Crystal Structure of DNA Cytidine Deaminase ABOBEC3G Catalytic Deamination Domain Suggests a Binding Mode of Full-length Enzyme to Single-stranded DNA*

Received for publication, November 6, 2014, and in revised form, December 18, 2014 Published, JBC Papers in Press, December 25, 2014, DOI 10.1074/jbc.M114.624262

Xiuxiu Lu1, Tianlong Zhang5, Zeng Xu1, Shanshan Liu9, Bin Zhao1, Wenxian Lan1, Chunxi Wang1, Jianping Ding5, and Chunyang Cao†1

From the ‡State Key Laboratory of Bio-organic and Natural Product Chemistry, Shanghai Institute of Organic Chemistry, Chinese Academy of Sciences, 345 Lingling Road, Shanghai 200032, China and the §Institute of Biochemistry and Cell Biology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, 320 Yue Yang Road, Shanghai 200031, China

APOBEC3G (A3G) is a DNA cytidine deaminase (CD) that demonstrates antiviral activity against human immunodeficiency virus 1 (HIV-1) and other pathogenic viruses. It has an inactive N-terminal CD1 virus infectivity factor (Vif) protein binding domain (A3G-CD1) and an actively catalytic C-terminal CD2 deamination domain (A3G-CD2). Although many studies on the structure of A3G-CD2 and enzymatic properties of full-length A3G have been reported, the mechanism of how A3G interacts with HIV-1 single-stranded DNA (ssDNA) is still not well characterized. Here, we report a crystal structure of a novel A3G-CD2 head-to-tail dimer (in which the N terminus of the monomer H (head) interacts with the C terminus of monomer T (tail)), where a continuous DNA binding groove was observed. By constructing the A3G-CD1 structural model, we found that its overall fold was almost identical to that of A3G-CD2. We mutated the residues located in or along the groove in monomer H and the residues in A3G-CD1 that correspond to those seated in or along the groove in monomer T. Then, by performing enzymatic assays, we confirmed the reported key elements and the residues in A3G necessary to the catalytic deamination. Moreover, we identified more than 10 residues in A3G essential to DNA binding and deamination reaction. Therefore, this dimer structure may represent a structural model of full-length A3G, which indicates a possible binding mode of A3G to HIV-1 ssDNA.

The APOBEC protein family are activation-induced deaminases, including APOBEC1, APOBEC2, APOBEC3, and APOBEC4 (1–8). Among them, the APOBEC3 subfamily consists of cellular proteins that prevent the mobilization of diverse retroviruses, retrotransposons, and other viral pathogens (4). In humans, seven members of this subfamily, APOBEC3A (A3A),2 APOBEC3B (A3B), APOBEC3C (A3C), APOBEC3D (A3D), APOBEC3F (A3F), APOBEC3G (A3G), and APOBEC3H (A3H), are encoded in a tandem array on chromosome 22 (9). These proteins are characterized by the presence of one or two conserved zinc-coordinated domains consisting of HXEX23–28CX2C motifs (10). Human A3G is a single-stranded DNA (ssDNA) cytidine deaminase that restricts replication of HIV-1 in strains lacking the virus infectivity factor protein (Vif) (11). HIV-1 Vif protein facilitates polyubiquitination of A3G, leading to proteasomal degradation (12–17). Without Vif, A3G is encapsulated in an HIV virus particle through interaction with the nucleocapsid portion of Gag and HIV RNA or 7SL RNA (18–27). A3G suppresses HIV-1 infectivity by deaminating the viral cDNA cytidine to uridine during the reverse transcription (25–27). Subsequent replication of the cDNA generates the hallmark G to A hypermutations, causing proviral inactivation (25–27).

Human A3G exists as monomer, dimer, and tetramer, depending on the DNA substrate and salt concentration. It possesses two homologous deaminase domains, an inactive N-terminal CD1 domain (i.e. A3G-CD1) required for Vif, DNA, and RNA binding and an active C-terminal CD2 domain (i.e. A3G-CD2) required for catalysis and motif specificity (28–30). The CD1 domain is also suggested to be required for the incorporation of A3G into virions (29). A3G deaminates cytidine processively 3’ → 5’ on ssDNA (31, 32). The processive deamination reactions have been decided in a non-random way (31, 32). To date, the three-dimensional structure of the free A3G-CD2 domain has been determined by NMR (33–35) and x-ray crystallography (36–38). The three-dimensional structures of free APOBEC2 (39) and other APOBEC3 sub-

* This work was supported by Ministry of Science and Technology of China Grant 2011CB966300; National Natural Science Foundation of China Grants 21472229, 20902103, and 21272261; and Science and Technology Commission of Shanghai Municipality Grant 15ZR1449300.

1 To whom correspondence should be addressed. Tel.: 86-21-54925491; Fax: 86-21-64166128; E-mail: ccao@mail.sioc.ac.cn.

2 The abbreviations used are: A3A, APOBEC3A; A3B, APOBEC3B; A3C, APOBEC3C; A3D, APOBEC3D; A3F, APOBEC3F; A3G, APOBEC3G; A3H, APOBEC3H; CD, cytidine deaminase; Vif, virus infectivity factor protein; A3G-CD1, inactive A3G N-terminal Vif binding domain; A3G-CD2, active A3G C-terminal catalytic domain; monomer H, head monomer; monomer T, tail monomer; Bistris propane, 1,3-bis[tris(hydroxymethyl)methylamino]propane; PDB, Protein Data Bank.
family members, such as A3A (40, 41), A3C (42), and A3F (43, 44), have also been reported. The structural basis for Vif hijacking CBF-β and CUL5 E3 ligase was recently revealed too (45). However, the mechanism of how A3G-CD2 or full-length A3G and other members of its family interact with ssDNA is still not well understood.

To address how A3G-CD2 interacts with HIV-1 ssDNA, we crystallized A3G-CD2 in the presence of ssDNA containing one target motif sequence 5'-CCG-3'. A novel head-to-tail dimer structure of A3G-CD2 was obtained. In its surface, a continuous groove for ssDNA binding was found. Our structural analysis and biochemical assays suggest that this dimer structure may represent a structural model of full-length A3G. From this model, we identified more than 10 new residues in both A3G-CD1 and A3G-CD2 critical to the deamination catalysis reaction.

EXPERIMENTAL PROCEDURES

Expression and Purification of A3G-CD2 and Its Variants—The DNA corresponding to gene of wild-type (WT) A3G-CD2 (residues 193–384) or its variant was cloned into a PET28a vector containing an N-terminal His tag and thrombin cleavage sites. A3G-CD2 and its variants were expressed in Rosetta (DE3)plys Escherichia coli cells. Cell cultures were grown to 600 Value equal to about 0.8 and induced with a final concentration of 0.5 mM isopropyl 1-thio-β-δ-galactopyranoside for 20 h at 18 °C. Cells were resuspended in nickel-binding buffer (20 mM Tris-HCl, pH 7.5, 500 mM NaCl, 10 μM ZnCl₂, and 0.5 mM dithiothreitol (DTT)) with protease inhibitor (DNase I) and lysed at 15,000 p.s.i. using a hydraulic cell disruption system (constant System JINBO Benchtop) (Guangzhou Juneng Biology and Technology Co., Ltd., Guangzhou, China). The lysate was centrifuged at 12,000 rpm and 4 °C for 55 min to remove cellular debris prior to loading into a nickel-nitrilotriacetic acid resin (GE Healthcare). The protein was washed with nickel-binding buffer 20 mM imidazole and then eluted by a five-step gradient of nickel-binding buffer with 40 mM imidazole, 100 mM imidazole, 250 mM imidazole, and 500 mM imidazole, respectively. Fractions containing A3G-CD2 or its variants were concentrated and purified on a Superdex75 16/600 GL column (GE Healthcare) previously equilibrated with buffer B (50 mM HEPES, pH 7.5, 50 mM NaCl, 50 μM ZnCl₂, 5 mM DTT).

X-ray Crystallization Screening and Data Collection—Purified A3G-CD2 and its D370A variant were quantified by A280 (19). The DNA corresponding to gene of wild-type (WT) A3G-CD2 protein and 91.8 and 8.2% of the residues of A3G-CD2 D370A variant were in most favored and allowed regions, respectively. The final model of WT protein contains residues 194–381 in monomer H and residues 196–246, 253–316, and 321–381 in monomer T. Weak electron density was observed for residues 246–253 and 316–321 in monomer T, whereas the final model of the D370A variant contains residues 193–381.

Real-time Studies of A3G-CD2-catalyzed Deamination by NMR—A series of one-dimensional 1H NMR spectra of the reported HIV-1 virus ssDNA with the sequence 5'-ATTC-CCAATT-3' (34) were acquired as a function of time at 20 °C in NMR buffer (50 mM Na₂HPO₄, 50 mM NaCl, pH 7.5, 50 μM ZnCl₀₂, 2 mM DTT), after adding concentrated A3G-CD2 solution. To accurately assign uridine NMR signals, three ssDNAs with sequences of 5'-ATTCUAATT-3', 5'-ATTCCUAATT-3', and 5'-ATTCCUAATT-3' were used. Concentrations of A3G-CD2 and its variants were fixed at 9.4 μM. The intensities of the 1H-NMR signal belonging to U² was used for quantification. Real-time monitoring of A3G-CD2 catalyzed deamination reaction by NMR was performed to extract initial rates (<5% dC → dU conversion) for a series of substrate concentration (37.5, 56.3, 93.8, 117.2, and 140.6 μM). Kₘ values were obtained using Michaelis-Menten module with the software Prism 5 (GraphPad Inc.).

E. coli-based Deaminase Activity Assays—The intrinsic DNA cytidine deaminase activity of full-length A3G and its variants was measured by expressing these proteins in ung-deficient E. coli BW310 and by quantifying the frequency of Rif⁶-confering rpoB mutations (33, 49), as described in the previous report (35). A lot of single-base mutations in rpoB lead to active site amino acid replacements that confer Rif⁶. In each case, five single colonies were grown at 37 °C in LB medium containing 100 μg/ml ampicillin and induced overnight by 1 mM isopropyl 1-thio-β-δ-galactopyranoside at 18 °C. Then appropriate volumes of cells were spread on plates containing 100 μg/ml rifampicin to select for Rif⁶ mutants and to plates containing 100 μg/ml ampicillin to determine the number of viable cells. Colonies were allowed to form overnight at 37 °C and then counted manually.

E. coli Immunoblots—The full-length A3G and its variant constructs were expressed in E. coli strain BW310. Proteins
were generated by overnight expression at 18 °C in LB medium containing 100 μg/ml ampicillin. To induce expression, cells were cultured 1:10 in LB medium containing 100 μg/ml ampicillin and 1 mM isopropyl 1-thio-β-D-galactopyranoside and grown for 1 h at 37 °C. Cells were pelleted and resuspended in SDS gel loading buffer (50 mM Tris-Cl, pH 6.8, 100 mM β-mercaptoethanol, 2% SDS (v/v) glycerol). Lysates were heated at 95 °C for 5 min and fractionated by SDS-PAGE. Proteins were transferred to a polyvinylidene difluoride (PVDF) membrane and probed with a rabbit anti-A3G polyclonal serum. The primary antibody was detected by incubation with donkey peroxidase-conjugated anti-rabbit IgG (Shanghai Sangon Biotech Co. Ltd.), followed by chemiluminescent imaging.

Size Exclusion Chromatography Assay—Size exclusion chromatography assays were performed on a Superdex 75 column (10/300 GL) (GE Healthcare), previously equilibrated with buffer B (50 mM HEPES, pH 7.5, 50 mM NaCl, 50 μM ZnCl₂, 5 mM DTT).

RESULTS

Overall Fold of Novel Head-to-tail A3G-CD2 Dimer Crystal Structure—In the presence of ssDNA during crystallization, two monomers of human A3G-CD2 domain (residues 193–384) occupy one asymmetric unit. The final crystal structure was determined at 1.8 Å resolution and solved by molecular replacement, using the structure of the A3G-CD2 2K3A variant (residues 191–384; PDB code 3IR2) (37) as a model. The final refinement statistics are summarized in Table 1. Different from the previously reported tail-to-tail or head-to-head dimer conformation (PDB code 3IR2) (37), the crystal structure demonstrates a new head-to-tail dimer conformation (to simplify discussion, the two monomers are referred to here as monomer H (head) and monomer T (tail), respectively), which means that the N terminus of the monomer H interacts with the C terminus of the monomer T, as shown in Fig. 1. The two monomers are identical to each other with a root mean square deviation value of 0.19 Å for backbone Cα atoms in the secondary structural regions. They contain a core sandwicht-like αβα sheet (36), significantly differing from those (discontinuous β2/β2’ sheets) in three NMR structures (PDB codes 2Y JW, 2KBO, and 2KEM) (33–35), the x-ray crystallographic tail-to-tail or head-to-head dimer structure (3IR2) (37), and the current dimer structure. The different regions include β1-β2, β2’-α2, and α2-β3, which are due to the ambiguity in electron density. The bulge between the β2-β2’ strands in the NMR structures and 3IR2 is obviously an intrinsic feature of A3G-CD2 structure rather than an experimental artifact (37). Thus, for correct comparison, monomer H in the current A3G-CD2 dimer was just overlapped with one monomer of the tail-to-tail A3G-CD2 2K3A variant dimer (Fig. 1), which produces root mean square deviation values of 1.64 Å for all of the backbone Cα atoms in the region of residues from 193 to 384 and of 0.30 Å for all of the backbone Cα atoms in the secondary structure regions. This observation indicates that the monomer conformation of the current A3G-CD2 dimer is almost identical to that of 3IR2. The main differences are located in loop 1 (residues 206–215), loop 3 (residues 245–256), and loop 7 (residues 315–320), which are all involved in ssDNA binding and stabilize the active deamination center. The orientations of the side chains of the residues in sequence P210/WVR213 in loop 1; of the residues His-248, His-249, Phe-252, and Glu-254 in loop 3; and of the residues in sequence Y315/DQO318 in loop 7 are apparently distinct from those in PDB entry 3IR2 (Fig. 2, A–C). Among these residues, Trp-211 (37),

| TABLE 1 | Summary of diffraction data and structure refinement statistics |
|---------|---------------------------------------------------------------|
| Data collection | A3G-CD2 | D370A |
| Wavelength (Å) | 0.9792 | 0.9792 |
| Space group | P21,21,2 | P21,21,2 |
| Resolution (Å) | 50.00–1.86 | 50.00–1.7 |
| Rfree (Ramchandran plot (%)) | 1.86–1.80 | (1.76–1.70) |
| No. of protein atoms | 2,983 | 1,591 |
| No. of water atoms | 441 | 190 |
| Average B factor of all atoms (Å²) | 33.2 | 29.2 |
| Protein | 27.4 | 20.0 |
| Water | 40.9 | 32.6 |
| Root mean square deviation | 0.007 | 0.007 |
| Bond angles (degrees) | 1.09 | 1.05 |
| Ramachandran plot (%) | 93.7 | 91.8 |
| Most favored regions | 6.3 | 8.2 |
| Allowed regions | 0.0 | 0.0 |
| Generously allowed regions | 0.0 | 0.0 |
| Disallowed regions | 0.0 | 0.0 |

Numbers in parentheses refer to the highest resolution shell.

a

Rfactor = Σ(|Fo| - |Fc|)/Σ|Fo|.

b

Rfree = Σ(|Fo| - |Fc|)/Σ|Fo|.

Differences between the Crystal Structures—The reported crystal structures of A3G-CD2 (residues 197–380) with PDB codes 3E1U and 3IQS have a continuous β2 sheet (36), significantly differing from those (discontinuous β2/β2’ sheets) in three NMR structures (PDB codes 2Y JW, 2KBO, and 2KEM) (33–35), the x-ray crystallographic tail-to-tail or head-to-head dimer structure (3IR2) (37), and the current dimer structure. The different regions include β1-β2, β2’-α2, and α2-β3, which are due to the ambiguity in electron density. The bulge between the β2-β2’ strands in the NMR structures and 3IR2 is obviously an intrinsic feature of A3G-CD2 structure rather than an experimental artifact (37). Thus, for correct comparison, monomer H in the current A3G-CD2 dimer was just overlapped with one monomer of the tail-to-tail A3G-CD2 2K3A variant dimer (Fig. 1), which produces root mean square deviation values of 1.64 Å for all of the backbone Cα atoms in the region of residues from 193 to 384 and of 0.30 Å for all of the backbone Cα atoms in the secondary structure regions. This observation indicates that the monomer conformation of the current A3G-CD2 dimer is almost identical to that of 3IR2. The main differences are located in loop 1 (residues 206–215), loop 3 (residues 245–256), and loop 7 (residues 315–320), which are all involved in ssDNA binding and stabilize the active deamination center. The orientations of the side chains of the residues in sequence P210/WVR213 in loop 1; of the residues His-248, His-249, Phe-252, and Glu-254 in loop 3; and of the residues in sequence Y315/DQO318 in loop 7 are apparently distinct from those in PDB entry 3IR2 (Fig. 2, A–C). Among these residues, Trp-211 (37),
Arg-213 (33, 34, 36, 37), Tyr-315 (36), and Asp-316 and -317 (36) have been confirmed to be crucial for the deaminase activities. Moreover, it has been suggested that the sequence Y317DDQ318 in loop 7 specifically recognizes the second cytosine in the target motif sequence of ssDNA (5’-CCC-3’), due to its polarity, whereas the sequence Y307YFW310 in loop 7 of A3F specifically identifies G or T in the ssDNA sequence 5’-(G/T)C-3’ (36, 43). Analysis on the surface of 3IR2 and the current dimer structure indicates that the conformational changes in loops 1, 3, and 7 result in a bigger ssDNA-binding groove in the monomer of 3IR2 than that in the current dimer structure (Fig. 2, D and E), thus probably enhancing DNA binding. Moreover, the positively charged side chain of Arg-213 points to the ssDNA binding groove in 3IR2 structure, strengthening DNA binding, but it deviates from the ssDNA binding groove in the current structure. These observations may account for the fact that the A3G-CD2 2K3A variant has deaminase activities ∼2.7-fold more than its WT protein (33).

*Intermolecular Interfaces in Head-to-tail A3G-CD2 Dimer—*

The obviously big difference between A3G-CD2 head-to-tail dimer and tail-to-tail dimer (or head-to-head dimer) (PDB code 3IR2) is the intermolecular interface. In monomer H of the current head-to-tail dimer, helix α2 and loop 3 form an L-shaped hook, which stabilizes the dimer conformation by interacting with helix α6’ and loop 1’ in monomer T. Thus, the current head-to-tail dimer contains three main interfaces, which are between loop 3 and loop 1’, between helix α2 and helix α6’, and between loop 3 and helix α6’ with surface areas of 223, 374, and 317 Å², respectively (Fig. 3A). The total surface area is 1032 Å², much larger than those in the tail-to-tail dimer conformation (901 Å²) and in the head-to-head dimer conformation (604 Å²), respectively. This indicates that the head-to-tail dimer conformation might be more stable than those of the tail-to-tail dimer and head-to-head dimer.

In the head-to-tail dimer conformation, the helix α2 in monomer H is almost perpendicular to the helix α6’ in monomer T, making up the largest interface. Residues Gly-373’, Ala-377’, and Gln-380’ in the C terminus of helix α6’ in monomer T form hydrogen bond nets by interacting with residues Leu-263 and Asp-264 in helix α2 in monomer H through several water molecules (Fig. 3, B and C). The side chain of Arg-376’ in monomer T has hydrophobic interactions with the side chain of Phe-268 in monomer H (Fig. 3D). To estimate the functional significance of this interface, four variants (D264A, F268A, R376A, and Q380A) were designed to disrupt the observed interactions. An NMR enzymatic assay was performed to measure the catalytic efficiency on the deamination at base C6 in the reported sequence of 5’-ATTCG-C6AATT-3’ (34). These variants show decreased DNA deaminase activity in vitro (Table 2), especially for the D264A variant. Compared with the WT A3G-CD2 (Kcat WT-CD2 = 0.11 ± 0.00049 min⁻¹, Km WT-CD2 = 20.42 ± 0.34 μM, Kcat WT-CD2/Km WT-CD2 = 5.42 × 10⁻³ min⁻¹ μM⁻¹), the catalytic efficiency Kcat D264A/Km D264A of the D264A variant is decreased about 20-fold (Kcat D264A/Km D264A = 0.11 ± 0.02 min⁻¹, Km D264A = 394.40 ± 85.43 μM, Kcat D264A/Km D264A = 2.78 × 10⁻⁴ min⁻¹ μM⁻¹). In the structure 3E1U, an intramolecular salt bridge was observed between the side chains of residues Asp-264 and Arg-256 in loop 3 (33)(Fig. 3E), which obviously stabilizes the conformation of loop 3. Therefore, the mutation from Asp-264 to Ala-264 not only disrupts

![Structure of APOBEC3G Catalytic Deamination Domain](image-url)
Structure of APOBEC3G Catalytic Deamination Domain

FIGURE 2. A–E, the differences in the side chain orientations of the residues in loop 1 (A), loop 3 (B), and loop 7 (C) of the current head-to-tail dimer (orange) and tail-to-tail dimer (green) result in the conformational changes in the ssDNA binding groove in the monomer of the 2K3A variant (PDB code 3IR2) (D) and of the current head-to-tail dimer (E), respectively. F, the mutation from Asp-370 to Ala-370 results in a wider ssDNA binding groove in the surface of the D370A variant than that in the WT protein.

the interactions between loop 3 in monomer H and the helix α6' in monomer T but also makes loop 3 more flexible, destabilizing the ssDNA binding center and thus further resulting in the reduced catalytic efficiency of the D264A variant. The mutation from Arg-376 to Ala-376 keeps its hydrophobic interactions with the side chain of Phe-268 (Fig. 3D) but destroys the hydrogen bond nets that stabilize the interactions between the helix α2 in monomer H and the helix α6’ in monomer T. Thus, the R376A variant has only 15% of the catalytic efficiency of the WT protein ($K_{cat}^{R376A} = 0.051 \pm 0.00039$ min$^{-1}$, decreased 2-fold; $K_m^{R376A} = 63.23 \pm 0.97$ μM, decreased about 3-fold; $K_{cat}^{R376A}/K_m^{R376A} = 8.10 \times 10^{-4}$ μM$^{-1}$, reduced 6.7-fold), consistent with the previous observation that Arg-376 is involved in the ssDNA binding activity and catalytic reaction (36, 37).

In the second interface, loop 3 in monomer H acts as a hook to catch the helix α6’ in monomer T. The side chains of the residues Gln-245 and Arg-256 form hydrogen bond nets that stabilize the interactions between the residues His-250 and Gly-251 in monomer H and the helix α6’ in monomer T (Fig. 3, F and G). To evaluate the contributions of these residues to the catalytic deamination, we replaced Gln-245, Arg-256, and Arg-374 with alanine. Compared with the WT protein, the Q245A and R256A variants nearly abolish the catalytic efficiency (for Q245A variant, $K_{cat}^{Q245A} = 0.083 \pm 0.019$ min$^{-1}$; $K_m^{Q245A} = 278.5 \pm 81.36$ μM, $K_{cat}^{Q245A}/K_m^{Q245A} = 2.99 \times 10^{-4}$ μM$^{-1}$, decreased about 20-fold; for R256A variant, $K_{cat}^{R256A} = 0.012 \pm 0.00018$ min$^{-1}$, $K_m^{R256A} = 163.90 \pm 4.17$ μM, $K_{cat}^{R256A}/K_m^{R256A} = 7.20 \times 10^{-5}$ μM$^{-1}$, reduced 75-fold). Obviously, these mutations destroy the hydrogen bond interactions observed above and thus account for the changes in the catalytic efficiency. These results are different from those observed in structure 3IR2, where a side chain of Gln-245 was coordinated to the zinc ion in the dimer interface (37).

The D370A variant has smaller $K_{cat}$ and $K_m$ values ($K_{cat}^{D370A} = 11.89 \pm 2.49$ μM, $K_{cat}^{D370A}/K_m^{D370A} = 0.010 \pm 0.00038$ μM$^{-1}$) than the WT protein, with only about 16% of the catalytic efficiency of the WT protein ($K_{cat}^{D370A}/K_m^{D370A} = 8.58 \times 10^{-4}$ μM$^{-1}$), suggesting that the D370A variant might have a stronger binding affinity to ssDNA than the WT protein (the DNA binding groove in the structure of the D370A variant becomes wider, as shown in Fig. 2F). We tried to crystallize the complex of D370A with ssDNA. Different from the WT protein, in the presence of ssDNA during crystallization, one asymmetric unit contains a D370A molecule. No electronic density was observed for ssDNA either. Its final crystal structure was determined at 1.7 Å resolution. One molecule of D370A forms a dimer with another molecule in an adjacent asymmetric unit in a head-to-tail way. The structure of the D370A variant reveals that the mutation from Asp-370 to Ala-370 not only directly results in the breakage of the hydrogen bond between Asp-370’ (in loop 7) and Gln-245 (in loop 3) (Fig. 3F) but also indirectly impairs the hydrophobic interaction between His-248 (in loop 3) and Trp-211’ (in loop 1) (Fig. 3F), and the intramolecular salt bridge interaction between Arg-374 (in...
helix a6) and Asp-316 (in loop 7) (the distances between the side chains of Arg-374 and Asp-316 become bigger in D370A than those in the WT protein, as shown in Fig. 3, K and L). Thus, the decrease in the catalytic efficiency of the A3G-CD2 D370A variant further reveals that the stability of the active center is important to the catalytic deamination reaction.

To assess the importance of Arg-374 in the cytidine deamination, we replaced Arg-374 by Ala-374. Compared with the WT protein, the R374A variant has a $K_m^{R374A}$ value almost equal to that of WT A3G-CD2 ($K_m^{R374A} = 27.05 \pm 7.48 \mu M$, $K_m^{WT-CD2} = 20.42 \pm 0.34 \mu M$), but its $K_{cat}^{R374A} = 0.005 \pm 0.00043 \text{ min}^{-1}$ is much less than that of WT protein ($K_{cat}^{WT-CD2} = 0.11 \pm 0.00049 \text{ min}^{-1}$), resulting in its catalytic efficiency ($K_{cat}^{R374A} / K_m^{R374A} = 1.85 \times 10^{-4} \text{ min}^{-1} \mu M^{-1}$) being decreased about 35-fold, consistent with previous observations (36, 37). On one hand, we think that the mutation from Arg-374 to Ala-374 may directly destroy the intramolecular salt bridge between Arg-374 (in helix a6) and Asp-316 (in loop 7) (Fig. 3K) and the intermolecular salt bridge between Arg-374 and Glu-209 (in loop 1) (which was observed in the tail-to-tail dimer conformation; Fig. 3M), both interactions making loops 1 and 7 more flexible. On the other hand, the replacement of Arg-374 by Ala-374 disrupts additional hydrogen bonds between the Arg-374 side chain and backbone oxygen atoms of residues Gly-251 and His-250 in loop 3 (Fig. 3G), which results in loop 3 being more flexi-

FIGURE 3. The interactions in three unique interfaces in the head-to-tail A3G-CD2 dimer conformation. A, the secondary structural elements involved in the interactions in the interfaces. Shown are detailed interactions in the first interface (B–D), the second interface (E–G), and the third interface (H), respectively. The salt bridge between side chains of the residues Asp-316 and Arg-374 was observed in the head-to-tail dimer conformer (K) but not found in D370A variant structure (L). The salt bridge between Arg-374 and Glu-209’ was observed in the tail-to-tail dimer conformer. The numbers in each panel indicate the distances between the specific atoms connected by dotted lines. The double-headed arrows indicate the hydrophobic interactions between the related groups. The red spheres represent water molecules.
ble. Thus, like the mutation from Asp-370 to Ala-370, the mutation from Arg-374 to Ala-374 destabilizes the active deamination center of A3G-CD2 by changing the conformations of loops 1, 3, and 7.

The interaction between loop 3 in monomer H and loop 1 in monomer T composes the third interface, in which the side chains of the residues His-250 and His-248 in loop 3 have weak hydrophobic interactions with the residues Pro-210 and Trp-211 in loop 1 as well as hydrogen bond interactions between the His-250 side chain nitrogen atom and the Pro-210 backbone oxygen atom through one water molecule (Fig. 3f). These interactions enhance the stability of the dimer conformation. To examine the significance of this interface, four variants (H248G, H250A, P210A, and P210G) of A3G-CD2 were designed. Beyond our expectations, compared with the WT protein, the catalytic efficiency values ($K_{cat}/K_m$) of the H248G and H250A variants are improved about 1.6- and 2.6-fold, respectively ($K_{cat}^{H248G} = 0.28 \pm 0.0011 \text{ min}^{-1}, K_{cat}^{H250A} = 33.06 \pm 0.37 \mu \text{M}, K_{cat}^{H248G}/K_m^{H248G} = 8.55 \times 10^{-3} \text{ min}^{-1} \mu \text{M}^{-1}, K_{cat}^{H250A}/K_m^{H250A} = 0.48 \pm 0.0094 \text{ min}^{-1} \mu \text{M}, K_{cat}^{H248G}/K_m^{H250A} = 33.25 \pm 0.18 \mu \text{M}, K_{cat}^{H248G}/K_m^{H250A} = 1.44 \times 10^{-2} \text{ min}^{-1} \mu \text{M}^{-1})$.

These observations are distinct from the previous studies on the structure 3IR2 (33,37), where the side chains of His-248 and His-250 were ligated to zinc ion, the decrease in DNA deaminate activity of the trisite H248A/H250A/C261A variant of A3G-CD2 was thought to be attributable to C261A alone, and the mutations on His-248 and His-250 had no impact on the DNA deamination. The P210A and P210G variants were devised to decrease hydrophobic interaction with the side chain of His-250; their catalytic efficiency $K_{cat}/K_m$ was thus nearly ablated (for the P210A variant, $K_{cat}^{P210A} = 0.0061 \pm 0.00078 \text{ min}^{-1}, K_m^{P210A} = 81.88 \pm 23.34 \mu \text{M}, K_{cat}^{P210A}/K_m^{P210A} = 7.45 \times 10^{-5} \text{ min}^{-1} \mu \text{M}^{-1}$, decreased about 100-fold). The P210G variant, $K_{cat}^{P210G} = 0.0099 \text{ min}^{-1}, K_m^{P210G} = 18.46 \pm 4.89 \mu \text{M}, K_{cat}^{P210G}/K_m^{P210G} = 5.37 \times 10^{-4} \text{ min}^{-1} \mu \text{M}^{-1}$, decreased about 10-fold).

Before measuring the catalytic efficiency of each variant of A3G-CD2, we tested their aggregation states by running an analytic Superdex™ G75 column (10/300) (Fig. 4). The results suggest that all mutations on the residues mentioned above have no effects on the aggregation state of A3G-CD2. Therefore, we can exclude the possibility that the differences in the catalytic efficiency of A3G-CD2 variants resulted from A3G-CD2 aggregation state changes.

We further investigated the contributions of the residues in the interface in the head-to-tail A3G-CD2 dimer to the deaminate activities of full-length A3G through an E. coli-based deaminase activity assay (Fig. 5). The expression of the full-length A3G and its variants was confirmed by running E. coli immunoblots (Fig. 5). The variants, including P210A, P210G, Q245A, R256A, D264A, D370A, R374A, R376A, and Q380A, demonstrate weaker deaminate activities. In addition, the residue Phe-252 (in loop 3) is not located in the interface, but it has hydrophobic interactions with the side chain of residue Arg-256 (in loop 3) within one monomer (Fig. 3E) of the current head-to-tail dimer conformation (this interaction was also observed in the x-ray structure 3E1U) (33), which may stabilize the active center by fixing the conformation of loop 3. Thus, the mutation from Phe-252 to Ala-252 impairs this hydrophobic interaction and decreases the activities of HIV-1 virus ssDNA deamination by A3G-CD2. The H248G, H250A, and H250G variants have higher deaminate activities than WT protein, consistent with in vitro enzymatic results from A3G-CD2 H248G, H250A, and H250G variants.

In summary, the interfaces in the current head-to-tail dimer conformation present new insights into the residues involved in HIV-1 virus ssDNA binding and the catalytic deamination. Mutagenesis studies on those residues further confirm that the stability of the active center is extremely important to catalytic C→U deamination. Nine new residues (Pro-210, Gln-245, His-248, His-250, Phe-252, Asp-264, Phe-268, Asp-370, and Glu-380) in the A3G-CD2 domain necessary for HIV-1 ssDNA binding and the catalytic deamination reaction were observed from this new dimer conformation.
DISCUSSION
The Head-to-tail Dimer Conformer of A3G-CD2 Reveals the Mode of Full-length A3G Binding to ssDNA—To well understand how the A3G CD1 and CD2 domains work together to facilitate cytidine deamination, a holoenzyme structure is a prerequisite. Although the three-dimensional structure of A3G-CD1 domain is not available, different models of full-length A3G were previously constructed. Two of them were predicted based on the APOBEC2 structure because APOBEC2 has amino acid sequence identical to A3G-CD1 (24\%) and A3G-CD2 (31\%) domains. One model, where A3G-CD1 and A3G-CD2 are tethered through the interactions between their $\beta$-strands, was successfully used to identify three residues (Arg-122, Trp-127, and Asp-128) important for packaging A3G into virions (50). The other model, however, produced from the extended NMR structure of A3G-CD2 domain (PDB code 2KEM) (35), suggested that the $\beta$-strands in A3G-CD1 and A3G-CD2 are tethered through the interactions between their $\beta$-strands, with two newly identified residues (Arg-122, Trp-127, and Asp-128) important for packaging A3G into virions (50). The other model, however, produced from the extended NMR structure of A3G-CD2 domain (PDB code 2KEM) (35), suggested that the $\beta$-strands in A3G-CD1 and A3G-CD2 were distant. Instead, an N-terminal pseudocatalytic domain, including the interdomain linker and some of helix $\alpha_6$ of A3G-CD1, packs A3G-CD1 and A3G-CD2 together. Unfortunately, in this model, the catalytic deaminase domains (i.e. A3G-CD2 and N-terminal pseudocatalytic domain) point away from each other and from the nucleic acid binding site, creating a topological dilemma. The third model, the tail-to-tail dimerization model of the full-length A3G-DR (A3G was treated with RNase), and tetramer model of the full-length A3G-D (A3G was not treated with RNase) were generated by small angle x-ray scattering and shape reconstruction methods. This model implied that the full-length A3G in either low molecular mass or high molecular mass is symmetrically associated (51). This model was further refined into the fourth model after the A3G-CD2 three-dimensional structure was reported, through an in-cell quenched fluorescence resonance energy transfer (FRET) assay, small angle x-ray scattering, and other techniques (52). In this model, A3G was self-associated via its CD2 domain, forming a dimer structure. It seems that this model is more reasonable than any other reported model. Its low resolution, however, limits its usage in the analysis of the biological functions of the full-length A3G.

It is well known that A3G deaminates ssDNA processively with a strong 3' $\rightarrow$ 5' bias (31, 32, 34). When there is more than one target motif in the sequence of ssDNA, A3G deaminates the 5'-CCC target motif 5-fold more rapidly than the 3'-CCC target motif (53). Unlike either WT full-length A3G or its monomer F126A/W127A mutant, the A3G-CD2 alone, which does not oligomerize, catalyzes ssDNA-dependent C $\rightarrow$ U deamination (33, 36, 49), displaying no deamination polarity and no dead zone. The non-catalytic A3G-CD1 plays an indispensable role in stabilizing ssDNA binding, enhancing the catalysis, and establishing 3' $\rightarrow$ 5' deamination polarity and processivity. To explain the polarity and processivity of the deamination by A3G, Chelico et al. (53) suggested the fifth structural model of the full-length A3G, in which A3G effectively deaminates the ssDNA 5'-proximal CCC motif only upon binding to ssDNA in an active orientation. However, it is still difficult to obtain information about how A3G-CD1 is involved in ssDNA binding from this model.

FIGURE 5. E. coli-based deaminase assay. The results of E. coli immunoblots (input) confirmed expression of the proteins. Error bars, S.E.
To address this, we predicted an A3G-CD1 model through the SWISS model server (54–57) on the basis of its high sequence identity (43.68%) to A3C (42) (Fig. 6A), which contains six α-helices and five β-sheets. The β2 sheet is continuous, different from that in the A3G-CD2 structure. This structural model has a root mean square deviation value of 3.27 Å for all backbone atoms upon superimposing with the monomer T in the current head-to-tail A3G-CD2 dimer, indicating that its overall fold is almost identical with that of the A3G-CD2 domain (Fig. 6B). Thus, we raise a question of whether the current head-to-tail A3G-CD2 dimer conformation is a potential three-dimensional structural model of the full-length A3G.

To answer this, first, we probed the surface of head-to-tail dimer conformation, where an obviously continuous groove was observed (Fig. 6C). Part of this groove is composed of the reported main ssDNA binding elements in A3G-CD2, such as loop 1 (33, 35–37), loop 3 (33, 36), loop 7 (33, 36), and helix α6 (36, 37). The residues located in or along this continuous groove in monomer H are involved in ssDNA binding, including the mostly reported side chain charged amino acids, such as Arg-213 (33, 36), Arg-215 (33, 36), Arg-256 (36), Arg-313 (33), Asp-316 (36, 43), Asp-317 (36, 43), Arg-320 (33), Arg-374 (36, 37), and Arg-376 (36, 37), and hydrophobic residues, such as Phe-202 (35), Phe-206 (35), Trp-211 (37), His-216 (36), His-257 (33, 36), Trp-269 (36), Leu-271 (36), Trp-285 (33, 36), Phe-289 (36), and Tyr-315 (36), as well as Asn-208 (35) and Asn-244 (36). The new active sites, including the residues Pro-210, Glu-254, His-257 (33, 36), Glu-259, Phe-262, Asp-264, Phe-268, Asp-370, and Glu-380, discussed above are also situated in or along this groove.

Second, the ssDNA binding sites in A3G-CD1 have not been completely identified. We thus tried to obtain the binding information basically from the sequence alignment (Fig. 6A). A3G-CD2 has 33.51% sequence identity to A3G-CD1. We replaced the residues in the A3G-CD1 domain by alanine or glycine, including Pro-25, Leu-27, Glu-61, His-65, Glu-67, Phe-70, His-72, Tyr-91, Trp-94, Arg-122, Tyr-124, Tyr-125, Trp-127, Asp-130, Leu-184, Met-188, and Glu-191 (Fig. 6D), which have positions similar to residues Pro-210, Val-212, Glu-254, His-257 (33, 36), Glu-259, Phe-262, Asp-264, Phe-282, Trp-285 (33, 36), Arg-313 (33), Tyr-315 (36), Asp-316, Asp-317, Arg-320, Asp-370, and Arg-374 (36, 37) seated along or in the
groove in monomer T (Fig. 6C), respectively, and tested the deaminase activities of these variants by the E. coli-based deaminase assay. As shown in Fig. 5, the variants E61A, H65G, E67A/D264H (the E67A variant was not successfully obtained), F70A, H72A, Y91A/D264H (the Y91A variant was not successfully obtained). W94A, R122A, Y124A, Y125A, W127A, D130A (and D130R), L184/185A (the L184A variant was not successfully generated), M188A, and E191A demonstrate weaker deaminase activities compared with the WT full-length A3G. Based on superimposing structures (Fig. 6B), we also suppose that the residues Pro-25, Trp-127, Leu-184, Ile-187, Met-188, and Glu-191 are also possibly involved in the interactions between A3G-CD1 and A3G-CD2 domain, as shown in Fig. 7.

Third, in what orientation does ssDNA upon its bind to full-length A3G? In our current A3G-CD2 head-to-tail dimer conformation, close to the catalytic center of the residue Glu-259 (ligated to Zn2+ ion), there is a deep pocket for accommodating the first base Cm in the 5′-proximal -Cm+2Cm+1Cm target motif in ssDNA (the ssDNA sequence is assumed to be 5′-Cm+2Cm+1Cm-Cm+1-Cm+2-Cm). By overlapping the structure of monomer H with that of the mouse cytokine deaminase in complex with cytidine in the active site (PDB code 2fr6), we found that the hot spot cytidine was really docked into this site near the residues -Y315DDQ318- in the conserved sequence -R312TYDDQ318- in loop 7 (Fig. 6C and D). This sequence was previously predicted to specifically identify the second base, Cm+1, in the 5′-proximal Cm+2Cm+1Cm target motif in ssDNA, which is different from the residues -Y307Y308FWD311- in loop 7 of A3F-CD2 (specifically binding to the second base Tm+1 in 5′-Tm+2Tm+1CmA-A-3′) (43, 44) and from the residues -Y130DYD133- in loop 7 of A3A-CD1 (specifically binding to the second base, Tm+1 or Cm+1, in target motif 5′-Tm+2Cm+1Cm+2-CmA-A-3′) (40, 41). Thus, upon binding to full-length A3G, the 5′ terminus of ssDNA is located in the A3G-CD2 side, whereas its 3′ terminus interacts with the A3G-CD1 side. In other words, ssDNA binds to full-length A3G in an active orientation, as shown in Fig. 6, C and D, which accords with the previous prediction (53). The ssDNA interaction cavity displays large hydrophobic regions and negatively charged regions in the electrostatic potential surface of A3G-CD2 plus the A3G-CD1 structural model (Fig. 6E). The main hydrophobic regions distribute near the 3′-end of the ssDNA, and the negatively charged regions mainly locate close to the catalytic center.

In conclusion, based on the structural analysis, NMR-based enzymatic assay, and E. coli-based deaminase activity assay, we suggest that the head-to-tail dimer conformation may represent a structural model of full-length A3G. This model indicates that A3G may bind to HIV-1 virus ssDNA in an active orientation. According to this structural model, several new residues in A3G-CD1 (including Pro-25, Glu-61, His-72, Asp-130, and Leu-184) were found to be necessary to ssDNA binding and catalytic deamination.
REFERENCES

1. Jarmuz, A., Chester, A., Bayliss, J., Gisbourne, J., Dunham, I., Scott, J., and Navaratnam, N. (2002) An anthropoid-specific locus of orphan C to U RNA-editing enzymes on chromosome 22. Genomics 79, 285–296
2. Sawyer, S. L., Emerman, M., and Malik, H. S. (2004) Ancient adaptive evolution of the primate antiviral DNA-editing enzyme APOBEC3G. PLoS Biol. 2, E275
3. Zhang, J., and Webb, D. M. (2004) Rapid evolution of primate antiviral enzyme APOBEC3G. Hum. Mol. Genet. 13, 1785–1791
4. Goila-Gaur, R., and Strebel, K. (2008) HIV-1 Vif, APOBEC, and intrinsic immunity. Retrovirology 5, 51
5. Chiu, Y. L., and Greene, W. C. (2008) The APOBEC3 cytidine deaminases: a single-stranded RNA facilitates nucleocapsid:APOBEC3G complex formation. Retrovirology 5, 51
6. Malim, M. H. (2009) APOBEC proteins and intrinsic resistance to HIV-1 replication. Retrovirology 6, 92
7. Sawyer, S. L., Emerman, M., Malik, H. S., and Landau, N. R. (2004) Ancient adaptive evolution of the primate antiviral DNA-editing enzyme APOBEC3G. PLoS Biol. 2, E275
8. LaRue, R. S., Andre´sdottir, V., Blanchard, Y., Conticello, S. G., Derse, D., NEGAYEN, D., ANDREWS, K. M., ANDERSON, B. D., YEN, L., STANLEY, D., MAHON, C., KANE, J., FRANKS-SKIBA, M. S., KLEEMAN, L., AND SNEDECOR, B. H. (2005) Evolution of the AID/APOBEC family of polynucleotide (deoxy)cytidine deaminases. Mol. Biol. Evol. 22, 367–377
9. Nachtergaele, E., Liddament, M. T., and Harris, R. S. (2005) The retroviral hydrolase APOBEC3G accounts for minus-strand deamination of the HIV genome. Nature 435, 374–386
10. Malim, M. H. (2009) APOBEC proteins and intrinsic resistance to HIV-1 infection. Philos. Trans. R. Soc. Lond. B Biol. Sci. 364, 689–703
11. Strebel, K., and Khan, M. A. (2008) APOBEC3G encapsidation into HIV-1 virions and intrinsic resistance. Annu. Rev. Immunol. 26, 317–353
12. Chiu, Y. L., and Greene, W. C. (2008) The APOBEC3 cytidine deaminases: an innate defensive network opposing exogenous retroviruses and endogenous retroelements. Annu. Rev. Immunol. 26, 317–353
13. Strebel, K., and Khan, M. A. (2008) APOBEC3G encapsidation into HIV-1 virions: which RNA is it? Retrovirology 5, 55
14. Cen, S., Guo, F., Niu, M., Saadatmand, J., Deflassieux, J., and Kleiman, L. (2004) The interaction between HIV-1 Gag and APOBEC3G. J. Biol. Chem. 279, 33177–33184
15. Douaisi, M., Dussart, S., Courcouil, M., Besson, G., Vigne, R., and Decroly, E. (2004) HIV-1 and MLV Gag proteins are sufficient to recruit APOBEC3G into virus-like particles. Biochem. Biophys. Res. Commun. 321, 566–573
16. Sivarova, S., Xu, H., Mbita, J. L., Arch, R., Gorelick, R. J., Ono, A., Freed, E. O., Hu, W. S., and Pathak, V. K. (2004) Human apolipoprotein B mRNA-editing enzyme-catalytic polypeptide-like 3G (APOBEC3G) is incorporated into HIV-1 virions through interactions with viral and nonviral RNAs. J. Biol. Chem. 279, 35822–35828
17. Yu, Q., Koenig, R., Pillai, S., Chiles, K., Kearney, M., Palmer, S., Richman, D., Coffin, J. M., and Landau, N. R. (2004) Single-strand specificity of APOBEC3G accounts for minus-strand deamination of the HIV genome. Nat. Struct. Mol. Biol. 11, 435–442
18. Mangeat, B., Tirelli, P., Caron, G., Friedli, M., Perrin, L., and Trono, D. (2003) Broad antiretroviral defence by human APOBEC3G through lethal editing of nascent reverse transcripts. Nature 424, 99–103
19. Zhang, H., Yang, B., Pomermanz, R. J., Zhang, C., Arunachalam, S. C., and Gao, L. (2003) The cytidine deaminase CEM15 induces hypermutation in newly synthesized HIV-1 DNA. Nature 424, 94–98
20. Hachè, 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
21. Navarro, F., Bollman, B., Chen, H., Koenig, R., Yu, Q., Chiles, K., and Landau, N. R. (2005) Complementary function of the two catalytic domains of APOBEC3G. Virology 333, 374–386
22. Harris, R. S., and Liddament, M. T. (2004) Retroviral restriction by APOBEC proteins. Nat. Rev. Immunol. 4, 868–877
23. 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
24. Chelico, L., Saco, E. J., Eire, D. A., and Goodman, M. F. (2008) A model for oligomeric regulation of APOBEC3G cytidine deaminase-dependent restriction of HIV. J. Biol. Chem. 283, 13780–13791
25. Chen, 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
26. Furukawa, A., Nagata, T., Matsugami, A., Habu, Y., Sugiyama, R., Hayashi, F., Kobayashi, N., Yokoyama, S., Takekawa, H., and Katahira, M. (2009) Structure, interaction and real-time monitoring of the enzymatic reaction of wild-type APOBEC3G. EMBO J. 28, 440–451
27. 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 structure. J. Mol. Biol. 389, 819–832
28. 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
29. Shandilya, S. M., Nalam, M. N., Nalaviaka, 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
30. Li, M., Shandilya, S. M., Carpenter, M. A., Rathore, A., Brown, W. L., Perkins, A. L., Harki, D. A., Solberg, J. H., Hoek, 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
31. Prochnow, C., Bransteitter, R., Klein, M. G., Goodman, M. F., and Chen, X. S. (2007) The APOBEC-2 crystal structure and functional implications for the deaminase AID. Nature 445, 447–451
32. Byeon, I. J., Ahn, J., Mitra, M., Byeon, C. H., Hercik, K., Hritz, J., Charlton, L. M., Levin, J. G., and Gronenborn, A. M. (2013) NMR structure of human restriction factor APOBEC3A reveals substrate binding and enzyme specificity. Nat. Commun. 4, 1890
33. Mitra, M., Hercik, K., Byeon, I. J., Ahn, J., Hill, S., Hinchee-Rodriguez, K., Singer, D., Byeon, C. H., Charlton, L. M., Nam, G., Heidecker, G., Gronen-
born, A. M., and Levin, J. G. (2014) Structural determinants of human APOBEC3A enzymatic and nucleic acid binding properties. Nucleic Acids Res. 42, 1095–1110
42. Kitamura, S., Ode, H., Nakashima, M., Imahashi, M., Naganawa, Y., Kurosawa, T., Yokomaku, Y., Yamane, T., Watanabe, N., Suzuki, A., Sugiuira, W., and Iwatani, Y. (2012) The APOBEC3C crystal structure and the interface for HIV-1 Vif binding. Nat. Struct. Mol. Biol. 19, 1005–1010
43. Siu, K. K., Sultana, A., Azimi, F. C., and Lee, J. E. (2013) Structural determinants of HIV-1 Vif susceptibility and DNA binding in APOBEC3F. Nat. Commun. 4, 2593
44. Bohn, M. F., Shandilya, S. M., Albin, J. S., Kouno, T., Anderson, B. D., McDougle, R. M., Carpenter, M. A., Rathore, A., Evans, L., Davis, A. N., Zhang, J., Lu, Y., Somasundaran, M., Matsuo, H., Harris, R. S., and Schiffer, C. A. (2013) Crystal structure of the DNA cytosine deaminase APOBEC3F: the catalytically active and HIV-1 Vif-binding domain. Structure 21, 1042–1050
45. Guo, Y., Dong, L., Qiu, X., Wang, Y., Zhang, B., Liu, H., Yu, Y., Zang, Y., Yang, M., and Huang, Z. (1997) Processing of x-ray diffraction data collected in oscillation mode. Methods Enzymol. 276, 307–326
46. Adams, P. D., Afonine, P. V., Bunkóczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. I., Hung, L. W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. I., Richardson, D. C., Richardson, J. S., Terwilliger, T. C., and Zwart, P. H. (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr. D Biol. Crystallogr. 66, 213–221
47. Emsley, P., and Cowtan, K. (2004) Coot: model-building tools for molecular graphics. Acta Crystallogr. D Biol. Crystallogr. 60, 2126–2132
48. Chen, K. M., Martemyanova, N., Lu, Y., Shindo, K., Matsuo, H., and Harris, R. S. (2007) Extensive mutagenesis experiments corroborate a structural model for the DNA deaminase domain of APOBEC3G. FEBS Lett. 581, 4761–4766
49. Zhang, K. L., Maneat, B., Ortiz, M., Zoete, V., Trono, D., Telenti, A., and Michielin, O. (2007) Model structure of human APOBEC3G. PLoS One 2, e378
50. Wedekind, J. E., Gillilan, R., Janda, A., Krucinska, J., Salter, J. D., Bennett, R. P., Raina, I., and Smith, H. C. (2006) Nanostructures of APOBEC3G support a hierarchical assembly model of high molecular mass ribonucleoprotein particles from dimeric subunits. J. Biol. Chem. 281, 38122–38126
51. Bennett, R. P., Salter, J. D., Liu, X., Wedekind, J. E., and Smith, H. C. (2008) APOBEC3G subunits self-associate via the C-terminal deaminase domain. J. Biol. Chem. 283, 33329–33336
52. 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
53. Arnold, K., Bordoli, L., Kopp, J., and Schwede, T. (2006) The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. Bioinformatics 22, 195–201
54. Kiefert, F., Arnold, K., Künnzi, M., Bordoli, L., and Schwede T (2009) The SWISS-MODEL Repository and associated resources. Nucleic Acids Res. 37, D387–D392
55. Guex, N., Peitsch, M. C., and Schwede, T. (2009) Automated comparative protein structure modeling with SWISS-MODEL and Swiss-PdbViewer: a historical perspective. Electrophoresis 30, S162–S173
56. biasini, M., Bienert, S., Waterhouse, A., Arnold, K., Studer, G., Schmidt, T., Kiefert, F., Casarino, T. G., Bertoni, M., Bordoli, L., and Schwede, T. (2014) SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information. Nucleic Acids Res. 42, W252–W258