# Molecular Recognition of CXCR4 by a Dual Tropic HIV-1 gp120 V3 Loop

Phanourios Tamamis and Christodoulos A. Floudas*
Department of Chemical and Biological Engineering, Princeton University, New Jersey

ABSTRACT HIV-1 cell entry is initiated by the interaction of the viral envelope glycoprotein gp120 with CD4, and chemokine coreceptors CXCR4 and CCR5. The molecular recognition of CXCR4 or CCR5 by the HIV-1 gp120 is mediated through the V3 loop, a fragment of gp120. The binding of the V3 loop to CXCR4 or CCR5 determines the cell tropism of HIV-1 and constitutes a key step before HIV-1 cell entry. Thus, elucidating the molecular recognition of CXCR4 by the V3 loop is important for understanding HIV-1 viral infectivity and tropism, and for the design of HIV-1 inhibitors. We employed a comprehensive set of computational tools, predominantly based on free energy calculations and molecular-dynamics simulations, to investigate the molecular recognition of CXCR4 by a dual tropic V3 loop. We report what is, to our knowledge, the first HIV-1 gp120 V3 loop:CXCR4 complex structure. The computationally derived structure reveals an abundance of polar and nonpolar intermolecular interactions contributing to the HIV-1 gp120:CXCR4 binding. Our results are in remarkable agreement with previous experimental findings. Therefore, this work sheds light on the functional role of HIV-1 gp120 V3 loop and CXCR4 residues associated with HIV-1 coreceptor activity.

## INTRODUCTION

The primary step of human immunodeficiency virus type 1 (HIV-1) cell entry is the interaction of the viral envelope glycoprotein (comprising subunits gp41 and gp120) with the host leukocyte glycoprotein receptor, CD4, and the two chemokine receptors CXCR4/CCR5 on the surface of the host cells (1–5). Specifically, the glycoprotein gp120 interaction with CD4 triggers conformational changes in gp120 that increase the exposure of the third variable region (V3) loop. Subsequently, the protein gp120, via its V3 loop, binds to chemokine receptors CXCR4 (infecting mostly T-cells) or CCR5 (infecting mostly macrophages) (6–11). The molecular recognition of chemokine receptors by the V3 loop results in a series of rearrangements in the envelope glycoprotein, leading to the fusion of the virus and the cell membranes (12).

At the beginning of the 1990s, the V3 loop was identified as the primary determinant of cell tropism in HIV-1 (13). Since the discovery of the key role of V3 loop in HIV-1 infection, with regard to the binding to chemokine receptors CXCR4 and CCR5 (6,14,15) and the determination of cell-tropism (13), recognizing CXCR4 or CCR5 or both (referred to as "dual tropic"), several experimental studies aimed at elucidating the key interacting residues of chemokine receptors involved in the V3 loop binding through the mapping of the chemokine receptors' binding sites (16–26). These studies employed site-directed mutagenesis or chimeric substitutions, and identified specific residues or residue moieties of the chemokine receptors that are critical to, or correlate with, viral infection.

The HIV-1 gp120 V3 loop is sustained in a loop conformation through a disulfide bridge between its N- and C-terminal ends, is encountered in a large sequence variability, is positively charged, and is predominantly composed of 35 residues (27–29). Owing to its highly dynamic character (27,29,30), the V3 loop is absent in the majority of gp120 crystallographic structures; nevertheless, it was resolved in two crystallographic Protein Data Bank (PDB) entries (4,5). Numerous studies aimed at understanding the physicochemical properties of the V3 loop and elucidating its viral tropism (5,11,19,26,31–34). It has been suggested that charge complementarity and electrostatic interactions among the N-terminal, extracellular loop 2 (ECL2) coreceptor domains, and the V3 loop (5,11,19,26,31–33), are associated with the viral tropism. Furthermore, it has been proposed that the interchange from coreceptor CCR5 to CXCR4, as the disease progresses, is linked to 1), The increase of the net charge of the V3 loop (10,31); 2), The presence of positively charged residues at one or more of positions 11, 24, and 25, known as the 11/24/25 rule (9); and 3), The absence of the glycosylation motif N6X7T8|S8X9 (where X = Pro) (8).

Recently, molecular dynamics (MD) simulations showed that V3 loops undergo common correlated motions, in association with specific charged interactions between residues on opposite stems (27). Understanding the unbound properties of gp120 domains is important for delineating the mechanism of conformational changes from unbound to bound structures, related to gp120:CD4 binding (35,36). Similarly, the identification of unbound V3 loop conformations associated with electrostatic-driven correlated motions (27) could prove significant for the elucidation of the gp120 (V3 loop):CXCR4 binding.

Despite the numerous studies related to the V3 loop and the chemokine receptors, the basic biological knowledge on the specific interactions between the V3 loop and the chemokine receptors is limited due to the absence of a

*Submitted May 23, 2013, and accepted for publication July 29, 2013.*

*Correspondence: floudas@titan.princeton.edu

Editor: Michael Feig.

© 2013 by the Biophysical Society

0006-3495/13/09/1502/13 $2.00

http://dx.doi.org/10.1016/j.bpj.2013.07.049

CrossMark

complete V3 loop:coreceptor complex structure (34). This could be associated with the high flexibility of the V3 loop leading to absence of electron density in the gp120 crystal structures, as in Liao et al. (37).

A comprehensive attempt to computationally derive a V3 loop:CXCR4 complex structure to enlighten the role of the key interacting V3 loop and CXCR4 residues has never before been reported, according to our knowledge. In this study, we exploit both the CXCR4 crystallographic structure (11) and one of the V3 loop crystallographic structures (5) to theoretically derive what is, to our knowledge, the first V3 loop:CXCR4 complex structure using a combination of primarily binding/interaction free-energy calculations and MD simulations. The computational protocol applied was not biased by any experimental evidence regarding the key interacting residues, and interestingly, our results are in remarkable agreement with previous experimental findings (see Table 1; marked in boldface are CXCR4 residues reported in experimental findings) (16–21,23–25). Thus, the reported V3 loop:CXCR4 complex structure sheds light on the functional role of V3 loop and CXCR4 residues, which are experimentally determined as critical for the HIV-1 coreceptor activity.

## METHODS

### Modeling, free energy calculations, and molecular dynamics simulations

The methodology used in this study to derive the V3 loop:CXCR4 complex structure consists of the seven following principal steps.

#### *Step 1: Modeling and selection of the initial V3 loop and CXCR4 structural conformations*

*V3 loop*. The initial backbone structural conformation of the V3 loop corresponds to PDB:2QAD, one of the two intact crystal structures of gp120 in the Protein Data Bank (PDB) (5). V3 loop residues 296–331 of PDB:2QAD were renumbered, starting from 1 and ending at 35. The gp120 residues outside the V3 loop were not considered for investigation in this study because the V3 loop region is the principal determinant of chemokine receptor specificity (2), and the presence of a disulfide bridge between the N- and C-terminal residues of V3 loop base constitutes a physical constraint for the preservation of the structure in the base region. In addition, according to a recent MD study, the V3 loop does not have any strong concerted motion with other parts of the gp120 protein but only within residues of the V3 loop (29). To identify a representative dual-tropic CXCR4/CCR5 V3 loop to examine in this study, we utilized the CD-HIT (38) to perform a sequence-based clustering on the dual tropic V3 loops, which was deposited in the Los Alamos National Laboratory (Los Alamos, NM) database (http://www.hiv.lanl.gov). According to the results, we identified the following sequence of subtype B as the centroid of the most populated dual tropic cluster (extracted from a Chinese patient),

CTRPNNNTRKRVSLGPGRVWYTTGQIVGDIRKAHC,

which obeys the 11/24/25 rule (8). In what follows, we use this particular sequence to derive the V3 loop:CXCR4 structure.

*CXCR4*. The initial structural templates used to construct the CXCR4 conformation correspond to PDB:3OE0 (11) and PDB:2K05 (39) (human fragment containing the missing N-terminal residues 1–24 of PDB:3OE0 (11)); the two fragments were combined using pairwise alignment on their overlapping regions. The FREAD loop modeling algorithm, was applied to model the missing loops (40), and the I-TASSER server was used to model the missing C-terminal residues (41).

#### *Step 2: Production of flexible templates for V3 loop and CXCR4 using MD simulations*

*V3 loop*. We performed replica exchange MD simulations of the V3 loop so as to enhance its conformational sampling (42–45). Solvent effects were taken into account by the FACTS19 implicit solvent model (46). The implementation of implicit solvent models (47), including the particular one used in this study, FACTS (46), in conjunction with the replica exchange MD scheme, successfully balances fast exploration of the conformational space with accuracy, to reproduce results from explicit solvent MD simulations (48,49). The simulations were conducted with the molecular mechanics program CHARMM, Ver. c35b6 (50). The average root mean-square deviation (RMSD) of the conformations produced at 300 K with regard to the starting conformation upon minimization is 5.1 $\pm$ 1.3 Å. The result demonstrates that this dual tropic V3 loop encounters a high flexibility in the unbound simulations, and the computational protocol used has successfully sampled the conformational space in order to produce representative flexible templates for docking. Additional information regarding the simulation protocol and parameterization is provided in the Supporting Material.

*CXCR4*. We employed MD simulations to produce multiple flexible templates for the human CXCR4 protein receptors and to refine the modeled regions of CXCR4. Because the goal was not only to refine the structure but also to produce flexible templates that could constitute docking receptors, we considered that a preliminary docking of V3 loop on CXCR4 would be beneficial for the subsequent docking procedure. Thus, we used the software CLUSPRO 2.0 (51) to primarily dock a V3 loop, corresponding to the PDB:2QAD conformation (5), on the modeled CXCR4, and we produced three conformations with different poses of the V3 loop proximal to the experimentally defined binding site. For each of the three complex structures yielded, we performed two independent MD simulations to produce flexible template structures for CXCR4. Within the MD simulations, the system was immersed in an implicit membrane represented by the switching-function generalized Born (GBSW) module (52,53). The membrane thickness $T_{\mathrm{memb}}$ was set to 36 Å with one-half of the membrane switching length equal to 2.5 Å; the choice of these values complies with the approximate intramembrane helical length of CXCR4 (11), and was shown to provide optimal protein stability compared to test simulations with larger or smaller. The surface tension coefficient was set to 0.03 kcal/(mol*Å2); all other parameters were set to default values (54). The simulations were conducted with the molecular mechanics program CHARMM, Ver. c35b6 (50). Additional information regarding the simulation protocol and parameterization is provided in the Supporting Material.

#### *Step 3: Docking of selected V3 loop structures on selected CXCR4 structures*

We clustered the structures produced in the V3 loop replica exchange MD simulations as well as the CXCR4 structures produced from the six independent aforesaid MD simulations using the quality clustering method of the command line utility WORDOM (55). We extracted the 20 most populated clusters for the V3 loop, and 17 clusters for CXCR4, including the initially modeled CXCR4 structure. Additional information is provided in the Supporting Material.

Subsequently, we used the parallel LINUX version of the ZDOCK Ver. 3.0.2 (56) to dock the 20 V3 loop clustered structures on the 17 CXCR4 clustered structures. For each run of ZDOCK, 2000 docked structures were produced with a dense rotational sampling and a masking applied on the region with protein coordinates $z < 0$ Å, so as to exclude the nonpotential binding region from the docking calculations. As a result, 680,000 complex structures were produced from docking.

### *Step 4: First round of energy minimization and binding free energy calculation of the docked complexes using the membrane-GBSA approximation*

All 680,000 complexes were subjected to 100 steps of steepest-descent minimization to alleviate bad contacts and, the binding free energy was evaluated subsequently for all complexes using the generalized Born (GB) solvent-accessible (SA) approximation in a heterogeneous water-membrane-water environment, modeled by GBSW (53). The binding free energy is evaluated via the expression

$$\Delta G = E_{PL} - E_P - E_L,$$

where $E_X$ is the total (free) energy of molecule X (complex PL:CXCR4:V3 loop, free protein P:CXCR4, or free ligand L:V3 loop) as in the literature (57–59). The protein and ligand conformations were assumed identical in the complex and in their free (unbound) states as in the literature (60–64). With this assumption, any bonded-energy contributions to $\Delta G$ are canceled in the equation. The solvation free energy components of the complex and unbound protein were computed in the heterogeneous membrane/water environment, while the solvation free energy component of the unbound ligand was computed in a homogeneous aqueous environment, also modeled by the implicit GBSW model (52,53).

### *Step 5: Second round of energy minimization and binding free energy calculations of the docked complexes using the membrane-PBSA approximation*

Out of the 680,000 complexes, we selected the 9000 V3 loop:CXCR4 complexes with the lowest GBSA binding free energy, and subsequently, we performed an additional round of 100 steps steepest-descent minimization and calculated the binding free energy using the Poisson-Boltzmann (PB) SA approximation (57). At the end of this procedure, we identified the complex structure with the lowest binding free energy −144.2 kcal/mol, and additionally, we selected all the complex structures within a 10-kcal/mol range of the lowest binding free energy (−144.2:−134.1 kcal/mol) for subsequent investigation. As a result, the total number of complex structures selected for subsequent investigation was 17. Table S1 in the Supporting Material presents the binding free energies of the 17 different complex structures produced in Step 5. Additional information regarding the PB calculations and the results is provided in the Supporting Material.

### *Step 6: MD simulations of the docked complexes acquiring the lowest binding free energy*

We performed 17 independent MD simulations of the complexes with the lowest PBSA binding free energies, as identified in the previous step. The MD simulations comprised a 400-ps heating procedure and an additional 700-ps equilibration procedure at which the harmonic restraints were gradually removed from the protein and the peptide. No restraints were imposed during the production run at 300 K, the duration of which was equal to 20 ns for every individual complex. The simulation methodology and force-field parameterization used was identical to Step 2, at which point we also performed MD simulations in implicit membrane to produce flexible templates for CXCR4.

### *Step 7: Binding free energy calculations of the complex structures produced in the MD simulations to identify the complex structure with the lowest average binding free energy*

We extracted 1000 snapshots, corresponding to 20-ps intervals, from the 17 MD 20-ns simulations and reevaluated the binding free energy by employing the MM PBSA approximation (57) (using the parameters presented in Step 5). Interestingly, the MM PBSA calculation of the average binding free energies for each complex expanded the free energy range from [−144.2:−134.1 kcal/mol] in step 5 to [−367.2:−275.2 kcal/mol] in Step 7 (see Table S1). The complex with the lowest average binding free energy in this step, according to both MM PBSA and supplementary MM GBSA calculations (58,59,60–64), corresponds to the complex structure of Step 5, which also possesses the lowest binding free energy (−144.2 kcal/mol; see Table S1). The change in the binding free energy ranking among Steps 5 and 7, as well as the improvement of interactions leading to lower binding free energies in Step 7 compared to Step 5, emphasize the key role of conducting MD simulations a posteriori to docking.

In what follows, we present the methodology used to investigate pairwise residue interaction free energies of the complex structure possessing the lowest average binding free energy.

## Analysis of interaction free energies of V3 loop:V3 loop and V3 loop:CXCR4 residue pairs

The interaction free energies between two residues ($R$ and $R'$) were computed by the relation

$$\Delta G^{\text{int}}_{RR'} = \underbrace{\sum_{i\in R}\sum_{j\in R'}\left(E^{\text{Coul}}_{ij}+E^{\text{GB}}_{ij}\right)}_{\Delta G^{\text{polar}}_{RR'}} + \underbrace{\sum_{i\in R}\sum_{j\in R'}E^{\text{vdW}}_{ij}+\sigma\sum_{i\in R,R'}\Delta S_i}_{\Delta G^{\text{nonpolar}}_{RR'}}. \quad (1)$$

The first and second group of terms on the right-hand side of Eq. 1 describe, respectively, polar and nonpolar interactions between $R$ and $R'$. For the investigation of V3 loop:CXCR4 intermolecular interactions, $R$ corresponds to a V3 loop residue and $R'$ to a CXCR4 residue. For the investigation of V3 loop intramolecular interactions, both $R$ and $R'$ correspond to V3 loop residues. To compute the $E_{ij}^{\text{GB}}$ term in Eq. 1, we included all protein (CXCR4) and ligand (V3 loop) atoms and set the charges of atoms outside the two residues $R$ and $R'$ under investigation to zero. The generalized-Born energies and the atomic accessible-surface areas (ΔSi) entering in Eq. 1 depend on the location of $R$ and $R'$ in the complex. The polar component contains a Coulombic term and a GB contribution, modeling the interaction between group R and the solvent polarization potential induced by $R'$. Similarly, the nonpolar component contains a van der Waals interaction between $R$, $R'$ and a surface term, expressing cavity contributions and nonpolar interactions with the surrounding solvent. The last term contains the difference in solvent-accessible surface areas of residues $R$ and $R'$ in the complex and unbound states (62–64). The analysis was performed using all 1000 snapshots from Complex 1. The nonpolar and polar solvation terms were calculated using the heterogeneous membrane-water GBSW using the same parameters of step 4. The sum of the two components reflects the total direct interaction between $R$ and $R'$ in the solvated complex (62,63,64).

We decomposed the polar and nonpolar interaction free energy contributions and present the results of the average interaction free energies of the lowest binding free energy complex in two-dimensional density maps in Fig. S1 and Fig. S2 in the Supporting Material. In addition, we calculated the average interaction energy per V3 loop residue, by summing up the interaction energies of the specific residue with all of its possible CXCR4 interacting residues, so as to investigate the critical gp120-V3 loop residues with regard to coreceptor binding. The results are presented in Table S4.

## RESULTS

We employed a comprehensive set of computational tools, primarily comprising binding free energy calculations and MD simulations, and identified Complex 1 as the MD simulation with the lowest average binding free energy (Fig. 1, and see Table S1). In all 17 complexes, the V3 loop residues lying outside the chemokine receptor experience higher flexibility; the average backbone RMSD without alignment

with respect to the starting simulation conformation is 3.9 ± 0.3 and 2.9 ± 0.1 Å for the entire V3 loop and the embedded region (8:26), respectively, in Complex 1 (see Table S2). In general, in the lowest binding free energy complexes (1, 4, and 12) the V3 loop possesses a relatively low RMSD average and standard deviation values, specifically within the residue moiety 8–26 (see Table S2). On the contrary, the unbound V3 loop is considered to be highly flexible as shown in Methods, in line with (27,29,30). Therefore, the results of this study indicate that the coreceptor binding stabilizes the conformation of the V3 loop, especially for the V3 loop residue moiety 8:26. In addition, at least for the lowest binding free energy complexes (1, 4, and 12), the V3 loop binding also provides stabilization for the N-terminal end of CXCR4 (see Table S2).

The nonpolar binding free energy of Complex 1 is comparable to Complex 4. However, according to both GB and PB approximations, the polar component of Complex 1 is more favored than Complex 4. According to the analysis of hydrogen-bond interactions of all complexes (see Table S3), this is mainly due to the absence of salt bridges between V3 loop:CXCR4 residues $Arg^{3}$:$Glu^{268}$, $Arg^{18}$:$Asp^{171}$, and $Arg^{31}$:$Glu^{14}$ in Complex 4 (see Table S3). Also, the lower polar binding free energy of Complex 12, compared to Complex 1, according to both GB and PB calculations, could mainly be attributed to the presence of additional salt bridges between V3 loop:CXCR4 residues $Arg^{11}$:$Asp^{10}$ and to a lesser extent $Arg^{9}$:$Asp^{10}$, $Lys^{10}$:$Tys^{12}$ (see Table S3); in addition, the salt bridge of $Arg^{31}$ with residue $Glu^{14}$ in Complex 1 is interchanged with a lower occupancy salt bridge with residue $Tys^{12}$ in Complex 12. On the contrary, $Arg^{3}$:$Asp^{22}$ and $Arg^{3}$:$Glu^{268}$ have lower occupancy in Complex 12 compared to 1. Nevertheless, despite the lower polar binding free energy of Complex 12 compared to Complex 1, the nonpolar component of Complex 1 is more favored compared to Complex 12 by ~24 kcal/mol, and overall Complex 1 is the most favored among all complexes. In what follows, we oriented our analysis to Complex 1, as its binding free energy is optimum compared to other complexes with regard to the sum of polar and nonpolar binding components, and provide a detailed investigation of the structural and physicochemical properties encountered within the 1000 frames extracted from the simulation.

## Structural properties of the bound V3 loop

Residues 8–26 of the V3 loop are buried within CXCR4, whereas residue moieties 1–7 and 27–35 mainly lie upon the N-terminal end of CXCR4 (all V3 loop residues are renumbered, starting from 1 and ending at 35). The V3 loop conformation is twisted, as shown in Fig. 1 and, two consecutive antiparallel $\beta$-sheets among the following residue moieties, 12–13:22–21 and 3–9:32–26, are observed in the trajectory. A $\beta$-turn or bend is observed within the core of the tip comprising residues 16:20, which is the mostly buried region of the V3 loop within the binding pocket, as shown in Fig. 1. The $\beta$-sheets provide a compact-thin shape and a relatively stable conformation of the V3 loop within the simulation. The two primary hydrogen bonds involving side chains of opposite V3 loop stems are $Arg^{9}$ NH1/2:$Gln^{25}$ OE1 and $Thr^{8}$ O:$Arg^{31}$ NH1/2. The relatively tight binding of a highly flexible loop can be attributed to both the cooperativity of intramolecular interactions in the bound structure, shown in Fig. S1, and the intermolecular interactions, analyzed below.

## Investigation of the V3 loop: CXCR4 complex structure

We present a detailed overview of the structural and physicochemical properties of the complex structure, which is based on the assessment of the intermolecular pairwise residue interaction free energies, shown in Fig. S2. Fig. 2 presents the most important intermolecular polar interactions in the trajectory, using the software VMD (65). Table 1 extracts information from Fig. S2 as well as Table S3 and summarizes the key interactions between V3 loop:CXCR4 residues, over all 1000 snapshots, and features in boldface the agreement with experimental findings.

The coordinates of structures in Complex 1, extracted every 2 ns, are provided (see Supporting Material). Within the simulation, the conformation of the coreceptor is very well retained with regard to the starting simulation structure. The average backbone RMSD of the intramembrane helical residues is equal to 1.3 ± 0.0 Å, and the average RMSD of the entire backbone is 2.2 ± 0.2 Å. The larger value of the latter is attributed to the higher flexibility of the non-intramembrane domains. It is worth noting that in all simulations, the average backbone RMSD of the intramembrane helical residues is low (see Table S2), showing the ability of the implicit solvent model used to preserve the integrity of the x-ray structure (11).

## Interactions of V3 loop residues 1:15 with CXCR4

V3 loop residues 3–15 are mainly engaged to intermolecular interactions with the N-terminal end of CXCR4 and to a lesser extent with ECL2 and ECL3 residues of CXCR4. $Cys^{1}$, $Thr^{2}$, and $Pro^{4}$ of the V3 loop are predominantly solvent-exposed. $Arg^{3}$ of the V3 loop forms a highly interacting salt bridge with $Glu^{268}$, a salt bridge with CXCR4 residue $Asp^{22}$ (Fig. 2 *A*), and is also proximal to the negatively charged $Asp^{20}$ and $Tys^{21}$. $Asn^{5}$ of the V3 loop intercalates between CXCR4 residues $Asp^{20}$, $Tys^{21}$, $Asp^{22}$, and $Ser^{23}$; two simultaneous hydrogen bonds are observed: $Asn^{5}$ ND2:$Tys^{21}$ O and $Asn^{5}$ ND2:$Asp^{20}$ OD1/2. $Asn^{6}$ of the V3 loop mainly interacts with CXCR4 residue $Ser^{23}$; as a result, two simultaneous hydrogen bonds are frequently formed among atom pairs $Asn^{6}$ N:$Ser^{23}$ OG and $Asn^{6}$ ND2/

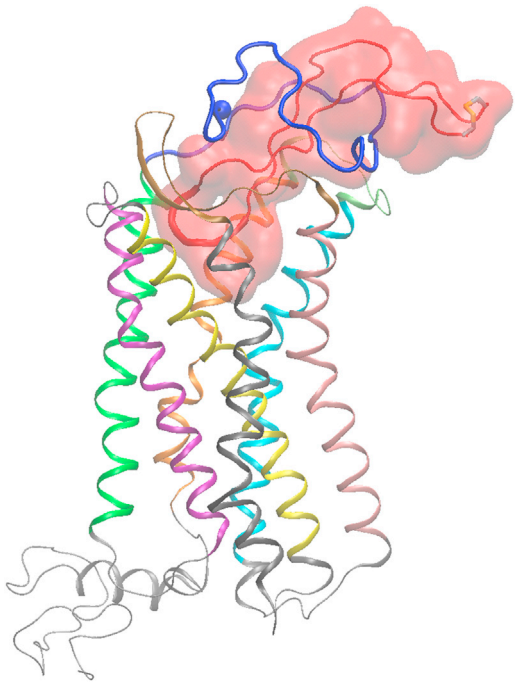

FIGURE 1 Molecular graphics image of the entire simulation system corresponding to the complex with the lowest average binding free energy. The V3 loop is shown in tube and transparent surface representation in red color, and its 16–20 residue moiety is shown in thick tube representation. The CXCR4 is shown in cartoon representation, and the coloring used for different protein domains is as follows: (*blue*) N-terminal domain; (*green*) intramembrane helix 1 (IH1); (*light gray*) intracellular loop 1 (ICL1); (*purple*) IH2L; (*light gray*) ECL1; (*yellow*) IH3; (*light gray*) ICL 2; (*medium gray*) IH4; (*ochre*) ECL2; (*pink*) IH5; (*light gray*) ICL3; (*cyan*) IH6; (*lime*) ECL3; (*orange*) IH7; and (*light gray*) C-terminal domain. (*van der Waals sphere*) N-terminal Cα atom of CXCR4. (*Thick transparent licorice representation*) V3 loop disulfide bridge. The definition of CXCR4 and V3 loop domains is presented in the Supporting Material (colors appear in the online version only).

OD1:Ser$^{23}$ OG. Because position 6 is an N-linked glycosylation site, we assume that apart from the loss of the aforementioned hydrogen-bond interactions, the N-linked glycosylation will not interfere with the binding as, according to the simulation, the side chain of Asn$^{6}$ points toward the aqueous environment. Asn$^{7}$ of the V3 loop is mainly interacting with CXCR4 residues Glu$^{14}$ and Asn$^{11}$ through weak Asn$^{7}$ OD1:Asn$^{11}$ ND2, Asn$^{7}$ ND2:Glu$^{14}$ OE1/2 hydrogen bonds. Thr$^{8}$ of the V3 loop intercalates between the backbone moieties of CXCR4 residues Tyr$^{21}$, Asp$^{22}$, Ser$^{23}$, and Lys$^{25}$; Thr$^{8}$ OG1 is hydrogen-bonded to Asp$^{22}$ O consistently during approximately the last half of the simulation. Arg$^{9}$ of the V3 loop interacts with the backbone moieties of residues Met$^{1}$, Tys$^{7}$, Thr$^{8}$, Ser$^{9}$, Asp$^{10}$, and Asn$^{11}$, as well as the side chain of CXCR4 residue Tys$^{12}$. Two hydrogen bonds are formed between the charged amide group of Arg$^{9}$ and the backbone carbonyl groups of Tys$^{7}$ (Fig. 2 *B*) and Asp$^{10}$; also, the proximity of the positively charged amide group of Arg$^{9}$ to the negatively charged Tys$^{12}$ results in a polar interaction between the two.

Lys$^{10}$ of the V3 loop participates in a highly interacting salt bridge with CXCR4 residue Asp$^{193}$ (Fig. 2 *A*) and in the following hydrogen bonds: Lys$^{10}$ N:Tys$^{12}$ SO(2/3), Lys$^{10}$ NZ:Asp$^{20}$ O, and Lys$^{10}$ NZ:Leu$^{266}$ O (Fig. 2 *B*). In addition, the nonpolar part of the Lys$^{10}$ side chain is proximal to the backbone of Ser$^{18}$, Gly$^{19}$, and the aromatic ring of Tys$^{21}$. Arg$^{11}$ of the V3 loop is buried in a diverse polar/nonpolar binding pocket comprising CXCR4 residues Met$^{1}$ (Fig. 2 *B*), Glu$^{2}$, Gly3, Tys7, Tys$^{12}$, Phe$^{189}$, and Pro$^{191}$; its nonpolar side-chain part is proximal to the aromatic rings of Tys$^{7}$, Tys$^{12}$, and Phe$^{189}$, and its backbone amide is hydrogen-bonded to Tys$^{12}$ SO3; moreover, its charged amide is close to the backbone of the first three CXCR4 N-terminal residues and is hydrogen-bonded to Met$^{1}$ O and Glu$^{2}$ O. Val$^{12}$ of the V3 loop forms nonpolar contacts with nearby CXCR4 residues Phe$^{189}$, Tyr$^{190}$, and Val$^{196}$. Ser$^{13}$ of the V3 loop is interacting mainly with Met$^{1}$ through the Ser$^{13}$ OG:Met$^{1}$ N hydrogen bond (Fig. 2 *B*). Leu$^{14}$ of the V3 loop is positioned between N-terminal and ECL2 of CXCR4 residues. Its backbone is mainly interacting with CXCR4 residues Glu$^{2}$, Gly$^{3}$ and, as a result, a low frequency hydrogen bond is observed between Leu$^{14}$ N and Glu$^{2}$ O; in addition the backbone carbonyl group of Leu$^{14}$ is hydrogen-bonded to the charged side-chain NH1 or NH2 Arg$^{183}$ in more than half of the simulation snapshots. The Leu$^{14}$ side chain intercalates around residues Ser$^{178}$, Ala$^{180}$, Ile$^{185}$, Asp$^{187}$, and Phe$^{189}$. The placement of the charged Asp$^{187}$ in the vicinity of the hydrophobic Leu$^{14}$ is attributed to the hydrogen bond among Asp$^{187}$ OD1 or OD2 with the V3 loop atom Gly$^{15}$ N (Fig. 2 *B*).

## Interactions of V3 loop residues 16:22 with CXCR4

Pro$^{16}$ of the V3 loop is buried within a binding pocket composed of CXCR4 residues Arg$^{30}$, Phe$^{36}$, Asn$^{37}$, and Leu$^{41}$, and its amide group forms a weak hydrogen bond with atom Asn$^{37}$ ND2. Gly$^{17}$ of the V3 loop is mainly involved in a hydrogen bond with Asp$^{97}$ OD1/2 in approximately the last-quarter of the simulation; in addition it is involved in a nonpolar interaction with residue Trp$^{94}$ and in a weak hydrogen bond with Tyr$^{45}$ OH. Arg$^{18}$ of the V3 loop forms the most highly interacting polar interactions with a group of CXCR4 residues because it is simultaneously engaged in two high occupancy salt bridges with residues Asp$^{171}$ and Glu$^{288}$ (Fig. 2 *A*). In addition, Arg$^{18}$ NH1/2 is hydrogen-bonded to His$^{203}$ NE2 (Fig. 2 *B*), Thr$^{117}$ OG1, Tyr$^{255}$ OH, and the Arg$^{18}$ backbone carbonyl group is hydrogen-bonded to the charged amide of Arg$^{188}$. The nonpolar part of the Arg$^{18}$ side chain is involved in nonpolar contacts with residues Trp$^{94}$ and Tyr$^{116}$, and the nonpolar moieties of Asp$^{187}$ and Arg$^{188}$. Val$^{19}$ is involved in mainly nonpolar interactions with its neighboring CXCR4 residues Ile$^{284}$ and Ser$^{285}$. Trp$^{20}$ of the V3 loop is embedded in a binding pocket that mainly consists of

**TABLE 1 Important intermolecular polar and nonpolar interaction free energies, hydrogen bonds, and salt bridges, between V3 loop and CXCR4 residue pairs within the MD simulation of the complex with the lowest average binding free energy**

| V3 loop residue[b] | CXCR4 residues (polar, nonpolar interaction free energies)[c] | Salt bridges[d] | Hydrogen bonds[e] |
|---|---|---|---|
| Arg$^{3}$ | **Asp$^{20}$** (−0.0, −1.6)[e] | Arg$^{3}$:**Asp$^{22}$** | |
| | **Tys$^{21}$** (1.1, −2,4) | Arg$^{3}$:**Glu$^{268}$** | |
| | **Asp$^{22}$** (−9.3, −2.5) | | |
| | **Glu$^{268}$** (−29.3, 2.6) | | |
| Asn$^{5}$ | **Asp$^{20}$** (−2.9, −1.9) | | Asn$^{5}$ ND2:**Asp$^{20}$** OD* |
| | **Tys$^{21}$** (−1.1, −1.5) | | Asn$^{5}$ ND2:**Tys$^{21}$** O |
| | **Ser$^{23}$** (−2.2, −1.3) | | |
| Asn$^{6}$ | **Ser$^{23}$** (−3.3, −2.3) | | Asn$^{6}$ N:**Ser$^{23}$** OG |
| | | | Asn$^{6}$ ND2:**Ser$^{23}$** OG |
| | | | Asn$^{6}$ OD1:**Ser$^{23}$** OG |
| Asn$^{7}$ | Asn$^{11}$ (−0.1, −0.8) | | Asn$^{7}$ OD1:Asn$^{11}$ ND2 |
| | **Glu$^{14}$** (−1.9, −0.9) | | Asn$^{7}$ ND2:**Glu$^{14}$** OE* |
| Thr$^{8}$ | **Tys$^{21}$** (0.3, −2.9) | | Thr$^{8}$ OG1:Asp$^{22}$ O |
| | **Asp$^{22}$** (−3.1, −0.7) | | |
| | **Lys$^{25}$** (0.5, −2.1) | | |
| Arg$^{9}$ | Met$^{1}$ (0.0, −1.9) | | Arg$^{9}$ NH*:**Tys$^{7}$** O |
| | **Tys$^{7}$** (−7.3, −1.2) | | Arg$^{9}$ NH*:**Asp$^{10}$** O |
| | Ser$^{9}$ (0.2, −1.6) | | |
| | **Asp$^{10}$** (−2.8, −1.0) | | |
| | **Tys$^{12}$** (−4.4, −5.0) | | |
| Lys$^{10}$ | **Tys$^{12}$** (−9.5, −3.0) | Lys$^{10}$:**Asp$^{193}$** | Lys$^{10}$ N:**Tys$^{12}$** SO* |
| | **Asp$^{20}$** (−8.5, −1.0) | | Lys$^{10}$ NZ:**Asp$^{20}$** O |
| | **Tys$^{21}$** (−1.0, −3.4) | | Lys$^{10}$ NZ:Leu$^{266}$ O |
| | **Asp$^{193}$** (−17.8, −0.9) | | |
| | Leu$^{266}$ (−6.7, −2.3) | | |
| Arg$^{11}$ | Met$^{1}$ (−7.9, −2.1) | | Arg$^{11}$ NH*/NE:Met$^{1}$ O |
| | **Glu$^{2}$** (−0.9, −1.9) | | Arg$^{11}$ NH*:**Glu$^{2}$ O** |
| | Gly$^{3}$ (−1.2, −1.7) | | Arg$^{11}$ N:**Tys$^{12}$** SO* |
| | **Tys$^{7}$** (−1.0, −4.5) | | |
| | **Tys$^{12}$** (−2.3, −6.3) | | |
| | **Phe$^{189}$** (0.1, −1.7) | | |
| Val$^{12}$ | **Phe$^{189}$** (0.1, −2.3) | | |
| | **Tyr$^{190}$** (−0.0, −3.2) | | |
| Ser$^{13}$ | Met$^{1}$ (−5.8, −1.9) | | Ser$^{13}$ OG:Met$^{1}$ N |
| Leu$^{14}$ | **Glu$^{2}$** (−1.2, −1.4) | | Leu$^{14}$ N:**Glu$^{2}$ O** |
| | Gly$^{3}$ (−0.2, −1.7) | | Leu$^{14}$ O:**Arg$^{183}$** NH* |
| | Ser$^{178}$ (−0.1, −1.7) | | |
| | Ala$^{180}$ (−0.0, −2.4) | | |
| | **Arg$^{183}$** (−1.6, −1.0) | | |
| | Ile$^{185}$ (0.1, −2.5) | | |
| | **Asp$^{187}$** (−3.8, −3.4) | | |
| | **Phe$^{189}$** (0.1, −1.6) | | |
| Gly$^{15}$ | Ile$^{185}$ (−0.1, −1.7) | | Gly$^{15}$ N:**Asp$^{187}$** OD* |
| | **Asp$^{187}$** (−6.3, 0.1) | | |
| Pro$^{16}$ | Arg$^{30}$ (−0.3, −2.0) | | Pro$^{16}$ N:Asn$^{37}$ ND2 |
| | Asn$^{37}$ (−0.1, −2.2) | | |
| Gly$^{17}$ | **Tyr$^{45}$** (−0.8, −0.6) | | Gly$^{17}$ N:**Tyr$^{45}$** OH |
| | Trp$^{94}$ (−1.2, −1.7) | | Gly$^{17}$ N:**Asp$^{97}$** OD* |
| | **Asp$^{97}$** (−2.2, −0.1) | | |
| | Glu$^{288}$ (−2.8, −0.1) | | |
| Arg$^{18}$ | Trp$^{94}$ (−0.1, −2.0) | Arg$^{18}$:**Asp$^{171}$** | Arg$^{18}$ NH*:Thr$^{117}$ OG1 |
| | His$^{113}$ (−3.2, −1.3) | Arg$^{18}$:**Glu$^{288}$** | Arg$^{18}$ O:**Arg$^{188}$** NH*/NE |
| | Tyr$^{116}$ (5.1, −2.5) | | Arg$^{18}$ NH*:His$^{203}$ NE2 |
| | Thr$^{117}$ (−2.1, −1.0) | | Arg$^{18}$ NH*:**Tyr$^{255}$** OH |
| | **Asp$^{171}$** (−37.2, −0.3) | | |
| | **Asp$^{187}$** (−1.1, −1.4) | | |
| | **Arg$^{188}$** (8.27, −3.1) | | |
| | His$^{203}$ (−18.1, 0.4) | | |
| | **Tyr$^{255}$** (2.8, −0.5) | | |
| | **Glu$^{288}$** (−65.6, 2.6) | | |
| Val$^{19}$ | **Glu$^{288}$** (−2.7,−1.0) | | |
| Trp$^{20}$ | **Arg$^{188}$** (−4.8, −3.2) | | Trp$^{20}$ NE1:**Arg$^{188}$** NH* |
| | **Tyr$^{190}$** (−1.5, −1.7) | | Trp$^{20}$ NE1:**Tyr$^{190}$** OH |
| | Val$^{196}$ (0.1, −2.2) | | Trp$^{20}$ O:His$^{281}$ ND1 |
| | Phe$^{199}$ (−0.0, −1.6) | | |
| | **Gln$^{200}$** (−1.5, −3.4) | | |
| | His$^{281}$ (−0.9, −0.9) | | |
| Tyr$^{21}$ | Arg$^{30}$ (0.2, −3.3) | | Tyr$^{21}$ OH:Asn$^{278}$ ND2 |
| | Glu$^{277}$ (−0.2, −2.0) | | |
| | Asn$^{278}$ (−1.2, −2.8) | | |
| | His$^{281}$ (0.2, −4.3) | | |
| Thr$^{22}$ | Leu$^{266}$ (−0.0, 1.8) | | Thr$^{22}$ OG1:Glu$^{277}$ OE* |
| | Glu$^{277}$ (−2.8, −0.8) | | |
| Thr$^{23}$ | Met$^{1}$ (−0.0, −2.1) | | Thr$^{23}$ OG1:**Cys$^{28}$** O |
| | **Cys$^{28}$** (−0.6, −1.8) | | |
| Gly$^{24}$ | **Lys$^{25}$** (−0.4, −1.8) | | Gly$^{24}$ N:**Cys$^{28}$** O |
| | **Cys$^{28}$** (−3.2, −2.1) | | Gly$^{24}$ O:**Cys$^{28}$** N |
| Gln$^{25}$ | Met$^{1}$ (0.1, −1.6) | | Gln$^{25}$ NE2:**Cys$^{28}$** O |
| | Pro$^{27}$ (0.0, −2.7) | | |
| | **Cys$^{28}$** (−0.4, −1.5) | | |
| Ile$^{26}$ | **Lys$^{25}$** (−0.0, −1.7) | | |
| Val$^{27}$ | Ser$^{9}$ (0.1, −1.7) | | |
| | **Asp$^{10}$** (−0.1, −1.7) | | |
| | Asn$^{11}$ (−0.0, −1.9) | | |
| Arg$^{31}$ | **Tys$^{12}$** (−1.8, −1.2) | Arg$^{31}$:**Glu$^{14}$** | Arg$^{31}$ NH*:Ser$^{18}$ OG |
| | **Glu$^{14}$** (−16.8, −0.5) | Arg$^{31}$:**Asp$^{20}$** | |
| | Ser$^{18}$ (0.2, −3.0) | | |
| | **Asp$^{20}$** (−8.4, −1.4) | | |
| Lys$^{32}$ | Gly$^{17}$ (−3.4, −0.8) | Lys$^{32}$:**Asp$^{20}$** | Lys$^{32}$ NZ:Gly$^{17}$ O |
| | Ser$^{18}$ (−3.1, −0.6) | | Lys$^{32}$ NZ:Ser$^{18}$ O |
| | **Asp$^{20}$** (16.4, 0.4) | | |
| His$^{34}$ | **Asp$^{20}$** (−1.0, −1.5) | | His$^{34}$ ND1:**Asp$^{20}$** OD2 |

CXCR4 residues marked in boldface are experimentally associated with HIV-1 coreceptor activity (see Discussion). The results presented correspond to analysis of 1000 snapshots. The asterisk (*) symbol used after any V3 loop/CXCR4 atom in the hydrogen bonding pair denotes that any of the atoms in the charged, carboxyl or amide, side-chain group can participate in the hydrogen-bond formation.

[a]Principal interacting V3 loop.

[b]Principal interacting CXCR4 residue pairs.

[c]For each pair, the average polar and nonpolar average interaction free energies (polar, nonpolar), are provided in parentheses next to each CXCR4 residue; all energies are in kcal/mol.

[d]Salt bridges between V3 loop and CXCR4 residue pairs.

[e]Hydrogen bonds between V3 loop and CXCR4 atom pairs.

residues Arg$^{188}$, Tyr$^{190}$, Val$^{196}$, Phe$^{199}$, Gln$^{200}$, and Ile$^{284}$; the interactions with Arg$^{188}$ and Tyr$^{190}$ are polarly driven due to the Trp$^{20}$ NE1:Arg$^{188}$ NH1/2 and Trp$^{20}$ NE1:Tyr$^{190}$ OH hydrogen bonds, whereas the latter interactions mainly involve nonpolar contacts. The backbone carbonyl group of Trp$^{20}$ is weakly hydrogen-bonded to CXCR4 atom His$^{281}$

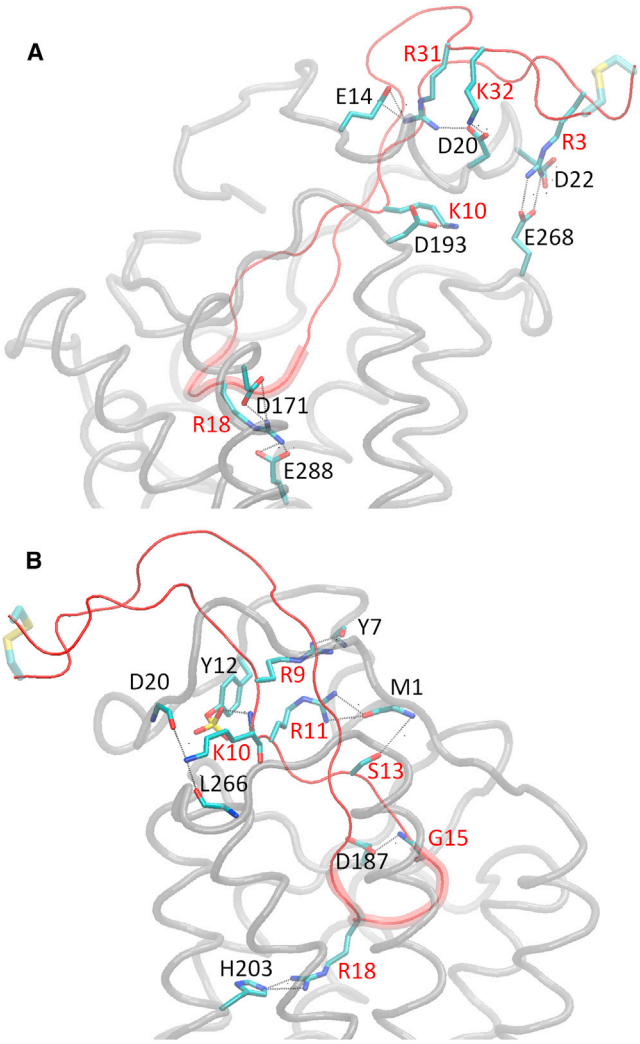


FIGURE 2 Molecular graphics images of important polar interactions corresponding to the complex with the lowest average binding free energy. Panels (*A*) and (*B*) depict the salt bridges and the most important hydrogen bonds, respectively, according to residue pair-wise interaction free energies. Panel (*B*) is rotated by approximately 180° with respect to (*A*) around the z (membrane) axis. The V3 loop is shown in tube and in red color, and its 16–20 residue moiety is shown in fat tube representation. The CXCR4 is shown in light gray transparent tube representation. The salt bridge and hydrogen bonds in panels (*A*) and (*B*) are denoted in dashed lines and the participating V3 loop and CXCR4 residue moieties are shown in licorice; V3 loop and CXCR4 residues are annotated in red and black, color respectively. Hydrogen atoms are omitted for clarity and the V3 loop disulfide bridge is shown in fat transparent licorice representation. (colors appear in the online version only).

ND1. $Tyr^{21}$ of the V3 loop is surrounded by CXCR4 residues $Phe^{29}$, $Arg^{30}$, $Asn^{278}$, $Glu^{277}$, and $His^{281}$; its side-chain hydroxyl group is hydrogen-bonded to CXCR4 atom $Asn^{278}$ ND2/OD1 and the aromatic ring predominantly participates in nonpolar interactions with the nonpolar moieties of these residues. $Thr^{22}$ of the V3 loop interacts with CXCR4 residues $Leu^{266}$ and $Glu^{277}$. The former residue interacts with the nonpolar moiety of $Thr^{22}$, whereas the latter residue is involved in a $Thr^{22}$ OG1:$Glu^{277}$ OE1 hydrogen bond during approximately the last-third of the simulation.

## Interactions of V3 loop residues 23:35 with CXCR4

V3 loop residues 23–35 interact solely with the CXCR4 N-terminal region. $Thr^{23}$ of the V3 loop forms nonpolar contacts with CXCR4 residues $Met^{1}$, the disulfide bridge residue $Cys^{28}$, and the backbone of $Phe^{29}$, and in addition, its side-chain hydroxyl group is weakly hydrogen-bonded to the CXCR4 atom $Cys^{28}$ O. $Gly^{24}$ of the V3 loop is in proximity with the nonpolar moiety of CXCR4 residue $Lys^{25}$ and interacts with the disulfide bridge CXCR4 residue $Cys^{28}$ via the $Gly^{24}$ O:$Cys^{28}$ N or $Gly^{24}$ N:$Cys^{28}$ O hydrogen bonds. $Gln^{25}$ of the V3 loop is in the vicinity of CXCR4 residues $Met^{1}$, $Pro^{27}$, $Cys^{28}$, and $Phe^{29}$; its interaction with residue $Cys^{28}$ is facilitated through a low-occupancy hydrogen bond between $Gln^{25}$ NE2 and $Cys^{28}$ O. $Ile^{26}$ of the V3 loop is primarily interacting with the nonpolar moieties of CXCR4 residues $Ser^{23}$ and $Lys^{25}$, and in a similar fashion, $Val^{27}$ of the V3 loop is buried within the nonpolar moieties of CXCR4 residues $Ser^{9}$, $Asp^{10}$, and $Asn^{11}$. Residues $Gly^{28}$, $Asp^{29}$, and $Ile^{30}$ of the V3 loop are solvent-exposed and are not engaged in intermolecular interactions. $Arg^{31}$ of the V3 loop forms two simultaneous salt bridges with CXCR4 residues $Glu^{14}$ and $Asp^{20}$, with the former being stronger (Fig. 2 *A*), and is proximal to oppositely charged residue $Tys^{12}$ and $Ser^{18}$ of CXCR4; as a result its charged amide is weakly hydrogen-bonded to $Ser^{18}$ OG. $Lys^{32}$ of the V3 loop forms a highly interacting salt bridge with CXCR4 residue $Asp^{20}$ (Fig. 2 *A*), and its charged amide group is hydrogen-bonded to the backbone carbonyl groups of CXCR4 residues $Glu^{17}$ and $Ser^{18}$. Residue $Arg^{31}$ of the V3 loop is proximal to the oppositely charged CXCR4 residue $Tys^{12}$, and residue $Lys^{32}$ of the V3 loop participates in a hydrogen bond through its charged amide group with the backbone carbonyl group of CXCR4 $Ser^{18}$. Residues $Ala^{33}$ and $Cys^{35}$ of the V3 loop are solvent-exposed and are not engaged in any interactions with CXCR4 residues, whereas the side-chain amide group of V3 loop $His^{34}$ is weakly hydrogen-bonded to the charged carboxyl group of CXCR4 $Asp^{20}$. The disulfide bridge points toward the aqueous environment throughout the simulation, as would be expected if it was covalently bonded to the entire gp120 protein.

## DISCUSSION

In the MD simulation, the bound V3 loop is twisted, and its compactness is associated with the formation of $\beta$-sheets and the $Arg^{9}$ NH1/2:$Gln^{25}$ OE1 hydrogen bond. These data are in agreement with recently published results from MD simulations of unbound V3 loops, according to which the compactness was linked to an electrostatic interaction

between the side chains of residues $Arg^{9}$ and $Glu^{25}$ in one of the V3 loops (27). The compact-thin shape of the V3 loop bound structure suggests it is necessary for it to adopt a compact-thin shape before coreceptor binding. The compactness is related the maximization of the tip-base distance, as suggested in López de Victoria et al. (27), and provides the ability to residue $Arg^{18}$ for the V3 loop to be buried deep inside the membrane and form highly interacting polar interactions with CXCR4 residues $Asp^{171}$ and $Glu^{288}$.

The average backbone RMSD of the unbound dual tropic V3 loop conformations of this study compared to the most representative, CCR5 tropic PDB:2QAD structures in solution (Fig. 7*B*, basins 1–3 of López de Victoria et al. (27)) is 6.2 ± 0.7, 5.4 ± 1.4, and 7.7 ± 0.8 Å, respectively. The corresponding values calculated for the residue moiety 8–26 are 5.9 ± 0.7, 4.6 ± 1.6, and 7.7 ± 1.0 Å. The relatively large values are in agreement with very recent results of Chandramouli et al. (30) showing, respectively overall structural dissimilarity between different tropic unbound V3 loops, at the level of crown region, defined as residues 11–25. Nevertheless, upon alignment, the average backbone RMSD between the bound dual tropic V3 loop conformations of Complex 1 for the last 5 ns, and the most representative, CCR5 tropic PDB:2QAD structure in solution with maximized tip-base distance (Fig. 7*B*, basin 2 of López de Victoria et al. (27)) is 3.8 ± 0.1 and 2.6 ± 0.1 Å, respectively, for the entire V3 loop and the 8–26 residues moiety. Thus, despite 1), the relatively large sequence dissimilarity between the two V3 loops in 9 out of 35 positions, and 2), the different simulation properties with regard to the V3 loop, bound versus unbound, the specific maximized tip-base conformation (which can both adopt and is required at least for CXCR4 binding) is interestingly comparable, specifically for the 8–26 residue moiety.

## Key interacting V3 loop residues for CXCR4 recognition

The sum of the interaction energies per V3 loop residue, presented in Table S4, presents the V3 loop residues associated with CXCR4 intermolecular interactions, in descending order. The 10 most highly interacting V3 loop residues, with average total interacting free energies lower than −20 kcal/mol are: $Arg^{18}$, $Lys^{10}$, $Arg^{3}$, $Arg^{31}$, $Arg^{11}$, $Arg^{9}$, $Trp^{20}$, $Leu^{14}$, $Lys^{32}$, and $Tyr^{21}$. Seven out of ten are positively charged and are highly interacting, predominantly owing to strong polar interactions that involve salt bridges and/or hydrogen bonds with CXCR4 residues. The specific dual tropic V3 loop possesses a +6 net charge, and within the simulation of this study we observe an abundance of highly interacting specific salt bridges and hydrogen bonds between positively charged V3 loop residues and negatively charged CXCR4 residues. This also supports: 1), The rule implying that the increase of the net charge of the V3 loop is associated with CXCR4 tropism (10,31); and 2), The suggestion that electrostatics likely dictate coreceptor choice (30).

Furthermore, it is worth noting that residues $Arg^{18}$, $Lys^{10}$, $Arg^{3}$, $Arg^{31}$, and $Arg^{9}$ possess the highest residue propensity for their corresponding positions regarding both CXCR4 and dual tropic 35-residue-V3 loops, while residues Arg11 and $Lys^{32}$ are the second most probable residues for positions 11 and 32 (see Table S4). Residues $Trp^{20}$, $Leu^{14}$, and $Tyr^{21}$ are highly interacting, predominantly owing to nonpolar interactions with CXCR4 residues. Residue $Tyr^{21}$ possesses the highest residue propensity for position 21 for both CXCR4 and dual tropic 35-residue-V3 loops, while residues $Trp^{20}$ and $Leu^{14}$ are the second most probable residue for positions 20 and 14, respectively, after residues $Phe^{20}$ and $Ile^{14}$ which are first, and possess similar physicochemical properties (see Table S4). In addition, the residues $Pro^{16}$, $Gly^{15}$, $Ser^{13}$, $Asn^{5}$, $Gly^{17}$, $Gln^{25}$, $Gly^{24}$, $Val^{12}$, $Thr^{8}$, $Val^{19}$, $Thr^{22}$, $Thr^{23}$, $Asn^{6}$, $Val^{27}$, $Ile^{26}$, and $Asn^{7}$ possess interaction free energies, in descending order of strength, within the range of −14 to −5 kcal/mol (see Table S4). Moreover, this study provides what is, to our knowledge, the first reported molecular recognition-based piece of evidence on the 11/24/25 rule (8), as $Arg^{11}$ NH1/2 of V3 loop is hydrogen-bonded to $Met^{1}$ O of CXCR4. Taking into consideration that $Met^{1}$ is the first N-terminal residue, and that interactions between the V3 loop and CXCR4 N-terminal are critical for the molecular recognition (19–21,24), this high-occupancy hydrogen bond could be invaluable for reducing the high flexibility of the N-terminal end (11) upon binding (see Table S2).

Since 1997, a series of experimental studies aimed at exploring the key CXCR4 residues related to the V3 loop interaction (16–22,24–26). A fraction of these studies also aimed at comparing the CXCR4 residues associated with V3 loop binding to the CXCR4 residues associated with the stromal cell-derived factor1 (SDF1$\alpha$) binding, because SDF1$\alpha$ constitutes a natural ligand of CXCR4 with blocking capacity of HIV-1 (66,67). In addition, some studies have compared the CXCR4 residues associated with the V3 loop binding to the CXCR4 residues involved in the binding of potential therapeutic candidates that act as V3 loop antagonists (68). The primary limitation of all studies reported until now is the lack of information regarding the interacting V3 loop:CXCR4 key residue pairs, due to the absence of a published complex structure.

## Role of the N-terminal of CXCR4

The deletion of CXCR4 residues 2–25 is associated with at least 60% reduction of HIV-1 coreceptor activity (17). The same residue moiety is also critical for SDF1$\alpha$ and DV1 dimer binding, where DV1 dimer is a CXCR4 HIV-1 inhibitor (68). Our results show that more than half of the V3 loop residues, mainly residues of the opposite stems, interact

with the 2–25 CXCR4 residue sector. Alanine substitutions to sulfated tyrosines 7 and 12 correlate with or markedly impair the coreceptor activity for HIV-1 for different dual tropic V3 loops (19,24). According to our results, $Tys^{7}$ is proximal to oppositely charged residues $Arg^{9}$ and $Arg^{11}$ of the V3 loop, and its backbone carbonyl is hydrogen-bonded to the charged side-chain group of $Arg^{9}$. It is worth noting that the functional role of some CXCR4 residues could be associated with their intramolecular interactions because they can be fundamental in stabilizing the coreceptor structure. Within the simulation, we observe that $Tys^{7}$ $SO_{2-4}$ is mainly hydrogen-bonded with the CXCR4 backbone amide groups of $Ile^{4}$ and $Ser^{5}$, contributing to the stabilization of the 3–6 $\beta$-turn in CXCR4. $Tys^{12}$ is attracted to oppositely charged residues $Arg^{9}$, $Lys^{10}$, and $Arg^{11}$ of the V3 loop, and is stabilized in proximity to $Lys^{10}$ and $Arg^{11}$ via simultaneous hydrogen bonds among its charged $SO_4$ side-chain group and the backbone amide groups of $Lys^{10}$ and $Arg^{11}$.

Additional alanine substitutions on CXCR4 N-terminal residues $Glu^{2}$, $Asp^{10}$, $Glu^{14}$, $Glu^{15}$, $Asp^{20}$, $Tys^{21}$, $Asp^{22}$, $Ser^{23}$, $Lys^{25}$, $Glu^{26}$, $Cys^{28}$, and $Glu^{32}$ lead to a decrease of HIV-1 coreceptor activity, showing that they are associated with V3 loop binding (19–21,24). In the MD simulation, $Glu^{2}$ of CXCR4 is proximal to the oppositely charged $Arg^{11}$ of the V3 loop, and owing to its position, it contributes to the stabilization of the polar interaction among the latter and CXCR4 residue $Met^{1}$. $Asp^{10}$ of CXCR4 is polarly attracted to oppositely charged $Arg^{9}$ of the V3 loop and its position is stabilized by a hydrogen bond through its backbone carbonyl group with the side-chain amide of V3 loop residue $Arg^{9}$. The charged side-chain carboxyl group of CXCR4 $Glu^{14}$ forms a highly interacting salt bridge with the V3 loop residue $Lys^{32}$ and a hydrogen bond with V3 loop atom $Asn^{7}$ ND2. $Glu^{15}$ of CXCR4 points toward the opposite side of the binding site and is not directly related to the binding with V3 loop; nevertheless, as both $Glu^{14}$ and $Glu^{15}$ correlate with HIV-1 coreceptor activity, the interactions formed by $Glu^{14}$ can be interchanged by its neighboring $Glu^{15}$. $Asp^{20}$ of CXCR4 forms two highly interacting salt bridges with residues $Arg^{31}$ and $Lys^{32}$, as well as hydrogen bonds with residues $Asn^{5}$, $Lys^{10}$, and $His^{34}$. Residue $Tys^{21}$ of CXCR4 is proximal to V3 loop residues $Thr^{8}$ and $Lys^{10}$, forming noteworthy nonpolar interactions. The critical role of CXCR4 residue $Asp^{22}$ could mainly be attributed to its high-occupancy salt bridge with V3 loop residue $Arg^{3}$. $Ser^{23}$ of CXCR4 participates in hydrogen bonds between its side-chain hydroxyl group with main- and side-chain atoms of V3 loop residue $Asn^{6}$. Residue $Asp^{25}$ of CXCR4 forms noteworthy nonpolar interactions with V3 loop residues $Thr^{8}$, $Gly^{24}$, and $Ile^{26}$. Residue $Glu^{26}$ of CXCR4 is weakly interacting with V3 loop residues $Gly^{24}$, $Gln^{25}$, and $Ile^{26}$, while residue $Glu^{32}$ of CXCR4 is not part of the binding pocket.

We observe the presence of intramolecular hydrogen bonds and salt bridges between the N-terminal side chains and ECL3 side/main chains of CXCR4 residue pairs $Tys^{21}$:$Gly^{272}$, $Asp^{22}$:$Lys^{271}$, $Glu^{26}$:$Gly^{273}$, and $Glu^{32}$:$Lys^{282}$. The aforementioned interactions involve the $Tys^{21}$, $Asp^{22}$, $Glu^{26}$, and $Glu^{32}$ N-terminal CXCR4 residues, which are associated with HIV-1 coreceptor activity. The interactions can be fundamental and constitute a sufficient condition for the proper connection of the N-terminal end and ECL3 of CXCR4, as the deletion of the $Cys^{28}$-$Cys^{274}$ disulfide bridge affects, but not to a large extent, the HIV-1 coreceptor activity (20). Therefore, the side chains of CXCR4 residues $Tys^{21}$, $Asp^{22}$, $Glu^{26}$, and $Glu^{32}$ can play a key role in the preservation of the appropriate CXCR4 structure required for the gp120 binding through its V3 loop.

The interactions reported in this study between the V3 loop and the CXCR4 N-terminal end are not similar to the predicted interactions between the V3 loop and the CCR5 N-terminal end by Huang et al. (5) using molecular docking. This variation can mainly be attributed to the facts that 1), the N-terminal conformations of CXCR4 (39) and CCR5 (5) are different, and 2) CXCR4 and CCR5 V3 loops mainly differ owing to their stem-region residues, which predominantly interact with the N-termini of the coreceptors. In line with this suggestion, a recent analysis indicated that the critical sites of features informative of viral tropism belong to stem regions (69).

## Role of intramembrane helices 1 and 2 of CXCR4

Alanine substitutions showed a decrease in HIV-1 coreceptor activity for $Tyr^{45}$ (17,68), $Asp^{97}$ (17,18,20), to a lesser extent for $His^{79}$ (17,18,68) and $Phe^{87}$ (18), but not for $Lys^{110}$ and $Tyr^{121}$ (18). In line with this, $Lys^{110}$ and $Tyr^{121}$, are not involved in interactions with the V3 loop in the simulation. On the contrary, both $Tyr^{45}$ and $Asp^{97}$ of CXCR4 interact through their side chains with mainly the carbonyl backbone group of $Gly^{17}$. According to our simulation, residue $Phe^{87}$ of CXCR4 is indirectly involved in the V3 loop binding by forming $\pi$-$\pi$ interactions with residue CXCR4 residue $Tyr^{116}$, which directly interacts with $Arg^{18}$; the $Phe^{87}$-$Tyr^{116}$ interaction is also present in the crystallographic structure (11). According to both the crystal structure (11) and our findings, $His^{79}$ is far from the binding site and it is hydrogen-bonded through its side-chain amide group to $Tyr^{157}$ OH of CXCR4.

## Role of ECL2 of CXCR4

Studies have shown that alanine substitution on residues $Trp^{161}$ and $Pro^{163}$ correlate with the HIV-1 activity (17,18). In both the crystallographic structure (11) and our simulation, residues $Trp^{161}$ and $Pro^{163}$ are positioned in the center of the membrane, away from the binding site; we hypothesize that these residues play a role in the preservation of the helical shape of intramembrane helix 4. $Asp^{171}$ of CXCR4 is experimentally considered one of the most

important residues involved in the HIV-1 coreceptor activity (17,18) because an alanine substitution at this position reduces by >60% the activity of both HIV-1 (17) and of DV1 (68). Within our simulation, $Asp^{171}$ of CXCR4 forms a highly interacting salt bridge with V3 loop $Arg^{18}$. Despite the fact that $Asp^{182}$ of CXCR4 is not part of the binding site, also according to the CXCR4 crystal structure (11), its alanine substitution results in a reduction of HIV-1 coreceptor activity (18). $Asp^{182}$ can play a key role in attracting the positively charged V3 loop before its binding, because $Asp^{182}$ possesses the peak position in ECL2 facing toward the aqueous extracellular region in our simulation and the crystallographic structure (11). Alanine substitutions revealed that $Arg^{183}$ (21,23), and to a lesser extent $Tyr^{184}$ (21,23) of CXCR4, are related to HIV-1 coreceptor activity. Within the simulation trajectory, the charged amide side chain of $Arg^{183}$ is involved in a hydrogen bond with the backbone carbonyl group of V3 loop residue $Leu^{14}$, while its nonpolar side-chain moiety is proximal to $Leu^{14}$ and $Gly^{15}$ of the V3 loop.

In line with the experimental structure (11), residue $Tyr^{184}$ is not in the binding site and it faces toward the aqueous extracellular environment, in similar fashion to $Asp^{182}$; consequently it can play a role in attracting the positively charged V3 loop before its binding through its side-chain OH group. Furthermore, alanine substitutions at CXCR4 residues $Asp^{187}$ and $Phe^{189}$ are experimentally associated with a >60% reduction for HIV-1 coreceptor activity, suggesting the critical role of both residues (17,23). In line with it, the charged carboxyl group of $Asp^{187}$ is hydrogen-bonded to atom $Gly^{15}$ N of the V3 loop throughout the entire simulation trajectory and its nonpolar moiety is interacting with the proximal V3 loop $Leu^{14}$ side chain. In the simulation, $Phe^{189}$ is buried in a nonpolar core composed by V3 loop side chains of residues $Arg^{11}$ (nonpolar moiety), $Val^{12}$ and $Leu^{14}$. Moreover, site-directed mutagenesis studies have depicted an involvement of the CXCR4 residues $Arg^{188}$, $Tyr^{190}$, and $Pro^{191}$ (17,18,20,21,23) in the HIV-1 binding. In our simulation, the $Arg^{188}$ charged amide group is hydrogen-bonded atoms $Arg^{18}$ O and $Trp^{20}$ NE1 atom of the V3 loop, the $Tyr^{190}$ side chain is also hydrogen-bonded to $Trp^{20}$ NE1 and forms nonpolar interactions with $Val^{12}$ of the V3 loop, and $Pro^{191}$ is also proximal to $Val^{12}$ of the V3 loop. The coreceptor activity is experimentally markedly impaired by an alanine substitution at 193 for a specific HIV-1 strain (19); in agreement with this, in our simulation $Asp^{193}$ forms a highly interacting salt bridge with V3 loop residue $Lys^{10}$.

### Role of intramembrane helices 5, 6, 7 and ECL3 of CXCR4

An alanine substitution on CXCR4 residue $Gln^{200}$ affects significantly the HIV-1 coreceptor binding (21), and our results show that $Gln^{200}$ forms a strong nonpolar interaction with the aromatic ring of the V3 loop residue $Trp^{20}$. This interaction is additionally stabilized by an intramolecular hydrogen bond between atoms $Gln^{200}$ NE2 and $Asp^{262}$ OD2, and an alanine substitution to $Asp^{262}$ also reduces the HIV-1 coreceptor activity (17,18). Furthermore, an alanine substitution of CXCR4 residue $Trp^{252}$ is related to a decrease of the binding affinity of the V3 loop, SDF1$\alpha$, and the DV1 inhibitor (17,68). The results of our simulation and the crystallographic structure (11) show that $Trp^{252}$ is positioned in the center of the membrane, being away from the binding site. Nevertheless, in both the experimental structure and the simulation, the aromatic group of $Trp^{252}$ forms aromatic nonpolar interactions with CXCR4 residues $Phe^{248}$, $Tyr^{255}$, $Tyr^{256}$, and $Phe^{192}$, which can be most important for the stabilization of the nucleus of aromatic residues and consequently, the hydrogen bond formation between $Tyr^{255}$ OH of CXCR4 and $Arg^{18}$ NH1/2 of the V3 loop. The latter interaction should be significant, owing to the experimentally determined critical role of $Tyr^{255}$ and $Tyr^{256}$ (17,18,68).

Our results justify the decrease of activity upon $Glu^{268}$ substitution to alanine, because $Glu^{268}$ participates in a highly interacting salt bridge with V3 loop residue $Arg^{3}$ (21). $Glu^{288}$ of CXCR4 is determined to be one of the most critical residues for HIV-1 coreceptor activity, because even an aspartic acid substitution at this position, which preserves the physicochemical properties of the former, results in a considerable 30–60% loss of HIV-1 coreceptor activity (17,18); also, alanine substitutions on $Glu^{288}$ resulted in a significant decrease of coreceptor activity related to SDF1$\alpha$, DV1 and the V3 loop (18,68). Within our simulation, $Glu^{288}$ forms the most highly interacting salt bridge with $Arg^{18}$ of the V3 loop. Because $Arg^{18}$ of the V3 loop is simultaneously forming two highly interacting salt bridges with CXCR4 residues $Asp^{171}$ and $Glu^{288}$ residues, our data suggest that a substitution to $Asp^{288}$ could not permit $Arg^{18}$ to form the two concurrent salt bridges due to the shorter size of aspartate compared to glutamate (Fig. 2 *A*). Alanine substitutions at CXCR4 residues $His^{294}$ and $Asn^{298}$ showed that these residues correlate with HIV-1 coreceptor activity (17,18). Their presence below the center of the membrane, far from the binding site, also in the crystallographic structure (11), suggests that their role should be attributed to their involvement in intramolecular interactions within CXCR4.

## CONCLUSIONS

We consider that the success of having remarkable accordance with experimental results is due to the systematic methodology employed, which includes the following features:

1. The modeling of the entire CXCR4 structure and the use of an extensive set of computational tools and methods to produce a variety of structural templates of V3 loop and CXCR4 for docking;

2. The large number of docked complexes investigated;
3. The heterogeneous dielectric solvation models used to rank the complex structures according to their binding free energy;
4. The employment of MD simulations for the most promising complexes, with regard to their binding free energy, in order to improve the conformational sampling and interactions; and
5. The selection of the final complex acquiring the lowest average binding free energy affinity throughout the MD simulations.

These steps constitute a systematic methodology, and our study suggests that a similar computational framework can be applied for the elucidation of subsequent V3 loop:CXCR4/CCR5 complexes, or more generally, a broader series of ligand-G protein coupled receptor complexes.

The identification of the most highly interacting V3 loop and CXCR4 residues, as well as the specificity of interacting residue pairs accomplished in this study, could provide gateways for the de novo design of novel inhibitors targeting key CXCR4 residues and serving as antagonists against HIV-1 (70). According to the results of recent experiments and computational modeling, the novel HIV-1 inhibitor, DV1, partially meets the above hypothesis as it participates in polar interactions with CXCR4 residues $Asp^{171}$ and $Glu^{288}$ (68), the two most highly interacting polar residues identified in this study. Guided by the findings of the present study, and within this direction, de novo design methods (62,71,72) in conjunction with binding free energy calculations and MD simulations, as in Tamamis et al. (62), can be implemented to design new peptides acquiring augmented interactions with the key CXCR4 residues for HIV-1 coreceptor activity.

We employed computational methods, predominantly comprising MD simulations, as well as binding and interaction free energy calculations, to investigate the molecular recognition of CXCR4 by a dual receptor HIV-1 gp120 V3 loop. The reported complex structure, despite acquiring absolutely no bias toward experimental results, is in a remarkable accordance with experiments. To our knowledge, it is the first reported V3 loop:CXCR4 complex structure to shed light on the functional role of the HIV-1 gp120 V3 loop and chemokine receptor residues, which are experimentally associated with HIV-1 coreceptor activity.

## SUPPORTING MATERIAL

Supporting Material comprises additional information on Methods for Step 2, Step 3, Step 5, coordinates of structures in Complex 1 (extracted every 2 ns), the definition of CXCR4 and V3 Loop Domains, two figures, four tables, and a movie are available at http://www.biophysj.org/biophysj/supplemental/S0006-3495(13)00873-4.

All MD simulations and free energy calculations were performed on SESAME and TIGER computer clusters at the TIGRESS high performance computer center at Princeton University.

We thank the anonymous referees for constructive comments and suggestions.

C.A.F. acknowledges funding from the National Institutes of Health (grant No. R01 GM052032).

## REFERENCES

1. Chan, D. C., D. Fass, …, P. S. Kim. 1997. Core structure of gp41 from the HIV envelope glycoprotein. *Cell.* 89:263–273.
2. Kwong, P. D., R. Wyatt, …, W. A. Hendrickson. 1998. Structure of an HIV gp120 envelope glycoprotein in complex with the CD4 receptor and a neutralizing human antibody. *Nature.* 393:648–659.
3. Kwong, P. D., R. Wyatt, …, W. A. Hendrickson. 2000. Structures of HIV-1 gp120 envelope glycoproteins from laboratory-adapted and primary isolates. *Structure.* 8:1329–1339.
4. Huang, C. C., M. Tang, …, P. D. Kwong. 2005. Structure of a V3-containing HIV-1 gp120 core. *Science.* 310:1025–1028.
5. Huang, C. C., S. N. Lam, …, P. D. Kwong. 2007. Structures of the CCR5 N terminus and of a tyrosine-sulfated antibody with HIV-1 gp120 and CD4. *Science.* 317:1930–1934.
6. Dittmar, M. T., A. McKnight, …, P. Simmonds. 1997. HIV-1 tropism and co-receptor use. *Nature.* 385:495–496.
7. Xiao, L. H., S. M. Owen, …, R. B. Lal. 1998. CCR5 coreceptor usage of non-syncytium-inducing primary HIV-1 is independent of phylogenetically distinct global HIV-1 isolates: delineation of consensus motif in the V3 domain that predicts CCR-5 usage. *Virology.* 240:83–92.
8. Pollakis, G., S. Kang, …, W. A. Paxton. 2001. N-linked glycosylation of the HIV type-1 gp120 envelope glycoprotein as a major determinant of CCR5 and CXCR4 coreceptor utilization. *J. Biol. Chem.* 276: 13433–13441.
9. Cardozo, T., T. Kimura, …, S. Zolla-Pazner. 2007. Structural basis for coreceptor selectivity by the HIV type 1 V3 loop. *AIDS Res. Hum. Retroviruses.* 23:415–426.
10. Edo-Matas, D., K. A. van Dort, …, N. A. Kootstra. 2011. Comparison of in vivo and in vitro evolution of CCR5 to CXCR4 coreceptor use of primary human immunodeficiency virus type 1 variants. *Virology.* 412:269–277.
11. Wu, B., E. Y. T. Chien, …, R. C. Stevens. 2010. Structures of the CXCR4 chemokine GPCR with small-molecule and cyclic peptide antagonists. *Science.* 330:1066–1071.
12. Chan, D. C., and P. S. Kim. 1998. HIV entry and its inhibition. *Cell.* 93:681–684.
13. Hwang, S. S., T. J. Boyle, …, B. R. Cullen. 1991. Identification of the envelope V3 loop as the primary determinant of cell tropism in HIV-1. *Science.* 253:71–74.
14. Weiss, R. A., and P. R. Clapham. 1996. Hot fusion of HIV. *Nature.* 381:647–648.
15. Doranz, B. J., M. J. Orsini, …, R. W. Doms. 1999. Identification of CXCR4 domains that support coreceptor and chemokine receptor functions. *J. Virol.* 73:2752–2761.
16. Dragic, T. 2001. An overview of the determinants of CCR5 and CXCR4 co-receptor function. *J. Gen. Virol.* 82:1807–1814.
17. Choi, W. T., S. Tian, …, Z. Huang. 2005. Unique ligand binding sites on CXCR4 probed by a chemical biology approach: implications for the design of selective human immunodeficiency virus type 1 inhibitors. *J. Virol.* 79:15398–15404.
18. Tian, S., W. T. Choi, …, Z. Huang. 2005. Distinct functional sites for human immunodeficiency virus type 1 and stromal cell-derived factor 1alpha on CXCR4 transmembrane helical domains. *J. Virol.* 79:12667–12673.
19. Brelot, A., N. Heveker, …, M. Alizon. 2000. Identification of residues of CXCR4 critical for human immunodeficiency virus coreceptor and chemokine receptor activities. *J. Biol. Chem.* 275:23736–23744.

20. Chabot, D. J., P. F. Zhang, …, C. C. Broder. 1999. Mutagenesis of CXCR4 identifies important domains for human immunodeficiency virus type 1 X4 isolate envelope-mediated membrane fusion and virus entry and reveals cryptic coreceptor activity for R5 isolates. *J. Virol.* 73:6598–6609.

21. Zhou, N., Z. Luo, …, Z. Huang. 2001. Structural and functional characterization of human CXCR4 as a chemokine receptor and HIV-1 coreceptor by mutagenesis and molecular modeling studies. *J. Biol. Chem.* 276:42826–42833.

22. Chabot, D. J., H. D. Chen, …, C. C. Broder. 2000. N-linked glycosylation of CXCR4 masks coreceptor function for CCR5-dependent human immunodeficiency virus type 1 isolates. *J. Virol.* 74:4404–4413.

23. Brelot, A., N. Heveker, …, M. Alizon. 1999. Effect of mutations in the second extracellular loop of CXCR4 on its utilization by human and feline immunodeficiency viruses. *J. Virol.* 73:2576–2586.

24. Kajumo, F., D. A. Thompson, …, T. Dragic. 2000. Entry of R5X4 and X4 human immunodeficiency virus type 1 strains is mediated by negatively charged and tyrosine residues in the amino-terminal domain and the second extracellular loop of CXCR4. *Virology.* 271:240–247.

25. Lu, Z., J. F. Berson, …, R. W. Doms. 1997. Evolution of HIV-1 coreceptor usage through interactions with distinct CCR5 and CXCR4 domains. *Proc. Natl. Acad. Sci. USA.* 94:6426–6431.

26. Liu, S., S. Fan, and Z. Sun. 2003. Structural and functional characterization of the human CCR5 receptor in complex with HIV gp120 envelope glycoprotein and CD4 receptor by molecular modeling studies. *J. Mol. Model.* 9:329–336.

27. López de Victoria, A., P. Tamamis, …, D. Morikis. 2012. Insights into the structure, correlated motions, and electrostatic properties of two HIV-1 gp120 V3 loops. *PLoS ONE.* 7:e49925.

28. Hartley, O., P. J. Klasse, …, J. P. Moore. 2005. V3: HIV's switch-hitter. *AIDS Res. Hum. Retroviruses.* 21:171–189.

29. Chandramouli, B., G. Chillemi, …, A. Desideri. 2012. Importance of V3 loop flexibility and net charge in the context of co-receptor recognition. A molecular dynamics study on HIV gp120. *J. Biomol. Struct. Dyn.* 29:879–891.

30. Chandramouli, B., G. Chillemi, …, A. Desideri. 2013. Structural dynamics of V3 loop with different electrostatics: implications on co-receptor recognition: a molecular dynamics study of HIV gp120. *J. Biomol. Struct. Dyn.* 31:403–413.

31. López de Victoria, A., C. A. Kieslich, …, D. Morikis. 2012. Clustering of HIV-1 subtypes Based on gp120 V3 loop electrostatic properties. *BMC Biophys.* 5:3.

32. Napier, K. B., Z. X. Wang, …, J. O. Trent. 2007. CCR5 interactions with the variable 3 loop of gp120. *J. Mol. Model.* 13:29–41.

33. Morikis, D., A. K. Rizos, …, E. Krambovitis. 2007. Electrostatic modeling of peptides derived from the V3-loop of HIV-1 gp120: implications of the interaction with chemokine receptor CCR5. *Int. J. Mol. Med.* 19:343–351.

34. Xiang, S. H., B. Pacheco, …, J. Sodroski. 2013. Characterization of a dual-tropic human immunodeficiency virus (HIV-1) strain derived from the prototypical X4 isolate HXBc2. *Virology.* 438:5–13.

35. Pan, Y., B. Ma, …, R. Nussinov. 2004. Characterization of the conformational state and flexibility of HIV-1 glycoprotein gp120 core domain. *J. Biol. Chem.* 279:30523–30530.

36. Pan, Y., B. Ma, and R. Nussinov. 2005. CD4 binding partially locks the bridging sheet in gp120 but leaves the $\beta 2/3$ strands flexible. *J. Mol. Biol.* 350:514–527.

37. Liao, H. X., R. Lynch, …, B. F. Haynes. 2013. Co-evolution of a broadly neutralizing HIV-1 antibody and founder virus. *Nature.* 496:469–476. http://dx.doi.org/10.1038/nature12053.

38. Li, W., and A. Godzik. 2006. CD-HIT: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics.* 22:1658–1659.

39. Veldkamp, C. T., C. Seibert, …, B. F. Volkman. 2008. Structural basis of CXCR4 sulfotyrosine recognition by the chemokine SDF-1/CXCL12. *Sci. Signal.* 1:ra4.

40. Choi, Y., and C. M. Deane. 2010. FREAD revisited: accurate loop structure prediction using a database search algorithm. *Proteins.* 78:1431–1440.

41. Roy, A., A. Kucukural, and Y. Zhang. 2010. I-TASSER: a unified platform for automated protein structure and function prediction. *Nat. Protoc.* 5:725–738.

42. Hansmann, U. 1997. Parallel tempering algorithm for conformational studies of biological molecules. *Chem. Phys. Lett.* 281:140–150.

43. Sugita, Y., and Y. Okamoto. 1999. Replica-exchange molecular dynamics method for protein folding. *Chem. Phys. Lett.* 314:141–151.

44. Nymeyer, H., S. Gnanakaran, and A. E. García. 2004. Atomic simulations of protein folding, using the replica exchange algorithm. *Methods Enzymol.* 383:119–149.

45. Pieridou, G., C. Avgousti-Menelaou, …, S. C. Hayes. 2011. UV resonance Raman study of TTR(105–115) structural evolution as a function of temperature. *J. Phys. Chem. B.* 115:4088–4098.

46. Haberthür, U., and A. J. Caflisch. 2008. FACTS: fast analytical continuum treatment of solvation. *J. Comput. Chem.* 29:701–715.

47. Tamamis, P., L. Adler-Abramovich, …, G. Archontis. 2009. Self-assembly of phenylalanine oligopeptides: insights from experiments and simulations. *Biophys. J.* 96:5020–5029.

48. Tamamis, P., E. Kasotakis, …, G. Archontis. 2009. Amyloid-like self-assembly of peptide sequences from the adenovirus fiber shaft: insights from molecular dynamics simulations. *J. Phys. Chem. B.* 113:15639–15647.

49. Tamamis, P., and G. Archontis. 2011. Amyloid-like self-assembly of a dodecapeptide sequence from the adenovirus fiber shaft: perspectives from molecular dynamics simulations. *J. Non-Cryst. Solids.* 357:717–722.

50. Brooks, B. R., C. L. Brooks, 3rd, …, M. J. Karplus. 2009. CHARMM: the biomolecular simulation program. *J. Comput. Chem.* 30:1545–1614.

51. Comeau, S. R., D. W. Gatchell, …, C. J. Camacho. 2004. CLUSPRO: an automated docking and discrimination method for the prediction of protein complexes. *Bioinformatics.* 20:45–50.

52. Im, W., M. S. Lee, and C. L. Brooks, 3rd. 2003. Generalized Born model with a simple smoothing function. *J. Comput. Chem.* 24:1691–1702.

53. Im, W., M. Feig, and C. L. Brooks, 3rd. 2003. An implicit membrane generalized born theory for the study of structure, stability, and interactions of membrane proteins. *Biophys. J.* 85:2900–2918.

54. Michino, M., J. Chen, …, C. L. Brooks, 3rd. 2010. FoldGPCR: structure prediction protocol for the transmembrane domain of G protein-coupled receptors from class A. *Proteins.* 78:2189–2201.

55. Seeber, M., M. Cecchini, …, A. Caflisch. 2007. WORDOM: a program for efficient analysis of molecular dynamics simulations. *Bioinformatics.* 23:2625–2627.

56. Pierce, B. G., Y. Hourai, and Z. Weng. 2011. Accelerating protein docking in ZDOCK using an advanced 3D convolution library. *PLoS ONE.* 6:e24657.

57. Massova, I., and P. A. Kollman. 2000. Combined molecular mechanical and continuum solvent approach (MM-PBSA/GBSA) to predict ligand binding. *Perspect. Drug Discov.* 18:113–135.

58. Gilson, M. K., and H. X. Zhou. 2007. Calculation of protein-ligand binding affinities. *Annu. Rev. Biophys. Biomed.* 36:21–42.

59. Zoete, V., M. B. Irving, and O. Michielin. 2010. MM-GBSA binding free energy decomposition and T cell receptor engineering. *J. Mol. Recognit.* 23:142–152.

60. Gohlke, H., and D. A. Case. 2004. Converging free energy estimates: MM-PB(GB)SA studies on the protein-protein complex Ras-Raf. *J. Comput. Chem.* 25:238–250.

61. Page, C. S., and P. A. Bates. 2006. Can MM-PBSA calculations predict the specificities of protein kinase inhibitors? *J. Comput. Chem.* 27:1990–2007.

62. Tamamis, P., A. López de Victoria, …, G. Archontis. 2012. Molecular dynamics in drug design: new generations of compstatin analogs. *Chem. Biol. Drug Des.* 79:703–718.

63. Tamamis, P., D. Morikis, …, G. Archontis. 2010. Species specificity of the complement inhibitor compstatin investigated by all-atom molecular dynamics simulations. *Proteins.* 78:2655–2667.

64. Tamamis, P., P. Pierou, …, G. Archontis. 2011. Design of a modified mouse protein with ligand binding properties of its human analog by molecular dynamics simulations: the case of C3 inhibition by compstatin. *Proteins.* 79:3166–3179.

65. Humphrey, W., A. Dalke, and K. Schulten. 1996. VMD: visual molecular dynamics. *J. Mol. Graph.* 14:33–38, 27–28.

66. Oberlin, E., A. Amara, …, B. Moser. 1996. The CXC chemokine SDF-1 is the ligand for LESTR/fusin and prevents infection by T-cell-line-adapted HIV-1. *Nature.* 382:833–835.

67. Bleul, C. C., M. Farzan, …, T. A. Springer. 1996. The lymphocyte chemoattractant SDF-1 is a ligand for LESTR/fusin and blocks HIV-1 entry. *Nature.* 382:829–833.

68. Choi, W. T., S. Kumar, …, J. An. 2012. A novel synthetic bivalent ligand to probe chemokine receptor CXCR4 dimerization and inhibit HIV-1 entry. *Biochemistry.* 51:7078–7086.

69. Bozek, K., T. Lengauer, …, F. S. Domingues. 2013. Analysis of physicochemical and structural properties determining HIV-1 coreceptor usage. *PLOS Comput. Biol.* 9:e1002977.

70. Cocchi, F., A. L. DeVico, …, P. Lusso. 1996. The V3 domain of the HIV-1 gp120 envelope glycoprotein is critical for chemokine-mediated blockade of infection. *Nat. Med.* 2:1244–1247.

71. Bellows, M. L., M. S. Taylor, …, C. A. Floudas. 2010. Discovery of entry inhibitors for HIV-1 via a new de novo protein design framework. *Biophys. J.* 99:3445–3453.

72. Bellows-Peterson, M. L., H. K. Fung, …, T. M. Woodruff. 2012. De novo peptide design with C3a receptor agonist and antagonist activities: theoretical predictions and experimental validation. *J. Med. Chem.* 55:4159–4168.