Published in final edited form as:

J Mol Biol. 2016 July 03; 428(13): 2661–2670. doi:10.1016/j.jmb.2016.03.031.

The *in vitro* Biochemical Characterization of an HIV-1 Restriction Factor A3F: Importance of Loop 7 on Both CD1 and CD2 for DNA binding and Deamination

Qihan Chen^{*}, Xiao Xiao[#], Aaron Wolfe[#], and Xiaojiang S. Chen^{*,§,1}

^{*}Molecular and Computational Biology Program, Departments of Biological Sciences and Chemistry, University of Southern California, Los Angeles, CA 90089, USA

[§]Center of Excellence in NanoBiophysics; University of Southern California, Los Angeles, CA 90089, USA; Norris Comprehensive Cancer Center; University of Southern California, Los Angeles, CA 90089, USA

[#]Genetic, Molecular and Cellular Biology, Department of Biological Sciences, University of Southern California, Los Angeles, CA 90089, USA

Abstract

APOBEC3F (A3F) is a member of the apolipoprotein B mRNA-editing enzyme catalytic polypeptide-like (APOBEC) family of proteins that can deaminate cytosine (C) to uracil (U) on nucleic acids. A3F is one of the four APOBEC members with two Zn-coordinated homologous cytosine deaminase domains (CD), with the others being A3G, A3D, and A3B. Here we report the *in vitro* characterization of DNA binding and deaminase activities using purified wild type (wt) and various mutant proteins of A3F from an *E. coli* expression system. We show that, even though CD1 is catalytically inactive and CD2 is the active deaminase domain, presence of CD1 on the N-terminus of CD2 enhances the deaminase activity by over an order of magnitude. This enhancement of CD2 catalytic activity is mainly through the increase of substrate ssDNA binding by the N-terminal CD1 domain. We further show that the loop 7 of both CD1 and CD2 of A3F plays an important role for ssDNA binding for each individual domain, as well as for the deaminase activity of CD2 domain in the full length A3F.

Graphical abstract

¹To whom correspondence should be addressed. Tel: +1 213 821 1255; FAX: +1 213 740 4340; xiaojiac@usc.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

The APOBEC family members contain one or two homologous cytosine deaminase (CD) domains that share a conserved sequence motif (H-X-E-X23-28-P-C-X-C) for Zn-coordination at the putative catalytic center^{1; 2; 3; 4; 5}. Four of the eleven family members (A3B, A3D, A3F, A3G) contain two homologous CD domains, and the others have only one CD domain. Nine of the eleven members are shown to catalyze deamination of cytosine to uracil on single-stranded (ss) DNA or RNA^{1; 4; 5; 6}. Most of the APOBEC3 family members constitute parts of the important immune system to safeguard the potential damage from the intrinsic retroviral elements and infectious retroviral and some other viral pathogens^{7; 8; 9; 10; 11; 12; 13; 14; 15; 16; 17}. However, some of these members are also implicated in genome-wide accidental mutations that may result in significant repercussions and biological consequences such as epigenetic remodeling, cancer, and even evolution^{18; 19; 20; 21; 22; 23; 24}.

Among the four APOBEC3 members containing double CD domains, A3F and A3G show potent anti-HIV activities^{25; 26; 27; 28; 29; 30}, which is through incorporation/encapsidation of the APOBEC3 enzymes into HIV virions^{31; 32; 33; 34}. The incorporation/encapsidation into HIV virions, which is mainly mediated through RNA binding by the A3 proteins, requires both CD1 and CD2 domains of A3F, whereas only CD1 of A3G is needed^{32; 34; 35; 36; 37; 38; 39}. On the other hand, while HIV Vif targets A3G for E3 ubiquitin ligase mediated degradation through binding to the CD1 of A3G, Vif targets A3F by binding to its CD2 domain^{25; 26; 40; 41; 42; 43}.

The anti-HIV activity is thought to be mostly deamination dependent, with some reports for deamination independent activity as well^{44; 45; 46; 47; 48; 49; 50}. Regardless, both CD1 and CD2 domains of A3F and A3G are required to display anti-HIV activity. For A3F and A3G, the CD1 is previously shown to be catalytically inactive, whereas CD2 is the catalytically active domain for both proteins^{42; 51; 52; 53}. Here we report *in vitro* biochemical studies of A3F and various mutant constructs containing loop 7 mutations on either the CD1 or CD2 domains using purified recombinant proteins from an *E. coli* expression system.

The role of CD1 and CD2 for A3F deaminase activity

To assess the *in vitro* biochemical activities of A3F and the role of loop 7 residues, we first cloned and purified wild type and mutant A3F proteins as MBP-fusions to near homogeneity (Figure 1A, 1B). Four constructs were tested, including the wild type full length (fl) A3F (A3F-fl) construct, as well as the mutants A3F-fl E67A (with the CD1 catalytic residue mutation at E67), A3F-fl E251A (with the corresponding CD2 catalytic residue mutation), and A3F-fl E67/E251A (with both catalytic residue mutations). These purified MBP fusion proteins of A3F were assayed for deaminase activity using a 30 nucleotide (nt) ssDNA substrate (Figure 1C). As expected, the *in vitro* activity assay result showed obvious deaminase activity for both A3F-fl and A3F-fl E67A, whereas no deamination activity was detected for A3F-fl E251A and A3F-fl E67/E251A (Figure 1D, Table 1). This *in vitro* assay result confirms that CD1 has no deaminase activity, with only CD2 acting as the catalytically active deaminase domain, and is consistent with previous finding. In addition, quantification of the dose response deamination activity (Figure 1E, 1F) revealed that the CD1 mutant A3F-fl E67A activity is more or less comparable to that of the wild type A3F-fl, suggesting that the E67A mutation within the CD1 domain did not have a significant impact on the deaminase activity of the full length protein. This is similar to what is observed for A3B and A3G, where the mutation of the CD1 Glu residue appears to have no obvious effect on the fl A3B and A3G deaminase activity^{51; 54}.

The effect of CD2 loop7 mutations on A3F deaminase activity

Since previous studies suggest a role of loop 7 for DNA binding and deaminase activity for some APOBEC3 proteins, such as A3G-CD2, A3B-CD2, and A3A^{29; 55; 56; 57}, we want to assess the functional role of specific residues on loop 7 of both the CD1 and CD2 of A3F. We made point mutations around the tip of CD2 loop 7 on the constructs of both A3F-fl and A3F-CD2 alone, including W310A, D311A, D311K, and T312P, which are located next to the Zn-coordinated active center (Figure 2A, 2B). All constructs were expressed as MBP fusions in *E. coli* and purified to near homogeneity (Figure 2C). For the four mutants in A3F-fl constructs, subsequent deamination assays revealed the W310A mutant showed the most reduced activity over the tested protein concentration range (Figure 2D), with about 57% reduction of deaminase activity when compared to wild type (Fig. 2E, Table 1). The A3F-fl D311A and A3F-fl T312P mutants showed very little effect on deaminase activity. Interestingly, the A3F-fl D311K mutant showed significantly increased activity, with approximately 38% increased activity over wild type A3F-fl.

The purified wild type A3F-CD2 alone also showed obvious deamination activity (Figure 2F, 2G, Table 1). However, specific deaminase activity of the A3F-CD2 construct by itself (~0.501 nM/μM/hr (product/substrate/hr)) is more than 10-fold lower than that of A3F-fl or A3F-fl E67A (Table 1), suggesting that the presence of CD1 on the N-terminus of CD2 greatly enhanced the activity of CD2. For the four loop 7 mutations on this A3F-CD2 construct, we observed a similar trend of deaminase activity changes as with the A3F-fl construct, with the exception of a dramatic activity drop on D311A (Figure 2F, 2G, Table 1). Again, the mutant A3F-CD2 W310A showed the most significant impact, reducing the deaminase activity to background level, followed by the A3F-CD2 D311A mutant, which

showed about an 80% reduction of deaminase activity compared to wild type A3F-CD2. As with the full length construct, A3F-CD2 T312P had little effect on deamination activity, and quite interestingly the D311K mutation once again showed an increased activity, with approximately 22% higher activity than wild type A3F-CD2 (Figure 2G, Table 1). To summarize, the above result showed that the four mutations of the three CD2 loop 7 residues showed similar effects on deaminase activity in the context of either the A3F-fl or the A3F-CD2 construct. Generally speaking, the W310A mutation had a significant negative impact on deaminase activity, whereas D311K had a positive impact.

The effect of CD2 loop7 mutations on substrate DNA binding

In order to examine whether the CD2 loop 7 mutations on A3F-fl and A3F-CD2 have altered substrate binding affinity, we performed the ssDNA binding assay for each protein construct using a gel shift assay as described in previous studies^{58; 59}, and the result is shown in Figure 2H and 2I and Table 2. For the A3F-fl constructs, the A3F-fl and A3F-fl D311A, D311K, and T312P constructs all had similar binding affinity, with a K_d of around 1.0 μM (Table 2). However, the A3F-fl W310A construct showed a greater than 3-fold reduction of the binding affinity compared to that of A3F-fl, with a K_d of $\sim 3.5 \mu\text{M}$ (Table 2). A similar observation was observed among the A3F-CD2 constructs, where the A3F-CD2 W310A mutant showed a much more dramatic reduction (over 10-fold) of DNA binding affinity compared with that of wt A3F-CD2, with a K_d of $\sim 152.5 \mu\text{M}$ (Table 2), whereas A3F-CD2 D311A showed about a 3-fold reduction. The other two A3F-CD2 mutants (D311K and T312P) showed no obvious change in DNA binding affinity. Thus, the significant reduction of DNA binding for W310A of A3F-CD2 and A3F-fl appears to correlate well with the reduced deaminase activity for their respective mutant constructs. The fact that the W310A mutation in A3F-CD2 alone has a much dramatic reduction of DNA binding than the same mutation in A3F-fl construct suggests that presence of CD1 in the A3F-fl construct compensates for the loss of DNA binding that is potentially controlled by this tryptophan on CD2 loop 7.

However, when comparing the DNA binding between wild type A3F-fl and A3F-CD2 alone, the binding affinity for A3F-CD2 (K_d of 12.8 μM) is about 13-fold lower than that for the A3F-fl (K_d of 1.1 μM) (Table 2), indicating that CD1 plays an important role in enhancing the DNA binding to the fl A3F. Therefore, this reduced DNA binding affinity of the truncated construct A3F-CD2 shows correlation with the lowered deaminase activity compared with that of full length A3F, and similarly, the lowered deaminase activity of the W310A mutants can be correlated to their decreased DNA binding affinity compared with their wild type constructs.

Previously, Ara et al reported a higher binding affinity of A3F-fl (K_d of 20 nM) and A3F-CD2 (K_d of 288 nM) to a 118 nt long ssDNA obtained by rotational anisotropy using proteins purified from insect cells⁶⁰. While there are multiple potential reasons for this discrepancy, the more likely causes could be stemmed from the different ssDNA lengths used (118 nt vs 30 nt reported here) and the different binding assay methods (anisotropy vs the gel shift assay reported here). Of interest, despite the different K_d values, both their

study and ours show that the binding affinity of full length A3F is about 13–14 fold higher than that of A3F CD2 alone.

Roles of CD1 loop 7 for A3F DNA binding and deaminase activities

Since the CD2 loop-7 W310A mutation significantly impacted both DNA binding and deaminase activity, we wondered whether the corresponding tryptophan on loop 7 of CD1, W126, has similar role (Figure 3A). Interestingly, the loop 7 sequence of A3F-CD1 is very similar to that of A3F-CD2 as well as A3G-CD1 (Figure 3B). We constructed a W126A mutant in A3F-fl (A3F-fl W126A) and in A3F-CD1 alone (A3F-CD1 W126A) (Figure 3C), and performed an additional DNA gel shift assay. The results demonstrated that, even though A3F-CD1 alone showed strong ssDNA binding, mutant A3F-CD1 W126A had no detectable ssDNA binding under the same condition (Figure 3D, Table 2), demonstrating that W126 on CD1 loop 7 plays an important role in binding DNA by the A3F-CD1 construct.

Despite the complete loss of DNA binding by the A3F-CD1 W126A mutant, it is worth noting that the wild type A3F-CD1 alone has a ssDNA binding affinity (K_d of 0.15 μ M, Figure 3F, Table 2) that is ~85-fold higher than that of wild type A3F-CD2 alone (K_d of 12.8 μ M), and ~7-fold higher than that of wild type A3F-fl (1.1 μ M) (Table 2). This result suggests that the three dimensional arrangement of CD1-CD2 in the context of a full length A3F construct may result in loss of the DNA binding affinity displayed by the CD1 domain alone, potentially through steric hindrances by the presence of CD2 that effectively reduces DNA accessibility and binding to CD1.

Considering the undetectable DNA binding by A3F-CD1 W126A, we wondered whether full length A3F with the CD1 W126A mutation (A3F-fl W126A) will have a DNA binding affinity comparable to that of the A3F-CD2 construct alone. Interestingly, the result revealed that the A3F-fl W126A construct displayed significantly stronger DNA binding than A3F-CD2 alone, with a K_d of ~3.9 μ M (Figure 3G, Table 2). This is over 3-fold higher than that of A3F-CD2 alone, which had a K_d of 12.8 μ M. However, this DNA binding affinity of A3F-fl W126A is still about 3-fold less than that of the wild type A3F-fl (K_d of 1.1 μ M) (Table 2). In other words, the A3F-fl W126A shows a DNA binding affinity that is somewhere between the wild type A3F-fl and A3F-CD2 alone.

Taken together, even though the CD1 mutation W126A resulted in loss of DNA binding with the CD1 domain alone, presence of the W126A mutant CD1 domain at the N-terminus of A3F-CD2 can still significantly enhance the DNA binding compared to wild type CD2 alone. This suggests that the three dimensional arrangement of CD1 and CD2 in the full length A3F construct may somehow increase the DNA binding affinity of CD2; it is possible that positively charged surface areas bridging the two domains may have additional impact beyond the observed contribution from loop 7 of the CD1 domain.

The deamination assay showed that the A3F-fl W126A construct had ~75% reduction of deaminase activity compared with wild type A3F-fl (Fig. 3H, Table 1), indicating the important role of the W126 residue on the loop 7 of CD1 for activity. On the other hand, the

deamination activity of A3F-fl W126A was about 4-fold higher than that of wild type A3F-CD2 alone (Figure 3H, Table 1), again suggesting that the presence of the mutated CD1 on loop 7 W126A can still effectively enhance the deaminase activity of CD2, likely benefitted from the enhanced DNA binding affinity facilitated by the CD1 as a whole.

In order to rule out the possibility that the A3F mutations in this study may cause structural alteration, we performed Circular Dichroism (CD) scan to compare the secondary structural profiles of wild type A3F and the selected mutants that showed reduced deaminase and nucleic acid binding activities. The result showed similar CD scan curves for the A3F-fl wt and its W310A and W126A mutants (Figure 4A), the A3F-CD2 wt and its W310A and D311A mutants (Figure 4B), and A3F-CD1 wt and its W126A mutant (Figure 4C). In addition the thermal scan showed similar heat stability for the wt A3F-fl and its W126A and W310A mutants (Figure 4D). These results indicated that the A3F mutations tested in this study showed no obvious structural alteration, which is expected from their surface location of the mutated residues.

Previous studies using A3F, A3G and AID revealed that loop 7 region of these proteins plays certain role in the deamination motif preference, with A3G loop 7 being the most sensitive to mutations in altering its motif preference^{56; 57; 61; 62; 63}. To investigate whether the reduced deaminase activity for the W310A and D311A mutants in the context of both A3F-fl and A3F-CD2 are due to the substrate hot spot change from “TC” to other motif sequences, we tested deaminase activity of these four mutants using ssDNA substrates containing the hotspot motif “AC”, “TC”, “GC” or “CC”. The results showed that “TC” motif remains to be the most preferred hot spot for these mutants on the loop 7 of CD2 (Figure 4E, 4F), suggesting that these loop 7 residue mutations did not have a significant impact on the preferred hot spot motif.

The mutation of W127A in A3G (equivalent of W126 of A3F) is previously reported to disrupt virion incorporation of A3G⁶⁴. Interestingly, the W126A mutation in A3F is also reported to disrupt virion incorporation of A3F³⁴, and later shown to result in loss of antiviral function³². These results suggest that CD1 W126 of A3F may also be involved in nucleic acid binding (RNA) that is important for virion incorporation. Our result of the reduced ssRNA and ssDNA binding by CD1 W126A mutation (Fig. 2D, 2E) is consistent with these previously reported functional loss in virion incorporation and anti-HIV activities. In addition, the reduced deaminase activity of W126A mutant in A3F-fl (Fig. 3H) may also affect the deaminase-dependent anti-HIV activity.

In summary, we report on the *in vitro* characterizations of DNA/RNA binding and deaminase activities using purified wild type and various mutant A3F proteins from an *E. coli* expression system. We showed that, despite the fact that CD1 is catalytically inactive and CD2 is the only active deaminase domain, presence of both a wild type and mutant CD1 domain in the context of full length A3F can greatly enhance the deaminase activity observed when compared to the second domain alone. This enhancement of catalytic activity is likely through an increase of overall substrate DNA binding due to the addition of the CD1 domain. We further showed that the loop 7 tryptophan residue of both CD1 and CD2 of

A3F are important for mediating substrate ssDNA binding of the individual A3F domains and play an important role in overall deamination activity of A3F.

Acknowledgments

We thank the core laboratory in the Center of Excellence in NanoBiophysics at University of Southern California for support in the related biophysics characterization of the purified protein. This work is supported in part by the NIH grant R01GM087986 to X.S.C.

Bibliography

1. Jarmuz A, Chester A, Bayliss J, Gisbourne J, Dunham I, Scott J, Navaratnam N. An anthropoid-specific locus of orphan C to U RNA-editing enzymes on chromosome 22. *Genomics*. 2002; 79:285–96. [PubMed: 11863358]
2. Bransteitter R, Prochnow C, Chen XS. The current structural and functional understanding of APOBEC deaminases. *Cell Mol Life Sci*. 2009; 66:3137–47. [PubMed: 19547914]
3. Conticello SG. The AID/APOBEC family of nucleic acid mutators. *Genome Biol*. 2008; 9:229.
4. Harris RS, Bishop KN, Sheehy AM, Craig HM, Petersen-Mahrt SK, Watt IN, Neuberger MS, Malim MH. DNA deamination mediates innate immunity to retroviral infection. *Cell*. 2003; 113:803–9. [PubMed: 12809610]
5. Yu XF. Innate cellular defenses of APOBEC3 cytidine deaminases and viral counter-defenses. *Curr Opin HIV AIDS*. 2006; 1:187–93. [PubMed: 19372807]
6. Prochnow C, Bransteitter R, Chen XS. APOBEC deaminases-mutases with defensive roles for immunity. *Sci China C Life Sci*. 2009; 52:893–902. [PubMed: 19911124]
7. Harris RS, Liddament MT. Retroviral restriction by APOBEC proteins. *Nat Rev Immunol*. 2004; 4:868–77. [PubMed: 15516966]
8. Duggal NK, Emerman M. Evolutionary conflicts between viruses and restriction factors shape immunity. *Nat Rev Immunol*. 2012; 12:687–95. [PubMed: 22976433]
9. Moris A, Murray S, Cardinaud S. AID and APOBECs span the gap between innate and adaptive immunity. *Front Microbiol*. 2014; 5:534. [PubMed: 25352838]
10. Koito A, Ikeda T. Intrinsic immunity against retrotransposons by APOBEC cytidine deaminases. *Front Microbiol*. 2013; 4:28. [PubMed: 23431045]
11. Thielen BK, McNevin JP, McElrath MJ, Hunt BVS, Klein KC, Lingappa JR. Innate immune signaling induces high levels of TC-specific deaminase activity in primary monocyte-derived cells through expression of APOBEC3A isoforms. *The Journal of biological chemistry*. 2010; 285:27753–66. [PubMed: 20615867]
12. Russell RA, Smith J, Barr R, Bhattacharyya D, Pathak VK. Distinct domains within APOBEC3G and APOBEC3F interact with separate regions of human immunodeficiency virus type 1 Vif. *J Virol*. 2009; 83:1992–2003. [PubMed: 19036809]
13. Farrow MA, Sheehy AM. Vif and Apobec3G in the innate immune response to HIV: a tale of two proteins. *Future Microbiol*. 2008; 3:145–54. [PubMed: 18366335]
14. Russell RA, Pathak VK. Identification of two distinct human immunodeficiency virus type 1 Vif determinants critical for interactions with human APOBEC3G and APOBEC3F. *J Virol*. 2007; 81:8201–10. [PubMed: 17522216]
15. Takaori-Kondo A. APOBEC family proteins: novel antiviral innate immunity. *Int J Hematol*. 2006; 83:213–6. [PubMed: 16720550]
16. Cho SJ, Drechsler H, Burke RC, Arens MQ, Powderly W, Davidson NO. APOBEC3F and APOBEC3G mRNA levels do not correlate with human immunodeficiency virus type 1 plasma viremia or CD4+ T-cell count. *J Virol*. 2006; 80:2069–72. [PubMed: 16439564]
17. Zheng YH, Irwin D, Kurosu T, Tokunaga K, Sata T, Peterlin BM. Human APOBEC3F is another host factor that blocks human immunodeficiency virus type 1 replication. *J Virol*. 2004; 78:6073–6. [PubMed: 15141007]

18. Burns MB, Lackey L, Carpenter MA, Rathore A, Land AM, Leonard B, Refsland EW, Kotandeniya D, Tretyakova N, Nikas JB, Yee D, Temiz NA, Donohue DE, McDougale RM, Brown WL, Law EK, Harris RS. APOBEC3B is an enzymatic source of mutation in breast cancer. *Nature*. 2013; 494:366–70. [PubMed: 23389445]
19. Burns MB, Temiz NA, Harris RS. Evidence for APOBEC3B mutagenesis in multiple human cancers. *Nat Genet*. 2013; 45:977–83. [PubMed: 23852168]
20. Nik-Zainal S, Wedge DC, Alexandrov LB, Petljak M, Butler AP, Bolli N, Davies HR, Knappskog S, Martin S, Papaemmanuil E, Ramakrishna M, Shlien A, Simoncic I, Xue Y, Tyler-Smith C, Campbell PJ, Stratton MR. Association of a germline copy number polymorphism of APOBEC3A and APOBEC3B with burden of putative APOBEC-dependent mutations in breast cancer. *Nat Genet*. 2014; 46:487–91. [PubMed: 24728294]
21. Roberts SA, Lawrence MS, Klimczak LJ, Grimm SA, Fargo D, Stojanov P, Kiezun A, Kryukov GV, Carter SL, Saksena G, Harris S, Shah RR, Resnick MA, Getz G, Gordenin DA. An APOBEC cytidine deaminase mutagenesis pattern is widespread in human cancers. *Nat Genet*. 2013; 45:970–6. [PubMed: 23852170]
22. Conticello SG. Creative deaminases, self-inflicted damage, and genome evolution. *Ann N Y Acad Sci*. 2012; 1267:79–85. [PubMed: 22954220]
23. Guo JU, Su Y, Zhong C, Ming GL, Song H. Hydroxylation of 5-methylcytosine by TET1 promotes active DNA demethylation in the adult brain. *Cell*. 2011; 145:423–34. [PubMed: 21496894]
24. Nik-Zainal S, Alexandrov LB, Wedge DC, Van Loo P, Greenman CD, Raine K, Jones D, Hinton J, Marshall J, Stebbings LA, Menzies A, Martin S, Leung K, Chen L, Leroy C, Ramakrishna M, Rance R, Lau KW, Mudie LJ, Varela I, McBride DJ, Bignell GR, Cooke SL, Shlien A, Gamble J, Whitmore I, Maddison M, Tarpey PS, Davies HR, Papaemmanuil E, Stephens PJ, McLaren S, Butler AP, Teague JW, Jonsson G, Garber JE, Silver D, Miron P, Fatima A, Boyault S, Langerod A, Tutt A, Martens JW, Aparicio SA, Borg A, Salomon AV, Thomas G, Borresen-Dale AL, Richardson AL, Neuberger MS, Futreal PA, Campbell PJ, Stratton MR. Mutational processes molding the genomes of 21 breast cancers. *Cell*. 2012; 149:979–93. [PubMed: 22608084]
25. Conticello SG, Harris RS, Neuberger MS. The Vif protein of HIV triggers degradation of the human antiretroviral DNA deaminase APOBEC3G. *Curr Biol*. 2003; 13:2009–13. [PubMed: 14614829]
26. Mariani R, Chen D, Schrofelbauer B, Navarro F, Konig R, Bollman B, Munk C, Nymark-McMahon H, Landau NR. Species-specific exclusion of APOBEC3G from HIV-1 virions by Vif. *Cell*. 2003; 114:21–31. [PubMed: 12859895]
27. Sheehy AM, Gaddis NC, Malim MH. The antiretroviral enzyme APOBEC3G is degraded by the proteasome in response to HIV-1 Vif. *Nature Medicine*. 2003; 9:1404–1407.
28. Wang X, Dolan PT, Dang Y, Zheng YH. Biochemical differentiation of APOBEC3F and APOBEC3G proteins associated with HIV-1 life cycle. *J Biol Chem*. 2007; 282:1585–94. [PubMed: 17142455]
29. Liddament MT, Brown WL, Schumacher AJ, Harris RS. APOBEC3F properties and hypermutation preferences indicate activity against HIV-1 in vivo. *Curr Biol*. 2004; 14:1385–91. [PubMed: 15296757]
30. Belanger K, Langlois MA. Comparative analysis of the gene-inactivating potential of retroviral restriction factors APOBEC3F and APOBEC3G. *J Gen Virol*. 2015; 96:2878–87. [PubMed: 26048885]
31. Wang X, Li X, Ma J, Zhang L, Ma L, Mi Z, Zhou J, Guo F, Kleiman L, Cen S. Human APOBEC3F incorporation into human immunodeficiency virus type 1 particles. *Virus research*. 2014; 191C:30–38.
32. Song C, Sutton L, Johnson ME, D'Aquila RT, Donahue JP. Signals in APOBEC3F N-terminal and C-terminal deaminase domains each contribute to encapsidation in HIV-1 virions and are both required for HIV-1 restriction. *J Biol Chem*. 2012; 287:16965–74. [PubMed: 22451677]
33. Zennou V, Perez-Caballero D, Gottlinger H, Bieniasz PD. APOBEC3G incorporation into human immunodeficiency virus type 1 particles. *J Virol*. 2004; 78:12058–61. [PubMed: 15479846]

34. Wang T, Tian C, Zhang W, Sarkis PT, Yu XF. Interaction with 7SL RNA but not with HIV-1 genomic RNA or P bodies is required for APOBEC3F virion packaging. *J Mol Biol.* 2008; 375:1098–112. [PubMed: 18067920]
35. Navarro F, Bollman B, Chen H, Konig R, Yu Q, Chiles K, Landau NR. Complementary function of the two catalytic domains of APOBEC3G. *Virology.* 2005; 333:374–386. [PubMed: 15721369]
36. Li X, Ma J, Zhang Q, Zhou J, Yin X, Zhai C, You X, Yu L, Guo F, Zhao L, Li Z, Zeng Y, Cen S. Functional analysis of the two cytidine deaminase domains in APOBEC3G. *Virology.* 2011; 414:130–6. [PubMed: 21489586]
37. Wang T, Tian C, Zhang W, Luo K, Sarkis PT, Yu L, Liu B, Yu Y, Yu XF. 7SL RNA mediates virion packaging of the antiviral cytidine deaminase APOBEC3G. *J Virol.* 2007; 81:13112–24. [PubMed: 17881443]
38. Apolonia L, Schulz R, Curk T, Rocha P, Swanson CM, Schaller T, Ule J, Malim MH. Promiscuous RNA binding ensures effective encapsidation of APOBEC3 proteins by HIV-1. *PLoS Pathog.* 2015; 11:e1004609. [PubMed: 25590131]
39. Donahue JP, Levinson RT, Sheehan JH, Sutton L, Taylor HE, Meiler J, D'Aquila RT, Song C. Genetic analysis of the localization of APOBEC3F to human immunodeficiency virus type 1 virion cores. *J Virol.* 2015; 89:2415–24. [PubMed: 25505075]
40. Liu B, Sarkis PT, Luo K, Yu Y, Yu XF. Regulation of Apobec3F and human immunodeficiency virus type 1 Vif by Vif-Cul5-ElonB/C E3 ubiquitin ligase. *J Virol.* 2005; 79:9579–87. [PubMed: 16014920]
41. Smith JL, Pathak VK. Identification of specific determinants of human APOBEC3F, APOBEC3C, and APOBEC3DE and African green monkey APOBEC3F that interact with HIV-1 Vif. *J Virol.* 2010; 84:12599–608. [PubMed: 20943965]
42. Nakashima M, Ode H, Kawamura T, Kitamura S, Naganawa Y, Awazu H, Tsuzuki S, Matsuoka K, Nemoto M, Hachiya A, Sugiura W, Yokomaku Y, Watanabe N, Iwatani Y. Structural Insights into HIV-1 Vif-APOBEC3F Interaction. *J Virol.* 2015
43. Siu KK, Sultana A, Azimi FC, Lee JE. Structural determinants of HIV-1 Vif susceptibility and DNA binding in APOBEC3F. *Nature communications.* 2013; 4:2593.
44. Holmes RK, Koning FA, Bishop KN, Malim MH. APOBEC3F can inhibit the accumulation of HIV-1 reverse transcription products in the absence of hypermutation. Comparisons with APOBEC3G. *J Biol Chem.* 2007; 282:2587–95. [PubMed: 17121840]
45. Aydin H, Taylor MW, Lee JE. Structure-guided analysis of the human APOBEC3-HIV restrictome. *Structure.* 2014; 22:668–84. [PubMed: 24657093]
46. Stenglein MD, Harris RS. APOBEC3B and APOBEC3F inhibit L1 retrotransposition by a DNA deamination-independent mechanism. *J Biol Chem.* 2006; 281:16837–41. [PubMed: 16648136]
47. Bélanger K, Savoie M, Rosales Gerpe MC, Couture JF, Langlois MA. Binding of RNA by APOBEC3G controls deamination-independent restriction of retroviruses. *Nucleic acids research.* 2013; 41:7438–52. [PubMed: 23761443]
48. Albin JS, Brown WL, Harris RS. Catalytic activity of APOBEC3F is required for efficient restriction of Vif-deficient human immunodeficiency virus. *Virology.* 2014; 450–451:49–54.
49. Mbisa JL, Bu W, Pathak VK. APOBEC3F and APOBEC3G inhibit HIV-1 DNA integration by different mechanisms. *J Virol.* 2010; 84:5250–9. [PubMed: 20219927]
50. Monajemi M, Woodworth CF, Zipperlen K, Gallant M, Grant MD, Larijani M. Positioning of APOBEC3G/F mutational hotspots in the human immunodeficiency virus genome favors reduced recognition by CD8+ T cells. *PLoS One.* 2014; 9:e93428. [PubMed: 24722422]
51. Hakata Y, Landau NR. Reversed functional organization of mouse and human APOBEC3 cytidine deaminase domains. *J Biol Chem.* 2006; 281:36624–31. [PubMed: 17020885]
52. Holden LG, Prochnow C, Chang YP, Bransteitter R, Chelico L, Sen U, Stevens RC, Goodman MF, Chen XS. Crystal structure of the anti-viral APOBEC3G catalytic domain and functional implications. *Nature.* 2008; 456:121–4. [PubMed: 18849968]
53. Newman EN, Holmes RK, Craig HM, Klein KC, Lingappa JR, Malim MH, Sheehy AM. Antiviral function of APOBEC3G can be dissociated from cytidine deaminase activity. *Curr Biol.* 2005; 15:166–70. [PubMed: 15668174]

54. Fu Y, Ito F, Zhang G, Fernandez B, Yang H, Chen XS. DNA Cytosine and Methylcytosine Deamination by APOBEC3B: Enhancing Methylcytosine Deamination by Engineering APOBEC3B. *Biochem J*. 2015
55. Logue EC, Bloch N, Dhuey E, Zhang R, Cao P, Herate C, Chauveau L, Hubbard SR, Landau NR. A DNA sequence recognition loop on APOBEC3A controls substrate specificity. *PLoS One*. 2014; 9
56. Rathore A, Carpenter MA, Demir O, Ikeda T, Li M, Shaban NM, Law EK, Anokhin D, Brown WL, Amaro RE, Harris RS. The local dinucleotide preference of APOBEC3G can be altered from 5'-CC to 5'-TC by a single amino acid substitution. *J Mol Biol*. 2013; 425:4442–54. [PubMed: 23938202]
57. Wang M, Rada C, Neuberger MS. Altering the spectrum of immunoglobulin V gene somatic hypermutation by modifying the active site of AID. *J Exp Med*. 2010; 207:141–53. [PubMed: 20048284]
58. Mitra M, Hercik K, Byeon IJ, Ahn J, Hill S, Hinchee-Rodriguez K, Singer D, Byeon CH, Charlton LM, Nam G, Heidecker G, Gronenborn AM, Levin JG. Structural determinants of human APOBEC3A enzymatic and nucleic acid binding properties. *Nucleic Acids Res*. 2014; 42:1095–110. [PubMed: 24163103]
59. Byeon IJ, Ahn J, Mitra M, Byeon CH, Hercik K, Hritz J, Charlton LM, Levin JG, Gronenborn AM. NMR structure of human restriction factor APOBEC3A reveals substrate binding and enzyme specificity. *Nat Commun*. 2013; 4:1890. [PubMed: 23695684]
60. Ara A, Love RP, Chelico L. Different mutagenic potential of HIV-1 restriction factors APOBEC3G and APOBEC3F is determined by distinct single-stranded DNA scanning mechanisms. *PLoS Pathog*. 2014; 10:e1004024. [PubMed: 24651717]
61. Langlois MA, Beale RC, Conticello SG, Neuberger MS. Mutational comparison of the single-domained APOBEC3C and double-domained APOBEC3F/G anti-retroviral cytidine deaminases provides insight into their DNA target site specificities. *Nucleic Acids Res*. 2005; 33:1913–23. [PubMed: 15809227]
62. Kohli RM, Abrams SR, Gajula KS, Maul RW, Gearhart PJ, Stivers JT. A portable hot spot recognition loop transfers sequence preferences from APOBEC family members to activation-induced cytidine deaminase. *J Biol Chem*. 2009; 284:22898–904. [PubMed: 19561087]
63. Kohli RM, Maul RW, Guminski AF, McClure RL, Gajula KS, Saribasak H, McMahon MA, Siliciano RF, Gearhart PJ, Stivers JT. Local sequence targeting in the AID/APOBEC family differentially impacts retroviral restriction and antibody diversification. *J Biol Chem*. 2010; 285:40956–64. [PubMed: 20929867]
64. 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; 81:3807–15. [PubMed: 17267497]
65. Bohn MF, Shandilya SM, Albin JS, Kouno T, Anderson BD, McDougale RM, Carpenter MA, Rathore A, Evans L, Davis AN, Zhang J, Lu Y, Somasundaran M, Matsuo H, Harris RS, Schiffer CA. Crystal structure of the DNA cytosine deaminase APOBEC3F: the catalytically active and HIV-1 Vif-binding domain. *Structure*. 2013; 21:1042–50. [PubMed: 23685212]

Highlights

- Highly purified full length and various mutational APOBEC3F proteins from *E. coli* expression.
- The N-terminal domain (CD1) of A3F greatly enhances the deamination activity of the C-terminal domain (CD2).
- The CD1 of A3F is catalytically inactive, but greatly enhances the ssDNA binding affinity of the CD2 domain.
- The loop 7 of A3F CD1 is critical for ssDNA binding by the CD1 domain.
- The loop 7 of A3F CD2 is important for ssDNA binding of CD2 domain.
- The conserved tryptophan residue on the loop 7 of both CD1 and CD2 regulates the substrate DNA binding and deaminase activities.

Figure 1. Deamination activities of purified wt and mutant full-length (fl) A3F proteins (A). Cartoon representation of the four A3F fl constructs in vector pMAL-c5X (New England Biolabs), all expressed as N-terminal MBP fusion proteins, with a PreScission protease cleavage site inserted between MBP and A3F constructs. (B) The SDS-PAGE analysis of the purified A3F proteins from *E. coli* expression. The MBP-A3F fusion proteins are purified as described for MBP-A3B fusions in Fu et al.⁵⁴. Briefly, after expressing the A3F proteins in *E. coli* cells at 16 °C for 18 hrs, clarified supernatant of cell lysates in a buffer containing 50 mM Tris-Cl pH 8.0, 150 mM NaCl, 0.5 mM TCEP was passed through an amylose resin column. After washing the column with 10 column volumes of wash buffer (50 mM Tris-HCL pH 8.0, 1.0 M NaCl, 0.5 mM TCEP, 0.1 μg/ml RnaseA), MBP-A3F proteins were then eluted with buffer containing 20 mM Maltose. The elution was concentrated and further purified using S200 gel filtration chromatography. The peak fractions containing the pure MBP-A3F proteins from gel filtration were collected, concentrated, aliquoted, and stored at -80 °C. (C). Flowchart of deamination assay on 3' FAM labeled ssDNA substrate containing a single target C. Briefly, various amount of purified proteins were incubated with 600nM FAM-labeled ssDNA (5'-ATT TAT ATT ATT TAT TCA TAT TTA TAT TTA- FAM -3') in deamination buffer (HEPEs pH 6.5, 5 mM CaCl₂, 1 mM DTT, 0.1 μg/ml RnaseA) at 37 °C for 3 hrs. The reaction was terminated by

heating to 95 °C for 10 mins. Then 2 units of UDG was added to the reaction mixture and incubated for 1 hr at 37 °C, followed by heating to 95°C for 10 mins to terminate the UDG reaction. After incubating with 0.1M NaOH at 95 °C for 10 mins, deamination products were separated on 20% urea-PAGE, visualized with Molecular Imager FX (Bio-Rad), and quantified with Quantity One 1-D analysis software (Bio-Rad). **(D)**. A typical gel analysis of the deamination assay of the four purified A3F proteins shown in panels-A, -B. Deaminase assay with purified A3A served as a positive control for the 13 nt deamination product. **(E, F)**. Dose-dependent deamination activity of wt A3F-fl and A3F-fl E67A mutant (panel-E), and the bar representation of the deaminase activity (see Table 1) for the two A3F constructs. Error bars represent standard deviation from the mean of three independent experiments.

Figure 3. Roles of CD1 loop 7 W126 for A3F DNA binding and deaminase activities

(A). A view of the A3F-CD1 structure that is modeled based on A3F-CD2 crystal structures^{42; 43; 65}, showing CD1 loop 7 and W126 (in yellow sticks) mutated in this study. (B). Sequence alignment of loop 7 of both CD1 and CD2 from A3F and A3G, showing the close homology between A3G-CD1, and A3F CD1 and CD2, with the conserved W underlined. (C). Three A3F constructs, and the SDS-PAGE analysis of their purified proteins. (D). Gel shift assay of ssDNA binding of A3F-CD1 (top) and A3F-CD1 W126A (bottom) constructs (see notes in Table 2 for details). The same 30 nt ssDNA in deamination assay was used for all ssDNA binding assay. The shifted band for the protein-DNA complex is clearly visible for the wt, and only slightly smear tailing was observed at high protein concentration for the W126A mutant, indicating loss of binding. (E). ssRNA binding of wt and W126A mutant of A3F CD1 (top) and full length (fl) constructs (bottom). A3F CD1 wt showed discrete band shift from 0.05μM. However, only band shift to larger complex started to show up at 0.45μM of A3F-CD1 W126A mutant, indicating disruption of RNA binding by W126A mutation. Even though the band shifting pattern is different when assayed using the A3F-fl and its W126A mutant (bottom), the results also confirmed that W126A mutation disrupted RNA binding in the context of full length. Proteins were incubated with 15 nM 50 nt ssRNA (5'FAM-AUU AUU AUU AUU UAU CCC UAU UUA UAU UUA UUG UUA UUG UUA UUG UUA UA 3') in the buffer (50mM Hepes 6.5, 5mM CaCl₂, 1mM DTT, 40 unit RNaseOUT Ribonuclease Inhibitor (Invitrogen)), and analyzed on 1% agarose gel electrophoresis. (F, G) The ssDNA binding curves of wt A3F-CD1 (panel-F), and a comparison of binding curves for A3F-fl, A3F-fl W126A, and A3F-CD2 (panel-G) (see notes in Table 2 for details). The calculated K_d values for each construct were listed in Table 2. (H) The comparison of the deamination activity for A3F-fl, A3F-fl W126A, and A3F-

CD2 (also see Table 1). Error bars represent standard deviation from the mean of three independent experiments.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Figure 4.

The Circular Dichroism (CD) profiles and the hotspot motif preference of deamination for A3F wt and its various mutants. (**A, B, C**). CD scan results for A3F-fl and its two mutants W310A and D311A (panel-A), for A3F-CD2 and its two mutants W310A and D311A (panel-B), for A3F-CD1 and its mutant W126A (panel-C). The results indicate very similar CD profiles for the wt and the mutants, suggesting similar structural content for each. Methods: CD experiments were conducted using a Jasco J-815 spectropolarimeter.

Wavelength scans were performed with 0.1 mg/mL proteins in buffer-S (50 mM Tris 8.0, 150 mM NaF, 0.5 mM TCEP) in a 0.1 cm cuvette in a water-jacketed cell at 25°C. Spectra were averaged over 3 consecutive scans collected from 200 to 250 nm. The individual scans were recorded using a scan rate of 30 nm/min, a bandwidth of 0.5 nm and a response time of 1 s per point. Buffer scans were accumulated and subtracted from the sample scans and the mean residue ellipticity was computed for each point based on the known concentration, molecular weight, and number of residues for each construct. (**D**). Thermal melt scan results for A3F-fl and its two mutants W126A and W310A. Shown is the estimated fraction of the protein present in the unfolded state based on values calculated for mean residue ellipticity. Subsequent regression analysis determined the melting temperature of all three constructs to be 63.9 ± 0.46 °C and the overall profile overlap suggests the three constructs have similar stability. Methods: Thermal melt experiments were conducted with 0.1 mg/mL A3F-fl, A3F-fl W126A and A3F-fl W310A in buffer-S in a 0.1 cm cuvette in a water-jacketed cell. Spectra were collected at 220nm. The individual scans were recorded every 0.3s, and the temperature increased 1°C/min from 25°C to 95°C. Buffer scans were accumulated and

subtracted from the sample scans and the mean residue ellipticity was computed. Regression analysis was completed using the Gibbs-Helmholtz equation to fit the unfolded state as a function of temperature; the derived values for ellipticity of the unfolded and folded states for each construct were then calculated for each point and fraction unfolded was determined as $\alpha = 1 - (\theta_o - \theta_u)/(\theta_f - \theta_u)$, with θ_o, θ_u , and θ_f being the observed, unfolded, and folded values for mean residue ellipticity at each point. (**E, F**). The deamination hotspot motif preference for the wt and mutants in the context of A3F-fl and A3F-CD2. The result showed that “TC” is the preferred motif for the wt and mutant proteins. The four ssDNA substrates differ only in the hotspot motifs are: “AC” (5′ FAM-ATT TAT ATT ATT TAT ACA TAT TTA TAT TTA 3′), “TC” (5′ FAM-ATT TAT ATT ATT TAT **TCA** TAT TTA TAT TTA 3′), “GC”(5′ FAM-ATT TAT ATT ATT TAT **GCA** TAT TTA TAT TTA 3′), “CC”(5′ ATT TAT ATT ATT TAT **TCC** TAT TTA TAT TTA-FAM 3′). Error bars represent standard deviation from the mean of three independent experiments.

Table 1

Deaminase activity of purified proteins of various A3F constructs.

Constructs	A3F-fl activity (nM/ μ M/hr)	A3F-CD2 activity (nM/ μ M/hr)
Wt	7.59 $\pm$ 0.981	0.501 $\pm$ 0.00434
E67A (in CD1)	5.89 $\pm$ 0.753	—
E251A (in CD2)	No	—
E67A/E251A (in CD1/CD2)	No	—
W310A (in CD2)	3.23 $\pm$ 0.435	0.00478 $\pm$ 0.000427
D311A (in CD2)	5.89 $\pm$ 0.292	0.105 $\pm$ 0.0155
D311K (in CD2)	10.50 $\pm$ 0.663	0.609 $\pm$ 0.0439
T312P (in CD2)	7.32 $\pm$ 0.494	0.523 $\pm$ 0.0283
W126A (in CD1)	1.90 $\pm$ 0.344	

Note: The deaminase activity data in this table were expressed as (Product (nM)/Enzyme (μ M)/Reaction time (hr)), which are obtained by linear regression fitting (using GraphPad Prism 6 software) based on the data at the beginning of linear range of the deaminase assay. Details for deaminase reaction conditions and substrate used are described in Figure 1C legend.

Table 2

ssDNA (30 nt) binding affinity of purified proteins of various A3F constructs.

Constructs	A3F-fl (μM)	A3F-CD2 (μM)	A3F-CD1 (μM)
wt	1.101 $\pm$ 0.3354	12.80 $\pm$ 2.028	0.146 $\pm$ 0.0047
W310A	3.447 $\pm$ 1.034	152.5 $\pm$ 30.96	—
D311A	1.796 $\pm$ 0.4253	32.12 $\pm$ 7.369	—
D311K	1.099 $\pm$ 0.3590	14.56 $\pm$ 2.346	—
T312P	0.917 $\pm$ 0.2173	13.65 $\pm$ 2.111	—
W126A	3.931 $\pm$ 0.4170	—	NO

Note: DNA binding assay was performed using native gel shift. Increasing amount of A3F proteins were incubated with 15 nM FAM-labeled ssDNA (5'-ATT TAT ATT ATT TAT TCA TAT TTA TAT TTA- FAM -3') at 25 °C for 5 mins, followed by native gel electrophoresis for 100 min at 4 °C. The FAM-labeled DNA and its complex with the protein on the gel were then visualized with Molecular Imager FX (Bio-Rad), and quantified with Quantity One 1-D analysis software (Bio-Rad). Based on the ratios between free ssDNA and ssDNA-protein complex of each reaction, a protein concentration dependent ssDNA binding curve was created by "one site specific binding" model from Prism 6. Dissociation Constant (Kd) of each A3F construct was calculated using the same software.