



Institute of Materia Medica, Chinese Academy of Medical Sciences
Chinese Pharmaceutical Association

Acta Pharmaceutica Sinica B

www.elsevier.com/locate/apsb
www.sciencedirect.com



ORIGINAL ARTICLE

Molecular modeling of human APOBEC3G to predict the binding modes of the inhibitor compounds IMB26 and IMB35

Zhixin Zhang^a, Congjie Zhai^a, Zeyun Mi^a, Jiwei Ding^a, Yongxin Zhang^a, Xing Shi^a, Xiaoyu Li^a, Liyan Yu^a, Zhuorong Li^a, Jiandong Jiang^b, Jinming Zhou^{a,*}, Shan Cen^{a,*}

^aInstitute of Medicinal Biotechnology, Chinese Academy of Medical Science and Peking Union Medical College, Beijing 100050, China

^bInstitute of Materia Medica, Chinese Academy of Medical Science and Peking Union Medical College, Beijing 100050, China

Received 22 March 2013; revised 3 May 2013; accepted 10 May 2013

KEY WORDS

Host restriction factor;
APOBEC3G;
HIV;
Molecular modeling;
Anti-HIV drug

Abstract APOBEC3G(A3G) is a host cytidine deaminase that incorporates into HIV-1 virions and efficiently inhibits viral replication. The virally encoded protein Vif binds to A3G and induces its degradation, thereby counteracting the antiviral activity of A3G. Vif-mediated A3G degradation clearly represents a potential target for anti-HIV drug development. Currently, there is an urgent need for understanding the three dimensional structure of full-length A3G. In this work, we use a homology modeling approach to propose a structure for A3G based on the crystal structure of APOBEC2 (APO2) and the catalytic domain structure of A3G. Two compounds, IMB26 and IMB35, which have been shown to bind to A3G and block degradation by Vif, were docked into the A3G model and the binding modes were generated for further analysis. The results may be used to design or optimize molecules targeting Vif–A3G interaction, and lead to the development of novel anti-HIV drugs.

© 2013 Institute of Materia Medica, Chinese Academy of Medical Sciences and Chinese Pharmaceutical Association. Production and hosting by Elsevier B.V. All rights reserved.

*Corresponding authors. Tel.: +86 10 63131011 (Jinming Zhou); +86 10 63037279 (Shan Cen).

E-mail addresses: zhou_jim@hotmail.com (Jinming Zhou); shancen@imb.pumc.edu.cn (Shan Cen).

Peer review under responsibility of Institute of Materia Medica, Chinese Academy of Medical Sciences and Chinese Pharmaceutical Association.



Production and hosting by Elsevier

1. Introduction

A3G is a single-stranded DNA (ssDNA) deoxycytidine deaminase that inhibits HIV-1 replication in strains that lack Vif^{1,2}. In the absence of Vif, virion encapsidated A3G can deaminate cytosine to uracil on newly formed viral DNA during reverse transcription of the viral genome, which leads to hypermutation and inactivation of the newly synthesized viral DNA^{3,4}. Furthermore, several lines of evidence showed that A3G also impaired viral DNA synthesis and integration through an antiviral mechanism distinct from deamination^{5,6}. As a counter-measure, HIV-1 Vif binds to A3G and recruits a cellular ubiquitin ligase complex containing cullin-5 (CUL5), elongin B (ELOB), elongin C (ELOC) and a RING-box protein. This leads to the ubiquitination of A3G and degradation by the 26S proteasome⁷⁻⁹.

A3G contains two cytosine deaminase domains at the N-terminal (CD1) and the C-terminal (CD2). The C-terminal domain (CD2) is responsible for the deaminase activity of A3G^{10,11}. The structure of the CD2 domain of A3G has been determined by X-ray crystallography and NMR¹²⁻¹⁴ and shown to fold into a five-stranded β sheet flanked by six α helices. While the CD1 domain is catalytically inactive, it is involved in virion encapsidation and mediates the oligomerization of A3G^{15,16}. Mutations in the CD1 domain affect multiple aspects of A3G function including dimerization, virion incorporation and interaction with Vif^{17,18}. The CD1 domain structure has not been determined, although several homology models have been proposed for this domain on the basis of the APO2 tetramer^{19,20}, and a similar head-to-head interface was proposed for the CD1 domain of A3G²¹.

As A3G is expressed in human cells infected by HIV-1, inhibition of Vif-mediated hA3G degradation represents a new anti-HIV-1 strategy for drug discovery. In our previous work we identified two small molecules (IMB26 and IMB35) that target the interface of Vif and A3G, thus protecting A3G from Vif-mediated degradation²². Although the inhibitors were shown to bind to A3G, the exact position at which the molecules bind, as well as binding mode, remains unclear. Here, a more accurate full-length A3G has been constructed based on APO2 dimer and the newly reported A3G CD2 domain structure (PDBID: 3IQS). We modeled the binding position of IMB26 and IMB35 at the A3G CD1 domain and the binding modes were also generated through molecular docking. The structural information obtained from this A3G model and the predicted binding mode will facilitate rational drug design targeting the A3G-Vif interaction.

2. Materials and methods

2.1. Homology modeling and model evaluation

The human A3G sequence (residues 1–384; Uniprot Entry: Q9HC16, <http://www.uniprot.org>) was defined as the target sequence. The crystallized human APO2 dimer²³ (PDBID: 2NYT, <http://www.pdb.org>) and the crystallized A3G CD2 domain structure (PDBID: 3IQS, <http://www.pdb.org>) served as template. The A3G sequence was aligned to the CD2 domain structure (3IQS) and APO2 dimer (2NYT) using the “align multiple sequences” module in the molecular modeling software Discovery Studio (DS).

The target-template alignment was used to build the A3G structure model using the “build homology” module in DS. To ensure that the crystal structure coordinates are not modified in the model building process, the reference template parameter in the protocol was set to the CD2 domain structure (3IQS). The PDF total energy and DOPE score were used to select the best model.

The Ramachandran Plot and Verify 3D in DS were used to evaluate the A3G structure model.

2.2. Molecular docking

The binding mode for IMB26 and IMB35 to the binding site of the homology model A3G was generated by molecular docking using the Genetic Optimization for Ligand Docking (GOLD) software from Cambridge Crystallographic Data Center, UK²⁴. The binding sites on the A3G model were located by LIGSITE (<http://projects.biotech.tu-dresden.de/pocket/>) and the site finder module implemented in the Molecular Operating Environment (MOE). The ligand-binding region was defined as a sphere of 10 Å radius around the binding site. Molecular interactions were observed using LigX implemented in MOE. The interactions of the ligand and the receptor were evaluated and ranked *via* GOLD score implemented in GOLD, and the orientation with the highest docking score was returned for further analysis.

3. Results and discussion

3.1. Establishment and evaluation of a homology model of full-length A3G

To construct the A3G full-length structure, the APO2 dimer (PDBID: 2NYT) and the CD2 domain structure (PDBID: 3IQS) served as the templates. We aligned the sequence of the CD2 domain structure to the A3G sequence. The sequence from the CD2 domain structure is almost identical to that of the A3G sequence. The identity was 29.1% and the similarity was 51.4% as a result of the APO2 and A3G sequence alignment (Fig. 1A). The model of full-length A3G was then built based on the structures, and the results are shown in Fig. 1. Both the CD1 and CD2 domains have a core platform composed of six α -helices and five β -strands (Fig. 1B). In the APO2 crystal structure, APO2 dimerizes *via* direct interactions between the β 2-strands²³. However, the CD1 domain in the A3G structure exhibits a discontinuous β 2-strand that is different from a continuous well-ordered β 2-strand observed in APO2 structure, which may be caused by the low sequence identity at the region between target protein and the templates. Molecular dynamics simulations predict that the β 2 strand of the CD2 domain may be able to adopt a more extended β -strand conformation to allow interactions between two β 2-strands²⁵. This result supports a different dimer configuration between the β 2-strands of the CD2 and CD1 domains. The new A3G structure captured information from the APO2 protein and the catalytic domain structure of A3G, and the CD2 domain remained stable during the homology modeling. The superimposition of the catalytic domain structure on the A3G model gave a RMSD value of 2.20 Å, which means that the two structures are very similar. Thus, consideration of the 3D information of A3G CD2 domain suggests that the A3G structural model proposed here is more reliable than that previously reported¹⁹, which was built based only on the APO2 dimer 3D structure.

The accuracy of the obtained A3G structural model was evaluated by Ramachandran Plot analysis and Verify 3D in DS (Fig. 1C). According to Ramachandran Plot analysis, most of the residues had ϕ and ψ angles in the core (93.9%) and allowed (4.5%) regions, and only a small part of non-glycine residues (1.6%) were in disallowed regions. Moreover, the Verify Score of the A3G structure (153.86) was very close to the Verify Expected High Score (173.24). These results suggest that the quality of the A3G structural model is high.

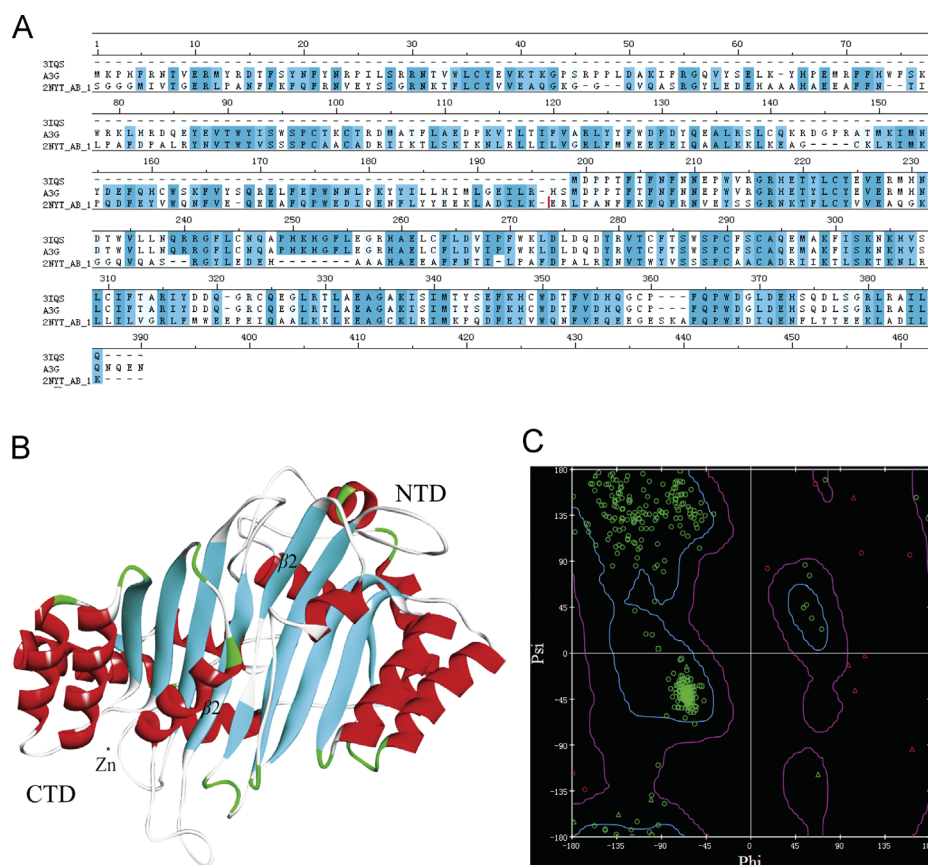


Figure 1 Homology modeling of the A3G structural model. (A) Alignment of the target sequence and templates. (B) The A3G structural model, shown as a ribbon; the Zn coordination is within the CD2 domain. A bulge occurs in the $\beta 2$ segment of the CD2 domain, and a discontinuous $\beta 2$ segment is shown in the CD1 domain. (C) Ramachandran plot of the proposed A3G structure.

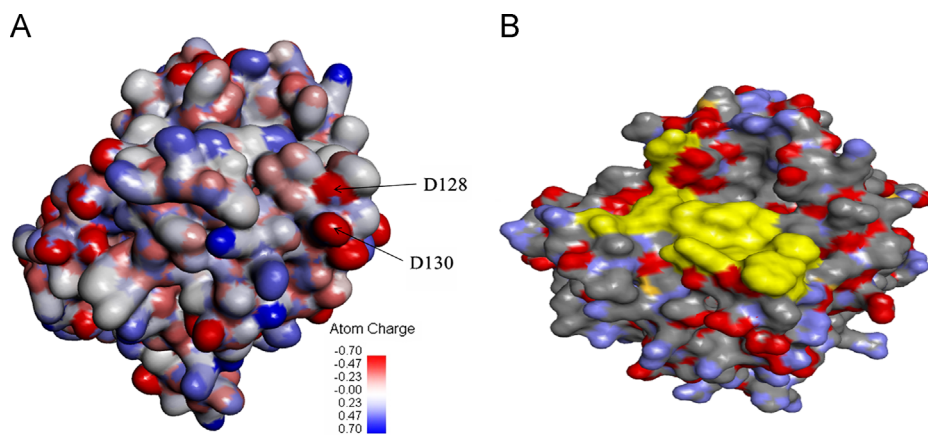


Figure 2 Structural mapping of the critical residues on the A3G model. (A) The electrostatic potential of A3G. The accessible surface area is colored according to the calculated electrostatic potential from -0.7 kT/e (red) to 0.7 kT/e (blue). The ASP128–ASP130 motif is crucial for the direct interaction with Vif. (B) Surface representation of the critical amino acids (Tyr19, Tyr22, Trp94, Tyr124, Tyr125, Phe126, and Trp127) for A3G dimerization as shown in yellow.

3.2. Structural mapping of the critical amino acids of the A3G structure model

To further evaluate the A3G model, we mapped the critical amino acids that are important to the biological function of A3G. Amino acids 128–130 of A3G have been identified as essential for the interaction of Vif and A3G using alanine-scanning mutations and

multiple different substitutions at key residues²¹. The electrostatic surface potential analysis revealed a negatively charged interface in this motif (Fig. 2A), which is an important feature for HIV-1 Vif binding. This structural region notably encompasses a series of amino acids, Tyr19, Tyr22, Trp94, Tyr124, Tyr125, Phe126, and Trp127 (Fig. 2B), that are involved in A3G dimerization and virion incorporation^{11,21,23}. The dimer interface was expected to be

a shallow cavity composed of hydrophobic or aromatic amino acids predicted to mediate the monomer contacts through a series of hydrophobic and π - π interactions. Thus, for both important interfaces within A3G involved in either binding to Vif or hA3G dimerization, the predictions of our model are quite reasonable.

3.3. Identification of binding sites

We predicted the binding sites for the A3G model using LIGSITE²⁶ and the site finder module implemented in MOE separately²⁷. The results are shown in Fig. 3. For the LIGSITE method, three binding sites were located (Fig. 3A). The Site Finder module identified the first-ranked binding site (Fig. 3B). Interestingly, this is rather similar to the one identified by the LIGSITE method, and the binding region was very close to the segment Tyr124-ASP130, which was critical for A3G dimerization and interaction with Vif. Thus, this binding site was selected as the binding site for the binding mode analysis of IMB26 and IMB35. The binding region for molecular docking was defined as a sphere of 12 Å radius around the binding site (shown as red ball in Fig. 3).

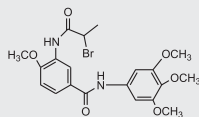
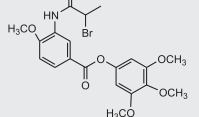
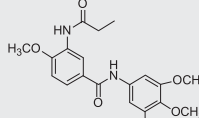
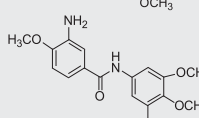
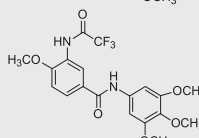
3.4. Binding mode prediction and targeting A3G/Vif interaction for novel anti-HIV-1 agents

IMB26 and IMB35 were shown to target the interface of Vif and A3G axis, thereby protecting A3G from Vif-mediated degradation²². These two small molecules were reported to bind to A3G using an SPR assay²². The exact position at which the two molecules bind as well as binding mode of the two molecules remains unclear. To predict their binding modes toward A3G, IMB26 and IMB35, as well as another 3 analogs with little or no inhibitory effect were docked into the predicted binding site of the A3G model using GOLD. The two active compounds have much better scores than the analogs (Table 1). The orientation of protein-ligand complex with the highest GOLD score was chosen for further analysis. The results are shown in Fig. 4. The binding mode of IMB26 (Fig. 4A and B) and IMB35 (Fig. 4C and D) was very similar. Both IMB26 and IMB35 interact with Tyr22 through a π - π interaction and have hydrophobic interactions with Phe17, Trp94, and Ala121. These aromatic or hydrophobic amino acids were predicted to mediate A3G dimerization, and also played an important role in interaction with Vif²¹. The region to which

IMB26 and IMB35 binds is close to the surface that interacts with Vif, and thus these compounds may block Vif binding to A3G and therefore protect A3G from Vif-mediated degradation.

Studies of these molecules targeting the A3G and Vif interaction may lead to the development of novel anti-HIV-1 agents. To date, in addition to IMB26 and IMB35, several small molecules have been reported to inhibit Vif-mediated A3G degradation through different mechanisms. The small molecular compound RN-18 could specifically degrade Vif in presence of A3G, which means RN-18 only degrades Vif in the Vif/A3G complex, thus blocking the Vif-mediated degradation of A3G and strongly inhibiting the growth of the HIV-1 virus. However, the precise mechanism remains unclear²⁸. A zinc chelate TPEN was shown to impair the ability of Vif to degrade A3G²⁹, but the essential role

Table 1 GOLD scores of the compounds docked into the predicted binding site.

Compound	Structure	GOLD score
IMB26		66.17
IMB35		67.89
IMB261		49.87
IMB262		51.54
IMB263		54.10

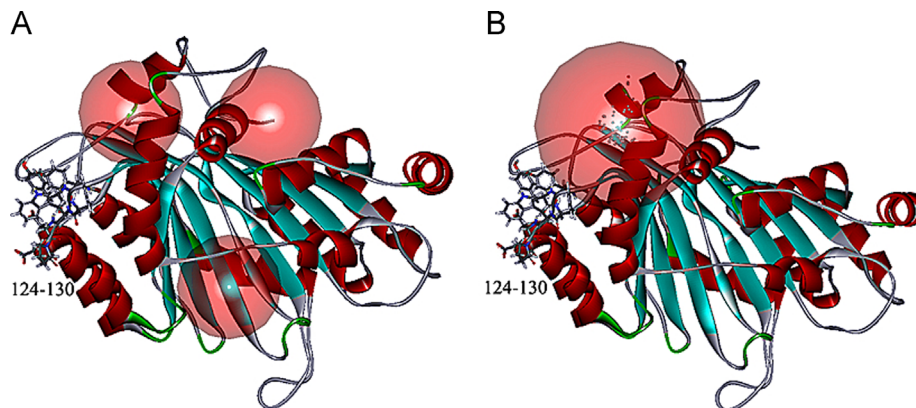


Figure 3 Prediction of the A3G binding region. (A) Defined by LIGSITE. One of these sites is close to the critical TYR124-ASP130 motif (stick depiction), and may be the most likely binding site. (B) Prediction of the A3G binding site as defined by site finder module in MOE. The predicted binding site is also close to the critical TYR124-ASP130 motif.

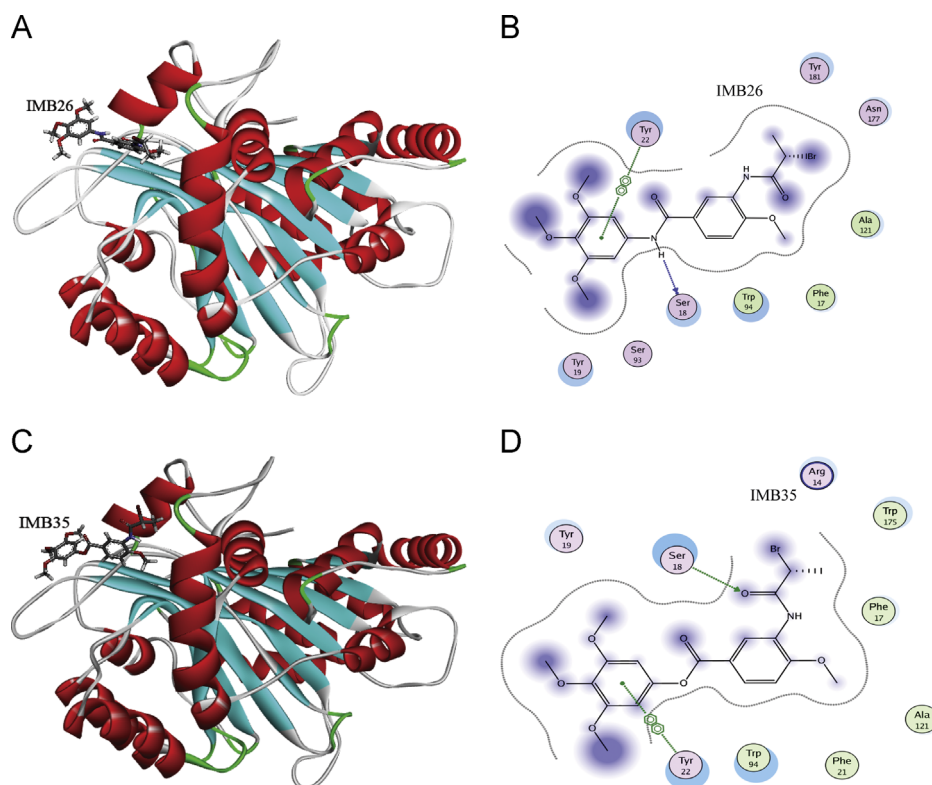


Figure 4 The predicted binding mode of IMB26 and IMB35. (A) The binding mode of the IMB26 (stick model) with A3G (ribbon); (B) the A3G–IMB26 interaction as predicted by the LigX module in MOE; (C) the binding mode of the IMB35 (stick) with A3G (ribbon) and (D) the A3G–IMB35 interaction as predicted by the LigX module in MOE.

for zinc in cells limits the opportunity to modulate its levels, and thus potential as anti-HIV drug. Recently Yu et al.³⁰ reported a small molecule, VEC-5, that targets the interface of Vif and ELOC, which inhibits Vif-mediated degradation of A3G and replication of HIV-1 in A3G-positive cells. It should be noted that a compound specifically targeting the BC-box of Vif might be very difficult to obtain, since the BC-box is a highly conserved sequence and is present in many important cell proteins. Therefore, the most efficient strategy, in principle, would be to use a small-molecular inhibitor that could selectively block the interaction between Vif and APOBEC3G, as IMB26 and IMB35 do.

4. Conclusions

The three-dimensional structure of A3G was predicted using the APO2 protein (2NYT) and the crystallized A3G CD2 domain structure (3IQS) as templates. Structure validation by Ramachandran plot and Verify 3D confirmed the reliability of the structural model. Molecular docking results for IMB26 and IMB35 toward the A3G model suggested the possible binding mode of these two active compounds. The predicted binding modes of IMB26 and IMB35 will be helpful for the better understanding of the mechanism of action of these two compounds, and thus will benefit the optimization of the lead compound and the search for other active compounds with novel scaffolds that target A3G–Vif interaction.

Acknowledgment

This work was supported in part by 973 program (2012CB911102), the National S&T Major Special Project on Major New

Drug Innovation (2012ZX09102101-018), the National S&T International Collaboration 2010DFA31580 (J.D.J.) and 2010DFB30870 (Q.J.).

References

1. Sheehy AM, Gaddis NC, Choi JD, Malim MH. Isolation of a human gene that inhibits HIV-1 infection and is suppressed by the viral Vif protein. *Nature* 2002;**418**:646–50.
2. Mangeat B, Turelli P, Caron G, Friedli M, Perrin L, Trono D. Broad antiretroviral defence by human APOBEC3G through lethal editing of nascent reverse transcripts. *Nature* 2003;**424**:99–103.
3. Zhang H, Yang B, Pomerantz RJ, Zhang C, Arunachalam SC, Gao L. The cytidine deaminase CEM15 induces hypermutation in newly synthesized HIV-1 DNA. *Nature* 2003;**424**:94–8.
4. Yu Q, Konig R, Pillai S, Chiles K, Kearney M, Palmer S, et al. Single-strand specificity of APOBEC3G accounts for minus-strand deamination of the HIV genome. *Nat Struct Mol Biol* 2004;**11**:435–42.
5. Li XY, Guo F, Zhang L, Kleiman L, Cen S. APOBEC3G inhibits DNA strand transfer during HIV-1 reverse transcription. *J Biol Chem* 2007;**282**:32065–74.
6. Iwatani Y, Chan DS, Wang F, Maynard KS, Sugiura W, Gronenborn AM, et al. Deaminase-independent inhibition of HIV-1 reverse transcription by APOBEC3G. *Nucleic Acids Res* 2007;**35**:7096–108.
7. Yu X, Yu Y, Liu B, Luo K, Kong W, Mao P, et al. Induction of APOBEC3G ubiquitination and degradation by an HIV-1 Vif-Cul5-SCF complex. *Science* 2003;**302**:1056–60.
8. Zhang W, Du J, Evans SL, Yu Y, Yu XF. T-cell differentiation factor CBF-beta regulates HIV-1 Vif-mediated evasion of host restriction. *Nature* 2012;**481**:376–9.

9. Jager S, Kim DY, Hultquist JF, Shindo K, LaRue RS, Kwon E, et al. Vif hijacks CBF-beta to degrade APOBEC3G and promote HIV-1 infection. *Nature* 2012;**481**:371–5.
10. Hache G, Liddament MT, Harris RS. The retroviral hypermutation specificity of APOBEC3F and APOBEC3G is governed by the C-terminal DNA cytosine deaminase domain. *J Biol Chem* 2005;**280**:10920–4.
11. Navarro F, Bollman B, Chen H, König R, Yu Q, Chiles K, et al. Complementary function of the two catalytic domains of APOBEC3G. *Virology* 2005;**333**:374–86.
12. Chen KM, Harjes E, Gross PJ, Fahmy A, Lu Y, Shindo K, et al. Structure of the DNA deaminase domain of the HIV-1 restriction factor APOBEC3G. *Nature* 2008;**452**:116–9.
13. Holden LG, Prochnow C, Chang YP, Bransteitter R, Chelico L, Sen U, et al. Crystal structure of the anti-viral APOBEC3G catalytic domain and functional implications. *Nature* 2008;**456**:121–4.
14. Furukawa A, Nagata T, Matsugami A, Habu Y, Sugiyama R, Hayashi F, et al. Structure, interaction and real-time monitoring of the enzymatic reaction of wild-type APOBEC3G. *EMBO J* 2009;**28**:440–51.
15. Huthoff H, Autore F, Gallois-Montbrun S, Fraternali F, Malim MH. RNA-dependent oligomerization of APOBEC3G is required for restriction of HIV-1. *PLoS Pathog* 2009;**5**:e1000330.
16. Bulliard Y, Turelli P, Rohrig UF, Zoete V, Mangeat B, Michielin O, et al. Functional analysis and structural modeling of human APOBEC3G reveal the role of evolutionarily conserved elements in the inhibition of human immunodeficiency virus type 1 infection and Alu transposition. *J Virol* 2009;**83**:12611–21.
17. Compton AA, Hirsch VM, Emerman M. The host restriction factor APOBEC3G and retroviral Vif protein coevolve due to ongoing genetic conflict. *Cell Host Microbe* 2012;**11**:91–8.
18. 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.
19. Zhang KL, Mangeat B, Ortiz M, Zoete V, Trono D, Telenti A, et al. Model structure of human APOBEC3G. *PLoS One* 2007;**2**:e378.
20. Chelico L, Prochnow C, Erie DA, Chen XS, Goodman MF. Structural model for deoxycytidine deamination mechanisms of the HIV-1 inactivation enzyme APOBEC3G. *J Biol Chem* 2010;**285**:16195–205.
21. Lavens D, Peelman F, van der Heyden J, Uyttendaele I, Cateeuw D, Verhee A, et al. Definition of the interacting interfaces of Apobec3G and HIV-1 Vif using MAPPIT mutagenesis analysis. *Nucleic Acids Res* 2010;**38**:1902–12.
22. Cen S, Peng ZG, Li XY, Li ZR, Ma J, Wang YM, et al. Small molecular compounds inhibit HIV-1 replication through specifically stabilizing APOBEC3G. *J Biol Chem* 2010;**285**:16546–52.
23. Prochnow C, Bransteitter R, Klein MG, Goodman MF, Chen XS. The APOBEC-2 crystal structure and functional implications for the deaminase AID. *Nature* 2007;**445**:447–51.
24. Jones G, Willett P, Glen RC, Leach AR, Taylor R. Development and validation of a genetic algorithm for flexible docking. *J Mol Biol* 1997;**267**:727–48.
25. Autore F, Bergeron JR, Malim MH, Fraternali F, Huthoff H. Rationalisation of the differences between APOBEC3G structures from crystallography and NMR studies by molecular dynamics simulations. *PLoS One* 2010;**5**:e11515.
26. Huang B, Schroeder M. LIGSITEcsc: predicting ligand binding sites using the Connolly surface and degree of conservation. *BMC Struct Biol* 2006;**6**:19.
27. Del CCA, Takahashi Y, Sasaki S. A new approach to the automatic identification of candidates for ligand receptor sites in proteins: (I). Search for pocket regions. *J Mol Graph* 1993;**11**:23–9, 42.
28. Nathans R, Cao H, Sharova N, Ali A, Sharkey M, Stranska R, et al. Small-molecule inhibition of HIV-1 Vif. *Nat Biotechnol* 2008;**26**:1187–92.
29. Xiao Z, Ehrlich E, Luo K, Xiong Y, Yu XF. Zinc chelation inhibits HIV Vif activity and liberates antiviral function of the cytidine deaminase APOBEC3G. *FASEB J* 2007;**21**:217–22.
30. Zuo T, Liu D, Lv W, Wang X, Wang J, Lv M, et al. Small-molecule inhibition of human immunodeficiency virus type 1 replication by targeting the interaction between Vif and Elongin C. *J Virol* 2012;**86**:5497–507.