Three-dimensional structure of HIV-1 VIF constructed by comparative modeling and the function characterization analyzed by molecular dynamics simulation†

Wei Lv,[a] Zhenming Liu,[a] Hongwei Jin,[a] Xianghui Yu,[b] Liangren Zhang*[a] and Lihe Zhang[a]

Received 21st August 2006, Accepted 18th December 2006
First published as an Advance Article on the web 18th January 2007
DOI: 10.1039/b612050d

VIF is one of the six accessory proteins of HIV-1. It has been shown to be necessary for the survival of HIV-1 in the human body and for the retention of viral infectivity. It is strongly expected that a new therapeutic strategy against HIV-1 infection could be realized by blocking the biological pathway to VIF. In this paper, a three-dimensional model of VIF was constructed by comparative modeling based on two templates, VHL and NarL, which were used to construct the C-terminal domain and N-terminal domain of VIF, respectively. A model of the VIF–ElonginB–ElonginC complex was constructed, and molecular dynamics simulations were used to investigate the interactions between VIF and ElonginB–ElonginC. Mutagenesis was used to identify the function of some conserved residues in the putative SOCS-box. The results showed that the mutations of the critical residues led to the disruption of the interactions between VIF and ElonginB–ElonginC, consistent with experimental observations. These novel models of VIF and its complex has therefore provided structural information for investigating the function of VIF at the molecular level.

Introduction

Despite the recent success of anti-HIV-1 therapy in controlling disease progression, human immunodeficiency virus type-1 infection is still one of the most serious infectious diseases across the world. In the past, therapeutic targets for HIV-1 infection have been focused on reverse transcriptase, protease, and HIV integrase. However, because drug resistance emerges during anti-HIV-1 treatment and no vaccine is currently available, new targets that can be used to disrupt the lifecycle of HIV-1 infection are continually sought. HIV-1 VIF protein, one of the six viral accessory proteins,[1,2] has been recently reported to play a crucial role in mediating the replications of HIV-1 in lymphocytes and macrophages. Furthermore, it is essential for the survival of HIV-1 *in vivo*,[3–6] suggesting that HIV-1 VIF protein could be a potential target for anti-HIV drug development.[7,8]

HIV-1 VIF is a small (192 residue) protein encoded by viral RNA. In the absence of VIF, the host anti-retroviral factor APOBEC3G, which is a member of the family of cytidine deaminase of nucleic acid-editing enzymes,[9–11] is packaged into the HIV-1 virus and down-regulates HIV-1 infection by causing extensive lethal deoxycytosine-to-deoxyuracil mutation in the newly synthesized HIV-1 negative-strand DNA.[12–16] However, the antiviral function of APOBEC3G is neutralized in the presence of VIF. VIF, through the specific binding to APOBEC3G and then targeting it for ubiquitination and degradation,[12,17–21] efficiently prevents APOBEC3G being packed into the newly synthesized virus, and ensures that the virus remains infectious.

VIF mediates neutralization of APOBEC3G through a process consisting of two steps[18,20] (Fig. 1). First, VIF irreversibly binds to APOBEC3G along with ElonginB, ElonginC, Cul5 and Rbx-1, forming a complex that has E3 ubiquitin ligase activity. Second, the complex mediates the ubiquitination and degrades APOBEC3G rapidly *via* a ubiquitin–proteasome pathway, resulting in the elimination of APOBEC3G and thus the loss of its anti-HIV

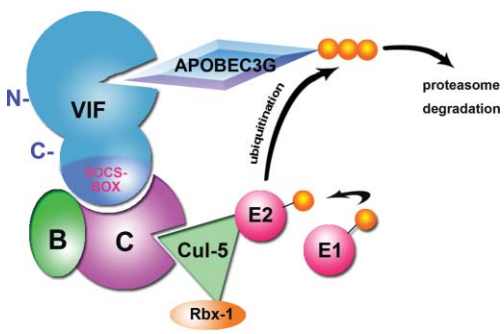

Fig. 1 VIF functions as a bridge that specifically recruits APOBEC3G into an E1–E2–E3 ubiquitin–ligase complex for ubiquitination and subsequent proteasome degradation. VIF consists of two domains: the N-terminal domain (which binds to APOBEC3G), and the C-terminal domain (where SOCS-box is located, which specifically interacts with the ElonginB–ElonginC complex).

[a]*State Key Laboratory of Natural and Biomimetic Drugs, School of Pharmaceutical Sciences, Peking University, Beijing, 100083, China*
[b]*School of Life Science, Jilin University, Changchun, China. E-mail: liangren@bjmu.edu.cn; Fax: +10-82822724; Tel: +10-82802567*

† The following abbreviations are used throughout this paper: HIV-1: human immunodeficiency virus type 1; VIF: viral infectivity factor; VHL: von Hippel–Lindau tumor suppressor protein; NarL: the nitrate/nitrite response regulator; SOCS: suppressor of cytokine signaling; APOBEC3G: apolipoprotein B mRNA-editing, enzyme-catalytic polypeptide-like domain.

functions. In this process, VIF acts as a bridge between its specific substrate (APOBEC3G) and the E3 ubiquitin ligase complex, and is thus central to the operation of the ubiquitin–proteasome pathway. Mutagenesis studies indicate that VIF contains at least two functionally crucial domains[18,20–22]: an N-terminal region that is important for binding to APOBEC3G, and a C-terminal region with a conserved SLQ(Y/F)LAΦΦΦΦ motif (Φ stands for a hydrophobic amino acid), which is responsible for its interaction with the ElonginB–ElonginC complex. Both domains are required for neutralization. The simple association of APOBEC3G with the C-terminal domain of VIF is insufficient to suppress the antiviral activity of APOBEC3G.

There are high similarities between the conserved SLQ(Y/F)LAΦΦΦΦ motif of VIF and the SOCS-box of SOCS proteins (suppressor of cytokine signaling). Both the SOCS-box motif and the SLQ(Y/F)LAΦΦΦΦ motif are located at the C-terminal regions of the corresponding proteins, and the sequences of these two motifs show strong homology and evolutionary conservation.[23,24] In addition, these motifs possess the same function. The SOCS-box in SOCS proteins, like the von Hippel–Lindau tumor suppressor (VHL), has been reported to be responsible for directly interacting with the ElonginB–ElonginC complex, and is thereby crucial for the formation of SCF-like E3–ubiquitin–ligase.[25,26] The SLQ(Y/F)LAΦΦΦΦ motif of VIF has also been reported to mediate its interactions with the ElonginB–ElonginC complex. In view of these similarities, it is speculated that the SLQ(Y/F)LAΦΦΦΦ motif is a novel type of SOCS-box and that VIF is an SOCS-box-containing protein. Therefore, the C-terminal region of VIF where the SLQ(Y/F)LAΦΦΦΦ motif is located is most likely to adopt a similar structure to the SOCS-box in SOCS proteins.[27] However, although the C-terminal region of VIF shows strong similarity to the SOCS-box of SOCS proteins, the N-terminal region shows no similarity to the corresponding parts of SOCS proteins. Furthermore, no report has been published regarding the structural characteristics of the N-terminal region of VIF.

To date, the function and structural features of VIF have been fairly well investigated. However, there still exist numerous aspects that need to be resolved. To achieve a clear understanding of the function and structure of VIF, more thorough biological studies are needed. Meanwhile, a three-dimensional structure could greatly help us to understand the behavior of VIF at the molecular level. Since the crystal structure of VIF is still unavailable, a three-dimensional model of VIF was constructed in this study based on the crystal structure of VHL and NarL.[28] VHL was used for the construction of the C-terminal domain and NarL used as the template for the N-terminal domain. A model of the VIF–ElonginB–ElonginC complex was also proposed based on the crystal structure of the VHL–ElonginB–ElonginC complex. The complex was subjected to molecular dynamics simulation to investigate the VIF–ElonginB–ElonginC interactions. Additionally, based on the constructed VIF model, the effects of some conserved amino acids on the function of VIF were investigated. The construction of a structural model and computational study of its interactions with other proteins could lead to further understanding of the function of VIF in the infectious cycle of HIV-1 and provide useful information for the development of new chemotherapy for HIV infection based on VIF as the therapeutic target.

Computational methods

Template selection and sequence alignment

The sequence of VIF was obtained from Swiss-Prot/TrEMBL (entry P69722), and its secondary structure was predicted using the PSI-PRED[29] program. The N-terminal (residue 1 to 143) of VIF was submitted to Threader 3,[30] a fold recognition method, for template searching. NarL (PDB code 1A04), which had the highest *Z*-score, was selected as the template for the construction of the N-terminal domain of VIF. VHL was selected as the template for the construction of the C-terminal (residue 142–177) domain of VIF based on the multiple sequence alignments of VIF and several SOCS proteins including VHL, SOCS-1, ASB2, WSB1 and RAR, using ClustalW.[31]

Homology model generation of the VIF (SOCS-box)–ElonginB–ElonginC complex

The homology module in Insight II was used to build the initial model of the SOCS-box region of VIF. The model backbone was obtained by transferring the backbone coordinates of the VHL to the corresponding residues of VIF (structurally conserved regions, SCRs) except for the variable regions (LOOPs), which were constructed using MODELLER[32,33] with satisfaction of spatial constraints. For the side chains, library values of rotamers are adopted.

The initial model of the VIF(SOCS-box)–ElonginB–ElonginC complex was constructed by replacing VHL in the crystal structure of the VHL–ElonginB–ElonginC complex with the SOCS-box region of VIF constructed above. Based on this model, a series of mutated complexes were obtained with the Biopolymer module in Insight II. To further refine the model of the complex, energy minimization was carried out by Discover3 with the CVFF force field. One thousand steps of energy minimization in combination with steepest descent and conjugated gradient were performed.

Residues from 50 to 57 in ElonginC were added using Insight II, and each resulting complex were solvated in a triclinic box with the simple point charge (SPC) water model. Before submission to 5 ns of molecular dynamics simulation, 500 steps of steepest-descent energy minimization were carried out to relieve possible atom bumps. All simulations were performed by GROMACS package 3.2.1[34,35] with the GROMOS96 force field.[36] To maintain the system at a constant temperature of 300 K, the Berendsen thermostat[37] was applied with a coupling time of 0.1 ps. Also, a constant pressure of 1 bar was applied with a coupling constant of 0.5 ps. The value of the isothermal compressibility was set to 4.5×10^{-5} bar^{-1} for water simulation. All bond lengths were restrained by the LINCS algorithm.[38] Electrostatic interactions between charged groups at distance less than 10 Å were calculated explicitly, and long-range electrostatic interactions were calculated using the particle-mesh Ewald algorithm.[39] A cutoff distance of 10 Å was applied for Lennard-Jones interactions. For each system, the simulation cell was a triclinic periodic box; the minimum distance between the protein and the box wall was set to more than 7 Å so that the protein would not directly interact with its own periodic image. To neutralize the charges of the modeled systems, 10 water molecules were replaced by 10 Cl^- ions. These ions were located at positions of the chosen water oxygen atoms.

Secondary structure analysis was carried out employing the DSSP (define secondary structure of proteins) method.[40]

Homology model generation of the VIF–ElonginB–ElonginC complex

A model of the whole VIF including both N and C-terminal domains was constructed using the Homology module in Insight II. The construction of the C-terminal domain of VIF was based on VHL, and the N-terminal domain of VIF was built with NarL as the template. VHL and NarL were pre-superposed on their C-terminal domains in order to give a rational orientation of the constructed N and C domains of VIF. Then, the initial 3D model of VIF was submitted to 1000 steps of energy minimization. Each step of energy minimization was performed by Discover3 with steepest descent and, subsequently, conjugated gradient optimization methods. Procheck3.4[41] was used to generate a Ramachandran plot for evaluating the minimized model, and Verify3D[42] was used for further evaluation.

The initial model of VIF–ElonginB–ElonginC complex was constructed by replacing VHL in the crystal structure of the VHL–ElonginB–ElonginC complex with the whole VIF model constructed above. This initial model was minimized by 1000 steps of energy minimization with the combination of steepest descent method and conjugated gradient optimization method. 5 ns of molecular dynamics simulation of the VIF–ElonginB–ElonginC complex was carried out by Gromacs3.2.1 to optimize the model further and to investigate its conformational space. The process and parameters employed were the same as in the simulation of the VIF (SOCS-box)–ElonginB–ElonginC complex, except that 2 water molecules were replaced by 2 Cl^- ions to neutralize the charges of the system. The electrostatic potential map of VIF was calculated by solving the Poisson–Boltzmann equation with the finite difference method (Delphi module in Insight II).

Results and discussion

Construction of VIF model

PSI-BLAST[43,44] was used to search the homologous sequence of VIF in the Protein Data Bank,[45] but the highest sequence similarity to any protein in the Data Bank was only 21%. The result implied that no template sequence was available for building the whole VIF directly. Recently, it was reported that proteins of the SOCS family, which are involved in forming a unique SCF-like E3–ubiquitin–ligase complex, act as a linker between their specific substrate and the ElonginB–ElonginC complex. Each SOCS protein contained at least two domains. One is a C-terminal domain that is responsible for interacting with ElonginB–ElonginC complex, and the other is an N-terminal domain that is responsible for binding to the specific substrate. VIF is a SOCS-box-containing protein and is also composed of two domains;[18] one is the SOCS-box located at the C-terminal region that interacts with the ElonginB–ElonginC complex, and the other is located in the N-terminal region that binds with APOBEC3G. Furthermore, it was reported that members of the SOCS family have similar C-terminal domains, whereas their N-terminal domains are diverse both in sequence and structure. Based on these reports, it is reasonable to divide VIF into two domains and construct the model of each domain separately.

C-terminal domain of VIF. It was reported recently[27] that the conserved SLQ(Y/F)LAΦΦΦΦ motif of VIF, which was located at the C-terminal region, possesses striking resemblance with the SOCS-box of SOCS proteins. The similarity did not only show in sequence homology, but also in function. The SOCS-box in SOCS proteins has been reported[24] to form a direct interaction with ElonginB–ElonginC, which is crucial for the formation of the SOCS–ElonginB–ElonginC complex, and thereby initiates the ubiquitin–proteasome degradation pathway. Since the SLQ(Y/F)LAΦΦΦΦ motif of VIF was also demonstrated to interact with ElonginB–ElonginC in the VIF–ElonginB–ElonginC complex, it was widely speculated[27] that the SLQ(Y/F)LAΦΦΦΦ motif of VIF was a novel type of SOCS-box and might adopt a 3D structure similar to the SOCS-box in SOCS proteins when interacting with the ElonginB–ElonginC complex.

To investigate further the similarity between VIF and the SOCS family members, a multiple sequence alignment was carried out. Sequences of several members of the SOCS family, including VHL, SOCS-1,[46] ASB2,[47] WSB1,[48] and RAR,[49] were aligned with VIF by ClustalW (Fig. 2a). The SOCS-box domains in these proteins (colored pink) that mediate interactions with ElonginB–ElonginC

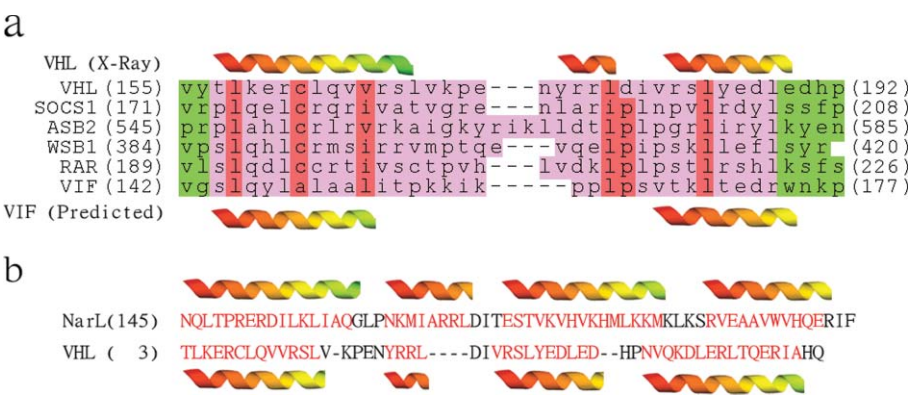

Fig. 2 (a) Sequence alignment of VHL and five SOCS family proteins. Residues in the SOCS-box are colored pink. Key point positions that are reported to make direct interactions to the ElonginB–ElonginC complex are colored red. The secondary structure of VIF is predicted by PSIPRED, and the secondary structure of VHL is taken from its crystal structure. (b) Sequence alignment between the SOCS-box of VIF and the C-terminal domain of NarL.

are fairly well conserved, especially the key positions (highlighted in red) that have been reported to make direct interaction with the ElonginB–ElonginC complex. One or two unique types of amino acids are conserved in these key point positions in the SOCS-box family of proteins. Therefore, mutation of these amino acids will lead to impairment of the ability of SOCS-box to bind to the ElonginB–ElonginC complex. The conservation of key point amino acid residues implies that the C-terminal domain of VIF likely have 3D structure similar to that of other SOCS-box domains. The 3D structure with conserved amino acids at crucial positions must be favored in the formation of the VIF–ElonginB–ElonginC complex. More evidence that supports the structure similarity between VIF and SOCS proteins comes from the secondary structure comparison between VIF and VHL (PSIPRED was used to predict the secondary structure of VIF). The results showed that two helices, which were highly similar to those in the SOCS-box region of VHL, existed in the putative SOCS-box region of VIF (Fig. 2a), suggesting that, at least in the SOCS-box regions, VIF and VHL have a similar structure. Since VHL protein (PDB code 1LM8)[50] of the SOCS protein family is the only one that currently has an available crystal structure, its structure was used as the template for building the structure of the C-terminal domain of VIF.

N-terminal domain of VIF. Although VIF and VHL display strong similarity in their C-terminal domains, their N-terminal domains bear no resemblance. When comparing the secondary structure of VIF predicted by PSIPRED with the crystal structure of VHL in their N-terminal regions, little similarity was found. Specifically, only β-sheets were found in the N domain in VHL, but there were at least three helices in the corresponding part of VIF. This indicates that VIF and VHL adopt different structure folds in their N-terminal domain, and it was inappropriate to use VHL as a single template for the construction of VIF.

It is well known that the 3D structures of homologous proteins are better conserved than their amino acid sequences. Remote homologs can have as few as 10–15% conserved residues and yet still have very similar 3D structures and often similar functions. The 3D structure is a more acceptable index for characterizing poorly homologous proteins than sequence. Threader3, which searches templates of domains based on a fold-recognition method, was used in this study to search for 3D structures similar to the N-terminal domain of VIF. The N-terminal domain sequence of VIF (1–143), along with its secondary structures predicted by the PSI-PRED, were submitted to Threader3 to search its redundant fold database. The top 10 proteins (according to their *Z*-scores) are shown in Table 1. Among the templates suggested by Threader3, four of them shared the same folds. They were the nitrate/nitrite response regulator from *E. coli* (NarL, PDB code 1A04), the hydrogenase from *E. coli* (PDB code 1CFZ), and the oxidoreductases from *Desulfovibrio africanus* (PDB code 1KEK) and *Desulfovibrio gigas* (PDB code 1E5D). All these proteins share a domain with similar folds, which consist of four (1KEK) or five α-helices and five parallel β-sheets, while other parts of these proteins differ remarkably.

NarL not only showed a strong similarity in secondary structures to VIF according to Threader3, but also is composed of an N-terminal domain and a C-terminal domain. While the N-terminal domain in NarL displayed good similarity to the N-terminal domain of VIF, its C-terminal domain also showed excellent structural resemblance to the SOCS-box in VHL (Fig. 2b, Fig. 3). Besides structural homology, the protein NarL and VIF share some functional similarities as well. It was reported[28] that phosphorylation and multimer formation are important for the function of NarL *in vitro*. Similar phenomena were also found for VIF. Based on these facts, the N-terminal domain of VIF was constructed by using NarL as the template.

VHL and NarL were used to construct the C-terminal putative SOCS-box region and the N-terminal region of VIF, respectively. In order to give a good orientation of these two domains, the alignment structure of VHL and NarL (Fig. 3) was used as the template. The last 15 residues (178–192) at the C-terminal region were neglected in the model construction, since they showed little structure similarity to both VHL and NarL. Furthermore, it was reported that the last 10–20 residues in the C-terminal are functionally unimportant, because the antiviral activities of VIF is still retained even these residues are deleted.[51] The quality of the constructed VIF model was evaluated by Procheck and Verify 3D. According to Ramachandran plot, most of the residues had Φ and Ψ angles in the core (82.7%) and allowed (14.7%) regions, only a small part of residues (2.6%) were in the generous allowed regions, and no residue was in disallowed region. Moreover, the verified 3D result showed that all residues were above the threshold.

The whole VIF–ElonginB–ElonginC complex was submitted to 5 ns of molecular dynamics simulation. The RMSD of the backbone atoms of VIF in the process of simulation was extracted and displayed in Fig. 4a. As shown in Fig. 4a, in the first 2 ns,

Table 1 Threader3 results for the N-terminal region (residues 1–143) of VIF. According to Threader3, a *Z*-score >4 indicates "very significant—probably a correct prediction", >3.5 indicates "significant—good chance of being correct", >2.7 indicates "borderline significant—possibly correct", and >2.0 indicates "poor score—could be right, but needs other confirmation"

PDB code	Description	Alignment length	*Z*-score	Rank
1a04A2	The nitrate/nitrite response regulator	117	3.91	1
1cfzA0	Hydrogenase maturating endopeptidase HYBD	121	3.13	2
1mfa01	Immunoglobulin	97	3.10	3
1kekA3	Pyruvate–ferredoxin oxidoreductase	129	3.04	4
1fdr02	Flavodoxin reductase	129	2.85	5
1v5uA0	Pleckstrin homology domain of SBF1	106	2.77	6
1bebA0	Bovine β-lactoglobulin	128	2.75	7
1jmxA4	Amine dehydrogenase	120	2.73	8
1e5dA1	Rubredoxin: oxygen oxidoreductase	117	2.59	9
2pth00	Peptidyl–tRNA hydrolase	125	2.55	10

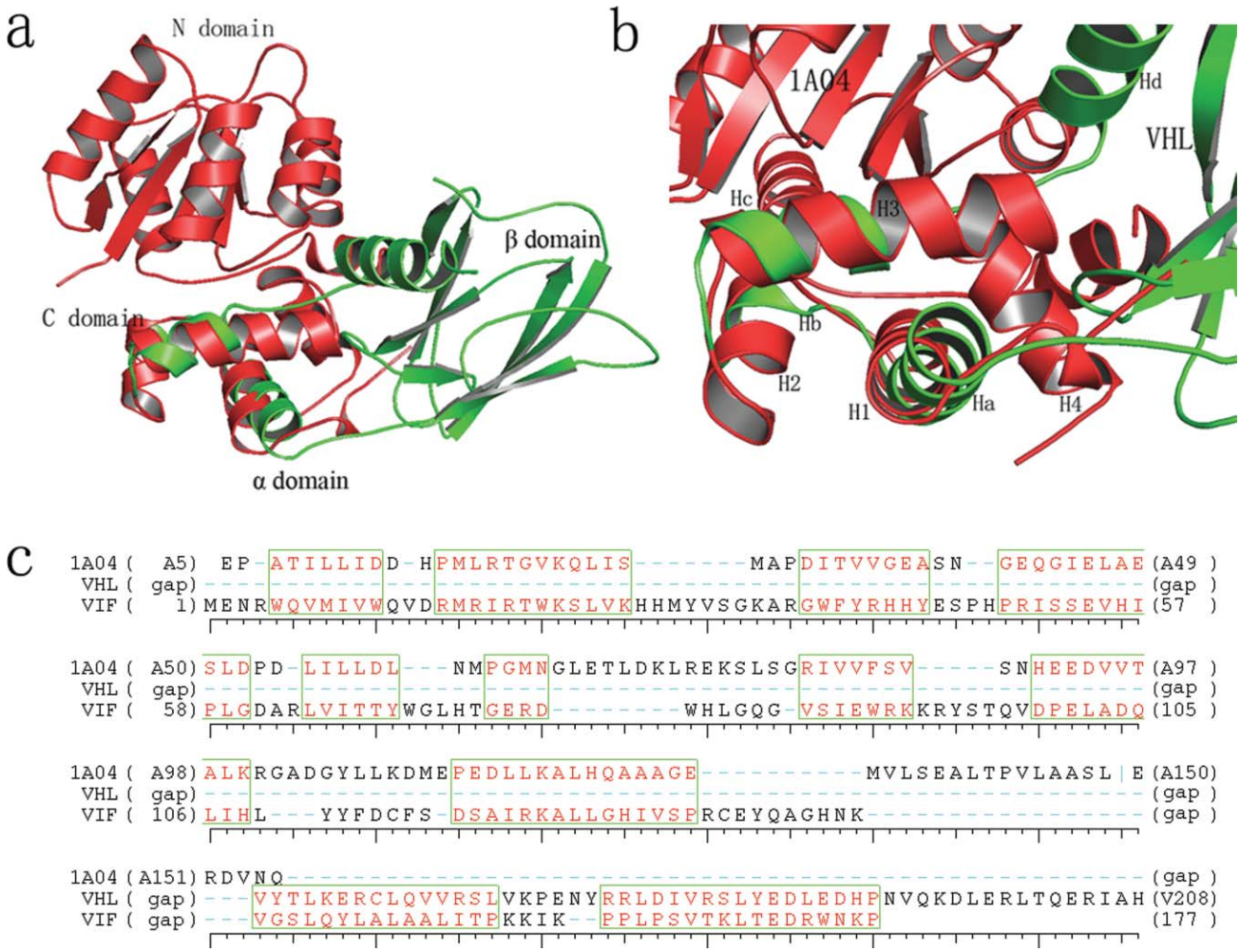

Fig. 3 (a) and (b) The superposition of VHL SOCS-box and the C-terminal domain of NarL. (c) The sequence alignment of VIF with NarL and VHL. Residues in the box are defined as structurally conserved regions (SCRs) and other residues are LOOPs.

the protein relaxed slowly, and then it reached a steady state. This simulation results indicated that an acceptable model was constructed. The steady state represented an optimized model of VIF, and a snapshot after 5 ns of simulation of the whole complex was extracted for investigation (Fig. 4b, 4c).

Model validation and function analysis

Interactions between VIF and the ElonginB–ElonginC complex. It has been revealed by mutational analysis[27,51] that some residues such as Ser144, Ala149, Leu163 and Leu169 in SLQ(Y/F)LAΦΦΦΦ motif were critical for the activity of VIF. Mutations of these residues would severely impair the ability of VIF to interact with ElonginC. The interaction between the SOCS-box of VIF and ElonginC and its comparison with the VHL(SOCS-box)–ElonginC interface are illustrated in Fig. 5. As shown in Fig. 5, the VIF–ElonginC interface involved two helices of VIF and two helices plus one loop region of ElonginC. Leu145, Ala149, Ala152 and Leu153 of VIF formed a hydrophobic cluster which made interaction with ElonginC. In addition, Leu163 and Leu169 constituted another hydrophobic cluster which contributed to the stabilization of the interaction between VIF and ElonginC. Based on this VIF(SOCS-box)–ElonginB–ElonginC model, mutations were performed (Table 2) and the stabilities of the mutant models were investigated by 5 ns of molecular dynamics simulation.

Although the SLQ(Y/F)LAΦΦΦΦ motif of VIF displays strong similarity to the SOCS-box of the SOCS family proteins, a major difference between the SOCS-box motif of VIF and other SOCS-box motifs is that the highly conserved cysteine (Cys162 in VHL) is replaced by an alanine (Ala149) in VIF (Fig. 2a). It was reported that this conserved cysteine in the SOCS box is critical for binding with ElonginC through a hydrophobic interaction[26] (Fig. 5). If VIF interacts with ElonginC through this SOCS-box motif, it is expected that both the alanine and cysteine are functional and exchangeable at this position. In this study, both the wild-type VIF model and the VIF-A149C mutant model remained stable during the whole simulation processes (Fig. 6). The initial structures of these two models were well conserved, especially the two helices (including helix-1 and helix-2), whose secondary structures were almost unchanged in the simulation. This conservation was also reflected in the RMSD values of the backbone atoms of the two helices. Both RMSD of helix-1 and helix-2 in these two models not only remained small but also

Table 2 Models used for molecular dynamics simulations

Models	Description	Biological activity
VIF	Wild type VIF (residue 142 to 177)	+
VIF-A149C	Ala 149 was mutated to Cys	+
VIF-A149L	Ala 149 was mutated to Leu	–
VIF-SLQ(AAA)	Ser 144, Leu 145, Gln 146 were mutated to Ala	–
VIF-Elc4	Ala 100, Leu 101, Leu 103, Leu 104 were mutated to Ser	–
VIF-L163S	Leu 163 was mutated to Ser	–
VIF-L169S	Leu 169 was mutated to Ser	–
VHL	Wild type VHL (residue 155 to 192)	+
VHL mutant	Cys 162 was mutated to Ala	+

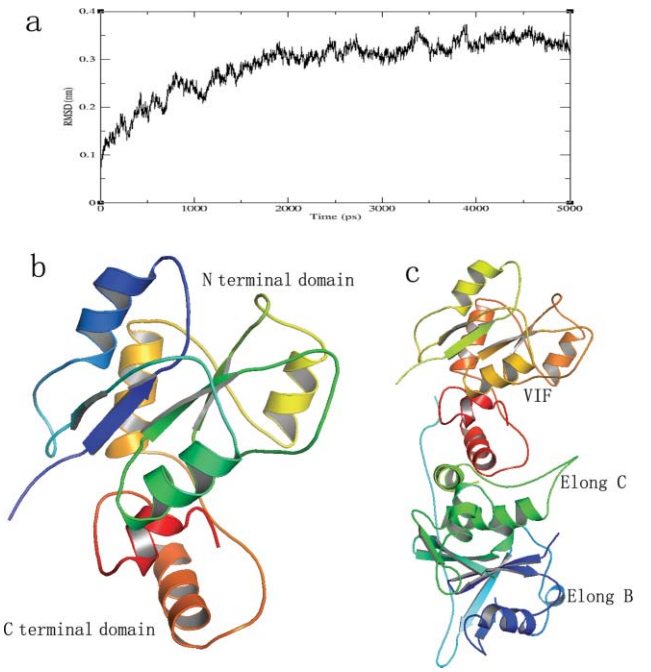

Fig. 4 (a) Time-dependent RMSD of VIF backbone atoms during 5 ns of molecular dynamics simulation. (b) and (c) The structures of VIF and the VIF–ElonginB–ElonginC complex after 5 ns of molecular dynamics simulation.

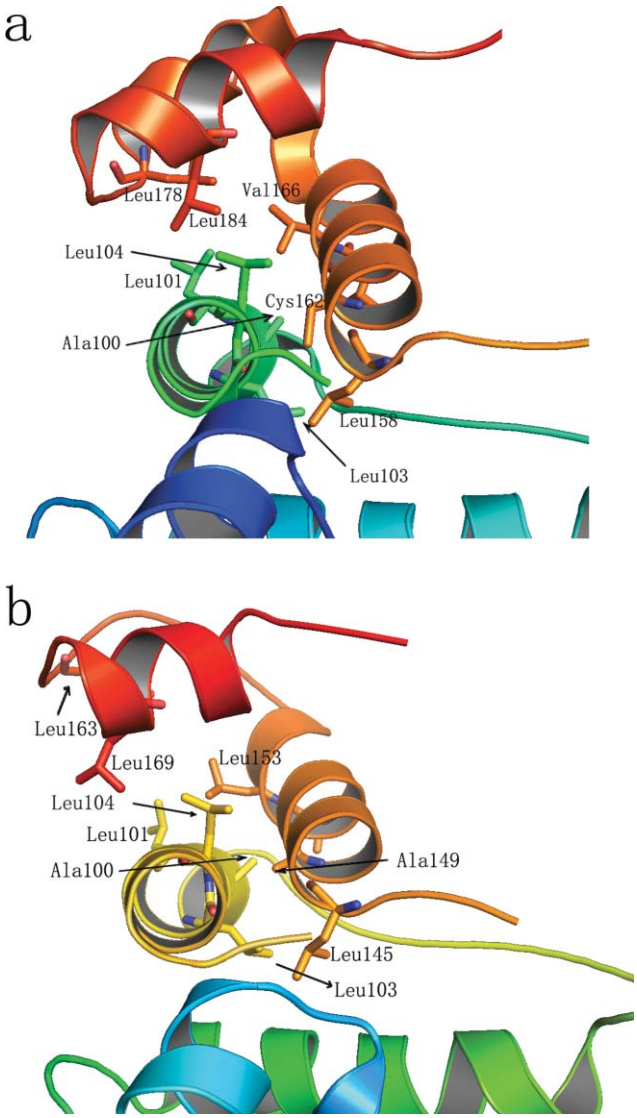

Fig. 5 The interfaces of (a) VHL(SOCS-box)–ElonginC and (b) VIF(SOCS-box)–ElonginC.

stable. These results implied that Ala149 of VIF could be replaced by cysteine and the replacement would not affect the binding of VIF with ElonginC. This exchangeable behavior was supported by experimental mutational analysis[27] that an A149C substitution in VIF did not significantly alter its interaction with ElonginC, and more importantly, its function of blocking the antiviral activity of APOBEC3G was retained. Moreover, experimental results demonstrated that since both C and A have short side chains, when A149 of VIF was changed to another hydrophobic amino acid, such as leucine, which has a longer side chain, VIF lost its ability to form a complex with ElonginC.[27] To test if our VIF model would show structural differences when A was mutated to L, a VIF-A149L mutant model was also constructed and submitted to a molecular dynamics simulation. In contrast to wild-type VIF and VIF149C models, the VIF-A149L mutant model showed instability in the simulation process and its initial structure was severely impaired (Fig. 6). Helix-1, where the mutant leucine is located, had almost the same secondary structure, and the RMSD was kept at about 0.9 Å. However, the structure of the helix-2 region was totally destroyed to a coil, and it was hard to reach a stable structure in the process of simulation. Our simulation study suggested that mutation of 149A to 149L of VIF led to the destruction of SOCS-box motif and loss of the binding ability to ElonginC, which was in good agreement with experimental observations.

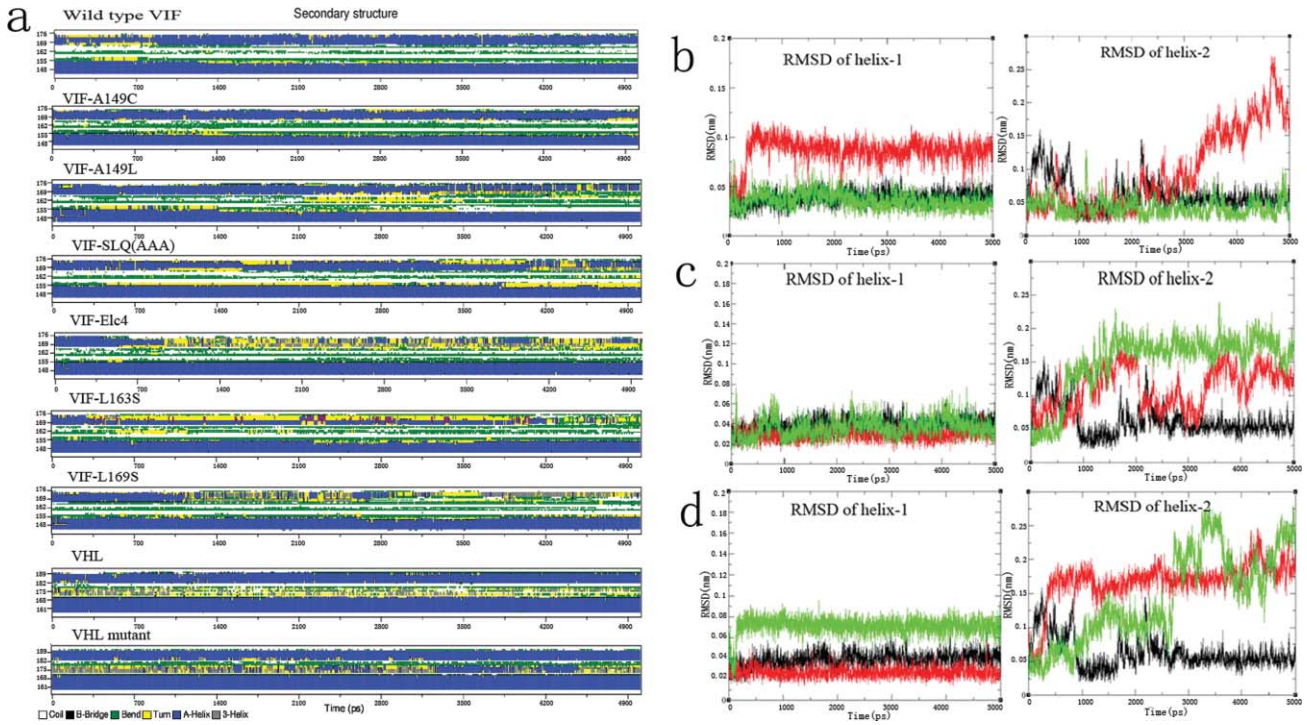

Fig. 6 (a) Secondary structure, analyzed using DSSP, as a function of time for SOCS-box region of VIF. (b) The RMSD of the backbone atoms of helix-1 (residues 145 to 152) and helix-2 (residues 167 to 174). The curve of wild-type VIF is colored black, while VIF-A149C is green, and VIF-A149L is red. (c) The RMSD helix-1 and helix-2 as described in (b). Wild-type VIF is colored black, while VIF-SLQ(AAA) is colored red and VIF-Elc4 is colored green. (d) The RMSD helix-1 and helix-2 as described in (b). Wild-type VIF is colored black, while VIF-L163S is red, and VIF-L169S is green.

The SLQ motif is important for VIF to interact with ElonginC.[20] In the simulated model, this motif fitted into a concave surface formed by two helices of ElonginC, and the interface was totally hydrophobic. The most significant van der Waals contact was made by Leu145, which protruded from helix-1 and into an ElonginC pocket (Fig. 5). Mutation of these residues (SLQ) to AAA, which possessed shorter side chains, significantly reduced their hydrophobic interactions. According to the VIF-SLQ(AAA) model, although helix-1 remained intact during the whole simulation process, one helix of ElonginC, which forms the hydrophobic pocket, tended to move out of the pocket. Moreover, the structure of helix-2 collapsed into a coil. The large fluctuation of RMSD values of helix-2 implied that the binding of this region to ElonginC was lost (Fig. 6). The simulation results were consistent with experimental mutation analysis, which indicated that the interaction of VIF SOCS-box (SLQ to AAA) and ElonginC was significantly reduced.[27,51]

Further investigation of the VIF–ElonginC interface revealed that four hydrophobic residues in ElonginC (Ala100, Leu101, Leu103, Leu104) played major roles in mediating its interactions with VIF. In the VIF-Elc4 model, these residues were all replaced by hydrophilic serine. Simulation results showed that the introduction of serine completely destroyed the VIF–ElonginC hydrophobic interface and that helix-2 of VIF collapsed to a coil (Fig. 6), agreeing with the experimental observation that the mutated ElonginC (Elc4) has a drastically reduced ability to interact with VIF.[27]

In addition to the SLQ motif, downstream Leu163 and Leu169 in the SOCS-box of VIF were also reported to be required for VIF function.[27] Mutations of these two residues to the more hydrophilic serine residues showed diminished activity of interaction with ElonginC. In the simulated models of wild-type VIF, one of these residues was located in helix-2 and the other was in the loop region. Both Leu163 and Leu169 contributed to the hydrophobic interactions with a helix in ElonginC. Molecular dynamics simulations of the VIF-L163S and VIF-L169S models verified the importance of these two residues, since both models displayed instability in the simulation process, especially in their helix-2 regions (Fig. 6). As in all the unstable models simulated, helix-2 of the VIF-L163S and VIF-L169S models turned into coils, and lost their ability to interact with ElonginC.

In order to ensure that the simulations were not in favor of protein unfolding, the same simulation procedures were performed on a complex of ElonginB–ElonginC along with the SOCS-box of VHL (155–192), whose crystal structure is known and can be obtained directly from the crystal structure 1LM8. To investigate whether alanine and cysteine were exchangeable in position 162, a mutated model VHL-C162A was constructed. Stable models were observed in the process of molecular dynamics simulations, in which the structure of helix-1 and helix-2 were maintained (Fig. 6). This result suggested that the simulation processes were not inclined to make protein chaos.

To sum up, 9 complex models were constructed, including both wild-type and mutated VIF/VHL(SOCS-box)–ElonginB–ElonginC complex models, and every model was submitted to a 5 ns molecular dynamics simulations. Through 9 independent molecular dynamics simulations, the dynamics behavior of the SOCS-box–ElonginC interface was characterized. Mutant

VIF-A149C and mutant VHL-C162A were reported to retain their activity to interact with ElonginC. Their models, along with wild-type VIF and VHL models, were constructed, appeared to be stable during the simulations, and the structures were maintained. In contrast, other models of mutated proteins, including VIF-A149L, VIF-SLQ(AAA), VIF-Elc4, VIF-L163S, VIF-L169S, which were reported to lose their ability to interact with ElonginC, were unstable, and (at least partly) disrupted structures were observed during simulations. In all simulations, the helix-2 of VIF appeared to be very sensitive to their structural stabilities. Unstable structures invariably led to the disappearance of the helix structure of helix-2, even when the mutated residue was located in helix-1. The fact that all simulation results were consistent with reported experimental mutational analysis supports the statement that VIF is a SOCS-box-containing protein, and the putative SOCS-box mediates its interaction with ElonginB–ElonginC.

Insight into Cys114 and Cys133. Two conserved cysteines, Cys114 and Cys133, play an important role in the function of VIF. Mutations of these two cysteines to serines destroy the ability of VIF to down-regulate APOBEC3G. Yu *et al.*[27] found that VIF with C114S or C133S mutation still binds to the ElonginB–ElonginC complex but not to Cul5. The inability to form a VIF–BC–Cul5 complex leads to the loss of function of VIF. These results indicated that these two cysteines might be responsible for the interaction with Cul5. However, according to the reports of Mehle *et al.*[52] and Masayuki *et al.*,[53] VIF with C114S or C133S mutations still interacts with Cul5 at least weakly, and forms the VIF–BC–Cul5 complex, but the C114S–BC–Cul5 or C133S–BC–Cul5 complex loses the E3 ligase activity toward APOBEC3G. Thus Masayuki *et al.* proposed that instead of interacting with Cul5, Cys114 and Cys133 might affect the conformation of VIF and perturb the position of APOBEC3G in the ligase complex, resulting in the loss of the E3 ligase activity of the VIF–BC–Cul5 complex. Based on the VIF–ElonginB–ElonginC model we constructed, Cys114 and Cys133 were located in two adjacent loop regions (Fig. 7). These two loops, along with the two helices surrounding them, constituted a groove. This groove was heavily negatively charged and forms the negatively charged center in VIF. Thus it is highly possible that Cys114 and Cys133 of VIF might be directly involved in the interaction with another protein, like CuI5 or APOBEC3G.

PPLP domain. According to our VIF–ElonginB–ElonginC model, the 161PPLP164 domain was located in the SOCS-box region of VIF and constituted a positively charged concave surface (Fig. 8). This concave face is important, since synthesized peptides containing PPLP domain were reported to potently block the function of VIF.[54,55] Thus, this interface could be targeted in a new strategy to develop diverse VIF inhibitors, such as peptidomimetics and other small organic molecules.

The importance of this domain is widely appreciated due to its role in mediating VIF multimerization, since multimerization is reported to be necessary for the function of VIF in the viral life cycle, and deletion of the PPLP domain significantly impairs both the ability of VIF to form multimers and its ability to down-regulate APOBEC3G.[54,55] However, in our model, the PPLP domain partly interacts with ElonginB. It is possible that the deletion of the PPLP domain might also affect the VIF–ElonginB interactions and inhibit the formation of a VIF–BC–Cul5 complex, therefore leading to the loss of VIF function.

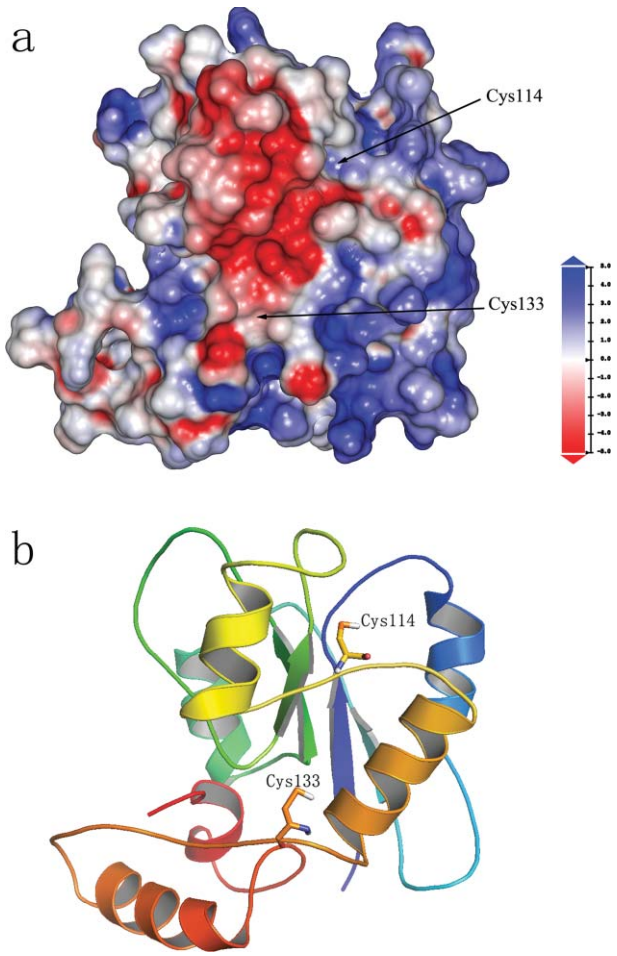

Fig. 7 (a) Solid surface of VIF showing negatively charged concave. Coloring is according to the electrostatic potential at the surface calculated with DELPHI. (b) 3D structure of VIF from the same viewpoint as (a).

Conclusion

In conclusion, a three-dimensional model of HIV-1 VIF proteins was constructed using homology modeling. Due to low sequence identity, the modeling process was performed using two templates. The C-terminal domain of VIF was reported to be a novel SOCS-box, so it was constructed using the SOCS-box of VHL as the template. The N-terminal domain of VIF was constructed using NarL as the template, because the N-terminal domain of NarL showed strong similarity to the N-terminal domain of VIF in its secondary structure. Molecular dynamics simulations revealed that this VIF(SOCS-box)–ElonginB–ElonginC model was stable, and mutation of the critical residue located in the VIF–ElonginC interface led to an unstable model. These results were consistent with experimental mutational analysis. The model obtained in this study provided the structural information on the molecular level for investigating the function of VIF and exploring its prospects as a novel drug target.

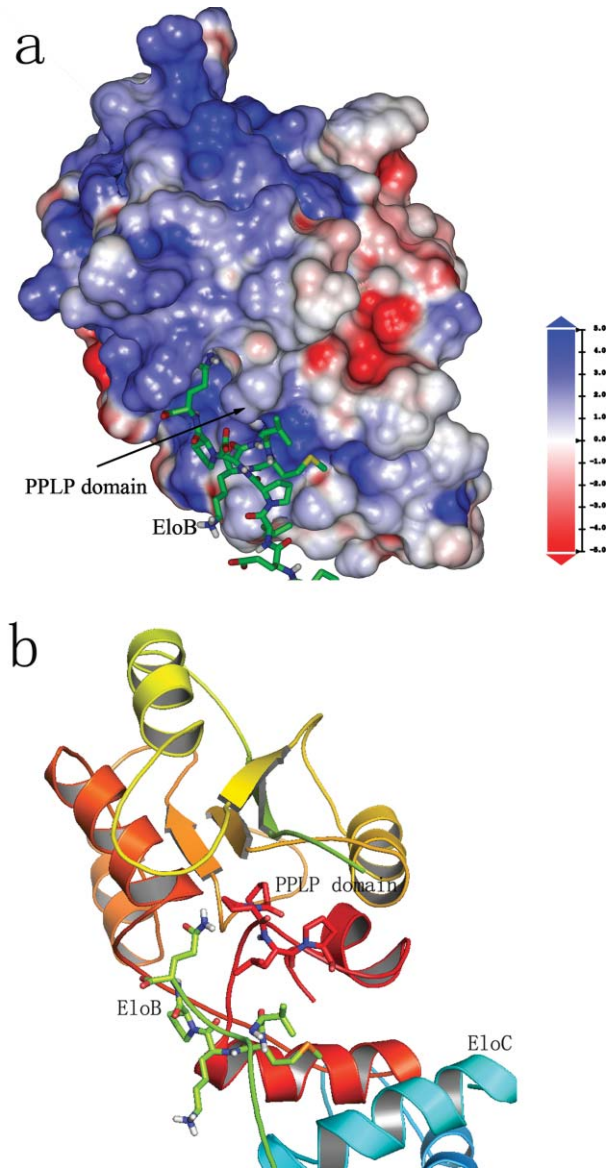

Fig. 8 (a) Solid surface of VIF showing positively charged concave surface. Coloring is according to the electrostatic potential at the surface calculated with DELPHI. (b) 3D structure of VIF from the same viewpoint as (a).

Acknowledgements

This study was supported by the Ministry of Science and Technology of China (Grant No. 2004CB518904). We thank Dr Jundong Zhang for help in revising the manuscript and Mr Chao Ma for help in preparing the figures.

References

1 D. Trono, *Cell*, 1995, **82**, 189–192.
2 M. Emerman and M. H. Malim, *Science*, 1998, **280**, 1880–1884.
3 N. Madani and D. Kabat, *J. Virol.*, 1998, **72**, 10251–10255.
4 D. H. Gabuzda, K. Lawrence, E. Langhoff, E. Terwilliger, T. Dorfman, W. A. Haseltine and J. Sodroski, *J. Virol.*, 1992, **66**, 6489–6495.
5 H. M. S. James, C. G. Nathan, A. M. F. Ron and H. M. Michael, *Nat. Med.*, 1998, **4**, 1397–1400.
6 M. S. Ann, C. G. Nathan, D. C. Jonathan and H. M. Michael, *Nature*, 2002, **418**, 646–650.
7 M. R. Kristaine, M. Mariana, L. K. Susan and K. David, *Trends Mol. Med.*, 2004, **10**, 292–297.
8 B. Lea and K. Moshe, *Curr. Med. Chem.*, 2003, **11**, 221–231.
9 A. Jarmuz, A. Chester, J. Bayliss, J. Gisbourne, I. Dunham, J. Scott and N. Navaratnam, *Genomics*, 2002, **79**, 285–296.
10 N. B. Kate, K. H. Rebecca, M. S. Ann, O. D. Nicholas, C. Soo-Jin and H. M. Michael, *Curr. Biol.*, 2004, **14**, 1392–1396.
11 P. Phuong, B. Ronda and F. G. Myron, *Biochemistry*, 2005, **44**, 2703–2715.
12 R. Mariani, D. Chen, B. Schröfelbauer, F. Navarro, R. König, B. Bollman, C. Münk, H. Nymark-McMahon and N. R. Landau, *Cell*, 2003, **114**, 21–31.
13 D. Lecossier, F. Bouchonnet, F. Clavel and A. J. Hance, *Science*, 2003, **300**, 1112.
14 M. Bastien, T. Priscilla, C. Gersende, F. Marc, P. Luc and T. Didier, *Nature*, 2003, **424**, 99–103.
15 H. Zhang, B. Yang, R. J. Pomerantz, C. Zhang, S. C. Arunachalam and L. Gao, *Nature*, 2003, **424**, 94–98.
16 K. Shindo, A. Takaori-Kondo, M. Kobayashi, A. Abudu, K. Fukunaga and T. Uchiyama, *J. Biol. Chem.*, 2003, **278**, 44412–44416.
17 K. Stopak, C. de Noronha, W. Yonemoto and W. C. Greene, *Mol. Cell*, 2003, **12**, 591–601.
18 M. Mariana, M. R. Kristine, L. K. Susan and K. David, *Nat. Med.*, 2003, **9**, 1398–1403.
19 A. M. Sheehy, N. C. Gaddis and M. H. Malim, *Nat. Med.*, 2003, **9**, 1404–1407.
20 X. H. Yu, Y. K. Yu, L. D. Liu, K. Luo, W. Kong, P. Y. Mao and X. F. Yu, *Science*, 2003, **302**, 1056–1060.
21 A. Mehle, B. Strack, P. Ancuta, C. Zhang, M. McPike and D. Gabuzda, *J. Biol. Chem.*, 2004, **279**, 7792–7798.
22 L. D. Liu, X. H. Yu, K. Luo, Y. K. Yu and X. F. Yu, *J. Virol.*, 2004, **78**, 2072–2081.
23 J. G. Zhang, A. Farley, S. E. Nicholson, T. A. Willson, L. M. Zugaro, R. J. Simpson, R. L. Moritz, D. Cary, R. Richardson, G. Hausmann, B. J. Kile, S. B. H. Kent, W. S. Alexander, D. Metcalf, D. J. Hilton, N. A. Nicola and M. Baca, *Proc. Natl. Acad. Sci. U. S. A.*, 1999, **96**, 2071–2076.
24 T. K. Benjamin, A. S. Brenda, S. A. Warren, A. N. Nicos, M. E. M. Helene and J. H. Douglas, *Trends Biochem. Sci.*, 2002, **27**, 235–241.
25 K. Iwai, K. Yamanaka, T. Kamura, N. Minato, R. C. Conaway, J. W. Conaway, R. D. Klausner and A. Pause, *Proc. Natl. Acad. Sci. U. S. A.*, 1999, **96**, 12436–12441.
26 E. S. Charles, G. K. William, Jr. and P. P. Nikola, *Science*, 1999, **284**, 455–461.
27 Y. K. Yu, Z. X. Xiao, E. S. Ehrlich, X. H. Yu and X. F. Yu, *Genes Dev.*, 2004, **18**, 2867–2872.
28 B. Igor, S. Imke, K. G. Maria, C. Duilio, P. G. Rovert and E. D. Richard, *Biochemistry*, 1998, **37**, 3665–3676.
29 D. T. Jones, *J. Mol. Biol.*, 1999, **292**, 195–202.
30 D. T. Jones, W. R. Taylor and J. M. Thornton, *Nature*, 1992, **358**, 86–89.
31 D. Higgins, J. Thompson, T. Gibson, J. D. Thompson, D. G. Higgins and T. J. Gibson, *Nucleic Acids Res.*, 1994, **22**, 4673–4680.
32 A. Sali and T. L. Blundell, *J. Mol. Biol.*, 1993, **234**, 779–815.
33 A. Sali, L. Potterton, F. Yuan, H. van Vlijmen and M. Karplus, *Proteins*, 1995, **23**, 318–326.
34 H. J. C. Berendsen, D. van der spoel and A. R. van Drunen, *Comput. Phys. Commun.*, 1995, **91**, 43–56.
35 E. Lindahl, B. Hess and D. van der Spoel, *J. Mol. Model*, 2001, **7**, 306–317.
36 D. van der Spoel, A. R. V. Buuren, D. P. Tieleman and H. J. Berendsen, *J. Biomol. NMR*, 1996, **8**, 229–238.
37 H. J. C. Berendsen, J. P. M. Postma, W. F. V. Gunsteren, A. DiNola and J. R. Haak, *J. Chem. Phys.*, 1984, **81**, 3684–3690.
38 B. Hess, H. Bekker, H. J. C. Berendsen and J. G. E. M. Fraaije, *J. Comput. Chem.*, 1997, **18**, 1463–1472.
39 T. Darden, D. York and L. Pedersen, *J. Chem. Phys.*, 1993, **98**, 10089–10092.
40 W. Kabsch and C. Sander, *Biopolymers*, 1983, **22**, 2577–2637.
41 R. A. Laskowski, M. W. MacArthur, D. S. Moss and J. M. Thornton, *J. Appl. Crystallogr.*, 1993, **26**, 283–291.
42 J. U. Bowie, R. Luthy and D. Eisenberg, *Science*, 1991, **253**, 164–170.
43 S. F. Altschul, T. L. Madden, A. A. Schaffer, J. Zhang, Z. Zhang, W. Miller and D. J. Lipman, *Nucleic Acids Res.*, 1997, **25**, 3389–3402.

44 A. A. Schaffer, L. Aravind, T. L. Madden, S. Shavirin, J. L. Spouge, Y. I. Wolf, E. V. Koonin and A. S. F. ltschul, *Nucleic Acids Res.*, 2001, **29**, 2994–3005.
45 H. M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T. N. Bhat, H. Weissig, I. N. Shindyalov and P. E. Bourne, *Nucleic Acids Res.*, 2000, **28**, 235–242.
46 S. Minamoto, K. Ikegame, K. Ueno, M. Narazaki, T. Naka, H. Yamamoto, T. Matsumoto, H. Saito, S. Hosoe and T. Kishimoto, *Biochem. Biophys. Res. Commun.*, 1997, **237**, 79–83.
47 J. Kohroki, S. Fujita, N. Itoh, Y. Yamada, H. Imai, N. Yumoto, T. Nakanishi and K. Tanaka, *FEBS Lett.*, 2001, **505**, 223–228.
48 D. Vasiliauskas, S. Hancock and C. D. Stern, *Mech. Dev.*, 1999, **82**, 79–94.
49 M. T. Ross, *Nature*, 2005, **434**, 325–337.
50 E. S. Charles, G. K. William, Jr. and P. P. Nikola, *Science*, 1999, **284**, 5413.
51 H. M. S. James, M. S. Ann, A. C. Elise, A. M. F. Ron and H. M. Michael, *J. Virol.*, 1999, **73**, 2675–2681.
52 A. Mehle, J. Goncalves, M. Santa-Marta, M. McPike and D. Gabuzda, *Genes Dev.*, 2004, **18**, 2861–2866.
53 K. Masayuki, T. K. Akifumi, M. Yasuhiro, I. Kazuhiro and U. Takashi, *J. Biol. Chem.*, 2005, **280**, 18573–18578.
54 S. C. Yang, Y. Sun and H. Zhang, *J. Biol. Chem.*, 2001, **276**, 4889–4893.
55 B. Yang, L. Gao, L. Li, Z. X. Lu, X. J. Fan, C. A. Patel, R. J. Pomerantz, G. C. DuBois and H. Zhang, *J. Biol. Chem.*, 2003, **278**, 6596–6602.

