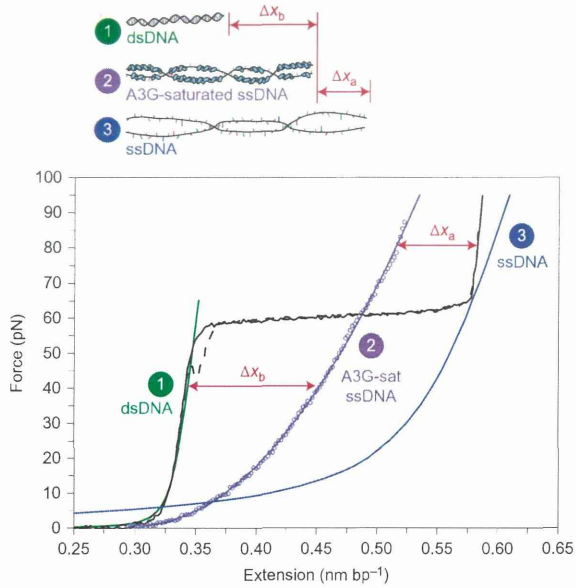


Figure 1 | Force-dependent difference in length between DNA and a saturated A3G–DNA complex allows us to measure A3G binding. Typical extension in nm per base pair (bp) (solid black) and return (dashed black line) of a single DNA molecule. At 61.0±0.5 pN, the molecule undergoes a force-induced melting transition from dsDNA (green line, Supplementary equation (1)) to ssDNA (blue line). A3G-saturated ssDNA (200 nM A3G, $t > 500$ s, data points fit to Supplementary equation (4), solid purple line) is longer than dsDNA (Δx_b, below the melting transition) and shorter than ssDNA (Δx_a, above the melting transition). A3G-saturated ssDNA is significantly shorter than ssDNA only (blue line, Supplementary equation (2)), which suggests that A3G may wrap ssDNA on binding.

peeling from the ends or by forming melting bubbles[39]. Although a stretched form of dsDNA called S-DNA may form on overstretching at monovalent salt concentrations above 150 mM, it is well-established that overstretching is force-induced melting in the presence of ssDNA binding proteins and at the ionic strength of 50 mM used in this work[39,40]. At the end of the overstretching transition, the bead movement reverses direction and the DNA tension is gradually released (Fig. 1, dashed black line). The minimal hysteresis, or mismatch between DNA extension and release, indicates that the ssDNA generated by force reanneals immediately into dsDNA during the return. At any given point along the transition, the molecule is a well-characterized combination of dsDNA and ssDNA (Supplementary equation (3))[40]. Pausing at fixed extension during the melting transition allows precise control of the fraction of ssDNA substrate available for protein binding.

The force–extension curve probes the length of the captured DNA molecule at a given force. At forces above 7 pN, ssDNA is longer than dsDNA. We exploit this force-dependent difference in length to measure the ssDNA binding properties of A3G. A3G-saturated ssDNA, obtained at a high protein concentration by first overstretching the DNA and then allowing A3G to fully bind and stabilize the DNA in its single-stranded form, is longer than dsDNA and shorter than ssDNA (Fig. 1). Accordingly, A3G bound to ssDNA increases the molecule length by Δx_b when below the melting transition, and reduces it by Δx_a when above the melting transition.

In the presence of 50 nM A3G, the extension curve follows the DNA-only curve before the melting transition, reflecting no measurable binding to dsDNA (Fig. 2a, solid line). A3G only binds after force-induced melting generates ssDNA. Based on the observed hysteresis (Fig. 2a, dashed line), most of the protein does not dissociate on DNA release, preventing the two strands from fully reannealing. A3G-bound ssDNA is longer than dsDNA (Fig. 1), so the change in length at a given force (Fig. 2a, Δx_t) describes the total fraction of A3G bound to ssDNA (f_{total}; Supplementary Fig. 1a).

A second stretch of the same molecule does not retrace the release curve, revealing that some fraction of the protein has dissociated during the 30 s incubation at zero force between stretch–release cycles (Fig. 2b). As soon as any A3G dissociates at forces below the force-induced melting transition, the two strands reanneal into dsDNA, which is shorter than A3G-bound ssDNA (Fig. 2b, Δx_f). Accordingly, the second stretch reflects the fraction of A3G that remains ssDNA-bound (f_{slow}), which allows the fraction that dissociates quickly (f_{fast}) to be quantified as $f_{fast} = f_{total} - f_{slow}$. The DNA was held at zero force for 30 s between the first release and the second stretch, but longer wait times of up to 120 min do not lead to further measurable dissociation (data not shown).

In this experiment, A3G was exposed to ssDNA for 50 s. However, A3G oligomerization observed in bulk experiments occurs on much longer timescales[37]. To measure slow binding, A3G was incubated for 250 s with ssDNA generated by force-induced melting (Fig. 2c). The DNA release curve obtained after this incubation exhibits a length increase relative to the initial release curve (Fig. 2c, Δx_i), because additional A3G binds ssDNA during incubation. This effect increases at longer incubation times (Fig. 2d), approaching the A3G-saturated ssDNA curve (Fig. 1). Fits to the DNA release curves at increasing incubation time yield $f_{total}(t)$, while the subsequent stretch (data not shown) quantifies $f_{slow}(t)$, and $f_{fast}(t)$ is the difference between the two. These measurements for 50 nM A3G are presented in Fig. 3a.

Quantitative binding model. The slow binding component increases at the expense of the fast component, suggesting that the A3G-ssDNA reaction may be modelled as a two-step process:

$$\text{A3G} + \text{ssDNA} \underset{k_{-1}}{\overset{k_1 c}{\rightleftharpoons}} (\text{A3G/ssDNA})_{\text{fast}} \underset{k_{-2}}{\overset{k_2}{\rightleftharpoons}} (\text{A3G/ssDNA})_{\text{slow}} \quad (1)$$

in which an initial bimolecular process leads to a fast complex that converts to a slow, more stable complex in the second unimolecular step. Binding rates were obtained from fits to this model (Supplementary equations (9)–(11)) at five A3G concentrations (Fig. 3b). (A3G precipitates at high concentrations[19], so experiments were calibrated using force–extension curves at known protein concentrations.) As expected, the observed fast rate k_{fast} and the on rate $k_1 c$ are both linear with A3G concentration (Fig. 3c). The bimolecular rate constant $k_1 = 1.5\ (\pm 0.1) \times 10^5\ \text{M}^{-1}\,\text{s}^{-1}$ and off rate $k_{-1} = 1.2\ (\pm 0.1) \times 10^{-2}\ \text{s}^{-1}$ are consistent with single-molecule Förster resonance energy transfer (FRET)[41] and fluorescence spectroscopy[37] measurements, considering differences in solution conditions. The observed slow rate k_{slow} saturates at high A3G concentrations (Fig. 3d), and both the on and off rates for the second, unimolecular step are concentration-independent ($k_2 = 6.7\ (\pm 0.6) \times 10^{-3}\ \text{s}^{-1}$ and $k_{-2} = 2.8\ (\pm 0.5) \times 10^{-5}\ \text{s}^{-1}$). Elementary reaction rates were obtained from the data in several different ways, and agreement of the resulting values (Supplementary Table 2) supports the binding model.

Oligomerization is responsible for slow binding. To determine whether slow binding is due to A3G oligomerization, we expressed and purified the F126A/W127A A3G mutant (A3G FW), which is severely defective in oligomerization[42]. When this mutant was incubated with ssDNA for 1,050 s, the release curve

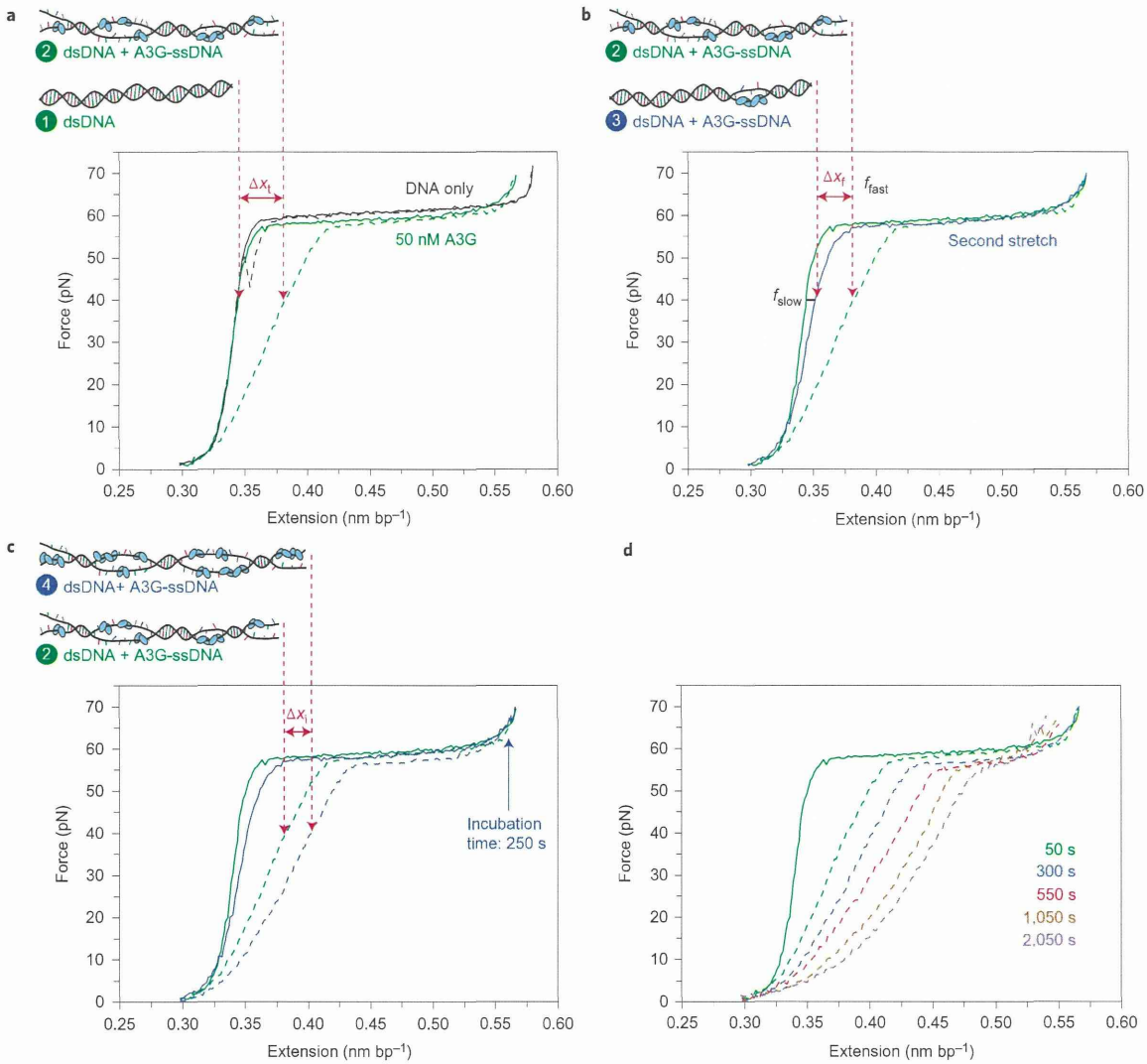


Figure 2 | Single-molecule method to measure fast and slow fractions of A3G binding. a, Without protein (black line), a single DNA molecule reanneals immediately on release, exhibiting minimal hysteresis, or mismatch between stretch (solid) and release (dashed) curves. In the presence of 50 nM A3G, the stretch curve (solid green line) follows the dsDNA-only curve, indicating negligible A3G–dsDNA binding. A3G binds the exposed ssDNA and prohibits the DNA strands from reannealing, resulting in hysteresis (dashed green line). For a given force (40 pN shown), there is a corresponding change in DNA length Δx_f between A3G-free dsDNA (left arrow, drawing 1) and partially A3G-bound ssDNA (right arrow, drawing 2). This force-dependent length change measures A3G–ssDNA binding (Supplementary Fig 1). **b**, The second stretch (solid blue line) lies between the first stretch and release curves, distinguishing the fraction of A3G that remains bound (f_{slow}) from the fraction that dissociated (f_{fast}) before the second stretch. The A3G that dissociates rapidly allows the strands to reanneal immediately into dsDNA (drawing 3), resulting in length decrease Δx_f. **c**, Pausing at fixed DNA extension after incubating ssDNA with 50 nM A3G results in additional binding (drawing 4), indicated by the corresponding length increase Δx_i measured during DNA release. **d**, A3G binding increases with total exposure time to ssDNA (dashed lines).

exhibited minimal hysteresis (Fig. 4a), and all the bound protein dissociated before the subsequent stretch (Fig. 4b). A direct comparison of the hysteresis observed for wild-type and mutant A3G is shown in Fig. 4c. The lack of a slow ssDNA bound fraction observed for the mutant, together with the striking difference between the hysteresis observed for the two proteins, shows that the oligomerization-defective mutant does not exhibit slow ssDNA binding kinetics. We therefore conclude that the slow kinetics observed for wild-type A3G is due to oligomerization.

Discussion

Here, we use a single-molecule method that allows us to quantify two distinct modes of A3G binding to ssDNA and characterize the conversion of a fast state into a slow state. These results suggest a binding mechanism in which monomers or dimers initially bind ssDNA and rapidly reach equilibrium ($1/k_{fast} = 24 \pm 1$ s at 200 nM), before slowly converting to oligomers ($1/k_{slow} = 206 \pm 20$ s) (Fig. 5a). Previous bulk solution experiments have established that A3G oligomerizes in the presence of single-stranded nucleic acids[42,43], which inhibits efficient deaminase activity[42]. We also

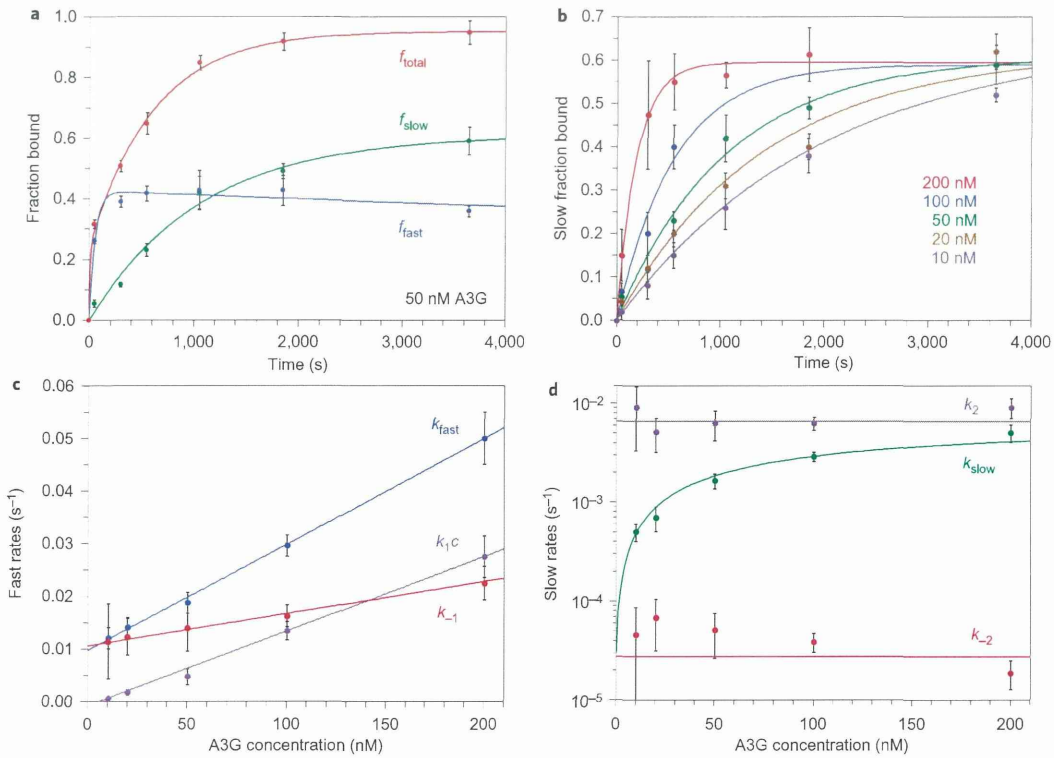


Figure 3 | Quantifying A3G binding reveals association and dissociation rates for fast and slow binding modes. a, Total binding at 50 nM A3G (f_{total}, red) separated into a fast fraction (f_{fast}, blue) and slow fraction (f_{slow}, green), as a function of ssDNA–A3G incubation time. Fits to the binding model (solid lines, Supplementary equations (9)–(11)) yield observed rates k_{fast} and k_{slow}. **b**, Slow fraction bound as a function of time for five A3G concentrations. Solid lines are fits to Supplementary equation (10). Error bars (**a**, **b**) are standard error ($N \geq 3$) for 50–200 nM A3G and propagated error for 10–20 nM A3G. **c**, Fast rates (k_{fast}, blue data points) obtained from fits to the binding model (shown in **a** for 50 nM A3G). The linear fit (solid blue line, Supplementary equation (13)) yields k_1 and k_{-1}. k_1c (purple data points) and k_{-1} (red data points) were also calculated from the binding model. Linear fits (solid lines, Supplementary equations (15) and (17)) yield consistent values of k_1 and k_{-1}. **d**, Slow rates (k_{slow}, green data points) from fits to the binding model (**b**). Fits to Supplementary equation (21) (solid green line, Supplementary Fig. 2b) yield k_2 and k_{-2}. Separate calculations of k_2 (purple) and k_{-2} (red) from the binding model (Supplementary equations (24) and (25)) are also shown. Error bars in **c**,**d** are from uncertainty in the fits to the binding model for total rates k_{fast} and k_{slow}, and propagated error for calculated elementary reaction rates.

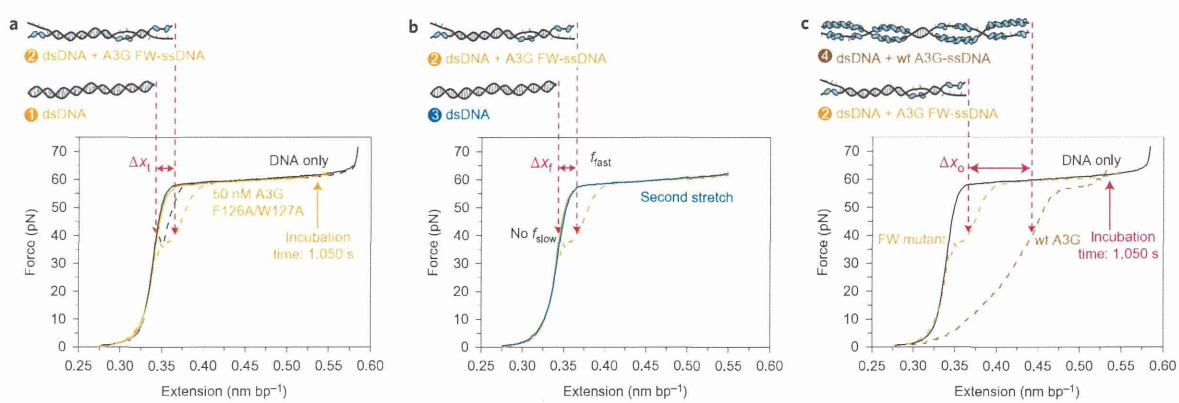


Figure 4 | Oligomerization-defective mutant F126A/W127A (FW) demonstrates that the slow kinetics observed for wild-type A3G is due to oligomerization. a, In the absence of protein (black line), a single DNA molecule reanneals immediately on release, exhibiting minimal hysteresis between extension (solid) and release (dashed). In the presence of 50 nM F126A/W127A A3G (orange line), the stretch curve (solid) follows the dsDNA-only curve, indicating no measurable A3G FW binding to dsDNA (drawing 1). Pausing at fixed DNA extension after the melting transition to incubate the ssDNA with the protein results in ssDNA binding (drawing 2), indicated by the corresponding increase in length Δx_t measured during DNA release at a given force (shown for 40 pN). **b**, The subsequent stretch (dark blue) follows the initial stretch curve (solid orange line), indicating that all the mutant A3G bound during incubation dissociates rapidly (drawing 3), resulting in a decrease in length Δx_f. **c**, Wild-type A3G (drawing 4) exhibits a greater change in length Δx_o relative to the FW mutant (drawing 2) at 1,050 s incubation due to oligomerization on ssDNA.

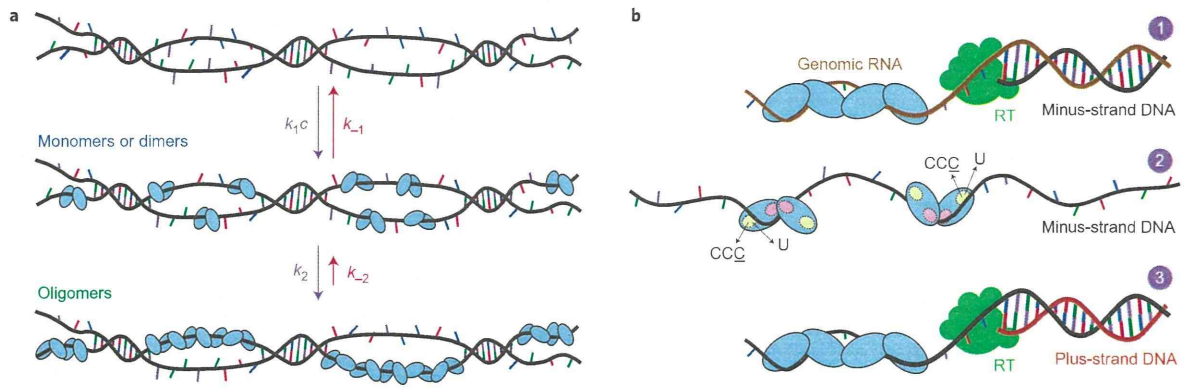


Figure 5 | Models for A3G oligomerization *in vitro* and *in virio*. a, *In vitro*, initially, monomers or dimers bind ssDNA with on rate k_1c and off rate k_{-1}. These forward and backward rates are on similar timescales ($1/k_1c = 33 \pm 1$ s at 200 nM A3G, and $1/k_{-1} = 85 \pm 5$ s), so fast binding reaches equilibrium before the monomers or dimers convert to oligomers ($1/k_2 = 149 \pm 13$ s) on ssDNA. Oligomer dissociation is significantly slower ($1/k_{-2} = 10 \pm 2$ h) *in vitro*. **b**, *In virio*, A3G oligomerizes on the RNA genome, blocking minus-strand DNA synthesis by RT (drawing 1). Once the oligomer dissociates, the monomers or dimers released bind ssDNA within a second, allowing rapid enzymatic activity (drawing 2) until A3G oligomerizes on the ssDNA template in 150 s and blocks plus-strand synthesis (drawing 3).

demonstrate that an A3G mutant (F126A/W127A) that is severely compromised in oligomerization, but retains deaminase activity, does not exhibit slow binding kinetics. Furthermore, a recent study shows that a similar oligomerization-defective mutant (W127A) inhibits deamination-independent viral restriction[44]. These observations support the hypothesis that A3G oligomerization is responsible for deaminase-independent inhibition of viral replication.

In light of these quantitative results *in vitro*, we propose a model for the effects of A3G on HIV-1 replication. A3G oligomerizes on viral RNA when it is packaged inside the virion[43,45], stalling RT during minus-strand synthesis. RT pauses until the oligomer dissociates, or switches to the other RNA template strand, circumventing the A3G roadblock and leading to partial inhibition of reverse transcription[35]. In contrast to our *in vitro* studies, the maximum size of the A3G oligomer may be limited by the small number of A3G molecules packaged in virions[36]. These limitations on oligomer size may lead to higher off rates *in virio*. Once the A3G oligomer dissociates during minus-strand synthesis, the monomers or dimers that are released have low affinity for the newly formed RNA–DNA duplex[14,19]. As RNase H activity exposes the minus-strand, all A3G molecules bind the ssDNA template within a second, as indicated by $1/k_1 = 0.7 \pm 0.1$ s when fast binding rates are extrapolated to the estimated 13 ± 8 μM A3G concentration in the virion[32]. A3G remains bound in a rapid sliding mode[37,41] for $1/k_{-1} = 85 \pm 5$ s, which allows high deamination rates[37,38] until oligomerization forms a roadblock to plus-strand synthesis after $1/k_2 = 149 \pm 13$ s (Fig. 5b).

The *in virio* model that we propose based on our measurements is consistent with the available data on A3G function. Recent cell-based experiments demonstrate that A3G uniformly blocks minus-strand synthesis in the absence of preferred RT termination sites along the viral template[35]. However, this does not necessarily indicate that A3G binds throughout the length of the genomic RNA. Oligomerization is nucleic acid sequence-independent, so we expect the roadblock to form at a small number of random sites along the viral genome, as observed *in virio*. Although A3G may bind RT in an RNA-independent interaction[46], a recent report concludes that there is no direct interaction between A3G and RT[31]. In any case, this type of binding is equally probable for any of the 100 RT heterodimers present in the virion[47], and the probability that one of 7 ± 4 A3G molecules binds and inhibits the catalytically active RT molecule is therefore less than 10%. Thus, a possible A3G–RT interaction appears unlikely to be primarily responsible for A3G-induced inhibition of reverse transcription.

The data we present resolve seemingly contradictory mechanisms for A3G inhibition of viral replication by demonstrating that A3G can function as both a fast deaminase and a slow nucleic acid binding protein. Regulation of enzymatic activity via protein oligomerization may be a general property of other APOBEC family members that inhibit replication of retroviruses and retrotransposons independent of deaminase activity. The single-molecule method described here does not require labelling of DNA or protein and is optimal for measuring the biologically important process of slow protein oligomerization on ssDNA.

Methods

A3G preparation and purification. Recombinant WT A3G and A3G FW were expressed in a baculovirus expression system using an *N*-terminal glutathione *S*-transferase (GST) tag and purified. The GST tag was then removed using a Novagen Enterokinase Cleavage Capture Kit. The deaminase activity of purified wild-type A3G was verified using a gel-based uracil DNA glycosylase assay[19] after removal of the GST tag. A similar method was used to verify the enzymatic activity of purified GST-tagged A3G FW. Additional protein preparation details are presented in the Supplementary Methods.

Single-molecule experiments. Biotin-labelled bacteriophage λ DNA was captured between two streptavidin-coated polystyrene beads, one on a fixed micropipette tip and the other held in an optical trap, as described in the Supplementary Methods. The DNA was extended by moving the micropipette tip at a rate of 100 nm s^{-1}, and the resulting force on the bead in the trap was recorded to obtain the force–extension curve for DNA alone. The buffer surrounding the DNA molecule was then exchanged for a solution of fixed protein concentration. Force–extension curves in the presence and absence of protein were fit to the 'Worm-Like Chain' model for dsDNA (Supplementary equation (1)) and the 'Freely-Jointed Chain' model (Supplementary equation (2)) for ssDNA. The A3G-saturated ssDNA force–extension curve was fit to Supplementary equation (4). Measurements of the fraction of ssDNA bound by protein were then determined by fitting force–extension curves to the linear combination of dsDNA extension and A3G-bound ssDNA extension (Supplementary equations (6) and (7)).

Two-step binding model. The kinetics of protein–DNA binding was fit to a two-step binding model, in which it was assumed that an initial fast bimolecular binding event was followed by a slow unimolecular binding event, as described by Supplementary equations (9) to (11), yielding measured fast and slow binding rates k_{fast} and k_{slow} (ref. 48). By then fitting the concentration-dependent measurements of k_{fast} and k_{slow} to the reaction shown in equation (1), we determine the elementary reaction rates for each step k_1, k_{-1}, k_2 and k_{-2}, as well as the elementary equilibrium constants for each step K_1 and K_2, as described in the Supplementary Methods.

Received 7 February 2013; accepted 11 October 2013; published online 24 November 2013

References

1. Malim, M. H. APOBEC proteins and intrinsic resistance to HIV-1 infection. *Philos. Trans. R. Soc. Lond. B* **364,** 675–687 (2009).
2. Harris, R. S. & Liddament, M. T. Retroviral restriction by APOBEC proteins. *Nature Rev. Immunol.* **4,** 868–877 (2004).
3. Duggal, N. K. & Emerman, M. Evolutionary conflicts between viruses and restriction factors shape immunity. *Nature Rev. Immunol.* **12,** 687–695 (2012).
4. Chiu, Y. L. & Greene, W. C. The APOBEC3 cytidine deaminases: an innate defensive network opposing exogenous retroviruses and endogenous retroelements. *Annu. Rev. Immunol.* **26,** 317–353 (2008).
5. Sheehy, A. M., Gaddis, N. C., Choi, J. D. & Malim, M. H. Isolation of a human gene that inhibits HIV-1 infection and is suppressed by the viral Vif protein. *Nature* **418,** 646–650 (2002).
6. Holmes, R. K., Malim, M. H. & Bishop, K. N. APOBEC-mediated viral restriction: not simply editing? *Trends Biochem. Sci.* **32,** 118–128 (2007).
7. Fisher, A. G. *et al.* The *sor* gene of HIV-1 is required for efficient virus transmission *in vitro*. *Science* **237,** 888–893 (1987).
8. Strebel, K. *et al.* The HIV 'A' (sor) gene product is essential for virus infectivity. *Nature* **328,** 728–730 (1987).
9. Goila-Gaur, R. & Strebel, K. HIV-1 Vif, APOBEC, and intrinsic immunity. *Retrovirology* **5,** 1–16 (2008).
10. Lecossier, D., Bouchonnet, F., Clavel, F. & Hance, A. J. Hypermutation of HIV-1 DNA in the absence of the Vif protein. *Science* **300,** 1112 (2003).
11. Mangeat, B. *et al.* Broad antiretroviral defence by human APOBEC3G through lethal editing of nascent reverse transcripts. *Nature* **424,** 99–103 (2003).
12. Zhang, H. *et al.* The cytidine deaminase CEM15 induces hypermutation in newly synthesized HIV-1 DNA. *Nature* **424,** 94–98 (2003).
13. Suspène, R. *et al.* APOBEC3G is a single-stranded DNA cytidine deaminase and functions independently of HIV reverse transcriptase. *Nucleic Acids Res.* **32,** 2421–2429 (2004).
14. Yu, Q. *et al.* Single-strand specificity of APOBEC3G accounts for minus-strand deamination of the HIV genome. *Nature Struct. Mol. Biol.* **11,** 435–442 (2004).
15. Harris, R. S. *et al.* DNA deamination mediates innate immunity to retroviral infection. *Cell* **113,** 803–809 (2003); erratum **116,** 629 (2004).
16. Levin, J. G., Mitra, M., Mascarenhas, A. & Musier-Forsyth, K. Role of HIV-1 nucleocapsid protein in HIV-1 reverse transcription. *RNA Biol.* **7,** 754–774 (2010).
17. Newman, E. N. C. *et al.* Antiviral function of APOBEC3G can be dissociated from cytidine deaminase activity. *Curr. Biol.* **15,** 166–170 (2005).
18. Holmes, R. K., Koning, F. A., Bishop, K. N. & Malim, M. H. APOBEC3F can inhibit the accumulation of HIV-1 reverse transcription products in the absence of hypermutation—comparisons with APOBEC3G. *J. Biol. Chem.* **282,** 2587–2595 (2007).
19. Iwatani, Y., Takeuchi, H., Strebel, K. & Levin, J. G. Biochemical activities of highly purified, catalytically active human APOBEC3G: correlation with antiviral effect. *J. Virol.* **80,** 5992–6002 (2006).
20. Luo, K. *et al.* Cytidine deaminases APOBEC3G and APOBEC3F interact with human immunodeficiency virus type 1 integrase and inhibit proviral DNA formation. *J. Virol.* **81,** 7238–7248 (2007).
21. Turelli, P., Mangeat, B., Jost, S., Vianin, S. & Trono, D. Inhibition of hepatitis B virus replication by APOBEC3G. *Science* **303,** 1829 (2004).
22. Bogerd, H. P., Wiegand, H. L., Doehle, B. P., Lueders, K. K. & Cullen, B. R. APOBEC3A and APOBEC3B are potent inhibitors of LTR-retrotransposon function in human cells. *Nucleic Acids Res.* **34,** 89–95 (2006).
23. Bogerd, H. P. *et al.* Cellular inhibitors of long interspersed element 1 and Alu retrotransposition. *Proc. Natl Acad. Sci. USA* **103,** 8780–8785 (2006).
24. Chen, H. *et al.* APOBEC3A is a potent inhibitor of adeno-associated virus and retrotransposons. *Curr. Biol.* **16,** 480–485 (2006).
25. Muckenfuss, H. *et al.* APOBEC3 proteins inhibit human LINE-1 retrotransposition. *J. Biol. Chem.* **281,** 22161–22172 (2006).
26. Kinomoto, M. *et al.* All APOBEC3 family proteins differentially inhibit LINE-1 retrotransposition. *Nucleic Acids Res.* **35,** 2955–2964 (2007).
27. Niewiadomska, A. M. *et al.* Differential inhibition of long interspersed element 1 by APOBEC3 does not correlate with high-molecular-mass-complex formation or P-body association. *J. Virol.* **81,** 9577–9583 (2007).
28. Bulliard, Y. *et al.* Structure–function analyses point to a polynucleotide-accommodating groove essential for APOBEC3A restriction activities. *J. Virol.* **85,** 1765–1776 (2011).
29. Narvaiza, I. *et al.* Deaminase-independent inhibition of parvoviruses by the APOBEC3A cytidine deaminase. *PLoS Pathog.* **5,** e1000439 (2009).
30. Bishop, K. N., Verma, M., Kim, E. Y., Wolinsky, S. M. & Malim, M. H. APOBEC3G inhibits elongation of HIV-1 reverse transcripts. *PLoS Pathog.* **4,** e1000231 (2008).
31. Adolph, M. B., Webb, J. & Chelico, L. Retroviral restriction factor APOBEC3G delays the initiation of DNA synthesis by HIV-1 reverse transcriptase. *Plos One* **8,** e64196 (2013).
32. Iwatani, Y. *et al.* Deaminase-independent inhibition of HIV-1 reverse transcription by APOBEC3G. *Nucleic Acids Res.* **35,** 7096–7108 (2007).
33. Li, X. Y., Guo, F., Zhang, L. & Kleiman, L. & Cen S. APOBEC3G inhibits DNA strand transfer during HIV-1 reverse transcription. *J. Biol. Chem.* **282,** 32065–32074 (2007).
34. Mbisa, J. L. *et al.* Human immunodeficiency virus type 1 cDNAs produced in the presence of APOBEC3G exhibit defects in plus-strand DNA transfer and integration. *J. Virol.* **81,** 7099–7110 (2007).
35. Gillick, K. *et al.* Suppression of HIV-1 infection by APOBEC3 proteins in primary human CD4+ T cells is associated with inhibition of processive reverse transcription as well as excessive cytidine deamination. *J. Virol.* **87,** 1508–1517 (2013).
36. Xu, H. Z. *et al.* Stoichiometry of the antiviral protein APOBEC3G in HIV-1 virions. *Virology* **360,** 247–256 (2007).
37. Chelico, L., Sacho, E. J., Erie, D. A. & Goodman, M. F. A model for oligomeric regulation of APOBEC3G cytosine deaminase-dependent restriction of HIV. *J. Biol. Chem.* **283,** 13780–13791 (2008).
38. Nowarski, R., Britan-Rosich, E., Shiloach, T. & Kotler, M. Hypermutation by intersegmental transfer of APOBEC3G cytidine deaminase. *Nature Struct. Mol. Biol.* **15,** 1059–1066 (2008).
39. King, G. A. *et al.* Revealing the competition between peeled ssDNA, melting bubbles, and S-DNA during DNA overstretching using fluorescence microscopy. *Proc. Natl Acad. Sci. USA* **110,** 3859–3864 (2013).
40. Chaurasiya, K. R., Paramanathan, T., McCauley, M. J. & Williams, M. C. Biophysical characterization of DNA binding from single molecule force measurements. *Phys. Life Rev.* **7,** 299–341 (2010).
41. Senavirathne, G. *et al.* Single-stranded DNA scanning and deamination by APOBEC3G cytidine deaminase at single molecule resolution. *J. Biol. Chem.* **287,** 15826–15835 (2012).
42. Chelico, L., Prochnow, C., Erie, D. A., Chen, X. S. & Goodman, M. F. Structural model for deoxycytidine deamination mechanisms of the HIV-1 inactivation enzyme APOBEC3G. *J. Biol. Chem.* **285,** 16195–16205 (2010).
43. Huthoff, H., Autore, F., Gallois-Montbrun, S., Fraternali, F. & Malim, M. H. RNA-dependent oligomerization of APOBEC3G is required for restriction of HIV-1. *PLoS Pathog.* **5,** e1000330 (2009).
44. Bélanger, K., Savoie, M., Rosales Gerpe, M. C., Couture, J-F. & Langlois, M-A. Binding of RNA by APOBEC3G controls deamination-independent restriction of retroviruses. *Nucleic Acids Res.* **41,** 7438–7452 (2013).
45. Soros, V. & Greene, W. APOBEC3G and HIV-1: strike and counterstrike. *Curr. HIV/AIDS Rep.* **4,** 3–9 (2007).
46. Wang, X. X. *et al.* The cellular antiviral protein APOBEC3G interacts with HIV-1 reverse transcriptase and inhibits its function during viral replication. *J. Virol.* **86,** 3777–3786 (2012).
47. Coffin, J. M., Hughes, S. H. & Varmus, H. E. *Retroviruses* (Cold Spring Harbor Laboratory Press, 1997).
48. Vo, M. N., Barany, G., Rouzina, I. & Musier-Forsyth, K. Mechanistic studies of mini-TAR RNA/DNA annealing in the absence and presence of HIV-1 nucleocapsid protein. *J. Mol. Biol.* **363,** 244–261 (2006).

Acknowledgements

The authors thank D. Pollpeter, M.H. Malim and D. Rueda for valuable discussions, and M.F. Goodman for his generous gift of the F126A/W127A mutant clone. This work was supported in part by the National Institutes of Health (GM072462 to M.C.W. and GM065056 to K.M-F.) and the National Science Foundation (MCB-1243883 to M.C.W.), the Japan Society for the Promotion of Science (JSPS; KAKENHI_24590568 to Y.I.), and in part by funds from the NIH Intramural Research Program (NICHD; to J.G.L.). K.R.C. was supported by the NSF IGERT Program (DGE-0504331).

Author contributions

M.C.W., K.R.C. and I.R. designed the experiments. K.R.C. performed experiments and analysed the data. M.M. performed experiments with the mutant. H.G. performed preliminary experiments. S.K., W.W., D.F.Q., T.W., Y.I., D.S.B.C. and A.H. prepared the proteins. I.R. developed the binding model. K.R.C., M.C.W., I.R., J.G.L. and K.M-F. wrote the manuscript.

Additional information

Supplementary information is available in the online version of the paper. Reprints and permissions information is available online at www.nature.com/reprints. Correspondence and requests for materials should be addressed to M.C.W.

Competing financial interests

The authors declare no competing financial interests.

nature chemistry

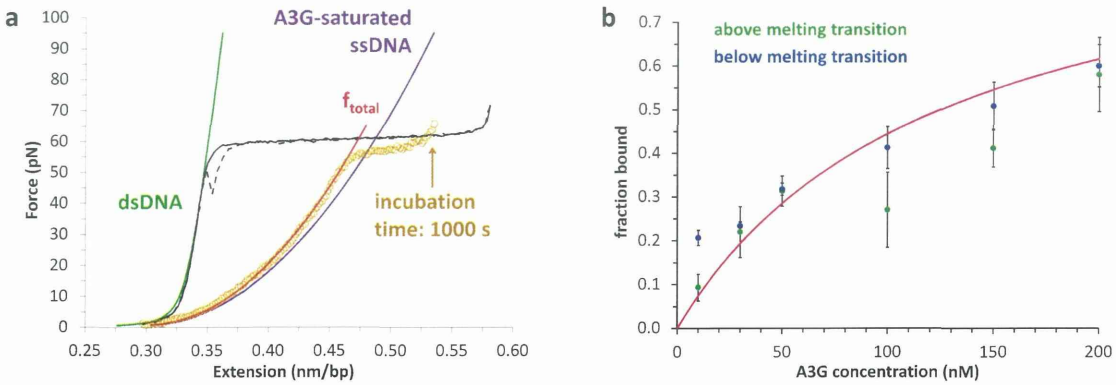
SUPPLEMENTARY INFORMATION

DOI: 10.1038/NCHEM.1795

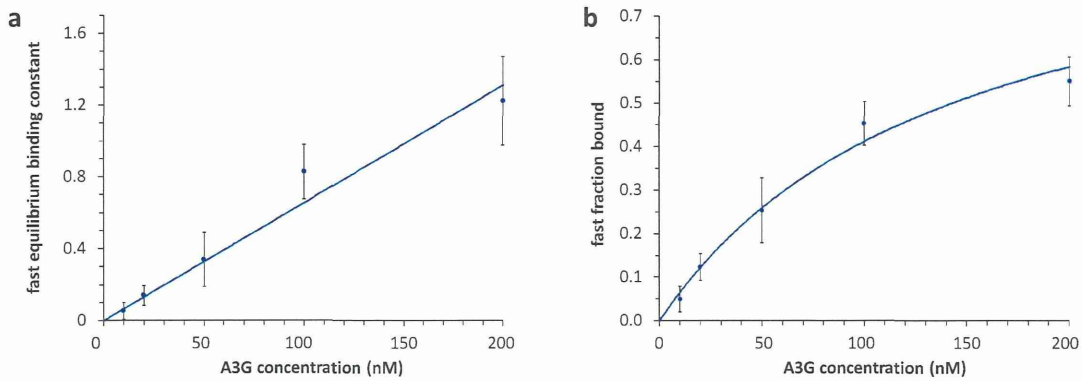
Oligomerization transforms human APOBEC3G from an efficient enzyme to a slowly dissociating nucleic acid-binding protein

Kathy R. Chaurasiya, Micah J. McCauley, Wei Wang, Dominic F. Qualley, Tiyun Wu, Shingo Kitamura, Hylkje Geertsema, Denise Chan, Amber Hertz, Yasumasa Iwatani, Judith G. Levin, Karin Musier-Forsyth, Ioulia Rouzina, and Mark C. Williams

Supplementary Figure 1	DNA force-extension curves in the presence of A3G are used to quantify A3G binding to ssDNA.
Supplementary Figure 2	Concentration dependence of fast equilibrium binding constant and fast fraction bound.
Supplementary Table 1	Values for fitting parameters in the steady-state binding model.
Supplementary Table 2	Concentration-independent reaction parameters are similar across multiple methods, which supports the binding model.
Methods	Detailed supplementary methods section



Supplementary Figure 1. DNA force-extension curves in the presence of A3G are used to quantify A3G binding to ssDNA. (**a**) DNA release curve (brown circles, data also shown in Fig. 2, panel d) after stretching dsDNA in the presence of 50 nM A3G followed by 1000 s incubation at 0.53 nm/bp (indicated by the brown arrow). Below the melting transition, the force-extension curve is a linear combination of dsDNA (green, Eq. S1) and A3G-saturated ssDNA (purple, Eq. S5) because any ssDNA generated by force-induced melting that is not bound to A3G will immediately reanneal into dsDNA. The fit yields the fraction bound f_{total} = 0.90 ± 0.03 (red, Eq. S6). Uncertainty from the fits is small relative to variations across multiple measurements, so only standard errors are reported here. This data may also be fit above the melting transition as a linear combination of A3G-saturated ssDNA (purple, Eq. S5) and ssDNA (Eq. S2) because the molecule is primarily ssDNA regardless of protein binding. A3G binding during incubation shifts the release curve toward the saturated A3G curve (Eq. S7, fit not shown for clarity). (**b**) Total A3G bound without incubation (t = 50 s) as a function of concentration. Fits above (green) and below (blue) the melting transition yield similar results, as expected. Data points are averages from at least three fits, and error bars reflect standard error. The values from two methods were averaged (points and propagated error bars not shown for clarity) and fit to a simple DNA binding isotherm (solid red line, Eq. S8). The fit yields K_d = 125 ± 25 nM, which is similar to K_d = 76 ± 21 nM [18] and K_d = 160 nM[37] obtained from ensemble measurements. This measured K_d value agrees with K_{d1} = 127 ± 6 nM for fast binding obtained from K_1 (Eq. S18, Supplementary Fig. 2, panel a, see Supplementary Table 2 online).



Supplementary Figure 2. Concentration dependence of fast equilibrium binding constant and fast fraction bound. (**a**) The equilibrium binding constant K_1 is linearly dependent on A3G concentration, as expected for a bimolecular process. The solid line is a linear fit to Eq. S18 with slope k_1/k_{-1} = 6.6 (± 0.4) x 10^6 M^{-1}, which yields a dissociation constant for the fast step (K_{d1} = 127 ± 6 nM) that agrees with the one obtained from a simple DNA binding isotherm without incubation (K_d = 125 ± 25 nM, see Supplementary Fig. 1, panel b). (**b**) Fast portion of A3G binding P_{fast} saturates at high protein concentration. Data points and their uncertainties are from fits to the binding model (see Supplementary Table 1), solid line is a fit to Eq. S12 that yields k_1 = 1.4 (± 0.1) x 10^5 $M^{-1}s^{-1}$ and k_{-1} = 2.1 (± 0.9) x 10^{-2} s^{-1}. These values are fixed in the k_{slow} versus c fit (Figure 3, panel d, solid green line) to limit the fit to two free parameters.

A3G (nM)	k_{fast} (x 10^{-2} s^{-1})	k_{slow} (x 10^{-3} s^{-1})	P_{fast}	P_{total}
10	1.2 ± 0.2	0.5 ± 0.1	0.05 ± 0.03	0.90 ± 0.01
20	1.4 ± 0.2	0.7 ± 0.2	0.12 ± 0.03	0.88 ± 0.02
50	1.9 ± 0.2	1.6 ± 0.3	0.25 ± 0.07	0.95 ± 0.02
100	3.0 ± 0.2	2.9 ± 0.3	0.45 ± 0.05	0.97 ± 0.05
200	5.0 ± 0.5	5.0 ± 1.0	0.55 ± 0.06	0.99 ± 0.15

Supplementary Table 1. Values for fitting parameters in the steady-state binding model. Total binding at a given A3G concentration (f_{total}) was separated into its fast (f_{fast}) and slow (f_{slow}) fractions and monitored as a function of ssDNA-A3G incubation time, as illustrated in Fig. 2. The three curves obtained at each protein concentration (shown in Fig. 3, panel a for 50 nM A3G) were fit to Eqs. S9-S11. The parameters from each fit were used to calculate weighted averages for the observed binding rates k_{fast} and k_{slow}, and the equilibrium fractions of fast and total binding P_{fast} and P_{total}.

Method	k_1 (x 10^5 $M^{-1}s^{-1}$)	k_{-1} (x 10^{-2} s^{-1})	Method	k_2 (x 10^{-3} s^{-1})	k_{-2} (x 10^{-5} s^{-1})	Method	K_1/c (x 10^6 M^{-1})	Method	K_2
A1	2.0 ± 0.1	1.0 ± 0.1	B1	7.0 ± 1.0	3.0 ± 2.0	C1	6.6 ± 0.4	D1	166 ± 17
A2	1.4 ± 0.1	1.1 ± 0.1	B2	6.6 ± 0.7	2.8 ± 0.5	C2	12.9 ± 0.8	D2	251 ± 54
A3	1.2 ± 0.1	2.0 ± 0.2							
A4	1.4 ± 0.1	2.1 ± 0.9							
Average	1.5 ± 0.1	1.2 ± 0.1		6.7 ± 0.6	2.8 ± 0.5		7.9 ± 0.4		174 ± 16

Supplementary Table 2. Concentration-independent reaction parameters are similar across multiple methods, which supports the binding model. Weighted averages and uncertainty are reported in the main text. Methods A1 and A2 are linear fits of k_{fast}, k_1c, and k_{-1} vs. c (Fig. 3, panel a). Calculations from Eqs. S14 and S16 were averaged across protein concentration in method A3. Method A4 is the P_{fast} vs. c fit (Fig. S2, panel b), and method B1 is the k_{slow} vs. c fit (Fig. 3, panel b). Calculations from Eqs. S24 and S25 were averaged across A3G concentration in method B2. Method C1 is the K_1 vs. c fit (Supplementary Fig. 2, panel a) and method C2 uses the weighted averages of k_1 and k_{-1} in Equation S18. The weighted average of calculations from Eq. S23 over all measured protein concentrations is method D1. Method D2 uses experimentally-determined parameters in Eq. S22.