J Mol Model (2012) 18:4517–4527
DOI 10.1007/s00894-012-1450-z

ORIGINAL PAPER

Understanding the molecular mechanism of the broad and potent neutralization of HIV-1 by antibody VRC01 from the perspective of molecular dynamics simulation and binding free energy calculations

Yan Zhang · Dabo Pan · Yulin Shen · Nengzhi Jin · Huanxiang Liu · Xiaojun Yao

Received: 31 October 2011 / Accepted: 30 April 2012 / Published online: 29 May 2012

Abstract VRC01 is one of the most broadly and potently neutralizing HIV-1 antibodies known—it has been shown to neutralize 91 % of the tested primary isolate Env pseudoviruses by recognizing the viral envelope glycoprotein gp120. To explore the mechanism of HIV-1 neutralization by VRC01 and thus obtain valuable information for vaccine design, we performed molecular dynamics simulations and binding free energy calculations for apo-VRC01, apo-gp120, and the gp120–VRC01 complex. For gp120, residue energy decomposition analysis showed that the hotspot residues Asn280, Lys282, Asp368, Ile371, and Asp457 are located in three primary loops, including the CD4-binding loop, loop D, and loop V5. For VRC01, the hotspot residues Trp47, Trp50, Asn58, Arg61, Gln64, Trp100, and Tyr91 mainly come from CDR2 of the heavy chain. By decomposing the binding free energy into different components, intermolecular van der Waals interactions and nonpolar solvation were found to dominate the binding process. Principal component analysis of loops D and V5, which are related to neutralization resistance, indicated that these two areas have a larger conformational space in apo-gp120 compared to bound gp120. A comparison of three representative structures from the cluster analysis of loops D and V5 indicated that changes primarily occur at the tip of loop V5, and are caused by fluctuations in the terminal Glu1 residue of the antibody. This information can be used to guide the design of vaccines and small molecule inhibitors.

Keywords gp120 · VRC01 · HIV-1 · Molecular dynamics simulation · MM-PBSA · MM-GBSA

Y. Zhang · D. Pan · H. Liu
School of Pharmacy, Lanzhou University,
Lanzhou 730000, China

H. Liu (✉) · X. Yao
State Key Laboratory of Applied Organic Chemistry
and Department of Chemistry, Lanzhou University,
Lanzhou 730000, China
e-mail: hxliu@lzu.edu.cn

Y. Shen · N. Jin
Gansu Computing Center,
Lanzhou 730030, China

Introduction

Human immunodeficiency virus type 1 (HIV-1) causes acquired immunodeficiency syndrome (AIDS), which is one of the three diseases that cause the most deaths worldwide [1]. According to estimates, more than 33 million people are currently living with HIV-1 after nearly 30 years of this highly complex epidemic [2]. Env of the human immunodeficiency virus (HIV) consists of a transmembrane glycoprotein (gp41) and a surface glycoprotein (gp120), which is a heterodimer [3]. The surface glycoprotein (gp120) consists of inner and outer domains, and each domain contributes two β strands to form a "bridging sheet." These two domains and bridging sheet comprise the binding site of co-receptor (CD4), the binding of which prompts the virus to infect host cells and then selectively bind a co-receptor: either CCR5 or CXCR4 [4–9]. Thus, gp120 is very important in mediating HIV-1 attachment, and is a promising target for research efforts aimed at developing HIV-neutralizing drugs.

The best long-term solution to the HIV-1 epidemic is to develop a safe, effective vaccine [10]. One strategy for identifying potential vaccine candidates is to isolate broadly neutralizing antibodies from the infected individuals and then attempt to elicit the same antibody response through

vaccination. Wu et al. isolated and identified three antibodies that were specific for the CD4-binding site of gp120 and exhibited great breadth and potency of neutralization [11]. Among the isolated and identified antibodies, VRC01 was the most effective, as it neutralized 91 % of the tested primary isolate env pseudoviruses. Zhou et al. solved the crystal structure of VRC01 in complex with its target gp120. The interaction of gp120 with VRC01 takes place primarily in the outer domain of gp120, unlike previously isolated and identified antibodies. Moreover, their experiments found that changes in the loop V5 and loop D areas of gp120 are the most common source of natural resistance to VRC01 [12]. The availability of the gp120–VRC01 structure offers a starting point in the development of broadly neutralizing antibodies and the design of an effective vaccine targeting HIV [13]. However, even though the high-resolution crystal structure can provide a great deal of atomic-level structural information, our understanding of the associated interaction, binding energy profile, and dynamic process based only on the static complex is still very limited. Increasing our comprehension of these elements will be necessary in order to develop new and effective vaccines or compounds that neutralize HIV-1 by interfering with protein–protein interactions.

As already done for an increasing number of crystal structures of protein–protein complexes, we can proceed with structural analysis based on computational methods to explore the thermodynamics and kinetics of the gp120–VRC01 interaction [14]. All-atom molecular dynamics (MD) simulations can be used as a complementary method to gain dynamic structural and energy information on protein complexes, including the binding free energy between protein partners, which is difficult to obtain experimentally [15–17]. The most rigorous MD-based approaches used to estimate binding free energies are the free energy perturbation and thermodynamic integration methods. However, because of their long convergence time, they are computationally intensive and thus difficult to apply to macromolecules and their complexes [18, 19]. A more commonly used and more tractable approach is the molecular mechanics Poisson–Boltzmann surface area (MM-PBSA) method [20, 21]. Using such a method [22], it has been possible to predict the binding affinities for a variety of macromolecular complexes and protein–ligand complexes [23–25]. The MM-PBSA method uses conventional molecular mechanics force fields such as AMBER [26, 27] to calculate the gas-phase energy, and use the Poisson–Boltzmann (PB) model to calculate the solvation free energy. Solute entropy can be calculated from statistical thermodynamics using normal-mode analysis. Recently, molecular mechanics generalized Born surface area (MM-GBSA), a variant of MM-PBSA, has also been widely applied [28–31]. The main differences between them are that GB electrostatics are calculated using the generalized Born (GB) approximation for electrostatics in water [20, 32, 33] instead of PB electrostatics. Another advantage of MM-GBSA is that it utilizes a fully pairwise potential that is useful for decomposing the total binding free energy into atomic/group contributions in a structurally nonperturbing formalism [29].

In this study, all-atom MD simulations of apo-gp120, apo-VRC01, and the gp120–VRC01 complex were performed to elucidate the detailed mechanism of the interaction of gp120 with VRC01. We analyzed the binding interface and calculated the binding free energy (including the energy contributions of individual residues), based on MM-PBSA and MM-GBSA, in order to characterize the binding hotspots. H-bond occupancy analysis was also applied to identify several important hydrogen bonds. Moreover, the origins of conformational changes in the important loop D and loop V5 areas were elucidated using principal component analysis and cluster analysis. The results obtained here provide useful information for designing vaccines and effective inhibitors against HIV-1.

Materials and methods

Structure preparation

The structural complex of the core of antigen-binding fragment (Fab) of VRC01 and HIV-1 gp120 from clade A/E recombinant 93TH057 was taken from the Protein Data Bank (PDB code 3NGB) [12]. This structure consists of four copies of the VRC01–gp120 complex. We retained chain A (gp120), chain B (VRC01 heavy chain), and chain C (VRC01 light chain), but deleted the other chains. The studied complex (shown in Fig. 1) consisted of 783 residues. Hydrogen atoms were added using the AMBER10 package [34]. Molecular dynamics simulation was performed on the treated crystal structure, retaining the crystal water molecules.

Molecular dynamics simulation

The molecular dynamics simulations, including energy minimization, system equilibration and production protocols, were performed with the AMBER10 package [34] using the AMBER99SB force field [35]. Counterions were added to keep the whole system neutral. The systems were then solvated using atomistic TIP3P water [36] in a cubic box with at least 10 Å around the structures, including a total of 106,567 atoms. Each simulated system was constructed using periodic boundary conditions with a rectangular box.

We first carried out an energy minimization using the steepest descent method in AMBER10 for 10,000 iterations, with a force constant of 0.1 kcal mol^{-1} $Å^{-2}$ applied to the proteins, and we then performed a further energy minimization

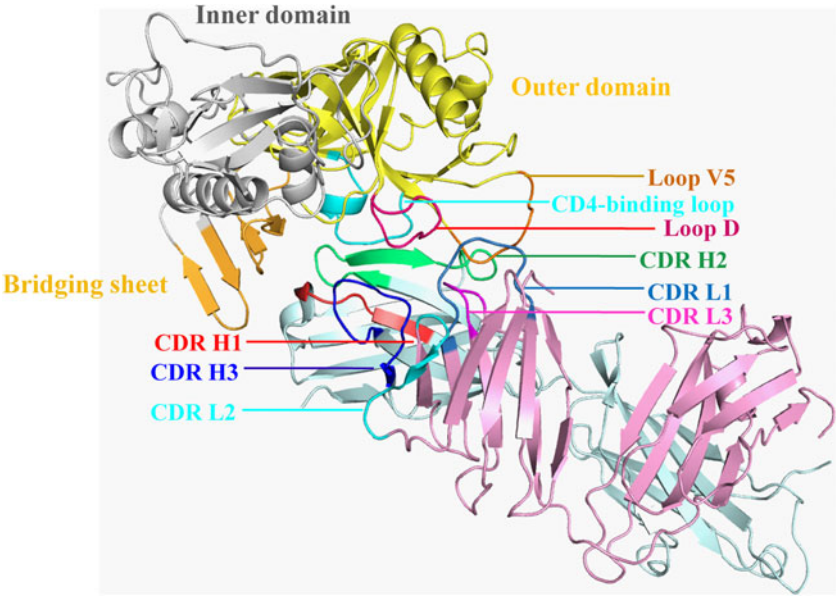

Fig. 1 Structure of HIV-1 gp120 in complex with antibody VRC01 (PDB 3NGB). Several important domains are shown in different colors

with the steepest descent method that was switched to conjugate gradient every 500 steps for a total of 25,000 steps, without any constraint. Particle mesh Ewald (PME) summation [37] was used to treat long-range coulombic interactions. The equilibration and subsequent production runs were carried out using the SHAKE algorithm [38] on all atoms covalently bonded to hydrogen atoms with an integration timestep of 2 fs. The systems were annealed from 0 to 300 K on a 50 ps timescale, applying a force constant of 2 kcal mol^{-1} $Å^{-2}$ to the proteins. All subsequent equilibration and production phases of the simulation were carried out in the isothermal isobaric (NPT) ensemble using a Berendsen barostat [39] in a target pressure of 1 bar and a pressure coupling constant of 2.0 ps, applying a force constant of 0.1 kcal mol^{-1} $Å^{-2}$ to the VRC01 partial residues that were more than 10 Å from gp120. The coordinate trajectories of all equilibration and production runs were recorded every 1 ps.

To study the conformational changes of apo-gp120 and bound gp120, three independent trajectories were performed for solvated systems of apo-gp120, apo-VRC01, and the gp120–VRC01 complex (15.0 ns each) in the NPT ensemble at 300 K.

Binding free energy calculation

The first step when applying MM-PBSA and MM-GBSA is to generate multiple snapshots from an MD trajectory of the protein–protein complex by stripping away water molecules and counterions. Snapshots at 10 ps intervals were extracted from the last 5 ns MD production runs. The free energy was calculated for the complex, the antigen, and the antibody in each snapshot, and the binding free energy was computed as follows:

$$\Delta G_{\text{bind}} = G_{\text{complex}} - G_{\text{antigen}} - G_{\text{antibody}}$$

The free energy G can be calculated by the following scheme, based on the MM-PBSA and MM-GBSA methods [20, 21]:

$$G = E_{\text{gas}} + G_{\text{sol}} - TS$$
$$E_{\text{gas}} = E_{\text{int}} + E_{\text{ele}} + E_{\text{vdw}}$$
$$E_{\text{int}} = E_{\text{bond}} + E_{\text{angle}} + E_{\text{torsion}}$$
$$G_{\text{sol}} = G_{\text{PB(GB)}} + G_{\text{SUR}}$$
$$G_{\text{SUR}} = \gamma\ \text{SAS}$$

Here, E_{gas} , E_{int} , E_{ele} , E_{vdw} , E_{bond} , E_{angle}, and E_{torsion} are, respectively, the gas-phase energy, the internal energy, the Coulomb energy, the van der Waals energy, the bond energy, the angle energy, and the torsion energy; G_{sol} is the solvation free energy, which can be decomposed into polar and nonpolar contributions. $G_{\text{PB(GB)}}$ is the polar solvation contribution, calculated by solving the PB or GB equation [20, 21]. The dielectric constants for the solute and solvent were set to 1 and 80, respectively. G_{SUR} is the nonpolar solvation contribution and is estimated via the solvent accessible surface area (SAS), determined using a water probe radius of 1.4 Å. The surface tension constant γ was set to 0.0072 kcal mol^{-1} $Å^{-2}$ [40]. S and T are the total solute entropy and the temperature, respectively. We did not calculate the vibrational entropy contributions here, since our aim was to identify hotspot residues at the binding interface and the detailed interaction features of the complex, rather than to obtain absolute values for the binding free energy. In fact, because of limited computational resources and the

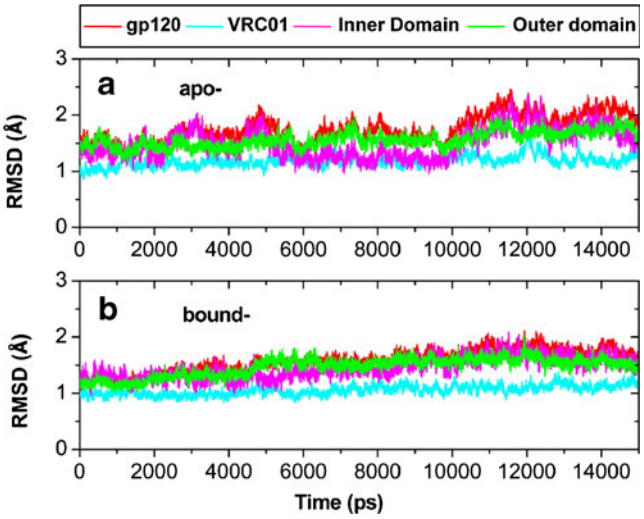

Fig. 2 RMSDs of the backbone atoms of gp120, VRC01, and the inner and outer domains of gp120, using the crystal structure as reference: **a** RMSDs of apo-gp120, apo-VRC01, and the inner and outer domains of apo-gp120; **b** RMSDs of bound gp120, bound VRC01, and the inner and outer domains of bound gp120

accumulation of round-off errors, it is extremely difficult to handle the Hessian matrix or the covariance fluctuation matrix of an all-atom model system with more than 400 residues [41, 42].

To obtain the detailed interaction profile of gp120 and VRC01, MM-GBSA was used to decompose the whole binding energy into components for each residue involved in the interaction by considering the molecular mechanics energies and solvation energies but not the contribution from entropy.

Principal component analysis

Principal component analysis (PCA) was performed using the PTRAJ module of AMBER10 [43] for loop D and loop V5 of the apo-gp120 and bound gp120 systems. The covariance matrix was calculated using the coordinates of the Cα carbon atoms and then diagonalized to obtain the principal component eigenvectors. Each structure obtained during the trajectory was then projected in the collective coordinate space, which was defined by the two largest principal component eigenvectors.

Cluster analysis

Cluster analysis of the MD trajectories was performed using the SOM algorithms in the PTRAJ module of AMBER [44]. Snapshots of bound gp120 obtained every 1 ps were collected for the cluster analysis. To investigate the conformational changes in loops D and V5, a clustering analysis was performed for the backbone atoms of the two loops. By superposing the two loops onto the crystal structure to remove rigid-body motions, including translations and rotations, a three-cluster solution was obtained.

Results and discussion

Monitoring the MD trajectories

Figure 2 shows the variations in the root-mean square deviation (RMSD) of the backbone atoms from gp120, VRC01, and specific domains of gp120 in the apo and bound forms during 15 ns MD trajectories. From Fig. 2a and b, it is apparent that the RMSDs of gp120 and VRC01 in the apo and bound forms converged during the last 5 ns. The inner domain of gp120 shows greater fluctuations than the outer domain in both the apo and bound forms, especially in the unbound form, which is consistent with previous reports that the inner domain is flexible while the outer domain is more rigid [45]. Figure 3 shows the root mean square fluctuations (RMSF) of the backbone atoms averaged for each residue of gp120 and VRC01 in the apo and bound forms during 15 ns MD trajectories. According to Fig. 3a,

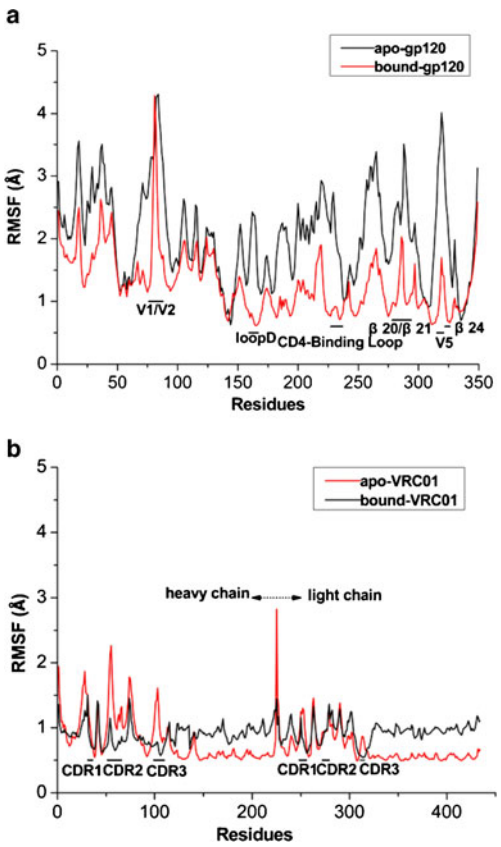

Fig. 3 RMSFs of the backbone atoms in gp120 (**a**) and VRC01 (**b**), as obtained from the crystal structure and from molecular dynamics simulation

several areas in the apo form of gp120 (including V1/V2, loop D, the CD4-binding loop, β20/β21, V5, and β24) show large fluctuations. The binding of VRC01 leads to a significant reduction in the flexibility of gp120. Figure 3b shows that the binding of gp120 reduces the flexibility of VRC01, while CDR2 and CDR3 of the heavy chain and CDR3 of the light chain of bound VRC01 show increased flexibility. The interacting areas of VRC01 are located on the sheets, while those of gp120 are located on the loops, leading to several CDRs of VRC01 that are influenced by interactions with loops in gp120.

Binding free energy calculation

MM-PBSA and MM-GBSA were used to calculate the binding free energy of gp120 with VRC01. First, we extracted snapshots from the last 5 ns of the molecular dynamics (MD) trajectory; 500 snapshots were extracted to calculate the enthalpy. The average binding free energies and the detailed contributions from various energy components are shown in Table 1. From Table 1, the calculated binding free energies are large: −70.22 kcal mol^{-1} according to the MM-PBSA method and −52.12 kcal mol^{-1} according to the MM-GBSA method, although MM-PBSA and MM-GBSA tend to overestimate free energy changes [30, 46, 47]. Although these two methods cannot precisely reproduce the absolute change in free energy, a good relationship between the change in free energy calculated by these methods and the experimentally determined value has been noted [46, 48]. In this work, our aim was not to calculate the absolute binding free energy but to explore the interaction features in detail and to identify the sections and residues that are crucial to the recognition of gp120 and VRC01. Since MM-PBSA and MM-GBSA can accurately rank the relative contributions of the residues, we were able to apply these two methods here. Based on the results obtained with the two methods, there is a difference of 18.1 kcal mol^{-1} due to a difference in the polar solvation energy, which is one component of electrostatic interaction contribution. In fact, at present, the greatest challenge to the accurate calculation of the binding free energy is the difficulty involved in estimating the contribution from the electrostatic interaction [19, 22, 30, 49]. Therefore, the large difference in the binding free energies calculated using MM-PBSA and MM-GBSA should also be attributed to the inaccurate estimation of the electrostatic interaction.

Table 1 Binding free energy components of the gp120–VRC01 complex (kcal mol^{-1})

Contribution*	gp120–VRC01	gp120	VRC01	Delta
E_{ele}	−23104.6	−10057	−12798.1	−249.47
E_{vdw}	−3324.83	−1443.18	−1781.22	−100.43
E_{int}	16643.89	7609.57	9034.32	0.00
E_{gas}	−9785.56	−3890.62	−5545.04	−349.9
G_{SUR}	261.74	125.22	153.44	−16.93
G_{PB}	−8164.01	−4000.95	−4459.68	296.61
G_{PBSOL}	−7902.28	−3875.72	−4306.23	279.68
$G_{nonpolar}$ [a]	−3063.09	−1317.96	−1627.78	−117.36
$G_{polar,PB}$ [b]	−31269	−14058	−17258	47.14
$H_{total,PB}$	−17688	−7766.4	−9851.3	−70.22
G_{GB}	−8240.42	−4033.65	−4521.49	314.71
G_{GBSOL}	−7978.7	−3908.4	−4368	297.78
$G_{polar,GB}$ [b]	−31345	−14091	−17320	65.24
$H_{total,GB}$	−17764	−7799.1	−9913.1	−52.12

[a] $\Delta G_{nonpolar}=E_{vdw}+\Delta G_{SUR}$; [b] $\Delta G_{polar\ ,PB}=E_{ele}+\Delta G_{PB}$; $\Delta G_{polar,GB}=E_{ele}+\Delta G_{GB}$

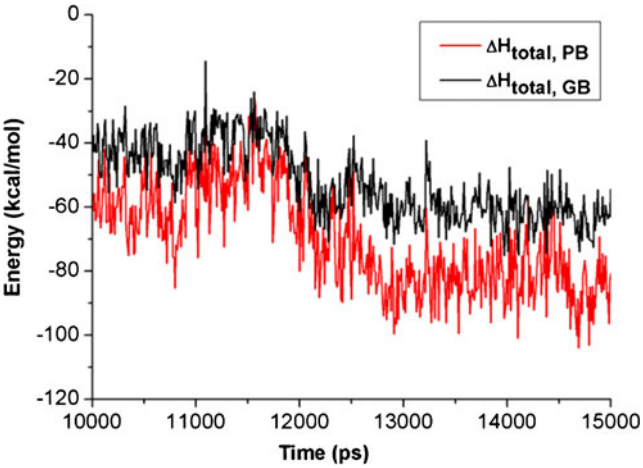

Fig. 4 Time evolution of the binding free energy of gp120–VRC01, calculated using the MM-PBSA and MM-GBSA methods

The favorable contribution from the direct electrostatic interactions between gp120 and VRC01 was compensated for by the electrostatic desolvation free energy upon binding, which led to an unfavorable contribution overall, consistent with other MM-GBSA and MM-PBSA studies [24, 31]. On the contrary, nonpolar interactions (including van der Waals interactions and nonpolar solvation) contribute −117.36 kcal mol^{-1}, which is very favorable to the binding process and consistent with the large hydrophobic binding surface between gp120 and VRC01. Therefore, the binding of gp120 and VRC01 is predominantly driven by nonpolar interactions, including van der Waals interactions and the nonpolar solvation contribution. The calculated enthalpy values for all of the snapshots from the last 5 ns are given in Fig. 4, to illustrate the changes in the enthalpy over time. From Fig. 4, it can be seen that the binding free energy shows small fluctuations, while the enthalpies calculated by MM-PBSA and MM-GBSA present similar trends in fluctuations, although the values obtained by MM-GBSA are larger.

Hotspot residues

The hotspot residues in gp120

MM-GBSA is advantageous for decomposing the binding free energy into contributions from individual residues and into different energy items, so we used this method to perform residue energy decomposition analysis, and the corresponding results are shown in Fig. 5. From Fig. 5a, it is apparent that several residues of gp120 contribute >2 kcal mol^{-1}: Asn280, Asp457, Lys282, Ala281, Gly459, Asp368, Gly458, and Ile371. In addition, some other residues (including Asn279, Thr278, Asn461, Gly367, and Gln362) contribute >1 kcal mol^{-1}. Figure 6 shows the distribution of these key residues on the molecular surface. From Fig. 6a and Table 2, it can be seen that these key residues of gp120 which interact with VRC01 are prominently from three areas: loop D, the CD4-binding loop, and

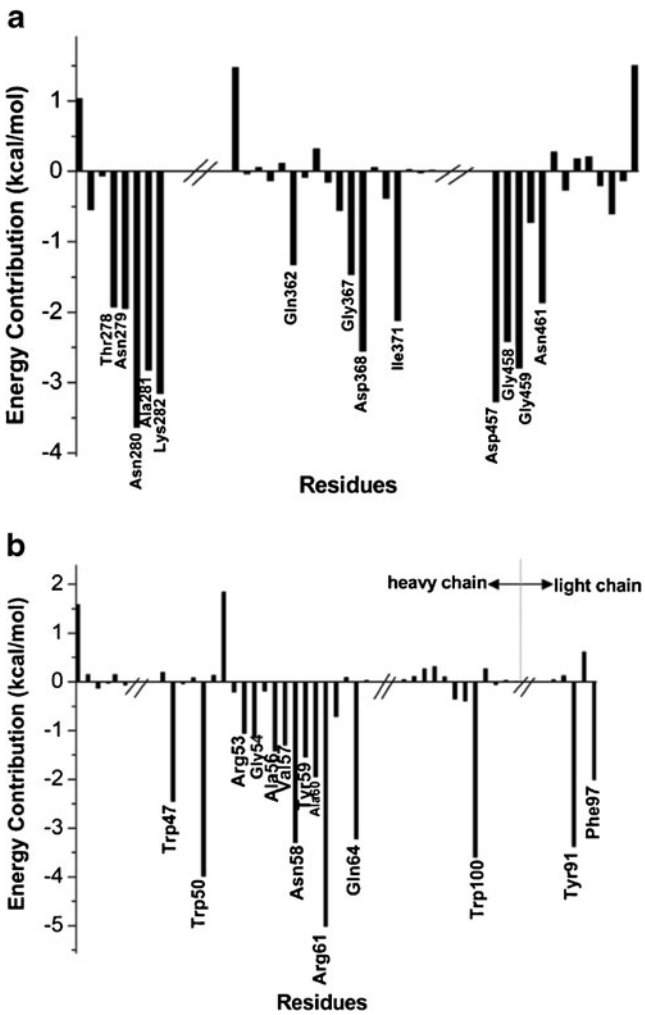

Fig. 5 Analysis of the pair interaction energy between gp120 and VRC01: **a** the contribution of each residue in gp120 to VRC01 binding; **b** the contribution of each residue in VRC01 to gp120 binding

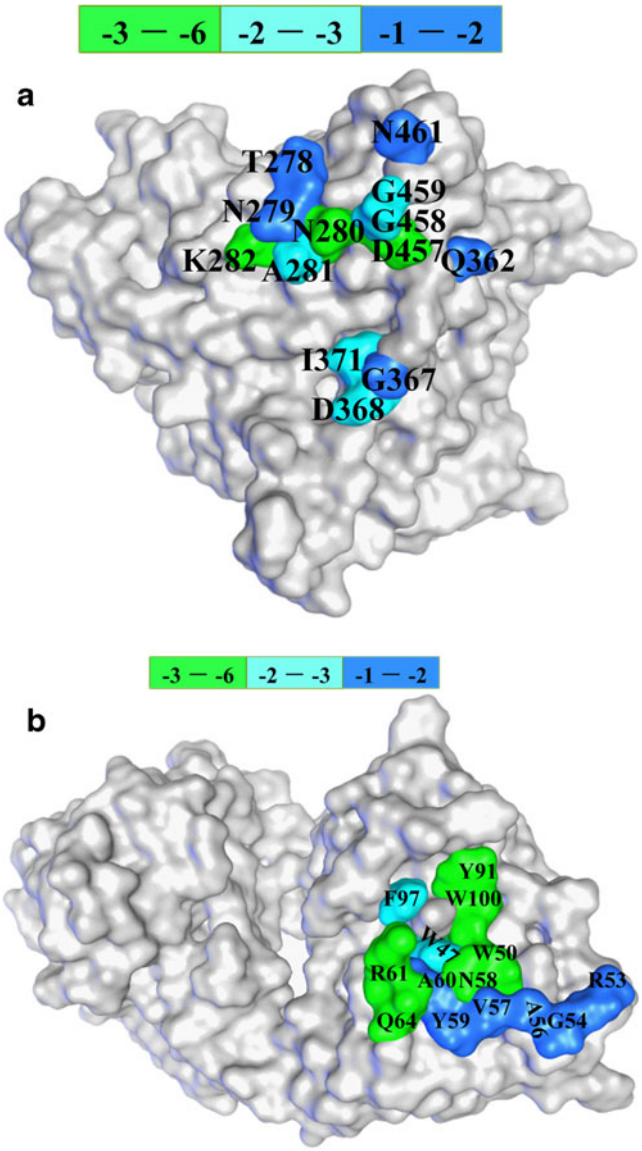

Fig. 6 Distributions of the identified hotspot residues on the **a** gp120 surface and **b** VRC01 surface

loop V5. These three areas make dominant contributions to the binding of VRC01, while the other areas provide almost negligible contributions. From Table 2, these three regions contribute 52 % of the total binding free energy; in particular, loops D and V5, which are related to the resistance of the virus to VRC01 [12], contribute 39 %. Moreover, it is also clear that van der Waals interactions and the nonpolar solvation contribution provide the driving force for the binding process according to Table 2.

The hotspot residues in VRC01

Residue energy decomposition results for VRC01 are given in Fig. 5b. From Fig. 5b, we can see that Arg61, Trp50,

Table 2 The energy contributions from different areas of the gp120–VRC01 complex (kcal mol^{-1})

	E_{ele}	G_{GBSOL}	E_{vdw}	G_{SUR}	$H_{total,GB}$
Loop D (E275–T283)	−40.67	41.98	−14.2	−2.64	−12.89
CD4-binding loop (P364–H374)	−108.47	111.54	−9.89	−1.57	−6.82
V5 (G458–S464)	−21.14	25.09	−11.55	−2.08	−7.6
gp120	−124.7	147.94	−49.98	−8.58	−27.06
CDR1 (D31–W36)[a]	2.92	−0.01	−1.29	−0.05	1.65
CDR2 (L51–G65)[a]	−70.54	75.93	−24.22	−4.33	−18.86
CDR3 (G95–H102)[a]	−35.61	37.69	−5.37	−1	−3.26
Heavy chain excluding CDRs	12.56	−7.05	−7.98	−0.96	−2.49
Heavy chain[a]	−90.67	106.56	−38.86	−6.34	−22.96
CDR1 (R24–A34)[b]	−1.31	4.2	−1.89	−0.26	0.98
CDR2 (S50–G57)[b]	−1.65	1.87	−0.05	0	0.16
CDR3 (Q89–F97)[b]	−26.91	29.1	−6.8	−1.21	−4.6
Light chain excluding CDRs	−4.2	8.14	−2.03	−0.51	1.38
Light chain[b]	−34.07	43.31	−11.22	−1.98	−2.08
VRC01	−124.74	149.87	−50.08	−8.32	−25.04
Total	−249.44	297.81	−100.06	−16.9	−52.1

[a]Residues from the heavy chain of VRC01; [b]Residues from the light chain of VRC01

Trp100, Asn58, Gln64, and Trp47 in the heavy chain of VRC01 as well as Tyr91 and Phe97 in the light chain of VRC01 each contribute >2 kcal mol^{-1} of the free energy, while Ala60, Tyr59, Ala56, Val57, Gly54, and Arg53 in the heavy chain of VRC01 contribute >1 kcal mol^{-1} of the free energy. From Figs. 5b and 6b as well as Table 2, it can be seen that these important residues on the VRC01 surface are relatively continuous. The long CDR2 of the heavy chain plays an important role in the interaction with gp120, and it can span the three identified areas of gp120. From Table 2, the CDR2 of the heavy chain provides 36 % of the total binding free energy. In addition, CDR3 of the heavy chain and CDR3 of the light chain also provide comparatively large energy contributions to gp120 binding. Identifying hotspot residues in the gp120–VRC01 complex can provide the valuable information for vaccine design.

Features of the interaction of gp120 and VRC01

In order to identify the energy contributions of the important residues mentioned above in detail, we further divided each residue's contribution into polar and nonpolar contributions, and the corresponding results are given in Fig. 7. From Fig. 7 as well as Tables 1 and 2, it can be seen that the energy contributions of the key residues relate primarily to the nonpolar interaction energy, including van der Waals interactions and the nonpolar desolvation energy. The identified hotspot residues on gp120 participate in important interactions with some residues of VRC01. For example, Asn280 on loop D participates in particularly favorable van der Waals interactions with the identified hotspot residues

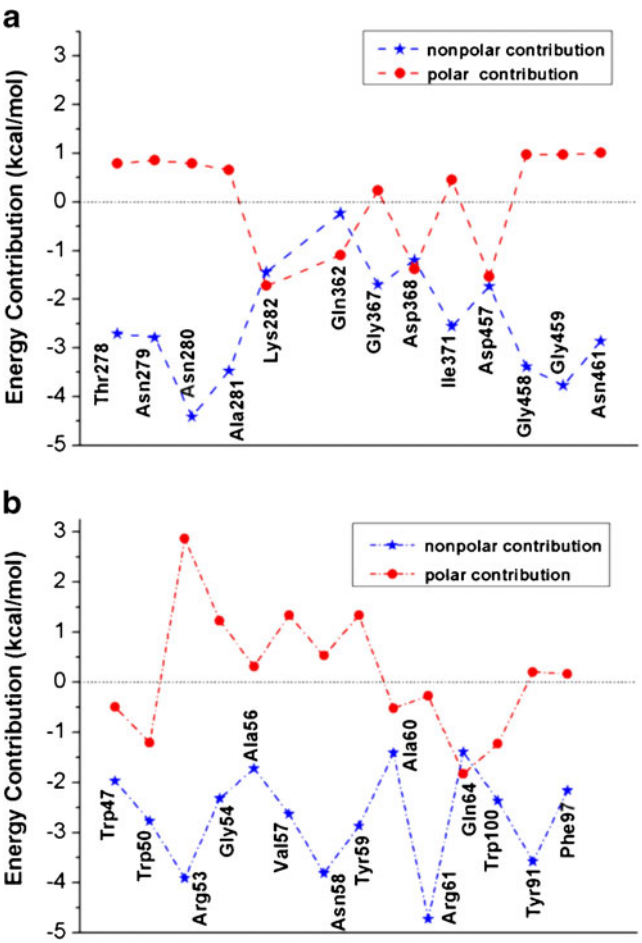

Fig. 7 Polar and nonpolar contributions of the hotspot residues from **a** gp120 and **b** VRC01

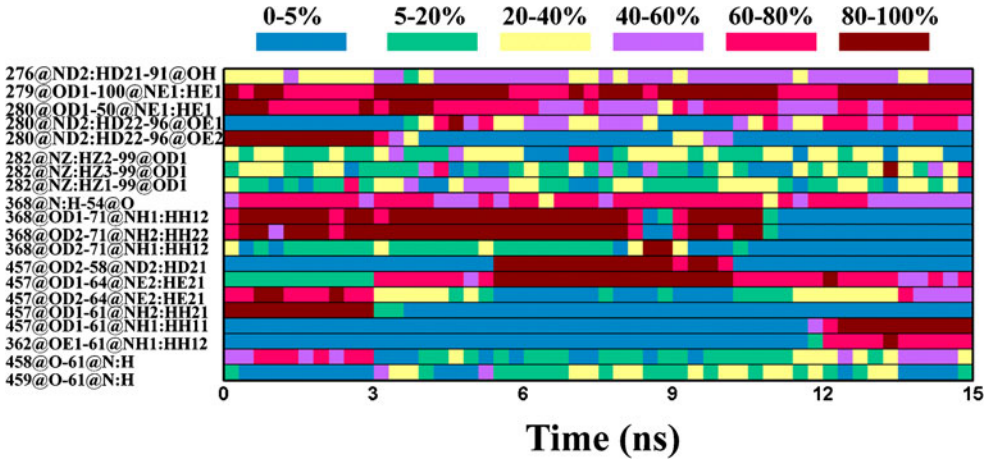

Fig. 8 Percentage occupancies of hydrogen bonds between gp120 and VRC01 (only hydrogen bonds with occupancies of >10 % are shown). The hydrogen bond was defined by the acceptor and the hydrogen atom; for example, 276@HD21:91@OH shows that the hydrogen atom of the amide group (denoted HD21) of residue 276 from gp120 formed a hydrogen bond with the oxygen atom of the hydroxyl group of residue 91 from VRC01 (the H-bond acceptor). A hydrogen bond was considered to be formed if the hydrogen–acceptor distance was ≦3.0 Å and the donor–hydrogen–acceptor angle was >120°

Trp47, Trp50, and Asn58 on the heavy chain of VRC01, and it also undergoes polar interactions with Trp50. Ile371 in the CD4-binding loop participates in hydrophobic interactions with Gly54 and Ala56 in CDR2 of the heavy chain. Gly459 in the V5 area undergoes favorable van der Waals interactions with Ala60 and Arg61 in CDR2 of the heavy chain and Phe97 in CDR3 of the light chain of VRC01. For VRC01, Arg61 provides the most favorable van der Waals interactions with residues in the V5 area of gp120 by side-chain packing.

Although van der Waals interactions provide a large contribution to the gp120–VRC01 binding energy, the polar interactions of some key residues also contribute significantly, such as Asp457, Lys282, Asp368, and Gln362 in gp120 as well as Trp50, Trp100, and Gln64 of the heavy chain of VRC01, each of which contribute >1 kcal mol^{-1}. Upon calculating the occupancies of the main hydrogen bonds between gp120 and VRC01, as shown in Fig. 8, it is clear that some important H-bond pairs are formed, including $Asp457_{gp120}$–$Gln64_{VRC01}$, $Asp457_{gp120}$–$Arg61_{VRC01}$, $Lys282_{gp120}$–$Asp99_{VRC01}$, $Asn279_{gp120}$–$Trp100_{VRC01}$, $Asn280_{gp120}$–$Trp50_{VRC01}$, $Asp368_{gp120}$–$Gly54_{VRC01}$, and $Gln362_{gp120}$–$Arg61_{VRC01}$. Among these hydrogen bonds, some provide favorable electrostatic contributions toward the formation of the complex. For example, the H-bond pairs $Asp457_{gp120}$–$Gln64_{VRC01}$ and $Asp457_{gp120}$–$Arg61_{VRC01}$ provide the most favorable electrostatic contributions. The residues on loop D of gp120 form hydrogen bonds with those of VRC01, including $Lys282_{gp120}$–$Asp99_{VRC01}$, $Asn279_{gp120}$–$Trp100_{VRC01}$, and $Asn280_{gp120}$–$Trp50_{VRC01}$. In addition, the H-bond pairs $Asp368_{gp120}$–$Gly54_{VRC01}$ and $Gln362_{gp120}$–$Arg61_{VRC01}$ also make favorable polar contributions. Another important H-bond pair, $Asp368_{gp120}$–$Arg71_{VRC01}$, can mimic the interaction between CD4 and gp120 [12].

Structural fluctuations in loop D and loop V5

Changes in loop D and loop V5 of gp120 are the source of most natural resistance to VRC01 [12]. According to our results from the binding free energy decomposition, these two areas also provide favorable contributions to the binding of VRC01. Therefore, studies of these two loops are of great significance in the design of vaccines and small molecule inhibitors. A principal component analysis of the Cα atoms of these two areas in apo-gp120 and bound gp120 was performed to describe the conformational space of

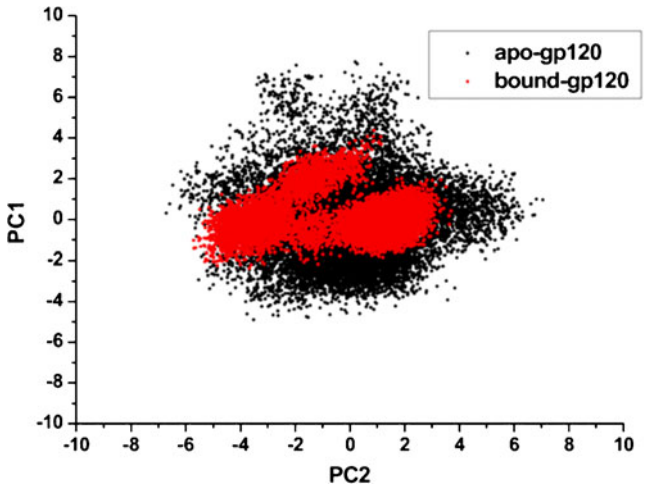

Fig. 9 Principal component analysis of loops D and V5

loops D and V5. The conformational space is shown in two-dimensional mode using the first two principal components in Fig. 9. From Fig. 9, the conformation of apo-gp120 shows much larger fluctuations than that of bound gp120, suggesting that the binding of VRC01 leads to a reduction in the flexibility of gp120. In bound gp120, the conformations of the two loops are primarily distributed into three classes, among which two of the classes dominate. Consistent with the PCA analysis, cluster analysis of the two loops in bound gp120, as shown in Fig. 10, indicates that the two areas have three classes of conformations, and that two classes of those three conformational classes were occupied >90 % of the time (Fig. 10a). A comparison of three representative structures indicates that loop V5 exhibits large conformational

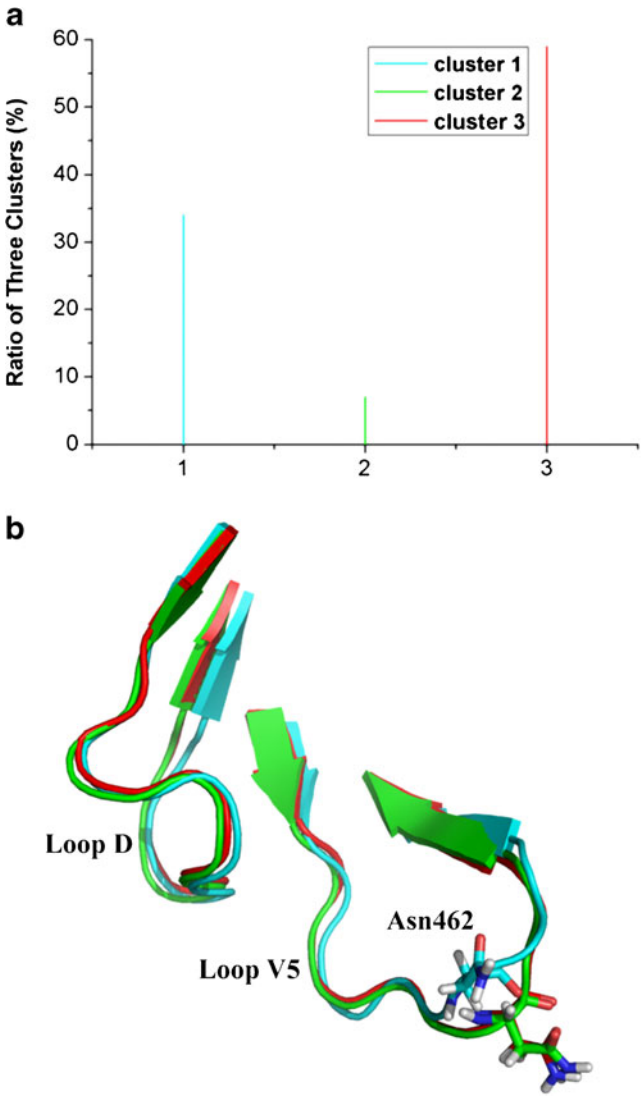

Fig. 10 Cluster analysis of loops D and V5: **a** the ratio of the three clusters; **b** the representative structures of the three clusters (cluster 1 shown in *cyan*, cluster 2 shown in *green*, and cluster 3 shown in *red*, Asn462 shown in *stick representation*)

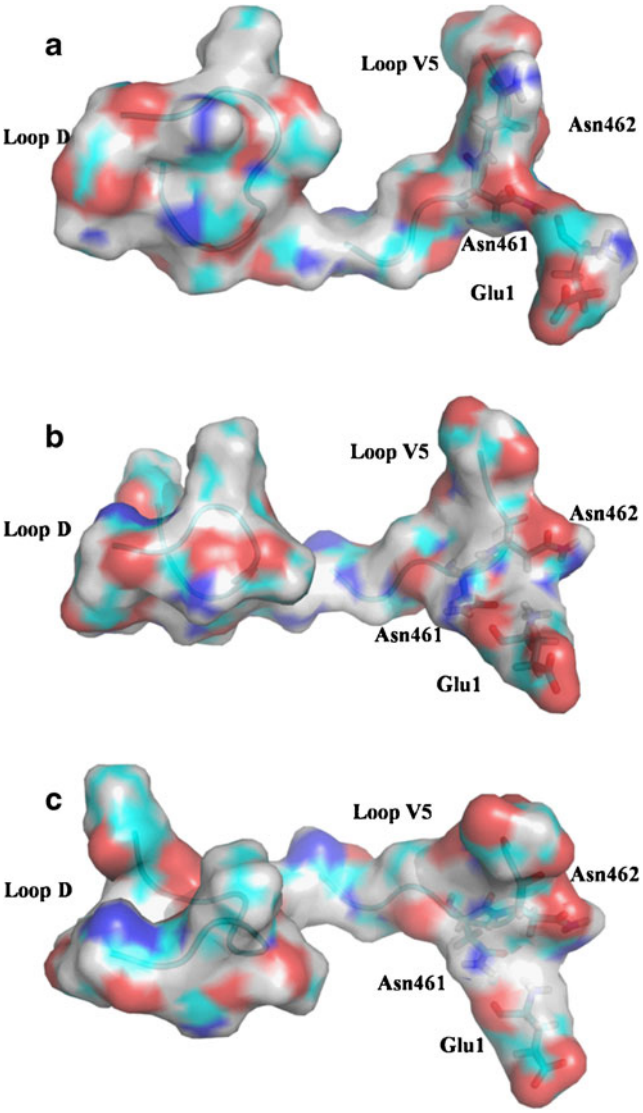

Fig. 11 Surface representations of the three clusters of loops D and V5 in gp120, colored by electrostatic potential: **a** cluster 1; **b** cluster 2; **c** cluster 3

fluctuations, mainly at Asn462, while conformational fluctuations of loop D are very small (Fig. 10b). Changes in loop V5 are caused by the terminal residue Glu1 in VRC01, which is shown in Fig. 11. Under the influence of conformational fluctuations in Glu1, especially the terminal amino group, the directions of the side chains of Asn461 and Asn462 in the antigen change together. Therefore, the N-terminal Glu1 of the antibody influences the flexibility of loop V5, especially the residue Asn462 of the antigen.

Conclusions

In this study, 15 ns molecular dynamics simulations were performed for the gp120–VRC01 complex, apo-gp120, and

apo-VRC01. The MM-PBSA and MM-GBSA methods were used to calculate the binding free energy and to analyze the binding interaction in detail. The results of binding free energy decomposition indicated that intermolecular van der Waals interactions and nonpolar solvation dominate the binding of gp120 and VRC01. By decomposing the binding free energy into the contribution from each residue, it was possible to identify the binding hot spots for gp120 and VRC01. For gp120, the residues Asn280, Lys282, Asp368, Ile371, and Asp457 from loop D, the CD4-binding loop, and the V5 area provide significant contributions (of >2 kcal mol^{-1}). Overall, loop D, the CD4-binding loop, and the V5 area contribute almost all of the binding free energy of gp120. For VRC01, Trp47, Trp50, Asn58, Arg61, Gln64, and Trp100 in the heavy chain and Tyr91 in the light chain were identified as hotspots. The contributions of VRC01 to binding process derive primarily from CDR2 of the heavy chain. Hydrogen-bond occupancy analysis indicated that several important hydrogen bonds stabilize the complex, while principal component analysis showed that the binding of VRC01 can stabilize loop D and the V5 area of gp120. Also, cluster analysis demonstrated that conformational fluctuations in the residue Glu1 in VRC01 change the orientations of Asn461 and Asn462. The information obtained in this study provide some useful insights that should aid the design of vaccines and small molecule inhibitors to neutralize HIV-1.

Acknowledgments This work was supported by the National Natural Science Foundation of China (grant nos. 21103075 and 21175063). We would like to thank the Gansu Computing Center for providing computing resources.

References

1. Flower DR, Macdonald IK, Ramakrishnan K, Davies MN, Doytchinova IA (2010) Computer aided selection of candidate vaccine antigens. Immunol Res 6:S1
2. UNAIDS (2010)Global report: UNAIDS report on the global AIDS epidemic 2010. Joint United Nations Programme on HIV and AIDS (UNAIDS), Geneva. http://www.unaids.org/en/media/unaids/contentassets/documents/unaidspublication/2010/20101123_globalreport_en[1].pdf
3. Liu J, Bartesaghi A, Borgnia MJ, Sapiro G, Subramaniam S (2008) Molecular architecture of native HIV-1 gp120 trimers. Nature 455:109–113
4. Wyatt R, Sodroski J (1998) The HIV-1 envelope glycoproteins: fusogens, antigens, and immunogens. Science 280:1884–1888
5. Dalgleish AG, Beverley PC, Clapham PR, Crawford DH, Greaves MF, Weiss RA (1984) The CD4 (T4) antigen is an essential component of the receptor for the AIDS retrovirus. Nature 312:763–767
6. Wu L, Gerard NP, Wyatt R, Choe H, Parolin C, Ruffing N, Borsetti A, Cardoso AA, Desjardin E, Newman W, Gerard C, Sodroski J (1996) CD4-induced interaction of primary HIV-1 gp120 glycoproteins with the chemokine receptor CCR-5. Nature 384:179–183
7. Trkola A, Dragic T, Arthos J, Binley JM, Olson WC, Allaway GP, Cheng-Mayer C, Robinson J, Maddon PJ, Moore JP (1996) CD4-dependent, antibody-sensitive interactions between HIV-1 and its co-receptor CCR-5. Nature 384:184–187
8. Feng Y, Broder CC, Kennedy PE, Berger EA (1996) HIV-1 entry cofactor: functional cDNA cloning of a seven-transmembrane, G protein-coupled receptor. Science 272:872–877
9. Singh A, Collman RG (2000) Heterogeneous spectrum of coreceptor usage among variants within a dualtropic human immunodeficiency virus type 1 primary-isolate quasispecies. J Virol 74:10229–10235
10. McElrath MJ, Haynes BF (2010) Induction of immunity to human immunodeficiency virus type-1 by vaccination. Immunity 33:542–554
11. Wu X, Yang ZY, Li Y, Hogerkorp CM, Schief WR, Seaman MS, Zhou T, Schmidt SD, Wu L, Xu L, Longo NS, McKee K, O'Dell S, Louder MK, Wycuff DL, Feng Y, Nason M, Doria-Rose N, Connors M, Kwong PD, Roederer M, Wyatt RT, Nabel GJ, Mascola JR (2010) Rational design of envelope identifies broadly neutralizing human monoclonal antibodies to HIV-1. Science 329:856–861
12. Zhou T, Georgiev I, Wu X, Yang ZY, Dai K, Finzi A, Do Kwon Y, Scheid JF, Shi W, Xu L, Yang Y, Zhu J, Nussenzweig MC, Sodroski J, Shapiro L, Nabel GJ, Mascola JR, Kwong PD (2010) Structural basis for broad and potent neutralization of HIV-1 by antibody VRC01. Science 329:811–817
13. Burton DR, Weiss RA (2010) A boost for HIV vaccine design. Science 329:770–773
14. Janin J (1995) Principles of protein–protein recognition from structure to thermodynamics. Biochimie 77:497–505
15. Newhouse EI, Xu D, Markwick PR, Amaro RE, Pao HC, Wu KJ, Alam M, McCammon JA, Li WW (2009) Mechanism of glycan receptor recognition and specificity switch for avian, swine, and human adapted influenza virus hemagglutinins: a molecular dynamics perspective. J Am Chem Soc 131:17430–17442
16. Zhou R, Das P, Royyuru AK (2008) Single mutation induced H3N2 hemagglutinin antibody neutralization: a free energy perturbation study. J Phys Chem B 112:15813–15820
17. Das P, Li J, Royyuru AK, Zhou R (2009) Free energy simulations reveal a double mutant avian H5N1 virus hemagglutinin with altered receptor binding specificity. J Comput Chem 30:1654–1663
18. Laitinen T, Kankare JA, Peräkylä M (2004) Free energy simulations and MM-PBSA analyses on the affinity and specificity of steroid binding to antiestradiol antibody. Proteins 55:34–43
19. Gouda H, Kuntz ID, Case DA, Kollman PA (2003) Free energy calculations for theophylline binding to an RNA aptamer: comparison of MM-PBSA and thermodynamic integration methods. Biopolymers 68:16–34
20. Srinivasan J, Cheatham TE, Cieplak P, Kollman PA, Case DA (1998) Continuum solvent studies of the stability of DNA, RNA, and phosphoramidate−DNA helices. J Am Chem Soc 120:9401–9409
21. Kollman PA, Massova I, Reyes C, Kuhn B, Huo S, Chong L, Lee M, Lee T, Duan Y, Wang W, Donini O, Cieplak P, Srinivasan J, Case DA, Cheatham TE (2000) Calculating structures and free energies of complex molecules: combining molecular mechanics and continuum models. Acc Chem Res 33:889–897
22. Simonson T, Archontis G, Karplus M (2002) Free energy simulations come of age: protein−ligand recognition. Acc Chem Res 35:430–437
23. Wang W, Kollman PA (2000) Free energy calculations on dimer stability of the HIV protease using molecular dynamics and a continuum solvent model. J Mol Biol 303:567–582
24. Lafont V, Schaefer M, Stote RH, Altschuh D, Dejaegere A (2007) Protein–protein recognition and interaction hot spots in an antigen–antibody complex: free energy decomposition identifies "efficient amino acids". Proteins 67:418–434
25. Liu HX, Yao XJ (2010) Molecular basis of the interaction for an essential subunit PA-PB1 in influenza virus RNA polymerase: insights from molecular dynamics simulation and free energy calculation. Mol Pharm 7:75–85

26. Duan Y, Wu C, Chowdhury S, Lee MC, Xiong G, Zhang W, Yang R, Cieplak P, Luo R, Lee T, Caldwell J, Wang J, Kollman P (2003) A point-charge force field for molecular mechanics simulations of proteins based on condensed-phase quantum mechanical calculations. J Comput Chem 24:1999–2012
27. Lee MC, Duan Y (2004) Distinguish protein decoys by using a scoring function based on a new AMBER force field, short molecular dynamics simulations, and the generalized born solvent model. Proteins 55:620–634
28. Zoete V, Meuwly M, Karplus M (2005) Study of the insulin dimerization: binding free energy calculations and per-residue free energy decomposition. Proteins 61:79–93
29. Gohlke H, Kiel C, Case DA (2003) Insights into protein–protein binding by binding free energy calculation and free energy decomposition for the Ras–Raf and Ras–RalGDS complexes. J Mol Biol 330:891–913
30. Gohlke H, Case DA (2004) Converging free energy estimates: MM-PB(GB)SA studies on the protein–protein complex Ras–Raf. J Comput Chem 25:238–250
31. Zoete V, Michielin O (2007) Comparison between computational alanine scanning and per-residue binding free energy decomposition for protein–protein association using MM-GBSA: application to the TCR-p-MHC complex. Proteins 67:1026–1047
32. Tsui V, Case DA (2000) Theory and applications of the generalized Born solvation model in macromolecular simulations. Biopolymers 56:275–291
33. Onufriev A, Bashford D, Case DA (2000) Modification of the generalized Born model suitable for macromolecules. J Phys Chem B 104:3712–3720
34. Case DA, Cheatham TE, Darden T, Gohlke H, Luo R, Merz KM, Onufriev A, Simmerling C, Wang B, Woods RJ (2005) The Amber biomolecular simulation programs. J Comput Chem 26:1668–1688
35. Hornak V, Abel R, Okur A, Strockbine B, Roitberg A, Simmerling C (2006) Comparison of multiple Amber force fields and development of improved protein backbone parameters. Proteins 65:712–725
36. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML (1983) Comparison of simple potential functions for simulating liquid water. J Chem Phys 79:926–935
37. Essmann U, Perera L, Berkowitz ML, Darden T (1995) A smooth particle mesh Ewald method. J Chem Phys 103:8577–9593
38. Ryckaert JP, Ciccotti G, Berendsen HJC (1977) Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. J Comput Phys 23:327–341
39. Berendsen HJC, Postma JPM, van Gunsteren WF, DiNola A, Haak JR (1984) Molecular dynamics with coupling to an external bath. J Chem Phys 81:3684–3690
40. Sitkoff D, Sharp KA, Honig B (1994) Accurate calculation of hydration free energies using macroscopic solvent models. J Phys Chem 98:1978–1988
41. Huo SH, Massova I, Kollman PA (2002) Computational alanine scanning of the 1:1 human growth hormone-receptor complex. J Comput Chem 23:15–27
42. Brooks BR, Jane D, Karplus M (1995) Harmonic analysis of large systems. I. Methodology. J Comput Chem 16:1522–1542
43. Prompers JJ, Brüschweiler R (2002) Dynamic and structural analysis of isotropically distributed molecular ensembles. Proteins 46:177–189
44. Shao J, Tanner SW, Thompson N, Cheatham TE (2007) Clustering molecular dynamics trajectories: 1. Characterizing the performance of different clustering algorithms. J Chem Theory Comput 3:2312–2334
45. Emileh A, Abrams CF (2011) A mechanism by which binding of the broadly neutralizing antibody b12 unfolds the inner domain α1 helix in an engineered HIV-1 gp120. Proteins 79:537–546
46. Hou T, Wang J, Li Y, Wang W (2011) Assessing the performance of the MM/PBSA and MM/GBSA methods: I. The accuracy of binding free energy calculations based on molecular dynamics simulations. J Chem Inf Model 51:69–82
47. Raju RK, Burton NA, Hillier IH (2010) Modelling the binding of HIV-reverse transcriptase and nevirapine: an assessment of quantum mechanical and force field approaches and predictions of the effect of mutations on binding. Phys Chem Chem Phys 12:7117–7125
48. Wang J, Morin P, Wang W, Kollman PA (2001) Use of MM-PBSA in reproducing the binding free energies to HIV-1 RT of TIBO derivatives and predicting the binding mode to HIV-1 RT of efavirenz by docking and MM-PBSA. J Am Chem Soc 123:5221–5230
49. Kukic P, Nielsen JE (2010) Electrostatics in proteins and protein–ligand complexes. Future Med Chem 2:647–666