEC3G oligomerization and dynamics. J Struct Biol. 2013
64. Stopak KS, et al. Distinct patterns of cytokine regulation of APOBEC3G expression and activity in primary lymphocytes, macrophages, and dendritic cells. J Biol Chem. 2007; 282(6):3539–3546. [PubMed: 17110377]
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(35):30655–30661. [PubMed: 21737457]
66. McDougall, WMaS; C, H. Direct Evidence that RNA Inhibits APOBEC3G ssDNA Cytidine Deaminase Activity. Biophysical and Biochemical Research Communications. 2011; 412(4):612–617.
67. Polevoda B, et al. RNA binding to APOBEC3G induces the disassembly of functional deaminase complexes by displacing single-stranded DNA substrates. Nucleic Acids Res. 2015; 43(19):9434–9445. [PubMed: 26424853]
68. Salter JD, et al. A hydrodynamic analysis of APOBEC3G reveals a monomer-dimer-tetramer self-association that has implications for anti-HIV function. Biochemistry. 2009; 48(45):10685–10687. [PubMed: 19839647]
69. Chelico L, et al. APOBEC3G DNA deaminase acts processively 3' -->5' on single-stranded DNA. Nat Struct Mol Biol. 2006; 13(5):392–399. [PubMed: 16622407]
70. Chen KM, et al. Structure of the DNA deaminase domain of the HIV-1 restriction factor APOBEC3G. Nature. 2008; 452(7183):116–119. [PubMed: 18288108]
71. Kreisberg JF, Yonemoto W, Greene WC. Endogenous factors enhance HIV infection of tissue naive CD4 T cells by stimulating high molecular mass APOBEC3G complex formation. J Exp Med. 2006; 203(4):865–870. [PubMed: 16606671]

72. Siriwardena SU, Guruge TA, Bhagwat AS. Characterization of the Catalytic Domain of Human APOBEC3B and the Critical Structural Role for a Conserved Methionine. J Mol Biol. 2015; 427(19):3042–3055. [PubMed: 26281709]
73. Smith HC. Analysis of protein complexes assembled on apolipoprotein B mRNA for mooring sequence-dependent RNA editing. Methods. 1998; 15(1):27–39. [PubMed: 9614650]
74. Smith HC. Measuring editing activity and identifying cytidine-to-uridine mRNA editing factors in cells and biochemical isolates. Methods Enzymol. 2007; 424:389–416. [PubMed: 17662851]
75. Shindo K, et al. A Comparison of Two Single-Stranded DNA Binding Models by Mutational Analysis of APOBEC3G. Biology (Basel). 2012; 1(2):260–276. [PubMed: 24832226]
76. Bransteitter R, et al. Activation-induced cytidine deaminase deaminates deoxycytidine on single-stranded DNA but requires the action of RNase. Proceedings of the National Academy of Sciences USA. 2003; 100:4102–4107.
77. Nabel CS, et al. AID/APOBEC deaminases disfavor modified cytosines implicated in DNA demethylation. Nat Chem Biol. 2012; 8(9):751–758. [PubMed: 22772155]
78. Nowarski R, et al. Hypermutation by intersegmental transfer of APOBEC3G cytidine deaminase. Nat Struct Mol Biol. 2008; 15(10):1059–1066. [PubMed: 18820687]
79. Coker HA, Morgan HD, Petersen-Mahrt SK. Genetic and in vitro assays of DNA deamination. Methods Enzymol. 2006; 408:156–170. [PubMed: 16793368]
80. Iwatani Y, et al. Biochemical activities of highly purified, catalytically active human APOBEC3G: correlation with antiviral effect. J Virol. 2006; 80(12):5992–6002. [PubMed: 16731938]
81. Larijani M, et al. AID associates with single-stranded DNA with high affinity and a long complex half-life in a sequence-independent manner. Mol Cell Biol. 2007; 27(1):20–30. [PubMed: 17060445]
82. Jarmuz A, et al. An Anthropoid-Specific Locus of Orphan C to U RNA-Editing Enzymes on Chromosome 22. Genomics. 2002; 79(3):285–296. [PubMed: 11863358]
83. 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(24):14762–14767. [PubMed: 7782342]
84. Navaratnam N, et al. Escherichia coli cytidine deaminase provides a molecular model for ApoB RNA editing and a mechanism for RNA substrate recognition. J Mol Biol. 1998; 275(4):695–714. [PubMed: 9466941]
85. Byeon IJ, et al. NMR structure of human restriction factor APOBEC3A reveals substrate binding and enzyme specificity. Nat Commun. 2013; 4:1890. [PubMed: 23695684]
86. Maris C, et al. NMR structure of the apoB mRNA stem-loop and its interaction with the C to U editing APOBEC1 complementary factor. RNA. 2005; 11(2):173–186. [PubMed: 15659357]
87. Wang J, et al. Identification of a specific domain required for dimerization of activation-induced cytidine deaminase. J Biol Chem. 2006; 281(28):19115–19123. [PubMed: 16687409]
88. Bennett RP, et al. APOBEC3G subunits self-associate via the C-terminal deaminase domain. J Biol Chem. 2008; 283(48):33329–33336. [PubMed: 18842592]
89. Chelico L, et al. A model for oligomeric regulation of APOBEC3G cytosine deaminase-dependent restriction of HIV. J Biol Chem. 2008; 283(20):13780–13791. [PubMed: 18362149]
90. Feng Y, et al. Natural polymorphisms and oligomerization of human APOBEC3H contribute to single-stranded DNA scanning ability. J Biol Chem. 2015
91. 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(13):11415–11426. [PubMed: 21300806]
92. 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. USA. 1994; 91:8522–8526. [PubMed: 8078915]
93. Luo K, et al. Cytidine deaminases APOBEC3G and APOBEC3F interact with human immunodeficiency virus type 1 integrase and inhibit proviral DNA formation. J Virol. 2007; 81(13):7238–7248. [PubMed: 17428847]

94. Hasler J, Rada C, Neuberger MS. Cytoplasmic activation-induced cytidine deaminase (AID) exists in stoichiometric complex with translation elongation factor 1alpha (eEF1A). Proc Natl Acad Sci U S A. 2011; 108(45):18366–18371. [PubMed: 22042842]

95. Huthoff H, Malim MH. Identification of amino acid residues in APOBEC3G required for regulation by human immunodeficiency virus type 1 Vif and Virion encapsidation. J Virol. 2007; 81(8):3807–3815. [PubMed: 17267497]

96. He Z, et al. Characterization of Conserved Motifs in HIV-1 Vif Required for APOBEC3G and APOBEC3F Interaction. J Mol Biol. 2008

97. Cen S, et al. The interaction between HIV-1 Gag and APOBEC3G. J Biol Chem. 2004; 279(32): 33177–33184. [PubMed: 15159405]

98. Bogerd HP, et al. A single amino acid difference in the host APOBEC3G protein controls the primate species specificity of HIV type 1 virion infectivity factor. Proc Natl Acad Sci U S A. 2004; 101(11):3770–3774. [PubMed: 14999100]

99. Yang Y, Sowden MP, Smith HC. Intracellular trafficking determinants in APOBEC-1, the catalytic subunit for cytidine to uridine editing of apolipoprotein B mRNA. Exp Cell Res. 2001; 267(2): 153–164. [PubMed: 11426934]

100. Chaudhuri J, Khuong C, Alt FW. Replication protein A interacts with AID to promote deamination of somatic hypermutation targets. Nature. 2004; 430(7003):992–998. [PubMed: 15273694]

101. Orthwein A, P A, Affar EB, Lamarre YJ, Di Noia JM. Regulation of activation-induced deaminase stability and antibody gene diversification by Hsp90. The Journal of Experimental Medicine. 2010; 207(12):2751–2765. [PubMed: 21041454]

102. Patenaude AM, O A, Hu Y, Campo VH, Kavli B, Buschiazzo A, Di Noia JM. Active nuclear import and cytoplasmic retention of activation-induced deaminase. nature structural and molecuklar biology. 2009; 16(5):517–527.

103. Wichroski MJ, Robb GB, Rana TM. Human retroviral host restriction factors APOBEC3G and APOBEC3F localize to mRNA processing bodies. PLoS Pathog. 2006; 2(5):e41. [PubMed: 16699599]

104. Galloway CA, et al. APOBEC-1 complementation factor (ACF) forms RNA-dependent multimers. Biochem Biophys Res Commun. 2010; 398(1):38–43. [PubMed: 20541536]

105. Sowden MP, et al. 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(Pt 5):1027–1039. [PubMed: 11870221]

106. Gallois-Montbrun S, et al. Antiviral protein APOBEC3G localizes to ribonucleoprotein complexes found in P bodies and stress granules. J Virol. 2007; 81(5):2165–2178. [PubMed: 17166910]

107. Gallois-Montbrun S, et al. Comparison of cellular ribonucleoprotein complexes associated with the APOBEC3F and APOBEC3G antiviral proteins. J Virol. 2008; 82(11):5636–5642. [PubMed: 18367521]

108. Apolonia L, et al. Promiscuous RNA binding ensures effective encapsidation of APOBEC3 proteins by HIV-1. PLoS Pathog. 2015; 11(1):e1004609. [PubMed: 25590131]

109. Huthoff H, et al. RNA-dependent oligomerization of APOBEC3G is required for restriction of HIV-1. PLoS Pathog. 2009; 5(3):e1000330. [PubMed: 19266078]

Highlights

- The APOBEC family of proteins share many structural features but are divergent in their cell-type expression, interactions with other macromolecules and functions.
- Methods for the functional analysis of APOBEC expressed in cells and as purified proteins have been developed.
- Protein-protein interactions of APOBEC and their binding to nucleic acids can be evaluated and quantified using an array of immunological-based assays and biophysical methods.
- Cross-linking methods combined with physical separate of APOBEC-containing complexes and mass spectroscopy can reveal structural interactions that are essential for function and regulation.

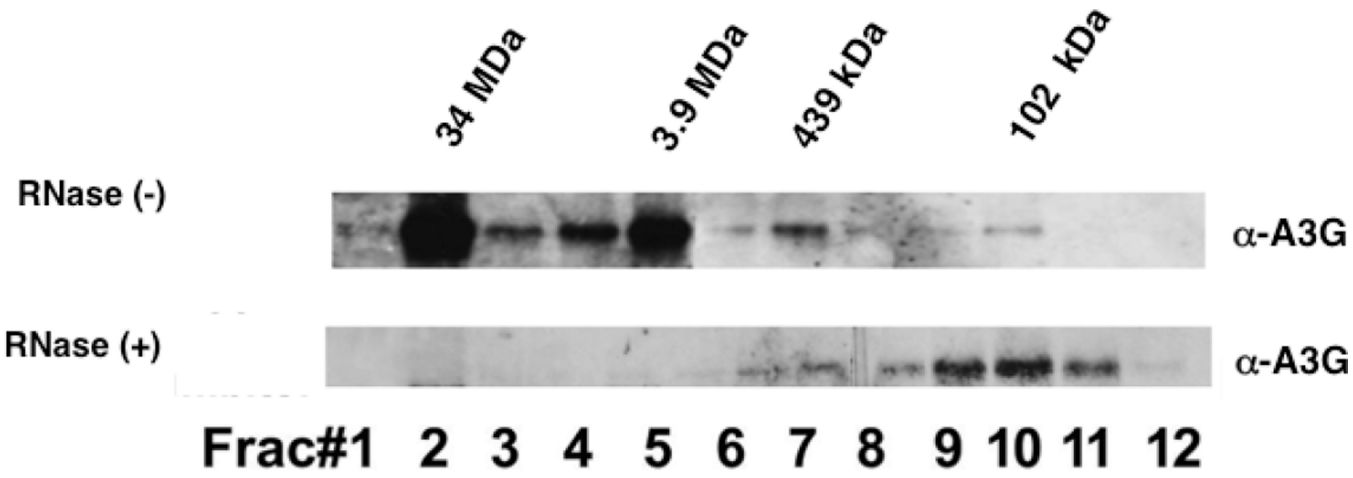

Figure 1. RNA-dependent Complexes of A3G

Size-exclusion HPLC chromatographic separation of A3G expressed in and isolated from Sf9 cells without (A) or following RNase A digestion (B) of the cell extract. Western blot of SDS-PAGE of chromatographic fractions were probed with A3G-specific primary antibodies. The protein molecular mass corresponding to fractions was determined based on the chromatographic separation of molecular mass standard proteins and indicated above the panels.

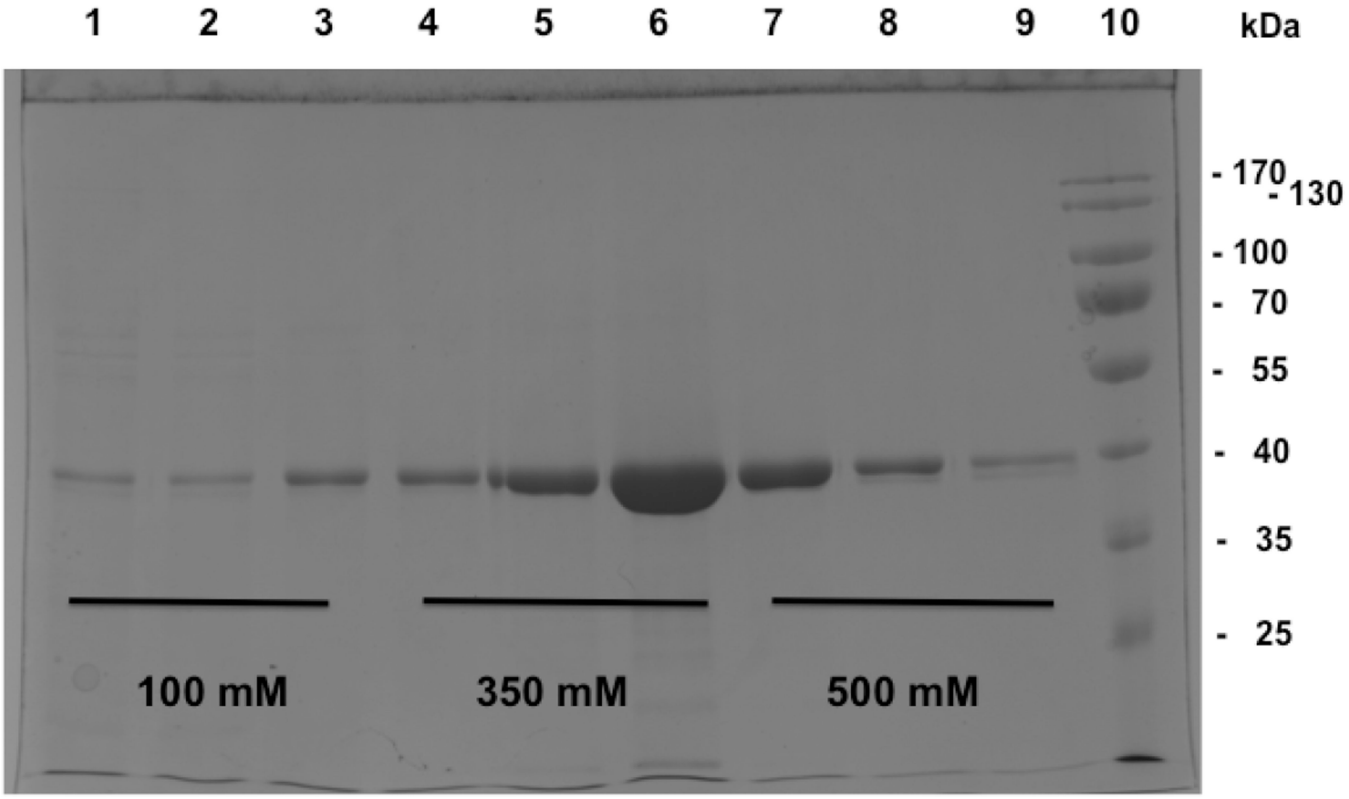

Figure 2. Affinity Chromatographic Enrichment of A3G Expressed in Sf9 Insect Cells

Coomassie blue was used to stain SDS-PAGE of A3G with a C-terminal 6xHis tag in eluted fractions from nickel affinity resin (Ni-NTA) with imidazole. Represented in lanes 1–9 are proteins in three equal aliquots of 100 mM imidazole wash, 350 mM and 500 mM imidazole elutions (concentrations shown at the bottom of the lanes). Lane 10 shows the relative migration of BioRad broad range molecular mass standard proteins (cat # 161-0363).

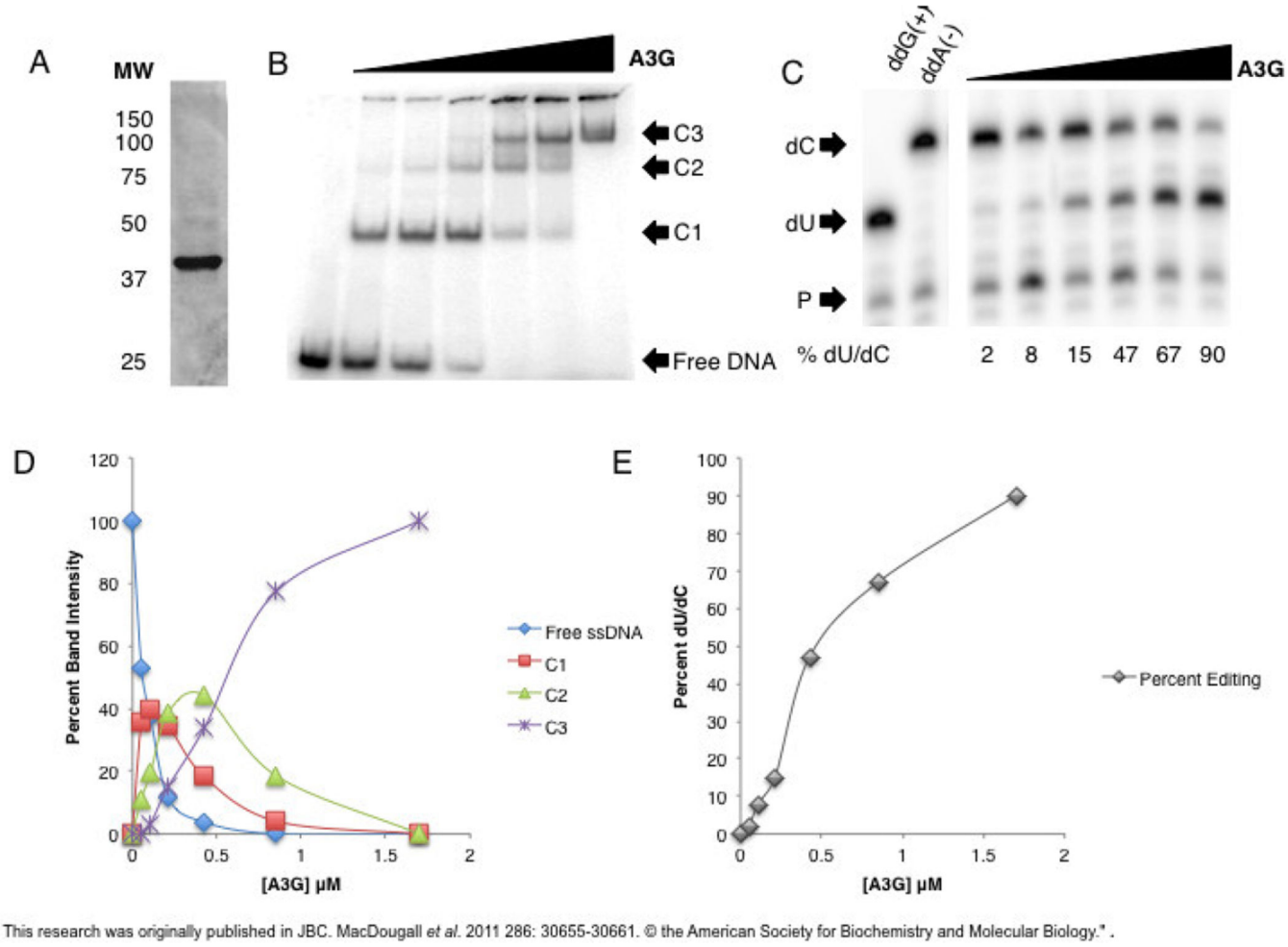

Figure 3. Complexes of A3G Bound to ssDNA and Their Corresponding Deaminase Activity

(**A**) Coomassie stain of 5 μg Sf9 cell-derived purified full length human A3G. (**B**) EMSA using [^{32}P] γ-ATP 5' end-labeled 41 nt substrate ssDNA. The concentration of ssDNA in each reaction was 0.6 μM whereas the concentration of A3G in each lane was 0 μM, 0.05 μM, 0.11 μM, 0.22 μM, 0.44 μM, 0.86 μM, 1.75 μM (Lanes 1–7 respectively). C1, C2, and C3 refer to A3G:ssDNA complex 1, 2, and 3 respectively. Complexes were visualized by phosphorimager scanning densitometry by virtue of their labeled ssDNA content. (**C**) Deaminase activity was assayed by primer extension on 41 nt ssDNA that was incubated with A3G. The concentration of ssDNA was 0.6 μM per reaction and the concentration of A3G in each lane was 0.05 μM, 0.11 μM, 0.22 μM, 0.44 μM, 0.86 μM, 1.75 μM (Lanes 1–6 respectively). The lowest arrow indicates free primer, the middle arrow indicates the primer extension product for deaminated ssDNA (dU) and the upper arrow indicates the primer extension product for unmodified ssDNA (dC). The percent deaminated ssDNA (%dU/dC) due to each reaction condition was determined by phosphorimager scanning densitometry and calculated as dU divided by (dU +dC) times 100. (**D**) Graphic representation of the quantification of individual bands in 'B' relative to input A3G in each reaction. (**E**) Graphic representation of deaminase activity in panel 'C' demonstrating the positive correlation of

deaminase activity with A3G concentration in the reaction and the appearance of complex C3 (as shown in panel ‘B’).

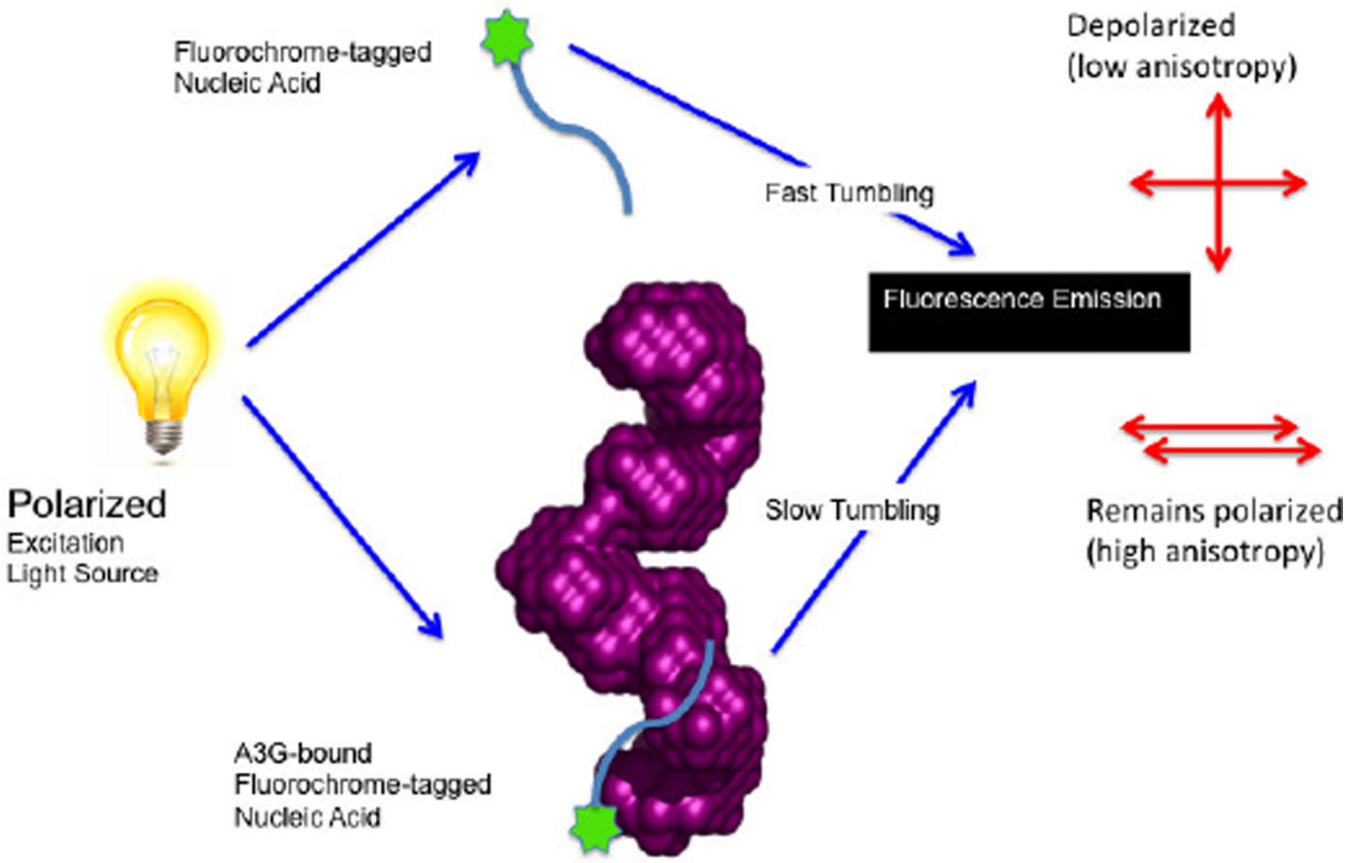

Figure 4. Schematic of Fluorescence Anisotropy Analysis of A3G Binding to Nucleic Acids

Polarized light is used to excite the fluorochrome bound to the 5' end of a short oligo nucleotide either in the absence of protein or bound to purified A3G (shown as a dimer based on Small Angle X-ray scattering determination [61]. The fast tumbling rate of free nucleic acid will result in anisotropic emission of fluorescent light. The tumbling rate of free nucleic acid in solution will decrease upon nucleic acid binding to A3G leading to a decrease in anisotropic emission. Quantification of the changes in fluorescence anisotropy can be used in the determination of binding affinity of A3G for nucleic acid.

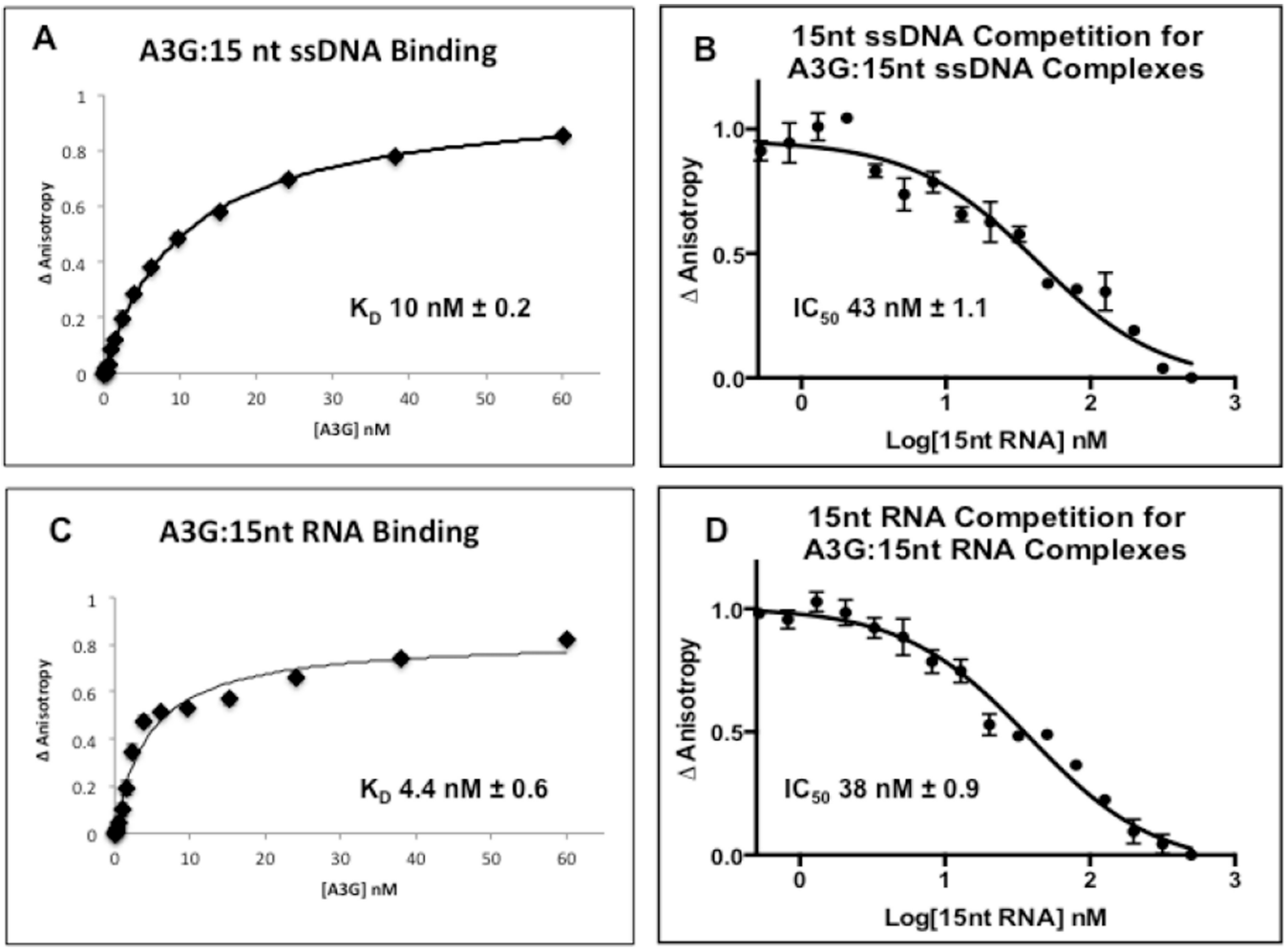

Figure 5. A3G Binding to Nucleic Acids and Competition Binding

Fluorescence anisotropy changes due to A3G binding to AlexFluor 647 labeled (**A**) 15 nt substrate ssDNA, 5'-TTATTUUUAAGGATT-3' (**C**) 15 nt substrate RNA (5'-UUAUUUUUAAGGAUU-3 were quantified in triplicate by titrating A3G into a binding buffer containing 2 nM of the indicated nucleic acid. The titration of A3G was 0–60 nM. The average normalized change in anisotropy (y-axis) of three individual experiments is plotted as a function of A3G concentration (x-axis) and error bars represent the Standard Error of the Mean (SEM) of the data. The data were fit to a single-site binding equation and binding constants (K_D) shown within each panel determined using GraphPad Prism 6.0. For competition analysis, A3G complexes assembled with AlexFluor 647 labeled 15 nt ssDNA (**B**) or AlexFluor 647 labeled 15 nt RNA (**D**) were subjected to competition with unlabeled 15 nt ssDNA or 15 nt RNA competitors (panels B and D, respectively) over a range of molar excess of competitor from 0 to 250. The average normalized change in anisotropy (y-axis) of three individual experiments was plotted versus the Log10 of the competitor molar excess (x-axis). Changes in fluorescence anisotropy were quantified. The error bars represent the Standard Error of the Mean (SEM) of the data. The data were fit to a single-site competition equation (see Methods) and the inhibitory concentration (IC_{50}) determined using GraphPad Prism 6.0.

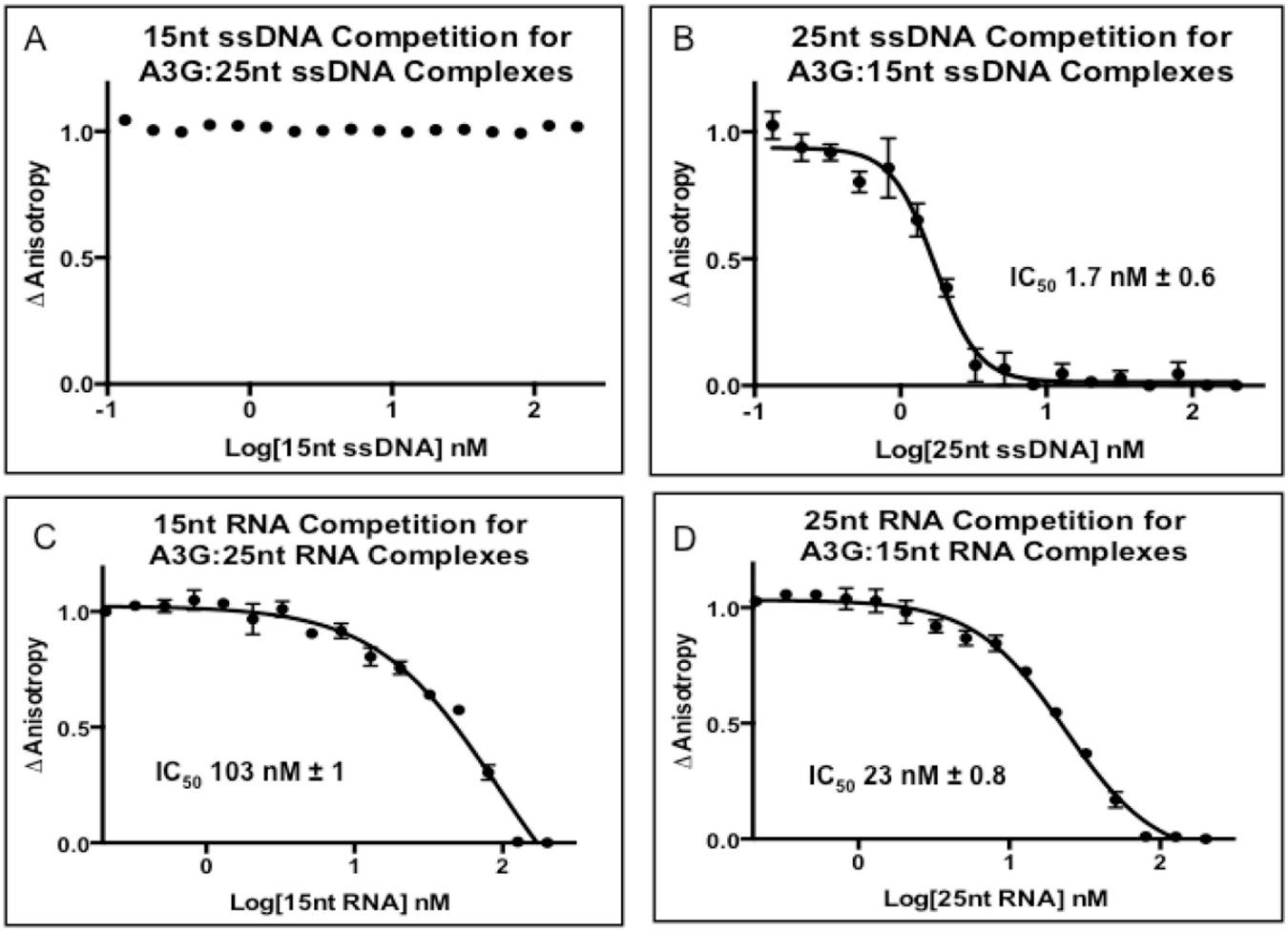

Figure 6. A3G Has a Higher Affinity for Longer ssDNA and RNA

For competition analysis, A3G complexes assembled with AlexFluor 647 labeled (**A**) 15 nt ssDNA, (**B**) 25 nt ssDNA, (**C**) 15 nt RNA or (**D**) 25 nt RNA were subjected to competition with unlabeled 25 nt ssDNA, 15 nt ssDNA, 25 nt RNA or 15 nt RNA in panels A through D respectively over a range of molar excess of competitor from 0 to 250-fold. The average normalized change in anisotropy (y-axis) of three individual experiments was plotted versus the Log10 of the competitor molar excess (x-axis). Changes in fluorescence anisotropy were quantified. The error bars represent the Standard Error of the Mean (SEM) of the data. The data were fit to a single-site competition equation (see Methods) and the inhibitory concentration (IC_{50}) determined using GraphPad Prism 6.0. 15 nt ssDNA and RNA sequences were as described in Fig. 5. 25 nt ssDNA sequences was: 5'-TTATTUUUAAGGATTTATTTATTTA-3' and RNA sequence was: 5'-UUAUUUUUAAGGAUUUAUUUAUUUA-3'.

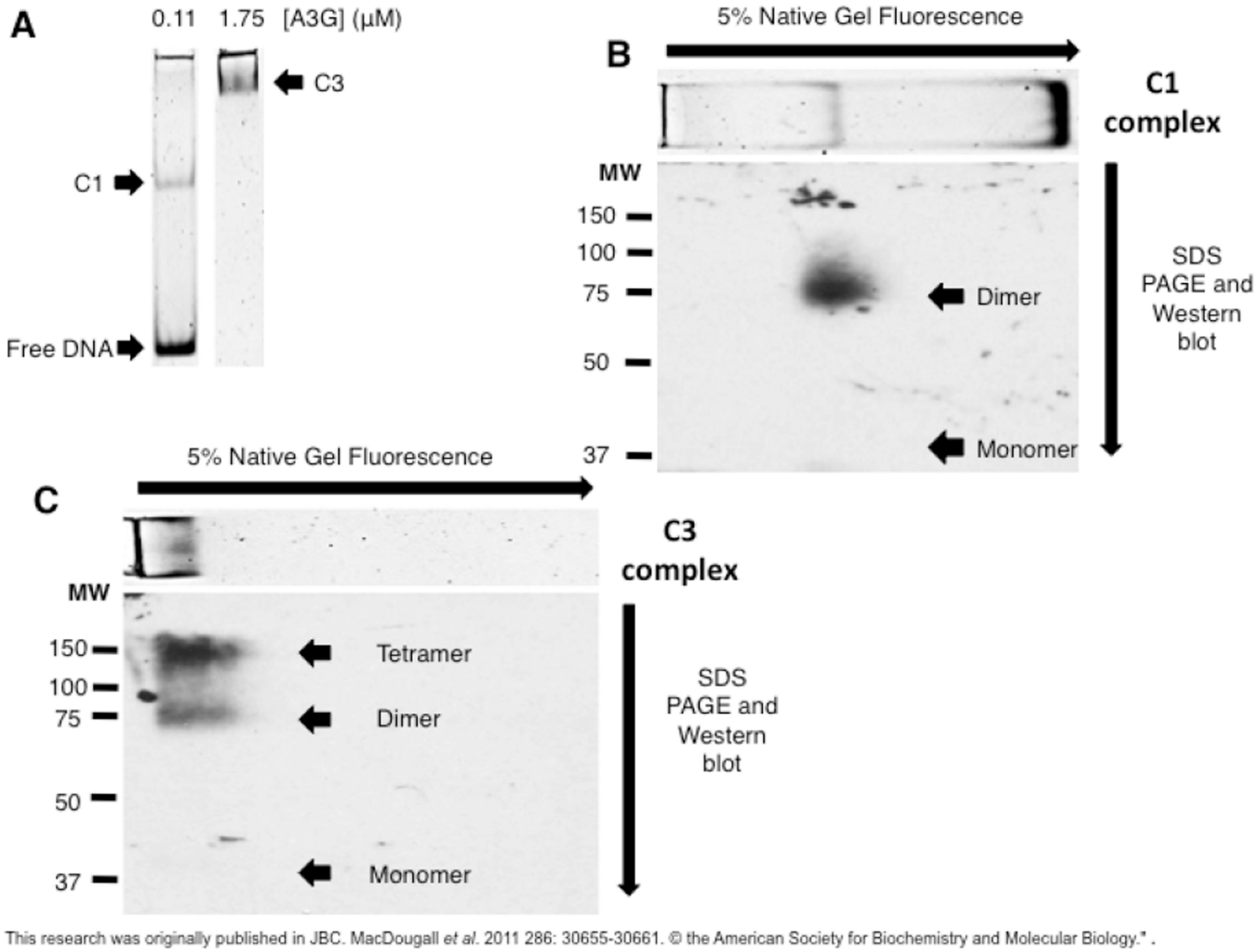

Figure 7. Chemical Cross-linking of A3G Homomultimers

(**A**) EMSA showing complexes that resulted from assembly reactions containing 0.6 μM of AlexaFluor® 647 labeled 30 nt substrate ssDNA and either 0.11 μM (left) or 1.75 μM (right) of A3G. These EMSA gel lane containing the C1 and C3 complexes of A3G with ssDNA (A3G homodimer and homotetramer, respectively) were excised in their entirety and incubated with DTBP cross-linking reagent. EMSA complexes were visualized by Typhoon Trio imager scanning densitometry by virtue of their labeled ssDNA content prior to cross-linking. Cross-linked complexes were resolved in the second dimension by denaturing SDS-PAGE and western blotted with A3G specific antibody to reveal A3G and its composite molecular mass in cross-linked C1 (**B**) and C3 (**C**) complexes. The image of the EMSA, based on fluorescently tagged ssDNA, has been graphically inserted at the top of the 2D-gel western image for reference and alignment of the EMSA with A3G detected by western blotting. Arrows on the top and side of the image show the directions of electrophoresis of the EMSA and SDS-PAGE, respectively. Arrows heads in B and C point to the migration of monomeric A3G or cross-linked multimers of A3G. The electrophoretic migration in the second dimension denaturing PAGE was calibrated using purified A3G and molecular mass standard proteins run on the same gel.

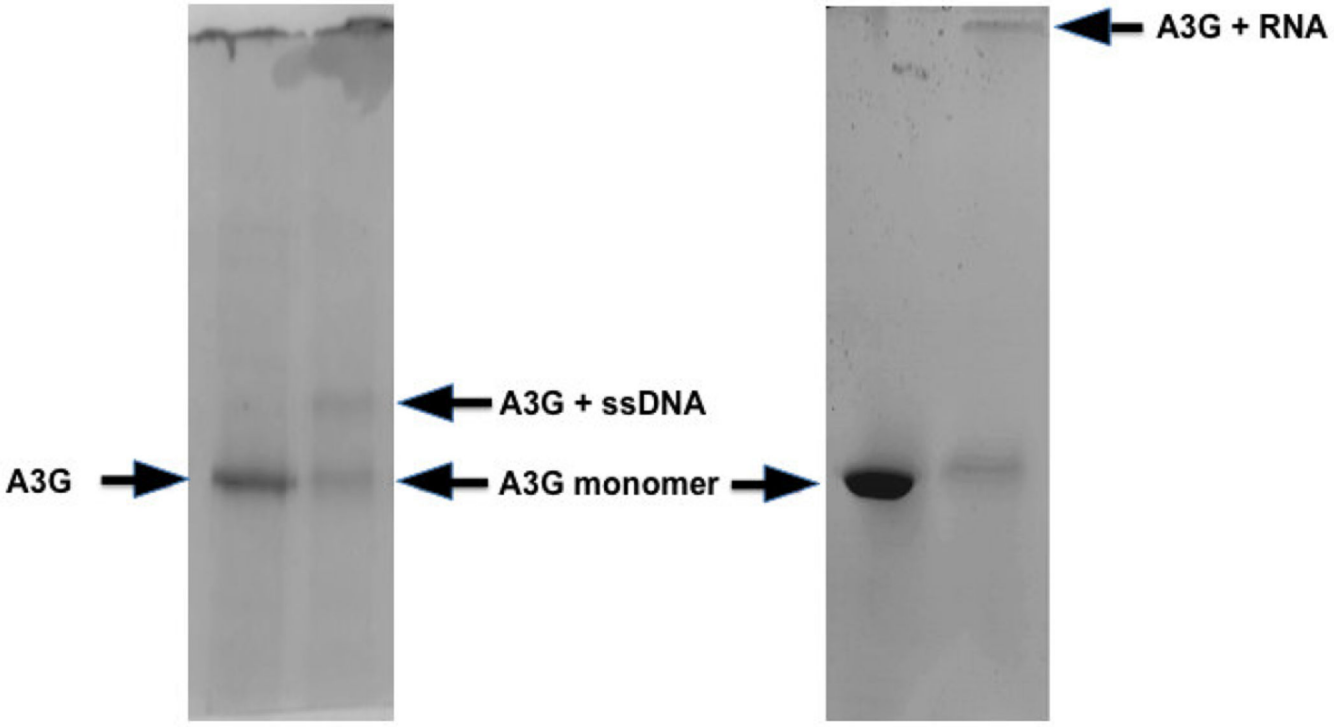

Fig. 8. SDS-PAGE separation of UV-induced cross-linked A3G to 25 nt ssDNA and RNA

A3G without UV cross-linkng or following UV crosslinking to 25 nt ssDNA (left panel) or 25 nt RNA (right panel) were resolved on 12% SDS PAGE and stained with Coomassie blue. The positions of monomeric and cross-linked A3Gs are indicated with arrows. Note that while A3G:ssDNA cross-linked product mobility is 5–6 kDa slower than monomeric A3G, A3G:RNA product is typically of much higher molecular mass.

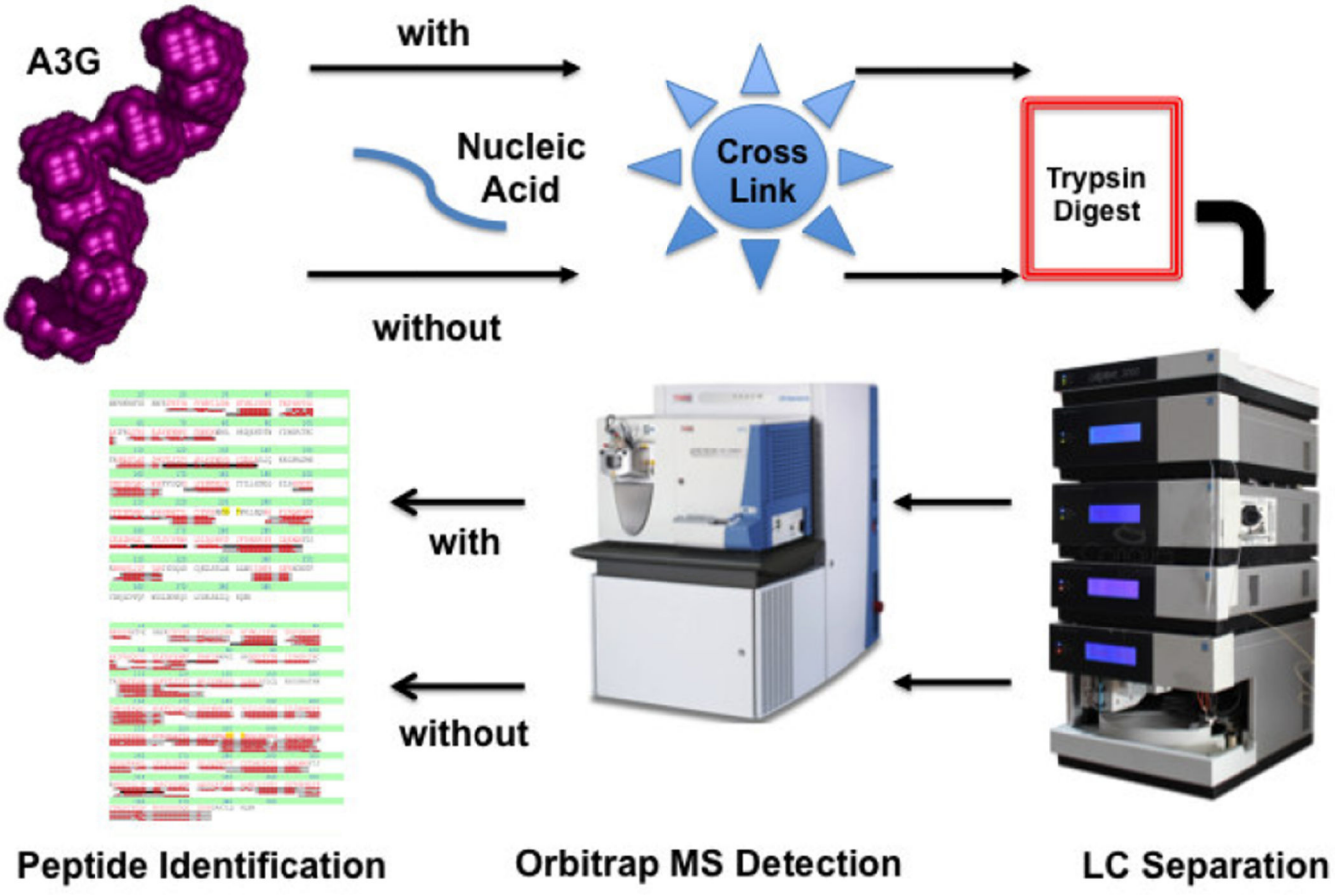

Fig 9. Proteomics approach to identify A3G peptides bound to ssDNA and RNA

This comparative approach includes cross-linking of nucleic acids to A3G and processing the samples as described in section 3.3 followed by identification of A3G peptides by MS and comparison of the peptide coverage for cross-linked ('with') and uncross-linked ('without') samples.

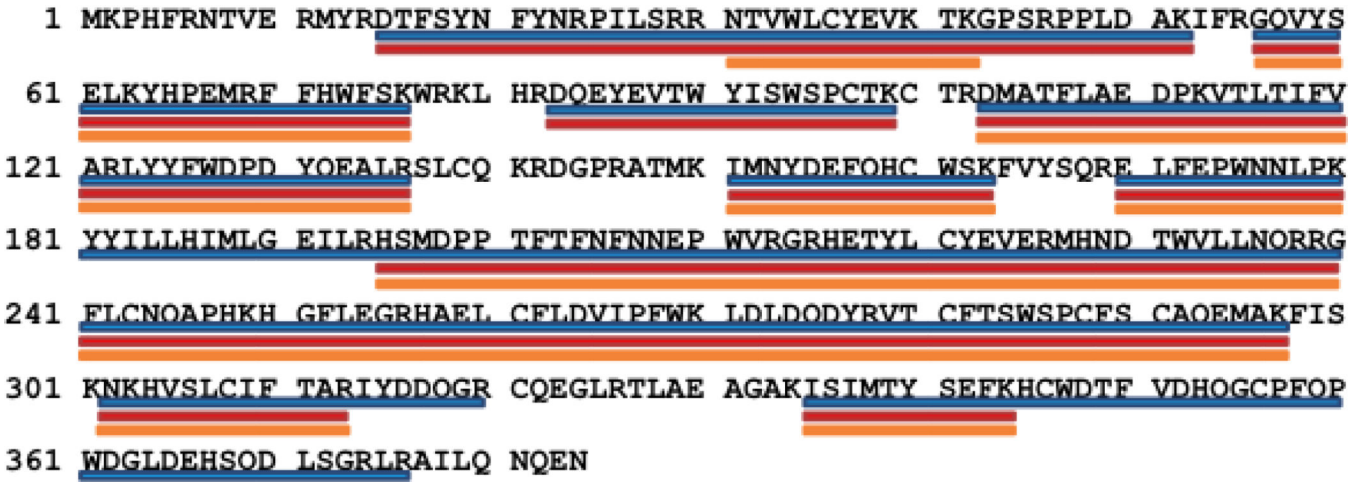

Fig 10. Mass Spectrometry Identification of Similar and Different A3G Peptides Bound to 25 nt ssDNA and RNA

Trypric peptides in uncross-linked, ssDNA cross-linked or RNA cross-linked samples initially prepared from gel bands as in Figure 9 are underlined with blue, red and orange colors, respectively. Three A3G peptides were reproducibly “missing” in cross-linked samples: aa 181–194, aa 314–320 and aa 345–374, therefore they are suggested to be involved in ssDNA and RNA binding. The RNA cross-linked sample also showed a unique set of peptides, aa 14–29, aa 41–52 and aa 83–99. It should be noted that regardless of which protein samples were analyzed. The following tryptic peptides were not detected by MS in any sample: aa 1–14, 53–55, 77–82, 99–102, 137–150, 164–169, 298–301, 321–334, 375–384. These tryptic peptides may have been too short and/or their signals may have been too low for MS detection. We cannot exclude that some of these peptides contained amino acid residues with posttranslational modifications and were missed because we performed our MASCOT searches set for unmodified peptides. However, peptide coverage that we typically observe in A3G control sample is reproducibly in a range of 76–77% or more. MASCOT search parameters for data from LTQ Obitrap XL included: trypsin as an enzyme; 3 missed cleavages (due to potential cross-linking to arginine or lysine residues that may cause trypsin to not cut protein at these positions); 0.5 Da for MS tolerance and MS/MS data; 1 for #13C, +1/+2;/+3 for charge state; acceptance criteria of minimum 1 peptide greater than identity score; minimum score of 15; and False Discovery Rate less than 5%. Also, A3G matched MS spectra were manually validated to ensure that the highest intensity peaks in the spectra belongs to matched peptides.

Table 1

Composition of Native the Gel

	mL	Final Concentration
30 % Acrylamide	3.4	5%
1× TBE	10	0.5×
50% glycerol	1.6	4%
Ammonium persulfate	0.2	0.1%
TEMED	0.02	0.1%
Water	4.85	N/A