BIOCHEMISTRY
including biophysical chemistry & molecular biology

Article

pubs.acs.org/biochemistry

Nanoscale Structure and Dynamics of ABOBEC3G Complexes with Single-Stranded DNA

Luda S. Shlyakhtenko,[†] Alexander Y. Lushnikov,[†] Atsushi Miyagi,[†] Ming Li,[‡] Reuben S. Harris,[‡] and Yuri L. Lyubchenko*,[†]

[†]Department of Pharmaceutical Sciences, College of Pharmacy, University of Nebraska Medical Center, 986025 Nebraska Medical Center, Omaha, Nebraska 68198-6025, United States

[‡]Department of Biochemistry, Molecular Biology, and Biophysics, Institute for Molecular Virology, Center for Genome Engineering, Masonic Cancer Center, University of Minnesota, 321 Church Street SE, 6-155 Jackson Hall, Minneapolis, Minnesota 55455, United States

Supporting Information

ABSTRACT: The DNA cytosine deaminase APOBEC3G (A3G) is capable of blocking retrovirus replication by editing viral cDNA and impairing reverse transcription. However, the biophysical details of this host−pathogen interaction are unclear. We applied atomic force microscopy (AFM) and hybrid DNA substrates to investigate properties of A3G bound to single-stranded DNA (ssDNA). Hybrid DNA substrates included ssDNA with 5′ or 3′ ends attached to DNA duplexes (tail-DNA) and gap-DNA substrates, in which ssDNA is flanked by two double-stranded fragments. We found that A3G binds with similar efficiency to the 5′ and 3′ substrates, suggesting that ssDNA polarity is not an important factor. Additionally, we observed that A3G binds the single-stranded region of the gap-DNA substrates with the same efficiency as tail-DNA. These results demonstrate that single-stranded DNA ends are not needed for A3G binding. The protein stoichiometry does not depend on the ssDNA substrate type, but the ssDNA length modulates the stoichiometry of A3G in the complex. We applied single-molecule high-speed AFM to directly visualize the dynamics of A3G in the complexes. We were able to visualize A3G sliding and protein association−dissociation events. During sliding, A3G translocated over a 69-nucleotide ssDNA segment in <1 s. Association−dissociation events were more complex, as dimeric A3G could dissociate from the template as a whole or undergo a two-step process with monomers capable of sequential dissociation. We conclude that A3G monomers, dimers, and higher-order oligomers can bind ssDNA substrates in a manner independent of strand polarity and availability of free ssDNA ends.

The human APOBEC3G (A3G) protein belongs to a family of DNA cytosine deaminases.[1,2] Several of these proteins, including A3G, have the capacity to block HIV-1 replication by editing viral cDNA and impairing cDNA synthesis (ref 3 and references cited therein). A3G has two domains: an N-terminal pseudocatalytic domain (PCD) and a C-terminal catalytic domain (CD).[4−6] Nuclear magnetic resonance and X-ray crystal structures have been obtained for the CD, but thus far, the PCD and the full-length apoenzyme have resisted structural interrogation.[7−12]

Several genetic and biochemical studies have concluded that the CD alone is responsible for deamination activity,[4−6,13−15] whereas the PCD is more important for binding RNA or single-stranded DNA (ssDNA), mediating oligomerization, and interacting with the natural A3G antagonist HIV-1 Vif.[4,5] In spite of the seemingly different functions of the CD and PCD in ref 6, it was suggested that the PCD indirectly contributes to catalysis by mediating A3G scanning of ssDNA. Moreover, it was reported in ref 16 that the antiviral function of A3G cannot be attributed solely to its cytosine deaminase function. Also in ref 17, using point mutations to the C- and N-terminal active sites, it was suggested that the entire protein structure is necessary for its antiviral function. Systematic biochemical studies[5,16,18,19] indicated that the deaminase reaction of A3G protein occurs predominantly in the 3′ to 5′ direction. Therefore, one asks the questions of whether the polarity of the DNA substrate is important for complex formation and how the availability of free 5′ or 3′ ssDNA ends might affect formation of the ssDNA−protein complex. The requirement of the DNA binding domain for enzyme processivity and polarity was proposed in ref 5, but so far, no evidence supporting this model has been obtained. In this study, we address the possible roles of DNA substrate polarity and free ends in formation of the A3G−ssDNA complex.

Atomic force microscopy (AFM) has already been instrumental in the characterization of complexes of A3G with ssDNA, specifically in the evaluation of protein stoichiometry.[5,18,20] The analysis of the AFM data led the authors to conclude that A3G, in complexes with ssDNA substrates, exists as dimers, tetramers, and higher-order oligomers, defining the directional deamination reaction of the enzyme.[5,18] They also showed that cations contribute to protein oligomerization but play a secondary role. However, the use of ssDNA substrates in these studies complicated the ability

Received: June 4, 2012
Revised: July 13, 2012
Published: July 18, 2012

   dx.doi.org/10.1021/bi300733d | *Biochemistry* 2012, 51, 6432−6440

to distinguish between complexes and free protein. Both the complexes and free proteins are morphologically similar and appear on the AFM images as round-shaped, globular features. The ambiguity issue was resolved in our recent paper,[20] in which AFM was used to analyze A3G bound to a hybrid ssDNA substrate. In this approach, ssDNA (69 nucleotides) is attached to the double-stranded DNA fragment, which acts as a built-in ruler. Protein bound to such a substrate appears on the AFM images as a globular feature (blob), located at the end of the DNA duplex. This hybrid substrate allowed us to use volume measurements to determine A3G stoichiometry in both free and ssDNA-bound states. We observed that A3G protein binds ssDNA mostly as a dimer. The A3G stoichiometry increased slightly with the addition of Mg^{2+}, but dimers remained predominate when Mg^{2+} was depleted.

We developed the hybrid DNA approach further to characterize the ssDNA binding properties of A3G. We designed 3′ and 5′ ssDNA ends (3′-tail-DNA and 5′-tail-DNA, respectively), attached to the dsDNA fragments, to understand whether the polarity of the single-stranded ends affects A3G binding. In addition, we designed the substrates such that the ssDNA region is located between DNA duplexes (gap-DNA), to see how critical the ssDNA end is for formation of the A3G complex. Note that DNA molecules of such types are formed during DNA replication and repair and may be substrates for deamination by one or more A3 family members.[21−23] Additionally, we made tail-DNA substrates with different lengths to answer the question of whether the length of the ssDNA contributes to A3G oligomerization. Using high-speed AFM (HS-AFM), we analyzed the dynamics of A3G in complexes with these hybrid DNAs. These studies indicate that both A3G monomers and dimers form stable complexes with ssDNA. We also directly observed dissociation−association events for A3G, suggesting that this protein utilizes a three-dimensional mechanism for the site search. Additionally, we visualized one-dimensional movement of A3G over ssDNA substrates. These findings suggest how A3G might engage ssDNA substrates inside retrovirus cores during reverse transcription.

MATERIALS AND METHODS

APOBEC3G Preparation. Details for preparing human A3G-myc-His have been reported previously.[12,20] A representative gel image and activity data are shown in Figure S5 of the Supporting Information.

Preparation of Hybrid DNA Substrates. *Preparation of the Hybrid dsDNA with Single-Stranded Tails (tail-DNA).* To construct the dsDNA with single-stranded tails attached (tail-DNA), we followed the procedures described in detail in refs 20 and 24. Figure S1 of the Supporting Information shows the schematic for preparation of tail-DNA. First, the polymerase chain reaction product digested with restriction enzyme was extracted and purified from agarose gels. Second, the synthetic oligonucleotide annealed with the adapter was ligated with the restriction fragment. Third, the ligated mixture was separated on an agarose gel, extracted, and purified using the Qiagen gel extraction kit. The final product was resuspended in 10 mM Tris-HCl (pH 7.4) and 1 mM EDTA (TE buffer). The conditions for the annealing and ligation procedures are described in detail in the Supporting Information and this section. SM1 and SM2 and the sequences are summarized in Table ST1 of the Supporting Information.

Preparation of Hybrid DNA with a Single-Stranded Region in the Middle (gap-DNA). Figure S2 of the Supporting Information shows the schematic for the preparation of hybrid dsDNA with a single-stranded DNA region in the middle. First, the two 5′- and 3′-tail-DNAs were mixed in a 1:1 ratio and annealed with the bridge oligos (see mixing−annealing steps). Then, the annealed product was ligated overnight at 16 °C. To remove the bridge oligos, the product after ligation was heated to 70 °C for 5 min and immediately put into ice. The final product, separated from the bridge oligos in an agarose gel, was extracted, purified, and resuspended in TE buffer. The detailed conditions for obtaining hybrid gap-DNA are described in the Supporting Information and this section. SM3 and all sequences are provided in Table ST2 of the Supporting Information.

Preparation of A3G−DNA Complexes. A3G was mixed with different DNA substrates at a 5:1 protein:DNA molar ratio in a reaction buffer containing 50 mM HEPES (pH 7.5), 100 mM NaCl, 5 mM Mg^{2+}, and 1 mM DTT, incubated at 37 °C for 10 min, and then purified using a Montage UFC spin column, as described in detail in ref 20.

Sample Preparation and AFM Imaging of Dry Samples. For AFM sample preparation, the 1-(3-aminopropyl)silatrane (APS)-functionalized mica was used as described in refs 24 and 25. Five microliters of the sample was deposited on APS mica for 2 min, rinsed with deionized water, and dried with Ar gas. Images were acquired in tapping mode in air using the Multimode Nanoscope IV system from Bruker (Santa Barbara, CA). Silicon probes with a spring constant of 42 N/m at resonance frequencies between 310 and 340 kHz were used.

High-Speed AFM in Liquid. The mica (1.5 mm × 1.5 mm) was glued to a glass cylinder attached to the AFM scanner, cleaved, and treated with an APS solution as described above. A drop of the A3G−hybrid DNA complex (2 μL), prepared as described in the section on preparation of A3G−hybrid DNA complexes, was deposited on the mica surface for 5 min. The scanner with the sample was immersed in 200 μL of a buffer solution. All images were taken in the buffer without drying the sample.

The HS-AFM images were acquired by using the HS-AFM instrument developed by the Ando group (Kanazawa University, Kanazawa, Japan, and distributed by RIBM Co., Tsukuba, Japan).[26] The data were acquired by operating the instrument in tapping mode in liquid. Silicon nitride AFM probes (BLAC10EGS, Olympus) were etched using the electron beam deposition (EBD) procedure. The spring constant of the AFM probes was between 0.1 and 0.2 N/m, with a resonance frequency between 400 and 1000 kHz in water. Continued scanning over the selected area (200 nm × 200 nm) was performed to follow the dynamics of A3G−DNA complexes. The scan rate varied between 720 and 990 ms/frame.

Data Analysis. For each protein−ssDNA complex, we measured the length of each double-stranded hybrid DNA arm. Protein volumes were calculated from measured AFM image heights and diameters using the cross-section option in Femtoscan Online (Advanced Technologies Center, Moscow, Russia), which is similar to the procedure described previously.[20] The conversion coefficient, as described in refs 20, 27, and 28, calculated from the dependence of the protein volume on the molecular weight (Figure S6 of the Supporting Information) was used to calculate the stoichiometry of

dx.doi.org/10.1021/bi300733d | Biochemistry 2012, 51, 6432−6440

proteins within complexes. To minimize the unavoidable tip convolution effect, the heights and widths of DNA duplexes were measured on the same images and the data with similar DNA parameters were used for the volume measurements. The data for measurements of protein volume and length of DNA arms were summarized into histograms using Origin 6.0 (Originlab, Northampton, MA).

RESULTS

A3G Binds Equivalently to 5′ and 3′ Ending ssDNAs. A3G, as a retroviral ssDNA deaminase, appears to have 3′ to 5′ polarity.[5] Therefore, we first sought to address the question of whether polarity affects binding of A3G to ssDNA, by testing the formation of the complex between A3G and two 69-nucleotide tail-DNA substrates with 5′ end and 3′ end polarities. Because of the differences in the lengths of the dsDNA arms attached to the 69ss-tail-DNA (3′-69ss-436bp and 5′-69ss-231bp), these different substrates could be visualized with AFM at the same image. The use of the 1:1 mixture of both DNA substrates eliminates potential concerns about the variability of conditions if the complexes are prepared with each substrate separately. Figure 1A shows a typical AFM image obtained for such a mixture. In this image, the complexes with 5′-tail-DNA are marked with (1) and 3′-tail-DNA complexes are marked with (2). Complexes formed with similar yields, 88% for A3G complexed with 5′-69-tail-DNA and 87% for A3G complexed with 3′-69ss-tail-DNA. Similar designs with shorter lengths of ssDNA were made, and the yields were 70% for 5′-42ss-tail-DNA, 63% for 3′-42ss-tail-DNA, 34% for 5′-27ss-tail-DNA, and 35% for 3′-27ss-tail-DNA, indicating that the polarity of the single-stranded regions has no effect on the efficiency of complex formation.

To determine the stoichiometry of the complexes formed between A3G and 3′- and 5′-tail-DNA, we calculated the volume of the protein in the complex. The volume distribution of complexes compiled from the analysis of more than 150 molecules is shown in panels B and C of Figure 1. The distributions for 3′-69ss and 5′-69ss-tail-DNA are similar, with the maxima on both distributions at approximately 120 nm^3. According to the conversion coefficient determined previously,[20] this value corresponds to A3G dimers. As in our previous work,[20] monomers, trimers, and tetramers are also present in each distribution, but the majority of the A3G–ssDNA complexes correspond to dimers.

A3G Binds Equivalently to Terminal and Internal ssDNA Regions. To understand whether the single-stranded end of the DNA is required for complex formation and if an internal ssDNA region might also bind to A3G, we designed a substrate with a 69-nucleotide ssDNA region flanked by double-stranded DNA arms (69ss-gap-DNA substrate). A3G complexes with this substrate were prepared in parallel with the 69ss-tail-DNA substrate and imaged using AFM. Typical images for complexes made with the 69ss-tail-DNA and 69ss-gap-DNA substrates are shown in panels A and B of Figure 2, respectively. A3G binds the substrate at the end of 69ss-tail-DNA (plate A) or inside of the 69ss-gap-DNA substrate (plate B). Galleries of more AFM images for complexes of A3G with tail-DNA and gap-DNA are shown in panels C and D of Figure 2, respectively. The bright protein blobs representing complexes of A3G and gap-DNA are located approximately one-third of the distance from the closest end. The length measurements confirmed that the position of A3G protein corresponds to the location of the single-stranded region within

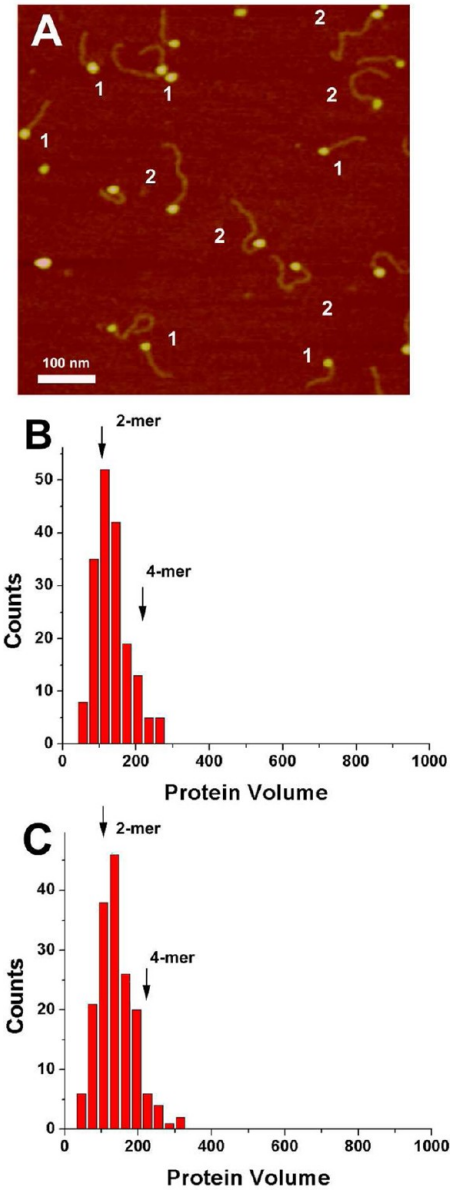

Figure 1. AFM image (A) and volume distribution histograms (B and C) corresponding to (B) a 5′-69ss-tail-DNA (N = 180, yield of the complexes) and (C) a 3′-69ss-tail-DNA (N = 170, yield of the complexes). In panel A, 1 denotes complexes with a 5′-69ss-tail-DNA (231 bp duplex) while 2 denotes 3′-69ss-tail-DNA (436 bp duplex). Arrows in panels B and C indicate the positions of dimers (2-mers) and tetramers (4-mers).

the gap-DNA. We also measured the yields of complexes in these two preparations. They were 50 and 40% for A3G complexes with the tail-DNA and gap-DNA, respectively, suggesting that the end of the single-stranded region is not essential for complex formation.

There is some variability in A3G complex size in the AFM image gallery. We measured the protein volumes in multiple images to estimate A3G stoichiometry, as performed in ref 20. On the basis of these measurements, in Figure 2A we marked dimers as 2 and trimers as 3. Similar heterogeneity in the protein sizes was observed for the complexes with the gap-DNA substrate (Figure 2B). In this figure, the complex as a dimer is marked with a 2 and the monomer with a 1. The protein stoichiometries determined from the volume measurements for

dx.doi.org/10.1021/bi300733d | Biochemistry 2012, 51, 6432–6440

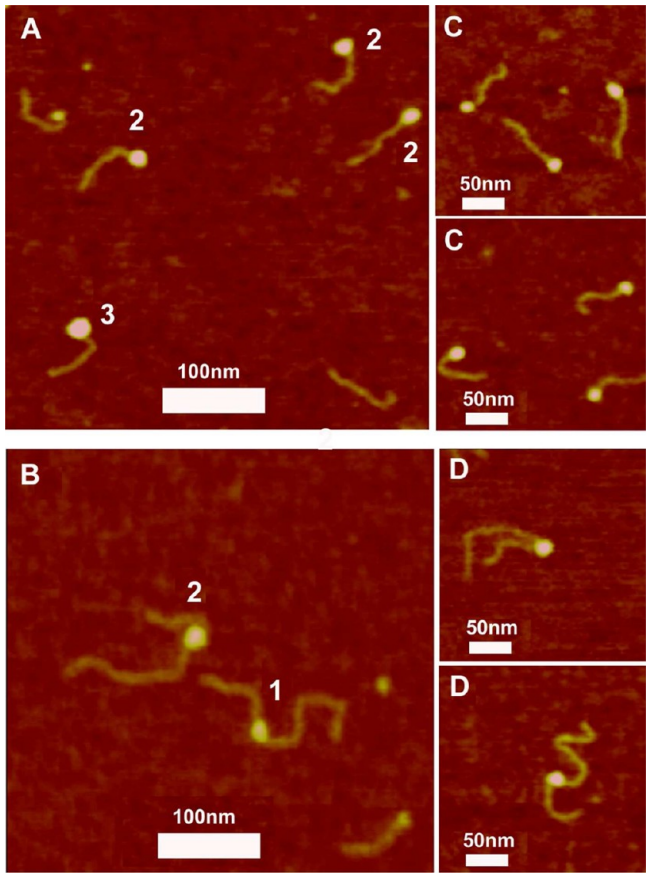

Figure 2. AFM images of complexes of A3G with tail-DNA and gap-DNA substrates containing a 69-nucleotide ssDNA. (A) AFM image of the A3G complex with 69ss-tail-DNA. Molecule labels: 2, dimers; 3, trimers. (B) AFM image of the A3G complex with 69ss-gap-DNA. Molecule labels: 1, monomer; 2, dimer. (C) Gallery of AFM images of complexes of A3G with the 69ss-tail-DNA. (D) Gallery of AFM images of complexes of A3G with the 69ss-gap-DNA.

the gap-DNA and tail-DNA complexes are shown in panels A and B of Figure 3, respectively. The histograms for tail-DNA

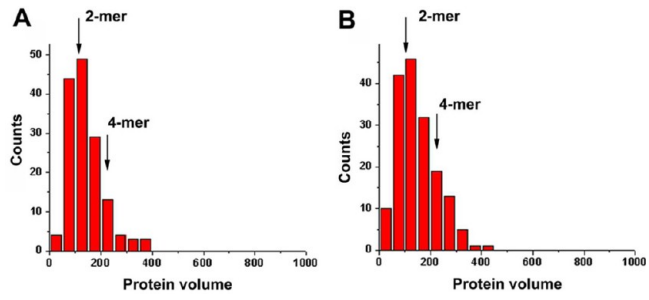

Figure 3. Histograms for the A3G volume measurements. (A) A3G in complexes with the 69ss-tail-DNA (N = 152). (B) A3G in complexes with the 69ss-gap-DNA (N = 170). The arrows show the positions of the A3G dimers and tetramers.

and gap-DNA are very close, with the maxima on both distributions at approximately 120 nm^3, which corresponds to an A3G dimer (indicated with an arrow in Figure 3A,B). A3G monomers and higher-order multimers are also present in the distributions.

Dependence of the Mode of Binding of A3G on the Length of ssDNA. To determine the effect of ssDNA length on A3G protein oligomerization, we examined A3G protein in complexes with 42-, 27-, and 18-nucleotide tail-DNA. The data for the protein volume distribution for the A3G complexes with 27 nucleotides are shown in Figure 4. Similar to longer tails, the

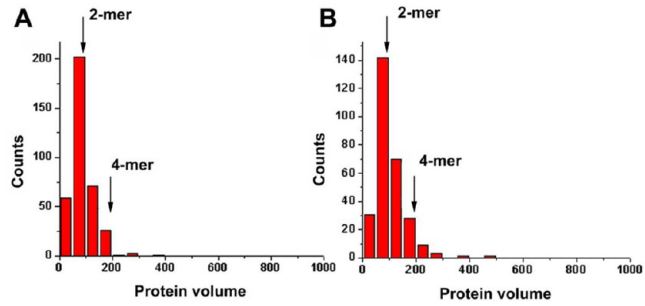

Figure 4. Histograms for the A3G volume measurements in complexes with the 27ss-tail-DNA: (A) 3′-tail-DNA (N = 362) and (B) 5′-tail-DNA (N = 284). The arrows point to the A3G dimers and tetramers. The yields of the complexes for 5′-tail-DNA and 3′-tail-DNA are 34 and 35%, respectively.

histogram for 27ss-tail-DNA complexes with A3G has the maximal distribution around dimers with traces of tetramers present. Similar observations were obtained for A3G complexes with 42ss-tail-DNA (see Figure S3A,B of the Supporting Information).

However, the A3G volume distribution for 18ss-tail-DNA complexes was different with higher frequencies of monomers: the amount of monomers was approximately equal to the amount of dimers (shown with yellow arrows in Figure 5A,B).

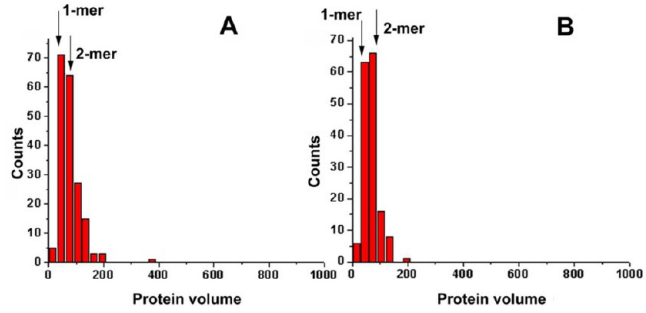

Figure 5. Histograms for the A3G volume measurements in complexes with the 180-nucleotide tail-DNA: (A) 3′-tail-DNA (N = 150) and (B) 5′-tail-DNA (N = 160). The arrows show the positions of the A3G monomers and dimers.

Moreover, there was a decrease in the number of higher-order oligomeric structures in the complex. These data clearly indicate that for short ssDNA substrates, such as 18ss-tail-DNA substrates, the A3G complexes were smaller with a tendency toward monomer binding. Notably, the yield of complexes also depended on ssDNA length, with the efficiency of complex formation decreasing ~3 times for the 18ss-tail-DNA compared to that for the longer 69ss-tail-DNA hybrid substrate.

HS-AFM Data for the Dynamics of A3G−ssDNA Complexes in Solution. *One-Step and Two-Step Protein Dissociation Pathways.* To follow directly the dynamics of the A3G−ssDNA complexes, we applied HS-AFM, which provides high-resolution time-lapse images of fully hydrated biological samples on a millisecond time scale (reviewed in refs 26, 29,

dx.doi.org/10.1021/bi300733d | *Biochemistry* 2012, 51, 6432−6440

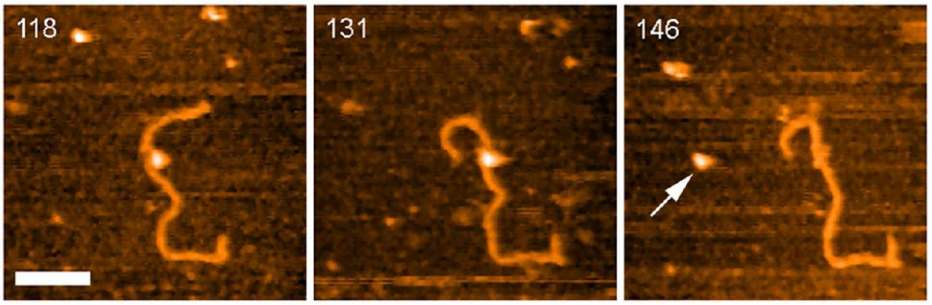

Figure 6. High-speed AFM images demonstrating A3G dissociation. Three frames from Movie S1 (Supporting Information) illustrate one-step dissociation of the A3G dimer (the arrow indicates the position of dissociated A3G). The bar is 50 nm; each frame is 720 ms.

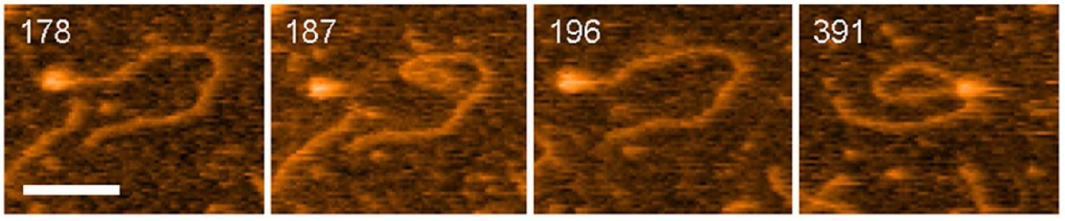

Figure 7. Selected frames demonstrating one-step dissociation followed by the formation of the complex of the A3G dimer with 27-nucleotide tail-DNA (Movie S2 of the Supporting Information). The bar is 30 nm; each frame is 720 ms.

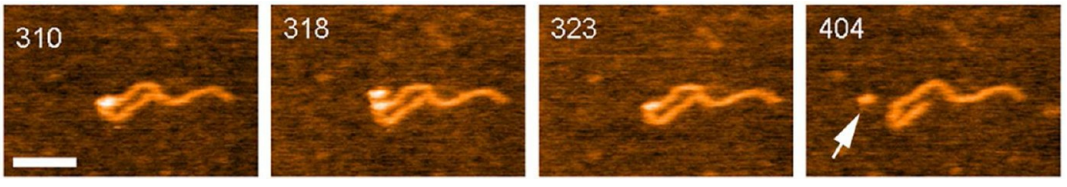

Figure 8. Two-step dissociation of A3G from the gap-DNA substrate (Movie S3 of the Supporting Information). The bar is 50 nm; each frame is 720 ms. The arrow indicates the position of dissociated A3G.

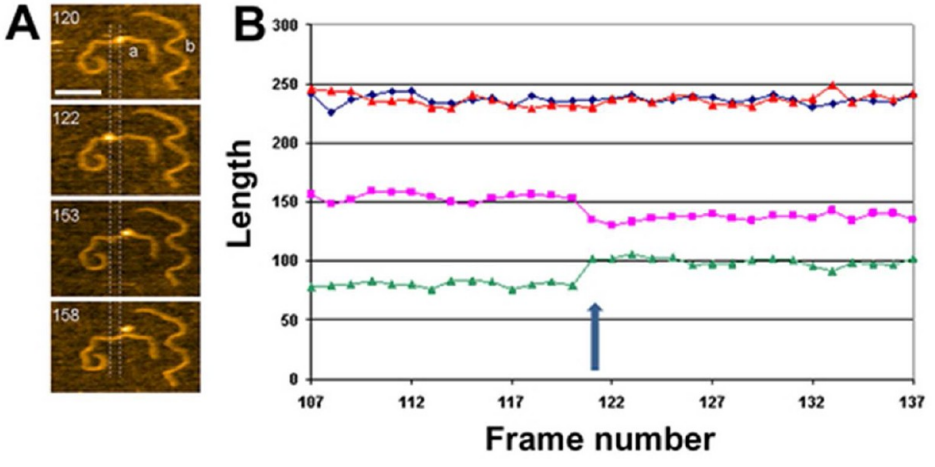

Figure 9. Sliding of A3G. (A) Selected frames for the A3G−gap-DNA complex (a) and gap-DNA (b) (Movie S4 of the Supporting Information). The bar is 50 nm; each frame is 720 ms. Vertical dotted lines show the distance between ends of DNA arms and the center of the protein. (B) DNA contour length in the complex with A3G (a, blue) and without protein (b, red). The lengths of the left (purple) and right (green) DNA arms were measured from the end to the center of the protein. The arrow points to the protein translocation event.

and 30), including protein−DNA complexes (e.g., refs 31 and 32). A3G complexes with both tail-DNA and gap-DNA substrates were analyzed with HS-AFM.

We first analyzed the dynamics of A3G bound to 69ss-gap-DNA substrates (Figure 6 shows key frames from Movie S1 of the Supporting Information). A3G in frame 118 appears to be located at approximately one-third from one end of the DNA, and this position corresponds to the position of ssDNA within the hybrid substrate. The short DNA arm changed shape in frame 131, but the protein remained bound regardless of the DNA segmental mobility. A3G then dissociated in one step in frame 146. The protein volume measurements were performed

dx.doi.org/10.1021/bi300733d | *Biochemistry* 2012, 51, 6432−6440

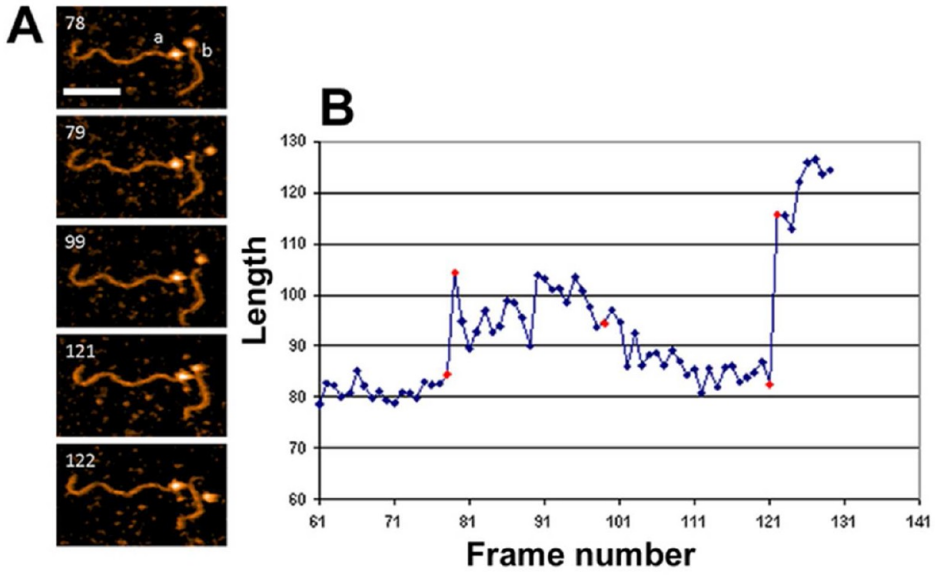

Figure 10. Protein dynamics for the end-bound complex (Movie S5 of the Supporting Information). (A) Selected frames for A3G–tail-DNA complexes. The bar is 50 nm; each frame is 720 ms. (B) Distances measured between the distant end of the DNA (molecule b) and the center of the protein for each frame. Red points on the graph correspond to the frames in panel A.

for each frame of the movie data set, and volume measurements are consistent with most substrates being bound by a dimeric form of A3G.

The dynamics of the end-bound complexes are illustrated in Figure 7 (Movie S2 of the Supporting Information). The complex shown in frame 178 dissociates after fewer than 10 frames (frame 187), but because of the DNA dynamics, it forms again (frame 196) and remains bound for a long time (almost 200 frames; frame 391) regardless of the ongoing DNA movement. Interestingly, the dissociation of the complex and its assembly after approaching the DNA occur fast, during the acquisition of one frame. According to the volume measurements, these dynamics mostly represent the dimeric form of A3G.

Previous results illustrated that A3G dimers can dissociate from and reassociate with ssDNA substrates. However, each protein monomer contains a ssDNA-binding domain, so theoretically each A3G monomer alone should be capable of intrinsically binding to ssDNA. HS-AFM was instrumental in identifying both binding modes. Figure 8 and Movie S3 of the Supporting Information demonstrate the dynamics of A3G bound to a 69ss-tail-DNA substrate (Movie S3). Frame 310 shows an A3G dimer bound to ssDNA. Frame 318 shows that indeed the protein dissociates into two blobs with one going away and another remaining bound to the same substrate. The protein remains bound (frame 323) and later dissociates from the substrate (frame 404) and goes away from the observation area. Volume measurements are consistent with the dissociated protein being an A3G monomer. This process was termed two-step dissociation.

A3G One-Dimensional Movement. Next we analyzed the dynamics of A3G bound to ssDNA substrates. It was proposed in refs 16 and 33 that A3G utilizes various mechanisms, including one-dimensional diffusion or sliding. The high temporal resolution of HS-AFM is sufficient for testing the sliding model. Frame 120 in Figure 9A shows a complex with the A3G monomer occupying the ssDNA gap in the hybrid DNA substrate. The protein shifts its position rather shortly (frame 122) and soon moves to the right, overshooting the original position (frame 153) followed by dissociation [frame 158 (see Movie S4 of the Supporting Information)]. To characterize this process quantitatively, we measured the contour lengths of the left and right DNA flanks (from the end of the DNA to the middle of the protein blob). The graphs (pink and green lines) show that both arms do not undergo changes in length until frame 120 (frame 120 indicated by the arrow on the graph). Between frames 120 and 122, the length of the left arm (pink line and squares) decreases by ~20 nm, whereas the length of the right arm (green triangles and line) increases by the same value. Interestingly, the overall contour length of the DNA molecule measured from the left end to the right end (blue line and diamonds) remains constant during the entire observation period, including the time interval between frames 120 and 122. This suggests that the change in the arm's length is due to the protein sliding over the ssDNA without a measurable change in ssDNA length. We estimated the protein sliding range by measuring the displacement of the protein center to avoid the contribution of the tip convolution effect. According to the length measurements, A3G can slide over 30 nm, and this range is consistent with the 69-nucleotide length of the ssDNA segment.

Another translocation event for a tail-DNA substrate is depicted in Figure 10A (Movie S5 of the Supporting Information). Frame 78 shows two end-bound complexes with rather stable complex "a" and dynamic complex "b". As one can see from frame 79, the protein dissociates from the ssDNA (complex b) but remains in the proximity of the end. Moreover, over time, the distance between protein and ssDNA changes, becoming shorter (frame 99) and disappearing in frame 121. Shortly after that, the protein fully dissociates and moves far from the DNA (frame 122). We measured the contour length along the ssDNA substrate up to the center of the protein, and the time dependence distance is shown in Figure 10B. The graph shows that the protein moves away after frame 78 but remained tethered to ssDNA during a long period of time. The average distance between the protein and DNA substrate is ~20 nm, which is close to that of an almost fully extended 27-nucleotide single-stranded region. The volume

dx.doi.org/10.1021/bi300733d | *Biochemistry* 2012, 51, 6432–6440

measurements show A3G protein as a monomer in complex b and a dimer in complex a.

DISCUSSION

Using a hybrid DNA approach, we have demonstrated that A3G binds 5′- and 3′-terminated ssDNA substrates with similar efficiency. These findings indicate that the polarity of the ssDNA substrate is not important for A3G binding. Additionally, using gap-DNA substrates, we have shown that the single-stranded DNA ends are dispensable for A3G binding, as the complexes between A3G and gap-DNA are formed with the same efficiency as complexes between A3G and the substrates containing 5′- or 3′-tail-DNA. Our data also show that under chosen conditions, namely the concentration of A3G in the reaction, dimeric A3G is the predominant form binding to ssDNA. The contribution of monomers and higher-order oligomers clearly depends on the lengths of the DNA substrates: the shorter the ssDNA region, the lower the fraction of higher-order oligomers.

A3G has two functionally distinct subunits, a C-terminal catalytic domain (CD) and an N-terminal pseudocatalytic domain (PCD). The latter is more positively charged under physiological conditions and is thought to function as the substrate-binding domain. It has been shown in refs 5, 16, and 18 that A3G catalyzes deamination on ssDNA with a 3′ → 5′ polarity, and it was further hypothesized in ref 5 that A3G binding should have the same polarity. Our data do not support this hypothesis. The DNA substrates with free 3′ and 5′ ends form complexes with similar efficiencies. Experiments with the length of ssDNA, which varied between 69 and 18 nucleotides, did not reveal any preferences for the polarity of ssDNA ends. Notably, in ref 16, the sequences of all ssDNAs in the hybrid substrates contained a 5′-AAA-CCC-AAA-3′ motif that supports maximal A3G deaminase activity. Therefore, if the DNA binding and deamination properties of A3G are synchronized, the polarity effect should be strongest for these sequences.

According to the model suggested in ref 5, monomeric A3G can have two orientations relative to the preferred 5′-CCC-3′ substrate sequence and only one orientation provides the active complex. To reconcile the strong polarity effect of the deaminase activity of A3G with the lack of such directionality for binding to the DNA substrate, we propose a dynamic model in which the protein flips between the two orientations, performing deamination when it adopts the active orientation. This model is consistent with the reported processive deamination activity of A3G.[5,16,18]

We also showed that A3G does not need ssDNA ends for complex formation. A3G binds with the same efficiency to DNA substrates without ends (gap-DNA) as compared to the tail-DNA substrates. For all designs, no binding to duplex DNA regions was observed. Overall, these findings demonstrate a very high specificity of A3G for the ssDNA, concordant with prior studies.[20,34−37] A major difference between ssDNA and dsDNA is flexibility. The high flexibility of ssDNA substrates may facilitate swapping between the active and nonactive orientations of A3G by transient dissociation−association events.

Our time-lapse HS-AFM data have provided additional insights into A3G dynamics. A3G spontaneously dissociates from ssDNA, suggesting that the protein−DNA complex is quite dynamic. We did not observe stronger stability for complexes formed by A3G dimers compared to that for the monomers, suggesting that dimerization is not required for complex stability. The dimer can dissociate from the DNA as a whole (one-step dissociation pathway) or sequentially in which each monomer can dissociate independently (two-step dissociation pathway). Notably, the remaining monomer can occupy the binding site for a long time, suggesting that the A3G monomer−ssDNA complex may be stable. These observations lead us to conclude that A3G oligomers do not increase the stability of the A3G−DNA complex and therefore should not be favored in models for biological activity.

It was proposed that, similar to other site-specific DNA binding proteins, A3G utilizes various pathways for the recognition of the deamination sites.[16,33] Time-lapse images allowed direct observations of jumping [three-dimensional (3D) diffusion] and sliding (one-dimensional diffusion). Although A3G dissociation events discussed above favor the 3D model for the site search mechanism, the ability to directly visualize the association step provides the justification for the 3D search pathway. The dissociation−association steps were directly observed in Figure 7 in which the protein after the dissociation returned to the same site (terminal ssDNA). Importantly, the reassembled complex was very stable and did not dissociate until the end of the observation period regardless of the long-range mobility of the DNA tag. Notably, the entire dissociation−association process was observed for the protein and DNA molecules moving nearby the mica surface, so their interaction with the surface did not prevent the site recognition process and the formation of the specific complex. We used a functionalized APS-mica to reliably visualize DNA,[25,31] and along with recent HS-AFM visualization of the enzymatic cleavage of DNA,[38] these observations provide additional credibility to the APS-mica methodology as a technique for the AFM visualization of protein−DNA complexes.

We were able to observe A3G sliding over ssDNA substrates. One of these observations is illustrated in Figure 9 in which A3G translocates without dissociation over the ssDNA region of the gap substrate. The translocation process is fast: the protein moves from one side of ssDNA to another during one scan frame, i.e., less than 1 s. Another example of the sliding process is shown in Figure 10 in which the protein slides over a terminal ssDNA segment. After the initial quick jump (∼1 s), A3G moves with small steps. The initial jump covers almost the entire ssDNA region, so the protein binds to the very end of the DNA, suggesting that few nucleotides may be sufficient for A3G binding. However, compared to the gap substrates, the analysis of the images with tail substrates is more complicated. Because of a low contrast of ssDNA on the AFM images, we were unable to distinguish protein sliding from the change of the length of ssDNA bound by the protein. Therefore, the observed change can be due to A3G sliding as well as to changes in the length of bound ssDNA. These data are consistent with A3G sliding events observed recently in a single-molecule FRET study.[39]

We also assessed how varying the length of the ssDNA substrate affected A3G binding and oligomerization. Results shown in Figures 3−5 demonstrate that the major binding mode under the chosen conditions is the dimeric form of A3G, but monomers and higher-order oligomers are also found. The level of larger oligomeric A3G forms is the highest for the 69-nucleotide substrate, decreasing to a very small percentage for the 27-nucleotide substrate. In the shortest 18-nucleotide substrate, monomers and dimers were present in equal amounts (Figure 5), with no detectable tetramers. Oligomerization of

A3G is considered a major factor in defining A3G deamination activity; oligomers as large as tetramers are considered active forms of the protein, while dimers have appeared to be associated with little activity.[19] In the framework of that paper, we suggest that even for the 69-nucleotide ssDNA in which the dimers are the predominant A3G forms, only a small percentage of the complexes are active. At the same time, very few tetramers are found associated with the 27ss-tail-DNA, suggesting that the A3G activity for this substrate should be very low. However, we did not find a dramatic change in the large oligomer populations for the 42ss-tail-DNA compared with those of the 69-nucleotide substrates (Figure 3 and Figure S4 of the Supporting Information), whereas ref 5 predicted an almost 3-fold decrease in deaminase activity. Note that the sequences of our DNA substrates were very close to those in ref 5. Therefore, a correlation between A3G deaminase activity and protein oligomeric state apparently exists; however, no clear consensus about its dependence can be reached at present.

This nanoscale AFM analysis of A3G−ssDNA complexes was possible because of the use of the hybrid DNA methodology and recombinant holoenzyme purified from human HEK293 cells.[20,40] The use of dsDNA segments as molecular rulers allowed us to unambiguously identify specific complexes in the AFM images. In this study, we designed various types of hybrid DNA substrates, allowing us to test a number of hypotheses regarding the A3G−ssDNA interaction. This approach can be applied to answer additional questions about the role(s) of both ssDNA and RNA polynucleotides in the A3G oligomerization and deoligomerization processes.

■ ASSOCIATED CONTENT

Supporting Information

Additional information about the methods used, schematics explaining the design of the DNA substrate, AFM images of A3G−DNA complexes, graphs characterizing A3G−DNA complexes, and movies of animated time-lapse AFM images of the complexes. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*Department of Pharmaceutical Sciences, College of Pharmacy, COP 1012, University of Nebraska Medical Center, 986025 Nebraska Medical Center, Omaha, NE 68198-6025. Phone: (402) 559-1971. Fax: (402) 559-9543. E-mail: ylyubchenko@unmc.edu.

Funding

The work was supported by the National Institutes of Health (Grants P01-GM091743 and R01-GM096039), U.S. Department of Energy Grant DE-FG02-08ER64579, the National Science Foundation (Grant EPS-1004094), and the Nebraska Research Initiative.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We thank P. Beal, M. Kotler, H. Matsuo, J. Mueller, and C. Schiffer for critical review of the data and Lyubchenko lab members for helpful discussions.

■ ABBREVIATIONS

A3G, APOBEC3G; AFM, atomic force microscopy; APS, 1-(3-aminopropyl)silatrane; HS-AFM, high-speed AFM; ssDNA, single-stranded DNA.

■ REFERENCES

(1) Albin, J. S., and Harris, R. S. (2010) Interactions of host APOBEC3 restriction factors with HIV-1 in vivo: Implications for therapeutics. *Expert Rev. Mol. Med. 12*, e4.

(2) Malim, M. H. (2009) APOBEC proteins and intrinsic resistance to HIV-1 infection. *Philos. Trans. R. Soc. London, Ser. B 364*, 675−687.

(3) Hultquist, J. F., Lengyel, J. A., Refsland, E. W., LaRue, R. S., Lackey, L., Brown, W. L., and Harris, R. S. (2011) Human and rhesus APOBEC3D, APOBEC3F, APOBEC3G, and APOBEC3H demonstrate a conserved capacity to restrict Vif-deficient HIV-1. *J. Virol. 85*, 11220−11234.

(4) 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 Pathog. 5*, e1000330.

(5) 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.

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

(7) 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.

(8) 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.

(9) 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.

(10) 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., Kono, T., Matsuo, H., Harris, R. S., Somasundaran, M., and Schiffer, C. A. (2010) Crystal structure of the APOBEC3G catalytic domain reveals potential oligomerization interfaces. *Structure 18*, 28−38.

(11) Autore, F., Bergeron, J. R., Malim, M. H., Fraternali, F., and Huthoff, H. (2010) Rationalisation of the differences between APOBEC3G structures from crystallography and NMR studies by molecular dynamics simulations. *PLoS One 5*, e11515.

(12) Li, M., Shandilya, S. M., Carpenter, M. A., Rathore, A., Brown, W. L., Perkins, A. L., Harki, D. A., Solberg, J., Hook, D. J., Pandey, K. K., Parniak, M. A., Johnson, J. R., Krogan, N. J., Somasundaran, M., Ali, A., Schiffer, C. A., and Harris, R. S. (2012) First-in-class small molecule inhibitors of the single-strand DNA cytosine deaminase APOBEC3G. *ACS Chem. Biol. 7*, 506−517.

(13) 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.

(14) Hache, G., Liddament, M. T., and Harris, R. S. (2005) The retroviral hypermutation specificity of APOBEC3F and APOBEC3G is governed by the C-terminal DNA cytosine deaminase domain. *J. Biol. Chem. 280*, 10920−10924.

(15) Schumacher, A. J., Hache, G., Macduff, D. A., Brown, W. L., and Harris, R. S. (2008) The DNA deaminase activity of human APOBEC3G is required for Ty1, MusD, and human immunodeficiency virus type 1 restriction. *J. Virol. 82*, 2652−2660.

(16) 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.
(17) Shindo, K., Takaori-Kondo, A., Kobayashi, M., Abudu, A., Fukunaga, K., and Uchiyama, T. (2003) The enzymatic activity of CEM15/Apobec-3G is essential for the regulation of the infectivity of HIV-1 virion but not a sole determinant of its antiviral activity. *J. Biol. Chem. 278*, 44412−44416.
(18) 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.
(19) 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.
(20) Shlyakhtenko, L. S., Lushnikov, A. Y., Li, M., Lackey, L., Harris, R. S., and Lyubchenko, Y. L. (2011) Atomic force microscopy studies provide direct evidence for dimerization of the HIV restriction factor APOBEC3G. *J. Biol. Chem. 286*, 3387−3395.
(21) Roberts, S. A., Sterling, J., Thompson, C., Harris, S., Mav, D., Shah, R., Klimczak, L. J., Kryukov, G. V., Malc, E., Mieczkowski, P. A., Resnick, M. A., and Gordenin, D. A. (2012) Clustered mutations in yeast and in human cancers can arise from damaged long single-strand DNA regions. *Mol. Cell 46*, 424−435.
(22) Nowarski, R., Wilner, O. I., Cheshin, O., Shahar, O. D., Kenig, E., Baraz, L., Britan-Rosich, E., Nagler, A., Harris, R. S., Goldberg, M., Willner, I., and Kotler, M. (2012) APOBEC3G enhances lymphoma cell radioresistance by promoting cytidine deaminase-dependent DNA repair. *Blood 120*, 366−375.
(23) 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., Menzies, A., Martin, S., Leung, K., Chen, L., Leroy, C., Ramakrishna, M., Rance, R., Lau, K. W., Mudie, L. J., Varela, I., McBride, D. J., Bignell, G. R., Cooke, S. L., Shlien, A., Gamble, J., Whitmore, I., Maddison, M., Tarpey, P. S., Davies, H. R., Papaemmanuil, E., Stephens, P. J., McLaren, S., Butler, A. P., Teague, J. W., Jonsson, G., Garber, J. E., Silver, D., Miron, P., Fatima, A., Boyault, S., Langerod, A., Tutt, A., Martens, J. W., Aparicio, S. A., Borg, A., Salomon, A. V., Thomas, G., Borresen-Dale, A. L., Richardson, A. L., Neuberger, M. S., Futreal, P. A., Campbell, P. J., and Stratton, M. R. (2012) Mutational Processes Molding the Genomes of 21 Breast Cancers. *Cell 149*, 979−993.
(24) Shlyakhtenko, L. S., Lushnikov, A. Y., Miyagi, A., and Lyubchenko, Y. L. (2012) Specificity of binding of single-stranded DNA-binding protein to its target. *Biochemistry 51*, 1500−1509.
(25) Lyubchenko, Y. L., and Shlyakhtenko, L. S. (2009) AFM for analysis of structure and dynamics of DNA and protein-DNA complexes. *Methods 47*, 206−213.
(26) Uchihashi, T., and Ando, T. (2011) High-speed atomic force microscopy and biomolecular processes. *Methods Mol. Biol. 736*, 285−300.
(27) Shlyakhtenko, L. S., Gilmore, J., Kriatchko, A. N., Kumar, S., Swanson, P. C., and Lyubchenko, Y. L. (2009) Molecular mechanism underlying RAG1/RAG2 synaptic complex formation. *J. Biol. Chem. 284*, 20956−20965.
(28) Shlyakhtenko, L. S., Lushnikov, A. Y., and Lyubchenko, Y. L. (2009) Dynamics of nucleosomes revealed by time-lapse atomic force microscopy. *Biochemistry 48*, 7842−7848.
(29) Yamamoto, D., Uchihashi, T., Kodera, N., Yamashita, H., Nishikori, S., Ogura, T., Shibata, M., and Ando, T. (2010) High-speed atomic force microscopy techniques for observing dynamic biomolecular processes. *Methods Enzymol. 475*, 541−564.
(30) Ando, T. (2012) High-speed atomic force microscopy coming of age. *Nanotechnology 23*, 062001.
(31) Lyubchenko, Y. L., Shlyakhtenko, L. S., and Ando, T. (2011) Imaging of nucleic acids with atomic force microscopy. *Methods 54*, 274−283.
(32) Miyagi, A., Ando, T., and Lyubchenko, Y. L. (2011) Dynamics of nucleosomes assessed with time-lapse high-speed atomic force microscopy. *Biochemistry 50*, 7901−7908.
(33) Chelico, L., Pham, P., and Goodman, M. F. (2009) Mechanisms of APOBEC3G-catalyzed processive deamination of deoxycytidine on single-stranded DNA. *Nat. Struct. Mol. Biol. 16*, 454−455 ; author reply 455−456.
(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) 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.
(36) Nowarski, R., Britan-Rosich, E., Shiloach, T., and Kotler, M. (2008) Hypermutation by intersegmental transfer of APOBEC3G cytidine deaminase. *Nat. Struct. Mol. Biol. 15*, 1059−1066.
(37) Harris, R. S., Bishop, K. N., Sheehy, A. M., Craig, H. M., Petersen-Mahrt, S. K., Watt, I. N., Neuberger, M. S., and Malim, M. H. (2003) DNA deamination mediates innate immunity to retroviral infection. *Cell 113*, 803−809.
(38) Suzuki, Y., Gilmore, J. L., Yoshimura, S. H., Henderson, R. M., Lyubchenko, Y. L., and Takeyasu, K. (2011) Visual analysis of concerted cleavage by type IIF restriction enzyme SfiI in subsecond time region. *Biophys. J. 101*, 2992−2998.
(39) Senavirathne, G., Jaszczur, M., Auerbach, P. A., Upton, T. G., Chelico, L., Goodman, M. F., and Rueda, D. (2012) Single-stranded DNA scanning and deamination by APOBEC3G cytidine deaminase at single molecule resolution. *J. Biol. Chem. 287*, 15826−15835.
(40) Nowarski, R., Britan-Rosich, E., Shiloach, T., and Kotler, M. (2008) Hypermutation by intersegmental transfer of APOBEC3G cytidine deaminase. *Nat. Struct. Mol. Biol. 15*, 1059−1066.