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.

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

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

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_{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 ΔG 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}^{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 β-sheets among the following residue moieties, 12–13:22–21 and 3–9:32–26, are observed in the trajectory. A β-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 β-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/

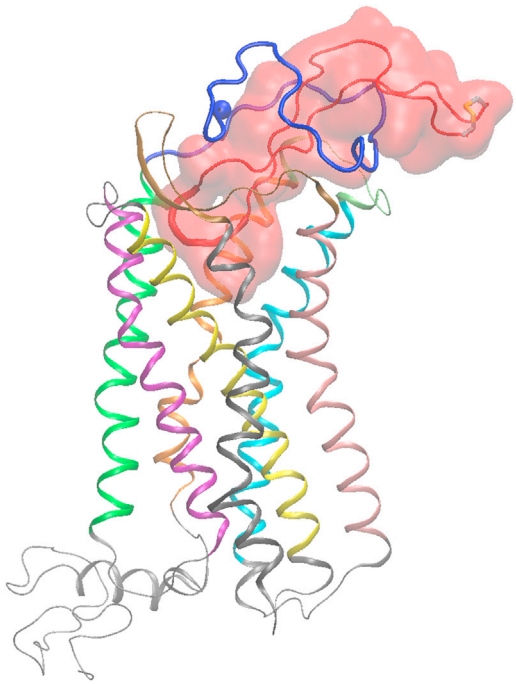


FIGURE 1 Molecular graphics image of the entire simulation system corresponding to the complex with the lowest average binding free energy. The V3 loop is shown in tube and transparent surface representation in red color, and its 16–20 residue moiety is shown in thick tube representation. The CXCR4 is 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}

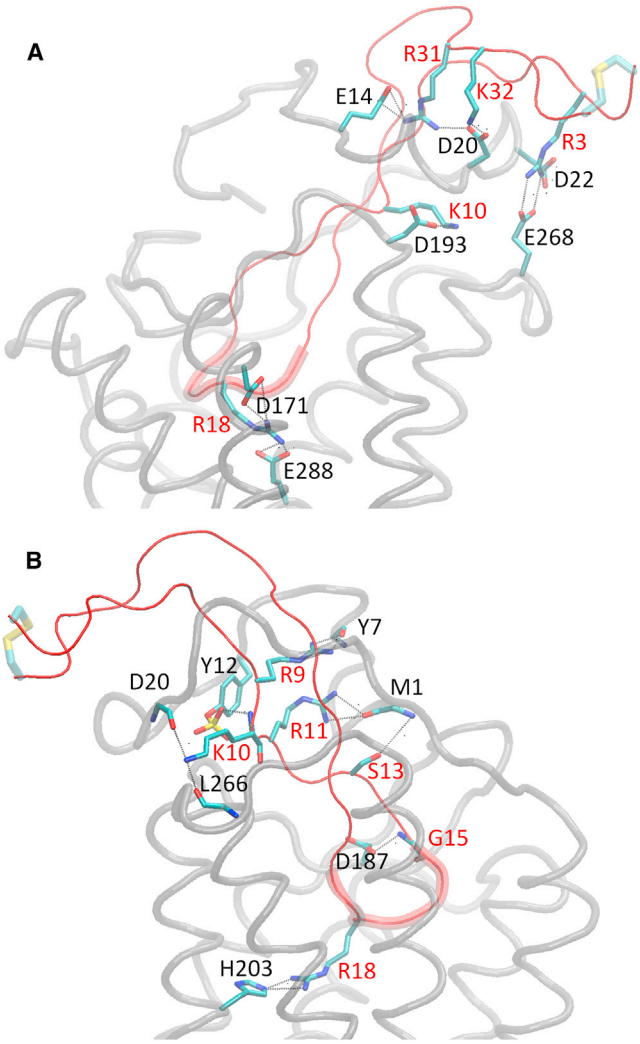


FIGURE 2 Molecular graphics images of important polar interactions corresponding to the complex with the lowest average binding free energy. Panels (*A*) and (*B*) depict the salt bridges and the most important hydrogen bonds, respectively, according to residue pair-wise interaction free energies. Panel (*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 β-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 (SDF1α) binding, because SDF1α 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 SDF1α 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 β-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 π-π 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, SDF1α, 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 SDF1α, 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.

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

Supporting Material

Phanourios Tamamis and Christodoulos A. Floudas

Department of Chemical and Biological Engineering, Princeton University, New Jersey

Supporting Text

Supporting Material on Methods, Step2 :

V3 loop: We employed 16 replicas with the following temperatures: 288, 300, 312, 325, 338, 352, 366, 381, 396, 412, 428, 445, 462, 480, 498, 517 K. Replica exchanges were attempted at 10-ps intervals and the total simulation length for all temperatures was equal to 0.8μs. We extracted 5000 snapshots (conformations), at 10-ps intervals, from the 50 ns replica simulations at 300 K. The peptide dielectric constant was set to $\varepsilon=2$ and a non-polar solvation energy surface term was included, with a surface-tension coefficient of 0.015 kcal/mol*Å^2; in addition, 7.5-Å cutoff was used for the non-bonded interactions in conjunction with the FACTS19 parametrization[1]. The peptide atomic charges, van der Waals and stereochemical parameters were taken from the CHARMM19 all-atom force field [2] in consistency with the FACTS 19 parametrization[1]. The equations of motion were integrated by the leap-frog algorithm, with a 2.0-fs timestep. Bonds involving hydrogen atoms were constrained to standard values with the SHAKE algorithm [3]. The temperature was controlled by the Langevin method; the friction coefficients were set to 5.0 ps^{-1} for heavy atoms and 0 ps^{-1} for hydrogen atoms.

CXCR4: Standard amino acids were described by the CHARMM27 all-atom topology and energy function [4] including the CMAP correction [5]. Sulfated tyrosine residues 7, 12, 21, denotes as "Tys", were parametrized using a combination of the standard tyrosine parameters and the methylsulfate parameters presented in CHARMM36 CGENFF force field [6]. A 16-Å cutoff distance was used for non-bonded interactions. The lengths of covalent bonds containing hydrogen atoms were constrained by the *SHAKE* algorithm [3], and the equations of motion were solved with an integration time step of 2.0 fs. The system was in contact with a Langevin heat bath at 300 K during the production and equilibration runs, and in addition, a friction coefficient of 5 ps^{-1} was used for heavy atoms. Prior to the production runs, four heating steps of total duration 400 ps were performed, and in addition, an equilibration procedure of a total duration of 1.7 ns was performed, during which the harmonic restraints were gradually removed. The production runs were performed at 300K with a total duration of 5 ns. The difference between the two independent simulations per complex was on the restraints imposed during the production runs: in the first simulation, no restraints were imposed on the system, whereas in the

second simulation, a weak harmonic force constant of 1 kcal/mol*A^2 was applied to the Cα atoms using the bestfit module in CHARMM [8].

Supporting Material on Methods, Step3 :

We merged the structures produced in the V3 loop replica exchange MD simulations in a single trajectory containing 5000 V3 loop structures. In addition, we merged the CXCR4 structures produced from the six independent aforesaid MD simulations in single trajectory containing 1500 CXCR4 structures. We employed the quality clustering method of WORDOM [7], to cluster independently the V3 loop and CXCR4 structures based on their Cα coordinates. For the V3 loop clustering, we included all Cα atoms and used a relatively large clustering radius of 2.0 Å, owing to the high flexibility of the V3 loop structures. We extracted the 20 most populated clusters for the V3 loop, with the smallest cluster acquiring a 0.09% percentage probability. For the CXCR4 clustering, we considered all Cα atoms with a z-coordinate value greater than 0 Å. The cluster analysis produced 17 clusters for CXCR4, including the initially modeled CXCR4 structure.

Supporting Material on Methods, Step5 :

In the PB calculations, we inserted the complex and protein in a membrane slab with a thickness of 31 Å (31≈36.0-2*(2.5), where 36.0 Å and 2.5 Å correspond to the membrane thickness and the half of membrane switching length, parameters used in the GBSW approximation). The membrane slab used in the PBSA calculations possessed a dielectric constant of 2, and was surrounded by water, with a dielectric constant of 80. The protein dielectric constant was set to 2. The PBSA calculations were performed with the Poisson-Boltzmann solver of the CHARMM program (PBEQ module) [8]. We used 250 grid points in each direction and a grid-spacing of 0.5 Å.

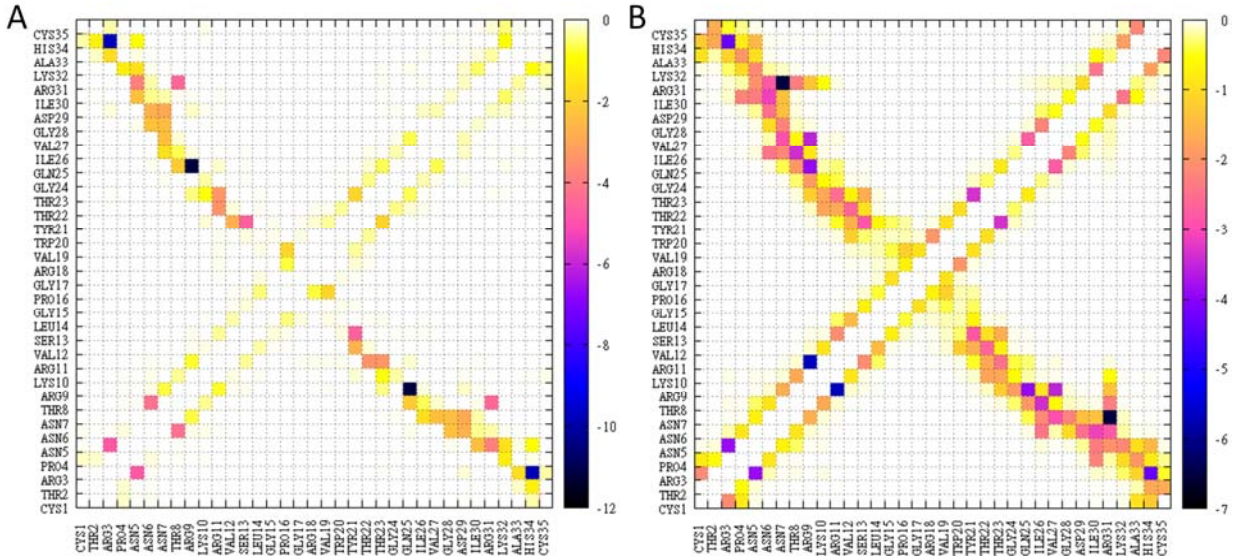
It is worth noting that in the initial set of 680,000 complexes, the docking of V3 loop performed by Zdock [9] was not in all cases in accordance with the experimentally defined binding pocket of CXCR4 [10]. Nevertheless, the binding free energy calculations showed that the binding free energies of the complex structures at which the V3 loop was outside of the binding pocket, was evidently higher compared to the binding free energy of the 17 complexes selected.

CXCR4 and V3 loop domains:

CXCR4 residues, denoted in brackets, can be approximately divided into the following domains: (i) N-terminal domain [1:37], (ii) Intramembrane helix 1 [38:65]; (iii) Intracellular loop 1 [66:71]; (iv) Intramembrane helix 2 [72:100], (v) Extracellular loop 1 (ECL1) [101:106]; (vi) Intramembrane helix 3 [107:137]; (vii) Intracellular loop 2 [138:146]; (viii) Intramembrane helix 4 [147:174]; (ix) Extracellular loop 2 (ECL2) [175:194]; (x) Intramembrane helix 5 [195:224]; (xi) Intracellular loop 3 [225:236]; (xii) Intramembrane helix 6 [237:265]; (xiii) Extracellular loop 3 (ECL3) [266:276]; (xiv) Intramembrane helix 7 [277:303]; (xv) C-terminal domain [304:352]. Regarding the V3 loop, the structure can be approximately divided into the following regions: (i) The base domain [1:4 and 31:35]; (ii) Stems [5:12 and 23:30] (iii) Tip [13:22]. According to the results of this study, the core region of the tip comprises residues 16-20.

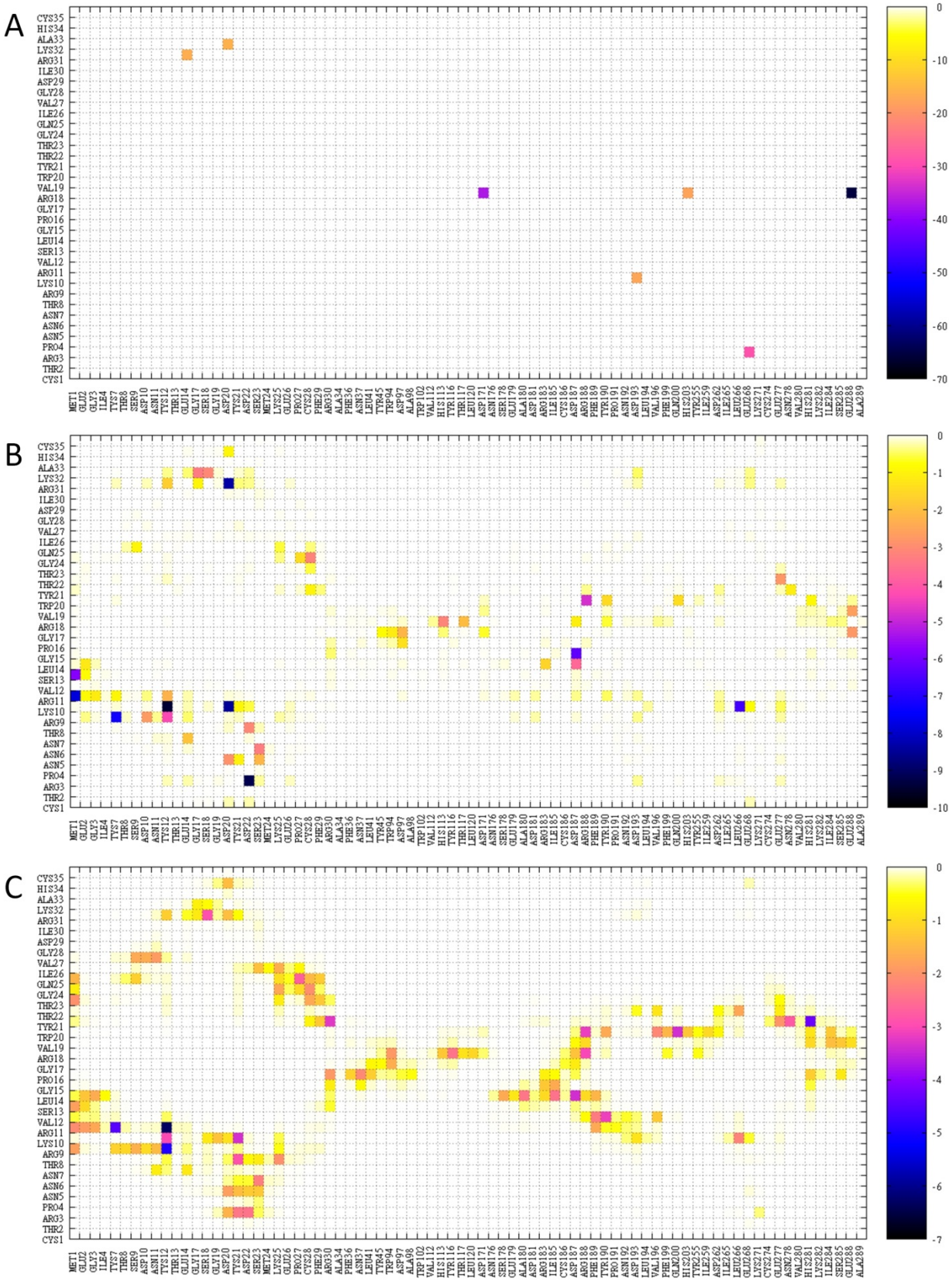
Supporting Figures

Supporting Figure 1:



Supporting Figure 1: Two dimensional density maps depicting the favorable (negative) average MM GBSA interaction free-energies for intramolecular V3 loop interacting residue pairs, within the simulation of the complex with the lowest average binding free energy. The left (A) and right (B) panels correspond, respectively, to polar and non-polar interactions. All energies are in kcal/mol. The color – interaction free energy correspondence is shown by the palette on the right-hand side of each panel. Interactions free energies were not calculated for pairs of covalently bonded residues. All values have been computed by analysis of 1000 snapshots, extracted from the 20-ns simulation of complex 1 at 20-ps intervals.

Supporting Figure 2:



Supporting Figure 2: Two dimensional density maps depicting the favorable (negative) average MM GBSA interaction free-energies for intermolecular V3 loop (y axis) : CXCR4 (x axis) interacting residue pairs, within the simulation of the complex with the lowest average binding free energy. The upper (A), middle (B) and bottom (C) panels correspond, respectively, to highly interacting polar [-70 kcal/mol : -10 kcal/mol], moderate interacting polar [-10 kcal/mol : 0 kcal/mol], and non-polar interactions. All energies are in kcal/mol. The color – interaction free energy correspondence is shown by the palette on the right-hand side of each panel. All values have been computed by analysis of 1000 snapshots, extracted from the 20-ns simulation of complex 1 at 20-ps intervals.

Supporting Tables

Supporting Table 1:

Binding free energies using the PBSA, MM GBSA and MM PBSA approximations for V3 loop : CXCR4 complexes in kcal/mol. All values have been computed by analysis of 1000 snapshots, extracted from the 20-ns simulation (of all complexes) at 20-ps intervals.

	Docking and Minimization[¶]			MD Simulations[†]				
	PBSA			MM PB/GBSA (Non-Polar)[‡]	MM GBSA[ǂ] (Polar and Total)		MM PBSA[#] (Polar and Total)	
	Non-Polar	Polar	Total	Non- polar	Polar	Total	Polar	Total
1[§]	-134.3	-10.0	-144.2	-269.0 (10.2)	-157.9 (19.8)	**-427.0 (22.4)**	-98.1 (11.1)	**-367.2 (12.9)**
2	-124.0	-19.8	-143.8	-230.7 (19.9)	-99.3 (11.4)	-330.1 (22.1)	-47.8 (16.1)	-278.5 (15.5)
3	-98.9	-44.7	-143.5	-210.8 (16.1)	-123.8 (13.4)	-334.6 (19.6)	-105.5 (9.3)	-316.3 (16.3)
4	-130.4	-12.2	-142.6	-269.3 (12.3)	-124.6 (13.8)	-393.8 (13.9)	-86.2 (10.5)	-355.5 (12.4)
5	-138.3	-4.0	-142.3	-225.5 (11.5)	-111.3 (22.0)	-336.9 (28.2)	-49.7 (9.8)	-275.2 (10.5)
6	-121.3	-19.9	-141.2	-225.7 (10.9)	-122.4 (14.7)	-348.2 (14.8)	-73.0 (9.1)	-298.7 (12.3)
7	-94.3	-46.6	-140.9	-229.2 (15.2)	-102.7 (16.7)	-331.9 (21.7)	-85.9 (10.2)	-315.1 (15.6)
8	-99.4	-39.5	-138.9	-231.8 (14.3)	-96.3 (15.2)	-328.2 (16.3)	-43.9 (15.4)	-275.8 (12.6)
9	-108.4	-28.4	-136.8	-233.6 (16.6)	-105.1 (13.3)	-338.6 (17.8)	-80.9 (12.5)	-314.4 (19.0)
10	-138.1	1.6	-136.5	-223.1 (10.0)	-62.3 (15.9)	-285.4 (19.5)	-25.8 (10.5)	-248.9 (13.6)
11	-119.6	-16.6	-136.2	-253.1 (10.4)	-107.3 (16.8)	-360.5 (16.0)	-80.1 (12.0)	-333.3 (13.9)
12	-123.5	-12.3	-135.8	-245.0 (12.4)	-134.9 (14.4)	-379.9 (16.2)	-109.6 (13.1)	-354.6 (16.2)
13	-137.8	2.0	-135.8	-235.0 (13.5)	-142.3 (18.3)	-377.3 (23.8)	-63.3 (21.0)	-298.2 (16.0)
14	-98.2	-37.3	-135.5	-217.2 (14.3)	-142.4 (16.8)	-359.6 (23.7)	-102.5 (10.4)	-319.7 (18.6)
15	-125.4	-9.9	-135.4	-239.7 (8.8)	-112.5 (15.8)	-352.2 (16.4)	-67.5 (8.7)	-307.2 (11.0)
16	-136.6	2.0	-134.6	-246.9 (12.8)	-109.1 (14.7)	-356 (18.0)	-49.1 (9.7)	-296.0 (15.1)
17	-136.5	2.5	-134.1	-233.0 (12.3)	-113.4 (14.3)	-346.4 (20.0)	-59.9 (9.8)	-292.9 (10.9)

The binding free energies are calculated as described in Methods. The numbering[§] (1-17) is sorted according to the total binding free energy, using the PBSA approximation, of the docked complexes after minimization, referred as step 5[¶]; the total binding free energy is the sum of non-polar and polar contributions of step 5. The MD simulation[†] results correspond the average binding free energies of the 17 complexes based on 20-ns MD simulation runs, referred as step 7, and were calculated both using MM PBSA[ǂ] and MM GBSA[#] methods.; the standard deviation is shown in parentheses. The non-polar component[‡] is the same in MM PBSA[ǂ] and MM GBSA[#]. For each calculation, MM GBSA[ǂ] or MM PBSA[#], the polar component is calculated using a different approach, thus, the total binding free energy, which is the sum of polar and non-polar

components for each approach is different, as well. According to both methods, the average total binding free energy of the first complex, marked in **bold** face, is the lowest.

Supporting Table 2:

Average and standard deviation values of Root Mean Square Deviation (RMSD) between the simulation coordinates of main-chain heavy atoms (N, Cα, and C) and their corresponding coordinates from the first simulation frame after equilibration. All values have been computed by analysis of 1000 snapshots (per complex), extracted from the 20-ns simulations at 20-ps intervals.

Complex	Protein Intramembrane		Protein N-terminal, residues 1 -37		V3 loop		V3 loop, residues: 8-26	
	Average	St. Deviation	Average	St. Deviation	Average	St. Deviation	Average	St. Deviation
1	1.31	0.03	3.65	0.29	3.86	0.27	2.9	0.14
2	1.09	0.01	3.34	0.28	4.99	0.91	3.57	0.31
3	1.32	0.07	4.5	0.65	5.61	0.75	3.15	1.44
4	1.16	0.02	2.92	0.1	2.12	0.16	1.56	0.04
5	0.93	0.01	2.06	0.11	5.14	0.36	1.79	0.07
6	1.11	0.02	3.3	0.4	3.5	0.29	2.77	0.18
7	1.33	0.03	7.43	0.59	7.02	1.76	3.04	0.42
8	1.19	0.02	2.96	0.25	8.01	0.86	2.52	0.15
9	1.17	0.01	3.96	0.35	6.8	0.48	5	0.29
10	0.95	0.02	3.51	0.45	4.34	2.41	1.71	0.15
11	1.06	0.01	3.76	0.15	4.48	0.24	3.53	0.2
12	1.09	0.02	2.39	0.1	4.27	0.51	2.73	0.09
13	1.22	0.02	3.78	0.43	5.42	3.69	1.58	0.09
14	1.29	0.05	4.63	0.39	5.27	0.47	4.62	0.52
15	1.04	0.01	2.99	0.14	5.43	1.29	1.78	0.08
16	1.16	0.01	2.04	0.07	8.05	2.36	1.32	0.04
17	1.09	0.01	1.87	0.05	3.82	0.16	2.23	0.07

The coordinates of the simulated systems are aligned with regard to the intramembrane backbone CXCR4 atoms of the first simulation frame. All values corresponding to protein N-terminal, V3 loop and V3 loop residues 8-26, are computed without any additional rotation/translation and are averaged over the 20 ns of the simulation trajectories. All values are reported in Å.

Supporting Table 3:

Percentage (%) hydrogen bond occupancy of important intermolecular hydrogen-bonding atom pairs within the 17 simulated complexes. Hydrogen bonding atom pairs with less than 10% are not reported. We present the hydrogen bond occupancies for each complex in a separate table and sort the hydrogen bonding atom pairs, firstly, with respect to the residue number of the V3 loop atom, and secondly, the residue number of the CXCR4 atom. All values have been computed by analysis of 1000 snapshots (per complex), extracted from the 20-ns simulations at 20-ps intervals. A hydrogen bond was present if the donor (D)–acceptor (A) distance was less than 3.5 Å and the corresponding angle (D–H ··· A) was larger than 90°. Hydrogen bond interactions associated with salt-bridge formation are highlighted in grey background.

Complex 1:

V3 loop Residue	Number	Group	Atom	CXCR4 Residue	Number	Group	Atom	Occupancy (%)
ARG	3	Side	NH2	ASP	22	Side	OD1	40.86
ARG	3	Side	NH2	ASP	22	Side	OD2	40.66
ARG	3	Side	NH2	GLU	268	Side	OE2	87.81
ARG	3	Side	NH1	GLU	268	Side	OE1	87.21
ARG	3	Side	NH2	GLU	268	Side	OE1	64.54
ARG	3	Side	NH1	GLU	268	Side	OE2	60.94
ASN	5	Side	ND2	ASP	20	Side	OD1	52.75
ASN	5	Side	ND2	ASP	20	Side	OD2	12.79
ASN	5	Side	ND2	TYS	21	Main	O	23.08
ASN	6	Main	N	SER	23	Side	OG	68.73
ASN	6	Side	OD1	SER	23	Side	OG	18.08
ASN	6	Side	ND2	SER	23	Side	OG	9.69
ASN	7	Side	OD1	ASN	11	Side	ND2	11.39
ASN	7	Side	ND2	GLU	14	Side	OE1	35.16
ASN	7	Side	ND2	GLU	14	Side	OE2	18.98
THR	8	Side	OG1	ASP	22	Main	O	61.04
ARG	9	Side	NH1	TYS	7	Main	O	84.62
ARG	9	Side	NH2	TYS	7	Main	O	57.14
ARG	9	Side	NH1	ASP	10	Main	O	65.13
LYS	10	Main	N	TYS	12	Side	OS3	85.71
LYS	10	Main	N	TYS	12	Side	OS2	11.09
LYS	10	Side	NZ	ASP	20	Main	O	95.3
LYS	10	Side	NZ	ASP	193	Side	OD1	97.8
LYS	10	Side	NZ	LEU	266	Main	O	94.21
ARG	11	Side	NH1	MET	1	Main	O	92.61
ARG	11	Side	NH2	MET	1	Main	O	48.05
ARG	11	Side	NE	MET	1	Main	O	13.69
ARG	11	Side	NH1	GLU	2	Main	O	10.79

ARG	11	Main	N	TYS	12	Side	OS3	23.88
SER	13	Side	OG	MET	1	Main	N	51.95
LEU	14	Main	N	GLU	2	Main	O	24.88
LEU	14	Main	O	ARG	183	Side	NH1	41.26
LEU	14	Main	O	ARG	183	Side	NH2	32.57
GLY	15	Main	N	ASP	187	Side	OD1	56.64
GLY	15	Main	N	ASP	187	Side	OD2	42.06
PRO	16	Main	N	ASN	37	Side	ND2	11.89
GLY	17	Main	N	TYR	45	Side	OH	11.79
GLY	17	Main	N	ASP	97	Side	OD2	25.47
ARG	18	Side	NH1	THR	117	Side	OG1	84.12
ARG	18	Side	NH1	THR	117	Side	OG1	29.47
ARG	18	Side	NH1	ASP	171	Side	OD1	69.63
ARG	18	Side	NH1	ASP	171	Side	OD2	55.04
ARG	18	Main	O	ARG	188	Side	NH2	99.1
ARG	18	Main	O	ARG	188	Side	NE	93.61
ARG	18	Side	NH2	HIS	203	Side	NE2	94.71
ARG	18	Side	NH1	HIS	203	Side	NE2	88.51
ARG	18	Side	NH2	TYR	255	Side	OH	71.23
ARG	18	Side	NH2	GLU	288	Side	OE2	99.4
ARG	18	Side	NE	GLU	288	Side	OE1	94.51
ARG	18	Side	NE	GLU	288	Side	OE2	62.34
ARG	18	Side	NH2	GLU	288	Side	OE1	34.17
TRP	20	Side	NE1	ARG	188	Side	NH2	53.95
TRP	20	Side	NE1	ARG	188	Side	NH1	10.69
TRP	20	Side	NE1	TYR	190	Side	OH	52.85
TRP	20	Main	O	HIS	281	Side	ND1	21.18
TYR	21	Side	OH	ASN	278	Side	ND2	52.95
TYR	21	Side	OH	ASN	278	Side	ND2	33.17
TYR	21	Side	OH	ASN	278	Side	OD1	21.58
THR	22	Side	OG1	GLU	277	Side	OE1	27.17
THR	23	Side	OG1	CYS	28	Main	O	18.08
GLY	24	Main	O	CYS	28	Main	N	94.61
GLY	24	Main	N	CYS	28	Main	O	15.58
GLN	25	Side	NE2	CYS	28	Main	O	11.59
ARG	31	Side	NE	GLU	14	Side	OE1	71.43
ARG	31	Side	NH2	GLU	14	Side	OE1	67.33
ARG	31	Side	NE	GLU	14	Side	OE2	57.84
ARG	31	Side	NH2	GLU	14	Side	OE2	30.17
ARG	31	Side	NH1	SER	18	Side	OG	12.69
ARG	31	Side	NH1	ASP	20	Side	OD1	47.25
LYS	32	Side	NZ	GLY	17	Main	O	53.55
LYS	32	Side	NZ	SER	18	Main	O	52.45
LYS	32	Side	NZ	ASP	20	Side	OD2	84.12
LYS	32	Side	NZ	ASP	20	Side	OD1	46.65
HIS	34	Side	ND1	ASP	20	Side	OD2	14.19

Complex 2:

V3 loop Residue	Number	Group	Atom	CXCR4 Residue	Number	Group	Atom	Occupancy (%)
ARG	3	Side	NH2	THR	13	Main	O	20.78
ARG	3	Side	NH1	GLU	14	Side	OE1	50.45
ARG	3	Side	NH2	GLU	14	Side	OE1	44.66
ARG	3	Side	NH2	GLU	14	Side	OE2	41.06
ARG	3	Side	NH1	GLU	14	Side	OE2	36.66
ASN	5	Side	ND2	ASN	11	Main	O	16.48
ASN	5	Side	OD1	ASN	11	Side	ND2	27.7
ASN	5	Side	ND2	TYS	12	Main	O	52
ASN	6	Main	N	TYS	12	Main	O	21.58
ASN	6	Main	O	TYS	12	Main	N	13.79
ASN	6	Side	ND2	TYS	12	Main	O	10.99
ASN	6	Side	ND2	GLU	14	Side	OE2	16.88
ASN	6	Side	ND2	GLU	14	Side	OE1	14.19
THR	8	Main	O	THR	8	Side	OG1	19.78
THR	8	Side	OG1	SER	9	Main	N	13.29
THR	8	Side	OG1	ASP	10	Main	N	29.67
THR	8	Side	OG1	ASP	10	Side	OD1	47.25
THR	8	Main	N	ASN	11	Side	OD1	12.19
THR	8	Side	OG1	ASN	11	Main	N	13.39
ARG	9	Side	NH2	ASP	22	Main	O	28.07
ARG	9	Side	NH2	SER	23	Main	O	13.39
ARG	9	Side	NH1	MET	24	Main	O	17.28
ARG	9	Side	NH2	MET	24	Main	O	12.09
ARG	9	Side	NH2	LYS	25	Main	O	62.14
ARG	9	Side	NH1	LYS	25	Main	O	44.66
ARG	9	Side	NE	LYS	25	Main	O	14.59
LYS	10	Side	NZ	GLU	2	Side	OE2	27.57
LYS	10	Side	NZ	GLU	2	Side	OE1	22.98
LYS	10	Side	NZ	ILE	6	Main	O	17.58
ARG	11	Side	NH1	TYS	21	Side	OS3	43.46
ARG	11	Side	NH1	TYS	21	Side	OS2	42.16
ARG	11	Side	NH2	TYS	21	Side	OS2	29.27
ARG	11	Side	NH2	TYS	21	Side	OS3	25.37
ARG	11	Side	NH1	TYS	21	Side	OS4	14.39
ARG	11	Side	NH2	TYS	21	Side	OS4	9.89
ARG	11	Side	NH1	GLU	277	Side	OE1	81.72
ARG	11	Side	NH1	GLU	277	Side	OE2	62.54
ARG	11	Side	NE	GLU	277	Side	OE1	14.29
VAL	12	Main	N	CYS	28	Main	O	80.32
LEU	14	Main	N	GLU	277	Side	OE2	40.66
LEU	14	Main	N	GLU	277	Side	OE1	22.28
GLY	15	Main	O	HIS	281	Side	ND1	84.12
GLY	17	Main	O	ASN	37	Side	ND2	66.13
GLY	17	Main	O	ARG	183	Side	NH2	46.35
GLY	17	Main	N	SER	285	Side	OG	75.02

GLY	17	Main	N	GLU	288	Side	OE2	19.37
ARG	18	Side	NE	TYR	45	Side	OH	12.29
ARG	18	Side	NH1	TRP	94	Side	NE1	24.68
ARG	18	Side	NH2	TYR	116	Side	OH	95
ARG	18	Main	O	ARG	183	Side	NH1	72.83
ARG	18	Main	O	ARG	183	Side	NH2	26.37
ARG	18	Side	NH2	TYR	255	Side	OH	55.14
ARG	18	Side	NH2	TYR	255	Side	OH	54.25
ARG	18	Side	NE	GLU	288	Side	OE1	99.7
ARG	18	Side	NH2	GLU	288	Side	OE2	99.5
ARG	18	Side	NE	GLU	288	Side	OE2	74.13
ARG	18	Side	NH2	GLU	288	Side	OE1	30.17
TRP	20	Side	NE1	ARG	183	Side	NH2	37.16
TRP	20	Main	N	ASP	187	Side	OD1	10.39
TRP	20	Side	NE1	HIS	281	Side	NE2	14.59
TYR	21	Side	OH	ARG	188	Side	NH2	45.05
TYR	21	Side	OH	TYR	190	Side	OH	10.79
TYR	21	Side	OH	GLN	200	Side	OE1	89.41
THR	22	Side	OG1	SER	178	Side	OG	70.53
THR	22	Main	N	ASP	187	Side	OD2	47.15
THR	22	Main	N	ASP	187	Side	OD1	29.27
THR	22	Side	OG1	ASP	187	Side	OD1	45.75
THR	22	Side	OG1	ASP	187	Side	OD2	44.36
THR	23	Side	OG1	ARG	188	Main	O	32.67
THR	23	Side	OG1	TYR	190	Main	N	69.43
GLN	25	Side	NE2	LYS	25	Main	O	16.08
GLN	25	Side	NE2	ASP	193	Side	OD1	10.29
ILE	26	Main	N	ASP	193	Side	OD2	26.07
ILE	26	Main	N	ASP	193	Side	OD1	18.68
GLY	28	Main	N	ASP	22	Main	O	52.05
GLY	28	Main	N	SER	23	Main	O	30.07
GLY	28	Main	N	SER	23	Side	OG	18.28
ASP	29	Main	N	SER	23	Main	O	33.97
ASP	29	Main	N	SER	23	Side	OG	15.58
ARG	31	Side	NH2	SER	23	Main	O	20.78
LYS	32	Side	NZ	GLU	14	Side	OE1	10.69

Complex 3:

V3 loop Residue	Number	Group	Atom	CXCR4 Residue	Number	Group	Atom	Occupancy (%)
ARG	3.00	Side	NH1	ASP	22.00	Main	O	43.36
ARG	3.00	Side	NH2	ASP	22.00	Side	OD2	99.30
ARG	3.00	Side	NH1	ASP	22.00	Side	OD1	95.70
ARG	3.00	Side	NH2	ASP	22.00	Side	OD1	77.42
ARG	3.00	Side	NH1	ASP	22.00	Side	OD2	22.78
ARG	3.00	Side	NH1	GLU	268.00	Side	OE1	81.62
ARG	3.00	Side	NE	GLU	268.00	Side	OE1	80.62
ARG	3.00	Side	NH1	GLU	268.00	Side	OE2	80.32
ARG	3.00	Side	NE	GLU	268.00	Side	OE2	14.39
ASN	6.00	Main	N	ASP	22.00	Side	OD2	90.21
ASN	6.00	Side	ND2	ASP	22.00	Side	OD1	14.29
ASN	7.00	Side	ND2	ASP	20.00	Side	OD2	19.28
ASN	7.00	Side	ND2	ASP	20.00	Side	OD1	15.68
THR	8.00	Main	N	ASP	20.00	Main	O	11.69
THR	8.00	Main	N	ASP	20.00	Side	OD1	41.96
THR	8.00	Main	N	ASP	20.00	Side	OD2	37.06
THR	8.00	Side	OG1	TYS	21.00	Main	O	58.94
THR	8.00	Side	OG1	TYS	21.00	Main	N	46.55
THR	8.00	Side	OG1	TYS	21.00	Main	N	18.58
ARG	9.00	Side	NH1	TYS	7.00	Main	O	13.09
ARG	9.00	Side	NH1	THR	8.00	Main	O	21.68
ARG	9.00	Side	NH2	ASP	10.00	Side	OD2	73.73
ARG	9.00	Side	NH2	ASP	10.00	Side	OD1	68.23
ARG	9.00	Side	NE	ASP	10.00	Side	OD1	64.54
ARG	9.00	Side	NE	ASP	10.00	Side	OD2	49.05
LYS	10.00	Side	NZ	GLY	19.00	Main	O	48.95
LYS	10.00	Side	NZ	TYS	21.00	Main	O	20.98
LYS	10.00	Side	NZ	TYR	190.00	Main	O	41.86
LYS	10.00	Side	NZ	ASN	192.00	Side	OD1	16.58
LYS	10.00	Side	NZ	ASP	193.00	Side	OD2	32.37
LYS	10.00	Side	NZ	ASP	193.00	Side	OD1	25.77
ARG	11.00	Side	NH1	GLU	2.00	Main	O	95.20
ARG	11.00	Side	NE	GLU	2.00	Main	O	84.72
ARG	11.00	Side	NH1	TYS	7.00	Side	OS2	54.05
ARG	11.00	Side	NH1	TYS	7.00	Side	OS4	43.76
ARG	11.00	Side	NH2	TYS	7.00	Side	OS2	38.06
ARG	11.00	Side	NH2	TYS	7.00	Side	OS4	12.99
ARG	11.00	Side	NH2	TYR	190.00	Main	O	32.37
ARG	11.00	Side	NH1	TYR	190.00	Main	O	26.37
SER	13.00	Side	OG	ASN	33.00	Side	ND2	27.27
SER	13.00	Side	OG	ASN	33.00	Side	OD1	10.09
SER	13.00	Side	OG	ASP	181.00	Side	OD2	11.19
SER	13.00	Side	OG	ARG	183.00	Side	NH1	11.29
LEU	14.00	Main	O	ARG	183.00	Side	NH1	78.62
LEU	14.00	Main	O	ARG	183.00	Side	NE	31.17

GLY	15.00	Main	N	ASP	187.00	Side	OD2	15.68
PRO	16.00	Main	O	SER	285.00	Side	OG	12.19
ARG	18.00	Side	NH1	HIS	113.00	Side	NE2	35.56
ARG	18.00	Side	NH1	TYR	116.00	Main	O	30.87
ARG	18.00	Side	NH2	TYR	116.00	Main	O	17.08
ARG	18.00	Side	NH1	THR	117.00	Side	OG1	76.92
ARG	18.00	Side	NH1	THR	117.00	Side	OG1	18.78
ARG	18.00	Side	NH1	ASP	171.00	Side	OD2	57.34
ARG	18.00	Side	NH1	ASP	171.00	Side	OD1	19.98
ARG	18.00	Main	O	ARG	188.00	Side	NH2	58.64
ARG	18.00	Main	O	ARG	188.00	Side	NE	42.36
ARG	18.00	Side	NH1	ARG	188.00	Side	NH1	18.88
ARG	18.00	Side	NH1	ARG	188.00	Side	NH1	10.79
ARG	18.00	Side	NH2	HIS	203.00	Side	NE2	52.75
ARG	18.00	Side	NH1	HIS	203.00	Side	NE2	16.18
ARG	18.00	Side	NH2	TYR	255.00	Side	OH	42.46
ARG	18.00	Side	NH2	TYR	255.00	Side	OH	23.68
ARG	18.00	Side	NE	TYR	255.00	Side	OH	20.48
ARG	18.00	Side	NE	TYR	255.00	Side	OH	19.08
ARG	18.00	Main	N	GLU	288.00	Side	OE2	67.73
ARG	18.00	Main	N	GLU	288.00	Side	OE1	26.07
ARG	18.00	Side	NE	GLU	288.00	Side	OE1	73.03
ARG	18.00	Side	NH2	GLU	288.00	Side	OE1	71.93
ARG	18.00	Side	NE	GLU	288.00	Side	OE2	44.56
ARG	18.00	Side	NH2	GLU	288.00	Side	OE2	28.37
TRP	20.00	Main	O	LYS	25.00	Side	NZ	34.67
TRP	20.00	Side	NE1	ASP	187.00	Side	OD1	11.49
TRP	20.00	Side	NE1	ARG	188.00	Main	O	59.54
TYR	21.00	Side	OH	CYS	28.00	Main	N	10.99
TYR	21.00	Side	OH	ASN	33.00	Side	ND2	20.08
TYR	21.00	Side	OH	SER	285.00	Side	OG	30.37
THR	22.00	Main	O	LYS	25.00	Main	N	10.39
THR	23.00	Side	OG1	CYS	28.00	Main	N	35.36
THR	23.00	Side	OG1	CYS	28.00	Main	O	17.08
GLY	24.00	Main	N	LYS	25.00	Main	O	24.38
GLY	24.00	Main	N	GLU	26.00	Main	O	56.44
ARG	31.00	Side	NH2	ASP	20.00	Main	O	84.32
ARG	31.00	Side	NE	ASP	20.00	Main	O	72.93
ARG	31.00	Side	NH2	TYS	21.00	Side	OS4	33.07
ARG	31.00	Side	NH2	TYS	21.00	Side	OS2	32.37
ARG	31.00	Side	NH1	TYS	21.00	Side	OS2	32.07
ARG	31.00	Side	NH1	TYS	21.00	Side	OS4	31.97
ARG	31.00	Side	NH1	TYS	21.00	Side	OS3	31.77
ARG	31.00	Side	NH2	TYS	21.00	Side	OS3	31.07
LYS	32.00	Side	NZ	TYS	21.00	Side	OS2	36.36
LYS	32.00	Side	NZ	TYS	21.00	Side	OS4	36.26
LYS	32.00	Side	NZ	TYS	21.00	Side	OS3	35.76

Complex 4:

V3 loop Residue	Number	Group	Atom	CXCR4 Residue	Number	Group	Atom	Occupancy (%)
ARG	3.00	Side	NH1	ASP	22.00	Main	O	36.16
ARG	3.00	Side	NH1	ASP	22.00	Side	OD2	10.49
ASN	5.00	Side	ND2	SER	23.00	Main	O	19.98
ASN	6.00	Main	N	SER	23.00	Main	O	35.76
ASN	7.00	Side	ND2	GLU	14.00	Side	OE2	29.77
ASN	7.00	Side	ND2	GLU	14.00	Side	OE1	17.18
THR	8.00	Side	OG1	LYS	25.00	Main	N	46.95
ARG	9.00	Side	NH1	ASP	10.00	Main	O	77.42
ARG	9.00	Side	NH1	ASN	11.00	Side	OD1	74.23
ARG	9.00	Side	NH2	ASN	11.00	Side	OD1	10.69
ARG	9.00	Side	NE	GLU	14.00	Side	OE2	50.15
ARG	9.00	Side	NH2	GLU	14.00	Side	OE2	49.45
ARG	9.00	Side	NE	GLU	14.00	Side	OE1	49.35
ARG	9.00	Side	NH2	GLU	14.00	Side	OE1	47.45
LYS	10.00	Main	N	TYS	12.00	Side	OS4	60.94
LYS	10.00	Main	N	TYS	12.00	Side	OS2	28.17
LYS	10.00	Side	NZ	GLY	19.00	Main	O	96.90
LYS	10.00	Side	NZ	ASP	20.00	Main	O	72.43
LYS	10.00	Side	NZ	ASP	193.00	Side	OD1	99.80
LYS	10.00	Side	NZ	GLU	268.00	Side	OE2	75.12
ARG	11.00	Side	NE	ASP	10.00	Side	OD2	95.10
ARG	11.00	Side	NH1	ASP	10.00	Side	OD2	91.51
ARG	11.00	Side	NH1	ASP	10.00	Side	OD1	10.19
ARG	11.00	Side	NH2	PRO	191.00	Main	O	89.71
SER	13.00	Side	OG	MET	1.00	Main	N	98.60
SER	13.00	Main	O	ARG	30.00	Side	NH1	18.08
LEU	14.00	Main	N	GLU	2.00	Main	O	75.12
LEU	14.00	Main	O	ARG	30.00	Side	NH1	36.56
LEU	14.00	Main	O	ARG	30.00	Side	NH2	19.28
GLY	15.00	Main	N	ASP	187.00	Side	OD2	53.75
GLY	15.00	Main	N	ASP	187.00	Side	OD1	44.96
ARG	18.00	Side	NH2	TYR	45.00	Side	OH	70.63
ARG	18.00	Side	NH1	TYR	45.00	Side	OH	16.88
ARG	18.00	Side	NH1	TRP	94.00	Side	NE1	76.12
ARG	18.00	Main	O	ARG	188.00	Side	NH2	94.31
ARG	18.00	Main	O	ARG	188.00	Side	NE	43.36
ARG	18.00	Side	NH2	TYR	255.00	Side	OH	10.39
ARG	18.00	Side	NH2	TYR	255.00	Side	OH	10.29
ARG	18.00	Main	N	GLU	288.00	Side	OE2	95.80
ARG	18.00	Side	NH2	GLU	288.00	Side	OE1	98.90
ARG	18.00	Side	NE	GLU	288.00	Side	OE1	80.32
ARG	18.00	Side	NE	GLU	288.00	Side	OE2	74.23
ARG	18.00	Side	NH2	GLU	288.00	Side	OE2	21.58
VAL	19.00	Main	N	GLU	288.00	Side	OE2	23.98
TRP	20.00	Side	NE1	ARG	188.00	Side	NE	15.78

TRP	20.00	Side	NE1	TYR	190.00	Side	OH	21.38
TYR	21.00	Side	OH	PHE	29.00	Main	O	20.78
TYR	21.00	Side	OH	ASN	278.00	Side	ND2	30.77
TYR	21.00	Side	OH	ASN	278.00	Side	ND2	27.57
THR	23.00	Side	OG1	MET	1.00	Main	N	54.95
GLY	24.00	Main	O	CYS	28.00	Main	N	30.67
GLN	25.00	Side	NE2	THR	8.00	Side	OG1	9.99
GLN	25.00	Side	OE1	SER	9.00	Main	N	73.83
GLN	25.00	Side	NE2	SER	9.00	Side	OG	20.98
ILE	26.00	Main	N	LYS	25.00	Main	O	54.75
ARG	31.00	Side	NH1	ASP	20.00	Side	OD1	50.55
ARG	31.00	Side	NH1	ASP	20.00	Side	OD2	38.56
ARG	31.00	Side	NH1	TYS	21.00	Main	O	81.72
ARG	31.00	Side	NH2	TYS	21.00	Main	O	70.63
ARG	31.00	Side	NH2	ASP	22.00	Main	O	10.59
ARG	31.00	Side	NH2	SER	23.00	Main	O	91.41
ARG	31.00	Side	NH1	SER	23.00	Main	O	12.29
LYS	32.00	Side	NZ	ASP	20.00	Side	OD1	45.85
LYS	32.00	Side	NZ	ASP	20.00	Side	OD2	44.76

Complex 5:

V3 loop Residue	Number	Group	Atom	CXCR4 Residue	Number	Group	Atom	Occupancy (%)
CYS	1	Main	N	MET	24	Main	O	33.57
CYS	1	Main	N	GLU	26	Side	OE1	71.83
CYS	1	Main	N	GLU	26	Side	OE2	69.53
ASN	7	Side	ND2	SER	23	Main	O	10.49
ARG	9	Side	NH1	ASP	20	Side	OD2	75.22
ARG	9	Side	NH1	ASP	20	Side	OD1	62.84
ARG	9	Side	NH2	ASP	20	Side	OD2	52.05
ARG	9	Side	NH2	ASP	20	Side	OD1	31.97
ARG	9	Side	NH1	TYS	21	Main	O	31.67
ARG	9	Side	NH1	ASP	22	Main	O	24.08
ARG	9	Side	NH2	ASP	193	Side	OD2	49.95
ARG	9	Side	NH2	ASP	193	Side	OD1	46.95
ARG	9	Side	NH1	ASP	193	Side	OD1	16.08
ARG	9	Side	NH1	ASP	193	Side	OD2	11.79
LYS	10	Side	NZ	CYS	28	Main	O	75.82
LYS	10	Side	NZ	PHE	29	Main	O	11.69
LYS	10	Side	NZ	GLU	277	Side	OE2	39.06
LYS	10	Side	NZ	GLU	277	Side	OE1	26.37
ARG	11	Side	NH2	ASP	10	Side	OD2	86.11
ARG	11	Side	NH1	ASP	10	Side	OD1	81.52
ARG	11	Side	NH1	ASP	10	Side	OD2	69.83
ARG	11	Side	NH2	ASP	10	Side	OD1	23.08
ARG	11	Side	NH2	PRO	191	Main	O	85.01
ARG	11	Side	NE	ASP	193	Side	OD2	12.79
ARG	11	Side	NH1	ASP	193	Side	OD1	9.69
SER	13	Side	OG	ARG	188	Main	O	97.10
SER	13	Side	OG	TYR	190	Main	N	61.84
LEU	14	Main	O	ARG	188	Side	NH2	99.90
LEU	14	Main	O	ARG	188	Side	NE	68.13
LEU	14	Main	N	GLN	200	Side	OE1	23.98
GLY	17	Main	O	SER	285	Side	OG	99.10
GLY	17	Main	N	GLU	288	Side	OE2	92.91
GLY	17	Main	N	GLU	288	Side	OE1	87.01
ARG	18	Side	NH1	ASP	181	Side	OD2	16.48
ARG	18	Side	NH2	ASP	181	Side	OD2	16.28
ARG	18	Side	NE	HIS	281	Side	NE2	42.56
ARG	18	Side	NH2	HIS	281	Side	NE2	20.58
VAL	19	Main	O	HIS	281	Side	ND1	11.29
TYR	21	Main	O	MET	1	Main	N	76.32
TYR	21	Side	OH	GLU	277	Main	O	57.74
TYR	21	Side	OH	ASN	278	Side	ND2	22.58
TYR	21	Side	OH	HIS	281	Side	ND1	12.39
THR	22	Side	OG1	MET	1	Main	N	77.42
THR	22	Side	OG1	ILE	6	Main	O	17.18
GLN	25	Main	N	THR	8	Side	OG1	20.48

GLN	25	Side	NE2	THR	8	Side	OG1	29.67
GLN	25	Side	NE2	SER	9	Side	OG	23.58
VAL	27	Main	N	TYS	12	Main	O	91.91
ILE	30	Main	O	SER	23	Side	OG	65.23
LYS	32	Side	NZ	ASP	22	Side	OD1	45.05
LYS	32	Side	NZ	ASP	22	Side	OD2	42.16
LYS	32	Main	N	SER	23	Side	OG	22.58
LYS	32	Main	N	SER	23	Side	OG	10.09
LYS	32	Side	NZ	SER	23	Side	OG	19.08

Complex 6:

V3 loop Residue	Number	Group	Atom	CXCR4 Residue	Number	Group	Atom	Occupancy (%)
CYS	1	Main	N	SER	23	Main	O	27.27
CYS	1	Main	N	SER	23	Side	OG	70.23
CYS	1	Main	O	SER	23	Side	OG	16.08
ASN	5	Side	ND2	SER	18	Side	OG	23.18
ASN	5	Side	ND2	ASP	20	Side	OD1	68.83
ASN	6	Main	N	ASP	20	Side	OD1	85.71
ASN	6	Main	N	ASP	20	Side	OD2	25.37
ASN	6	Side	ND2	ASP	20	Side	OD1	14.89
ASN	6	Side	ND2	TYS	21	Main	O	28.47
ASN	7	Side	ND2	TYS	12	Main	O	97.50
ASN	7	Side	ND2	GLU	14	Side	OE1	64.14
ASN	7	Side	ND2	GLU	14	Side	OE2	33.77
THR	8	Side	OG1	ASN	11	Main	N	35.26
THR	8	Side	OG1	ASN	11	Side	OD1	44.56
ARG	9	Main	O	SER	9	Main	N	39.26
ARG	9	Side	NH2	TYS	12	Side	OS2	93.11
ARG	9	Side	NH2	TYS	12	Side	OH	33.97
ARG	9	Side	NH2	TYS	12	Side	OS3	33.67
ARG	9	Side	NH1	TYS	12	Side	OS3	20.68
ARG	9	Side	NH1	TYS	12	Side	OS2	12.39
ARG	9	Side	NH2	GLY	19	Main	O	10.29
ARG	9	Side	NH1	ASP	20	Side	OD2	87.31
ARG	9	Side	NH2	ASP	20	Side	OD2	14.69
ARG	9	Side	NH2	ASP	193	Side	OD1	98.50
ARG	9	Side	NE	ASP	193	Side	OD2	97.00
ARG	9	Side	NH2	ASP	193	Side	OD2	73.43
ARG	9	Side	NE	ASP	193	Side	OD1	41.16
ARG	11	Main	N	TYS	7	Main	O	56.74
ARG	11	Side	NH1	ASP	10	Side	OD2	88.61
ARG	11	Side	NH1	ASP	10	Side	OD1	36.96
ARG	11	Side	NH2	ASP	10	Side	OD1	30.07
ARG	11	Side	NH2	TYR	190	Main	O	63.54
ARG	11	Side	NH1	TYR	190	Main	O	14.89
ARG	11	Side	NH2	PRO	191	Main	O	18.38
ARG	11	Side	NH1	PRO	191	Main	O	16.38
SER	13	Side	OG	MET	1	Main	N	14.09
SER	13	Side	OG	ASP	181	Main	N	14.79
SER	13	Side	OG	ASP	181	Side	OD1	30.47
SER	13	Side	OG	ASP	181	Side	OD2	19.88
SER	13	Side	OG	ARG	183	Side	NH2	20.58
GLY	15	Main	O	ASN	37	Side	ND2	27.17
PRO	16	Main	O	ASN	37	Side	ND2	15.48
PRO	16	Main	O	ARG	183	Side	NH1	39.76
GLY	17	Main	N	TYR	45	Side	OH	67.33
GLY	17	Main	N	GLU	288	Side	OE1	12.49

ARG	18	Side	NH2	TRP	94	Side	NE1	74.53
ARG	18	Side	NH1	HIS	113	Side	NE2	93.01
ARG	18	Side	NE	TYR	116	Side	OH	21.98
ARG	18	Side	NE	TYR	116	Side	OH	18.58
ARG	18	Side	NH2	GLU	288	Side	OE1	99.90
ARG	18	Side	NE	GLU	288	Side	OE2	99.80
ARG	18	Side	NE	GLU	288	Side	OE1	61.84
ARG	18	Side	NH2	GLU	288	Side	OE2	52.75
VAL	19	Main	N	ASP	187	Side	OD2	14.79
VAL	19	Main	O	TYR	190	Side	OH	30.97
TRP	20	Side	NE1	ARG	188	Side	NH1	18.98
TRP	20	Side	NE1	ARG	188	Side	NH1	16.68
TRP	20	Side	NE1	TYR	190	Side	OH	9.39
TYR	21	Side	OH	TYS	7	Side	OH	11.29
TYR	21	Side	OH	ASP	187	Side	OD1	96.60
TYR	21	Side	OH	ASP	187	Side	OD2	9.69
THR	22	Side	OG1	GLU	277	Side	OE2	93.41
THR	22	Side	OG1	GLU	277	Side	OE1	29.97
THR	23	Main	O	LYS	25	Side	NZ	90.71
THR	23	Side	OG1	LYS	25	Side	NZ	68.23
ARG	31	Side	NH1	GLU	14	Side	OE1	94.41
ARG	31	Side	NH1	GLU	14	Side	OE2	92.11
ARG	31	Side	NH2	GLU	14	Side	OE1	70.93
ARG	31	Side	NH2	GLU	14	Side	OE2	45.95

Complex 7:

V3 loop Residue	Number	Group	Atom	CXCR4 Residue	Number	Group	Atom	Occupancy (%)
ARG	3	Side	NH2	GLY	19	Main	O	15.08
ARG	3	Side	NH1	ASP	22	Side	OD1	30.57
ARG	3	Side	NH1	ASP	22	Side	OD2	24.78
ARG	3	Side	NE	ASP	22	Side	OD1	17.28
ARG	3	Side	NE	ASP	22	Side	OD2	16.58
ARG	3	Side	NH2	GLU	268	Side	OE1	52.95
ARG	3	Side	NH1	GLU	268	Side	OE2	48.45
ARG	3	Side	NH2	GLU	268	Side	OE2	44.26
ARG	3	Side	NH1	GLU	268	Side	OE1	18.38
ASN	7	Side	OD1	THR	13	Side	OG1	44.96
ASN	7	Side	ND2	THR	13	Side	OG1	12.99
THR	8	Main	N	THR	13	Side	OG1	21.78
THR	8	Main	O	THR	13	Side	OG1	14.39
THR	8	Main	O	GLU	14	Main	N	73.63
THR	8	Side	OG1	GLU	14	Side	OE2	38.46
ARG	9	Side	NH2	ASN	11	Side	OD1	17.88
ARG	9	Side	NE	TYS	12	Main	O	95.80
ARG	9	Side	NH2	TYS	12	Main	O	62.84
LYS	10	Side	NZ	GLU	14	Side	OE1	89.11
LYS	10	Side	NZ	GLU	14	Side	OE2	39.36
LYS	10	Side	NZ	SER	23	Main	O	28.27
LYS	10	Side	NZ	SER	23	Side	OG	10.29
LYS	10	Side	NZ	ASP	193	Side	OD1	63.04
LYS	10	Side	NZ	ASP	193	Side	OD2	44.56
ARG	11	Side	NH2	GLU	2	Side	OE2	35.66
ARG	11	Side	NH2	GLU	2	Side	OE1	12.89
ARG	11	Side	NH1	THR	8	Side	OG1	10.29
VAL	12	Main	N	GLU	2	Main	O	17.28
SER	13	Side	OG	ARG	30	Side	NE	10.19
GLY	15	Main	O	ASN	37	Side	ND2	19.58
GLY	15	Main	N	ASP	187	Side	OD2	65.03
GLY	15	Main	N	ASP	187	Side	OD1	16.18
GLY	17	Main	N	GLU	288	Side	OE2	68.03
GLY	17	Main	N	GLU	288	Side	OE1	47.45
ARG	18	Side	NH2	HIS	113	Side	ND1	16.98
ARG	18	Side	NH2	HIS	113	Side	ND1	16.88
ARG	18	Side	NE	TYR	116	Side	OH	26.57
ARG	18	Side	NH2	TYR	116	Side	OH	18.68
ARG	18	Side	NH1	THR	117	Side	OG1	85.81
ARG	18	Side	NH1	THR	117	Side	OG1	47.75
ARG	18	Side	NH2	THR	117	Side	OG1	25.27
ARG	18	Side	NH1	ASP	171	Side	OD1	95.40
ARG	18	Side	NH2	ASP	171	Side	OD2	92.81
ARG	18	Side	NH1	ASP	171	Side	OD2	73.93
ARG	18	Side	NH2	ASP	171	Side	OD1	39.16

ARG	18	Side	NH1	HIS	203	Side	NE2	51.65
ARG	18	Side	NH1	HIS	203	Side	ND1	33.37
ARG	18	Main	N	GLU	288	Side	OE1	64.34
ARG	18	Main	N	GLU	288	Side	OE2	62.14
TRP	20	Main	N	GLN	200	Side	OE1	13.89
TYR	21	Side	OH	CYS	28	Main	O	18.78
TYR	21	Side	OH	PHE	29	Main	O	19.58
TYR	21	Side	OH	GLU	277	Side	OE2	13.59
TYR	21	Side	OH	ASN	278	Side	ND2	58.34
TYR	21	Side	OH	ASN	278	Side	ND2	19.58
THR	22	Main	O	LYS	25	Side	NZ	35.46
THR	22	Side	OG1	LYS	25	Side	NZ	85.91
THR	23	Side	OG1	CYS	28	Main	O	86.21
THR	23	Side	OG1	CYS	28	Main	N	69.63
THR	23	Side	OG1	CYS	28	Main	N	44.36
GLY	24	Main	N	LYS	25	Main	O	20.78
GLY	24	Main	N	GLU	26	Main	O	51.65
GLN	25	Side	NE2	CYS	28	Main	O	9.79
ARG	31	Side	NH1	SER	18	Side	OG	17.78
ARG	31	Side	NH2	GLY	19	Main	O	21.08
ARG	31	Side	NH1	ASP	20	Side	OD1	75.12
ARG	31	Side	NH2	ASP	20	Side	OD2	72.73
ARG	31	Side	NH2	ASP	20	Side	OD1	57.24
ARG	31	Side	NH1	ASP	20	Side	OD2	49.15
LYS	32	Side	NZ	ASP	20	Side	OD2	60.64
LYS	32	Side	NZ	ASP	20	Side	OD1	54.35

Complex 8:

V3 loop Residue	Number	Group	Atom	CXCR4 Residue	Number	Group	Atom	Occupancy (%)
CYS	1	Main	N	MET	24	Main	O	53.85
CYS	1	Main	N	GLU	26	Side	OE1	26.47
CYS	1	Main	N	GLU	26	Side	OE2	23.18
ARG	3	Main	N	SER	23	Side	OG	30.87
ARG	3	Main	O	SER	23	Side	OG	14.49
THR	8	Side	OG1	TYS	12	Main	N	30.87
THR	8	Side	OG1	TYS	12	Main	O	22.28
THR	8	Side	OG1	ASP	20	Side	OD1	11.39
THR	8	Side	OG1	ASP	20	Side	OD2	9.69
ARG	9	Side	NH2	THR	8	Side	OG1	68.13
ARG	9	Side	NH1	THR	8	Side	OG1	61.04
ARG	9	Side	NH2	SER	9	Side	OG	78.02
LYS	10	Side	NZ	TYS	7	Main	O	13.19
LYS	10	Side	NZ	ASP	10	Side	OD1	55.64
LYS	10	Side	NZ	ASP	10	Side	OD2	42.36
LYS	10	Side	NZ	ASN	11	Main	O	52.45
ARG	11	Side	NH2	TYS	21	Side	OS4	25.97
ARG	11	Side	NH1	TYS	21	Side	OS4	25.77
ARG	11	Side	NH1	TYS	21	Side	OS3	15.08
ARG	11	Side	NH2	TYS	21	Side	OS3	14.09
ARG	11	Side	NH2	GLU	26	Main	O	10.59
VAL	12	Main	O	CYS	28	Main	N	82.62
SER	13	Side	OG	LYS	25	Side	NZ	13.69
SER	13	Side	OG	GLU	277	Side	OE2	14.99
SER	13	Side	OG	GLU	277	Side	OE1	12.29
GLY	15	Main	O	HIS	281	Side	ND1	97.40
PRO	16	Main	O	HIS	281	Side	ND1	60.94
GLY	17	Main	O	ARG	30	Side	NH1	69.23
GLY	17	Main	N	GLU	288	Side	OE1	77.02
ARG	18	Side	NH2	ASN	37	Side	OD1	87.11
ARG	18	Side	NH1	ASN	37	Side	OD1	29.67
ARG	18	Side	NE	TRP	94	Side	NE1	61.24
ARG	18	Side	NH1	ASP	97	Side	OD1	18.38
ARG	18	Main	N	GLU	288	Side	OE2	48.15
ARG	18	Main	N	GLU	288	Side	OE1	22.48
ARG	18	Side	NH2	GLU	288	Side	OE2	97.30
ARG	18	Side	NE	GLU	288	Side	OE2	40.56
TRP	20	Main	N	ASP	187	Side	OD1	45.15
TRP	20	Main	N	ASP	187	Side	OD2	14.69
TYR	21	Side	OH	GLN	200	Side	NE2	58.54
TYR	21	Side	OH	ASP	262	Side	OD1	53.85
TYR	21	Side	OH	ASP	262	Side	OD2	47.95
THR	22	Side	OG1	MET	1	Main	N	88.71
THR	22	Side	OG1	TYR	190	Main	O	14.09
GLY	24	Main	N	TYS	12	Side	OS4	86.21

GLY	24	Main	O	TYS	21	Main	N	29.77
GLY	24	Main	N	ASP	193	Side	OD2	21.28
GLN	25	Side	NE2	TYS	21	Main	O	12.99
GLN	25	Side	NE2	ASP	22	Main	O	41.36
GLN	25	Side	NE2	LYS	25	Main	O	19.98
ILE	26	Main	N	TYS	21	Main	O	50.65
ARG	31	Main	O	SER	9	Side	OG	13.19
HIS	34	Side	NE2	THR	8	Side	OG1	25.27
HIS	34	Side	ND1	LYS	25	Main	O	14.49

Complex 9:

V3 loop Residue	Number	Group	Atom	CXCR4 Residue	Number	Group	Atom	Occupancy (%)
ARG	3	Side	NH1	ASP	22	Side	OD2	83.62
ARG	3	Side	NE	ASP	22	Side	OD2	77.12
ARG	3	Side	NH1	ASP	22	Side	OD1	52.85
ARG	3	Side	NE	ASP	22	Side	OD1	14.19
ARG	3	Side	NH1	SER	23	Side	OG	56.34
ARG	3	Side	NH2	SER	23	Side	OG	39.26
ARG	3	Side	NH2	LEU	267	Main	O	54.95
ARG	3	Side	NH2	GLU	268	Side	OE1	32.97
ARG	3	Side	NH2	GLU	268	Side	OE2	21.48
THR	8	Main	N	GLY	19	Main	O	89.91
ARG	9	Side	NH2	ILE	4	Main	O	16.88
ARG	9	Side	NE	SER	5	Side	OG	0.10
ARG	9	Side	NH1	GLU	14	Side	OE2	63.94
ARG	9	Side	NH2	GLU	14	Side	OE2	56.94
ARG	9	Side	NH1	GLU	14	Side	OE1	55.64
ARG	9	Side	NH2	GLU	14	Side	OE1	44.26
ARG	9	Side	NH2	PRO	191	Main	O	64.24
ARG	9	Side	NE	PRO	191	Main	O	16.78
LYS	10	Side	NZ	GLY	19	Main	O	18.68
LYS	10	Side	NZ	TYS	21	Main	O	26.47
LYS	10	Side	NZ	SER	23	Main	O	89.01
LYS	10	Main	N	PRO	191	Main	O	14.89
LYS	10	Side	NZ	ASP	193	Side	OD2	99.10
ARG	11	Side	NH2	MET	1	Main	O	20.28
ARG	11	Side	NH2	GLU	2	Side	OE1	12.89
ARG	11	Side	NH2	GLU	2	Side	OE2	9.89
ARG	11	Side	NE	GLY	3	Main	O	24.18
VAL	12	Main	N	GLY	3	Main	O	17.68
LEU	14	Main	O	ARG	30	Side	NH2	11.79
PRO	16	Main	O	ARG	30	Side	NH1	32.07
PRO	16	Main	O	ASN	37	Side	ND2	21.28
GLY	17	Main	N	ASP	97	Side	OD1	32.97
GLY	17	Main	N	ASP	97	Side	OD2	24.78
GLY	17	Main	N	ALA	98	Main	O	9.89
ARG	18	Side	NH1	HIS	113	Side	NE2	58.34
ARG	18	Side	NH1	THR	117	Side	OG1	11.89
ARG	18	Side	NH2	HIS	203	Side	NE2	20.38
ARG	18	Side	NH1	HIS	203	Side	NE2	13.39
ARG	18	Side	NH2	TYR	255	Side	OH	30.37
ARG	18	Side	NH2	TYR	255	Side	OH	22.08
ARG	18	Side	NH2	GLU	288	Side	OE2	82.22
ARG	18	Side	NH2	GLU	288	Side	OE1	80.02
ARG	18	Side	NE	GLU	288	Side	OE1	77.52
ARG	18	Side	NE	GLU	288	Side	OE2	61.04
TRP	20	Side	NE1	HIS	113	Side	NE2	35.16

TRP	20	Side	NE1	ARG	188	Side	NH2	18.88
TRP	20	Side	NE1	ARG	188	Side	NH2	17.68
TRP	20	Side	NE1	ARG	188	Side	NE	12.09
TRP	20	Side	NE1	TYR	190	Side	OH	39.56
TRP	20	Main	O	HIS	281	Side	ND1	54.45
TYR	21	Side	OH	ARG	30	Side	NH1	29.47
TYR	21	Side	OH	GLU	31	Main	N	19.58
TYR	21	Side	OH	GLU	31	Main	O	9.69
TYR	21	Side	OH	ASN	33	Side	ND2	10.19
TYR	21	Side	OH	ASN	37	Side	ND2	44.66
THR	22	Main	O	LYS	25	Side	NZ	57.54
THR	22	Side	OG1	LYS	25	Side	NZ	85.51
THR	23	Side	OG1	GLU	26	Main	O	12.09
THR	23	Side	OG1	CYS	28	Main	N	67.43
THR	23	Side	OG1	CYS	28	Main	O	50.95
THR	23	Side	OG1	CYS	28	Main	N	19.88
GLY	24	Main	N	LYS	25	Main	O	26.77
GLY	24	Main	N	GLU	26	Main	O	44.46
GLN	25	Side	NE2	GLU	2	Main	O	29.97
ARG	31	Main	N	ASP	20	Side	OD1	46.05
ARG	31	Side	NH2	ASP	20	Side	OD2	99.30
ARG	31	Side	NE	ASP	20	Side	OD2	98.20
ARG	31	Side	NH2	TYS	21	Side	OS3	58.44
ARG	31	Side	NH2	TYS	21	Side	OS4	46.25
ARG	31	Side	NH2	TYS	21	Side	OS2	41.26
ARG	31	Side	NH1	TYS	21	Side	OS3	38.96
ARG	31	Side	NH1	TYS	21	Side	OS4	34.57
ARG	31	Side	NH1	TYS	21	Side	OS2	21.88
LYS	32	Side	NZ	TYS	21	Side	OS3	29.47
LYS	32	Side	NZ	TYS	21	Side	OS2	27.57
LYS	32	Side	NZ	TYS	21	Side	OS4	13.09
HIS	34	Side	ND1	ASP	20	Side	OD2	42.06

Complex 10:

V3 loop Residue	Number	Group	Atom	CXCR4 Residue	Number	Group	Atom	Occupancy (%)
CYS	1	Main	N	GLU	15	Side	OE1	26.87
CYS	1	Main	N	GLU	15	Side	OE2	23.28
CYS	1	Main	O	MET	16	Main	N	47.05
THR	2	Main	O	MET	16	Main	N	16.08
ASN	5	Side	ND2	TYS	12	Side	OS2	25.47
ASN	5	Side	ND2	TYS	12	Side	OS3	13.09
ASN	5	Main	N	GLU	14	Side	OE2	91.01
ASN	5	Main	N	GLU	14	Side	OE1	30.87
ASN	5	Side	ND2	GLU	14	Side	OE2	47.75
ASN	5	Side	ND2	GLU	14	Side	OE1	9.69
ASN	5	Side	ND2	MET	16	Main	O	14.19
ASN	5	Side	ND2	GLY	17	Main	O	15.38
THR	8	Main	O	ASP	10	Main	N	70.83
THR	8	Main	N	ASP	10	Side	OD1	18.38
THR	8	Side	OG1	TYS	12	Side	OS4	26.07
THR	8	Side	OG1	GLU	14	Side	OE1	13.99
ARG	9	Side	NH2	THR	8	Side	OG1	15.68
ARG	9	Side	NH2	SER	9	Side	OG	58.64
ARG	9	Side	NE	SER	9	Side	OG	13.59
ARG	9	Side	NH2	ASP	10	Side	OD1	16.58
ARG	9	Side	NH2	ASN	11	Side	OD1	10.19
LYS	10	Side	NZ	TYS	12	Side	OS4	31.97
LYS	10	Side	NZ	TYS	12	Side	OS3	24.08
LYS	10	Side	NZ	TYR	190	Main	O	47.05
LYS	10	Side	NZ	PRO	191	Main	O	25.37
LYS	10	Side	NZ	ASP	193	Side	OD2	54.75
LYS	10	Side	NZ	ASP	193	Side	OD1	40.96
ARG	11	Side	NH2	LYS	25	Main	O	38.36
ARG	11	Side	NH2	GLU	26	Main	O	60.94
ARG	11	Side	NH1	GLU	26	Main	O	34.67
ARG	11	Side	NH2	CYS	28	Main	O	17.38
SER	13	Side	OG	GLU	277	Side	OE2	35.86
SER	13	Side	OG	GLU	277	Side	OE1	20.28
GLY	15	Main	O	GLN	200	Side	NE2	32.77
GLY	17	Main	N	TYR	255	Side	OH	24.78
GLY	17	Main	N	TYR	255	Side	OH	12.39
ARG	18	Side	NH2	ASN	37	Side	ND2	71.73
ARG	18	Side	NH2	ASN	37	Side	ND2	20.88
ARG	18	Side	NH1	HIS	281	Side	NE2	96.60
ARG	18	Side	NH2	HIS	281	Side	NE2	27.77
ARG	18	Side	NH1	SER	285	Side	OG	99.80
ARG	18	Main	N	GLU	288	Side	OE2	98.20
TRP	20	Side	NE1	ASP	181	Side	OD1	25.47
TRP	20	Side	NE1	ASP	181	Side	OD2	23.48
TRP	20	Main	O	ARG	188	Main	N	73.63

TYR	21	Side	OH	GLN	200	Side	NE2	51.05
TYR	21	Side	OH	ASP	262	Side	OD2	69.73
TYR	21	Side	OH	ASP	262	Side	OD1	26.27
THR	22	Main	N	ARG	188	Main	O	18.98
THR	22	Side	OG1	TYR	190	Main	N	51.25
THR	23	Main	O	LYS	25	Side	NZ	10.69
THR	23	Side	OG1	LYS	25	Side	NZ	38.26
GLY	24	Main	O	LYS	25	Side	NZ	10.59
GLY	24	Main	N	ASP	193	Side	OD2	18.28
GLN	25	Side	NE2	SER	23	Main	O	15.38
GLN	25	Side	OE1	LYS	25	Main	N	80.72
ILE	26	Main	O	TYS	21	Main	N	29.97
ARG	31	Side	NH2	ASP	10	Side	OD1	71.73
ARG	31	Side	NH1	ASP	10	Side	OD2	61.44
ARG	31	Side	NH1	ASP	10	Side	OD1	42.26
ARG	31	Side	NH2	ASP	10	Side	OD2	20.68
ARG	31	Side	NH1	THR	13	Side	OG1	13.19

Complex 11:

V3 loop Residue	Number	Group	Atom	CXCR4 Residue	Number	Group	Atom	Occupancy (%)
CYS	1	Main	N	SER	23	Side	OG	55.64
CYS	1	Main	O	SER	23	Side	OG	11.19
ARG	3	Side	NH2	SER	23	Side	OG	12.99
ASN	5	Side	ND2	SER	18	Main	O	39.66
ASN	5	Side	ND2	GLY	19	Main	O	36.26
ASN	6	Main	N	ASP	20	Side	OD1	84.72
ASN	6	Main	N	ASP	20	Side	OD2	23.38
ASN	6	Side	ND2	ASP	20	Side	OD2	53.65
ASN	6	Side	ND2	ASP	20	Side	OD1	53.35
ASN	6	Side	ND2	ASP	22	Main	O	9.99
ASN	6	Side	OD1	SER	23	Side	OG	25.07
ASN	6	Side	ND2	SER	23	Side	OG	16.48
ASN	7	Side	ND2	TYS	12	Main	O	87.71
ASN	7	Side	ND2	TYS	12	Side	OS4	41.26
ASN	7	Side	ND2	TYS	12	Side	OS2	31.37
ASN	7	Side	ND2	TYS	12	Side	OS3	26.37
THR	8	Side	OG1	ASN	11	Side	ND2	94.51
ARG	9	Main	O	SER	9	Main	N	64.54
ARG	9	Side	NH1	TYS	12	Side	OH	50.45
ARG	9	Side	NH1	TYS	12	Side	OS3	37.16
ARG	9	Side	NH2	TYS	12	Side	OS4	27.87
ARG	9	Side	NH1	TYS	12	Side	OS4	24.98
ARG	9	Side	NH1	TYS	12	Side	OS2	24.38
ARG	9	Side	NH2	TYS	12	Side	OS2	20.68
ARG	9	Side	NH2	TYS	12	Side	OS3	17.58
ARG	9	Side	NH2	GLY	19	Main	O	79.82
ARG	9	Side	NH1	GLY	19	Main	O	30.27
ARG	9	Side	NH2	ASP	20	Main	O	13.79
ARG	9	Side	NH2	ASP	20	Side	OD1	34.47
ARG	9	Side	NH1	ASP	193	Side	OD1	88.91
ARG	9	Side	NH1	ASP	193	Side	OD2	78.72
LYS	10	Side	NZ	THR	8	Side	OG1	23.58
ARG	11	Main	N	TYS	7	Main	O	42.76
ARG	11	Side	NH1	TYS	7	Main	O	30.47
ARG	11	Side	NH2	ASP	10	Side	OD1	89.61
ARG	11	Side	NH1	ASP	10	Side	OD2	88.01
ARG	11	Side	NH2	ASP	10	Side	OD2	69.43
ARG	11	Side	NH1	ASP	10	Side	OD1	56.74
ARG	11	Side	NH2	PRO	191	Main	O	94.41
ARG	11	Side	NE	PRO	191	Main	O	44.76
SER	13	Side	OG	ASP	181	Main	N	15.78
SER	13	Side	OG	ASP	181	Side	OD1	74.53
SER	13	Side	OG	ASP	181	Side	OD2	61.54
GLY	15	Main	O	ARG	30	Side	NH2	80.82
GLY	15	Main	O	ARG	30	Side	NH1	69.13

PRO	16	Main	O	ARG	183	Side	NH1	35.86
GLY	17	Main	N	TYR	45	Side	OH	21.08
GLY	17	Main	N	GLU	288	Side	OE1	49.75
ARG	18	Side	NH2	TYR	45	Side	OH	13.49
ARG	18	Side	NH1	TRP	94	Side	NE1	22.48
ARG	18	Side	NH2	TRP	94	Side	NE1	13.79
ARG	18	Side	NH1	HIS	113	Side	NE2	82.82
ARG	18	Side	NE	TYR	116	Side	OH	23.78
ARG	18	Side	NE	TYR	116	Side	OH	22.58
ARG	18	Main	O	TYR	190	Side	OH	14.89
ARG	18	Main	N	GLU	288	Side	OE1	31.37
ARG	18	Side	NE	GLU	288	Side	OE2	97.80
ARG	18	Side	NH2	GLU	288	Side	OE1	97.40
ARG	18	Side	NE	GLU	288	Side	OE1	68.63
ARG	18	Side	NH2	GLU	288	Side	OE2	41.26
VAL	19	Main	O	TYR	190	Side	OH	14.59
TRP	20	Side	NE1	ARG	188	Side	NH2	22.08
TRP	20	Side	NE1	ARG	188	Side	NH2	20.28
TRP	20	Side	NE1	TYR	190	Side	OH	34.47
TYR	21	Side	OH	SER	178	Side	OG	9.59
TYR	21	Side	OH	ASP	187	Side	OD1	47.25
THR	22	Side	OG1	GLU	277	Side	OE2	89.21
THR	23	Main	O	LYS	25	Side	NZ	89.71
THR	23	Side	OG1	LYS	25	Side	NZ	64.44
GLY	24	Main	O	CYS	28	Main	N	95.80
ILE	26	Main	O	LYS	25	Main	N	88.81
ILE	26	Main	N	LYS	25	Main	O	63.84
GLY	28	Main	N	SER	23	Main	O	26.87
ASP	29	Side	OD2	ASN	11	Side	ND2	65.23
ASP	29	Side	OD1	ASN	11	Side	ND2	63.14
ARG	31	Side	NE	ASN	11	Side	OD1	23.88
ARG	31	Side	NH1	GLU	14	Side	OE1	81.92
ARG	31	Side	NH1	GLU	14	Side	OE2	80.52
ARG	31	Side	NH2	GLU	14	Side	OE1	56.64
ARG	31	Side	NH2	GLU	14	Side	OE2	56.14
LYS	32	Side	NZ	GLU	14	Side	OE1	17.28
LYS	32	Side	NZ	GLU	14	Side	OE2	10.69

Complex 12:

V3 loop Residue	Number	Group	Atom	CXCR4 Residue	Number	Group	Atom	Occupancy (%)
ARG	3	Side	NH2	ASP	22	Side	OD1	64.44
ARG	3	Side	NH2	ASP	22	Side	OD2	24.38
ARG	3	Side	NH1	ASP	22	Side	OD1	17.98
ARG	3	Side	NH2	GLU	268	Side	OE1	32.97
ARG	3	Side	NH2	GLU	268	Side	OE2	27.77
ARG	3	Side	NH1	GLU	268	Side	OE1	26.47
ARG	3	Side	NH1	GLU	268	Side	OE2	23.48
ASN	5	Side	ND2	TYS	21	Main	O	23.48
ASN	5	Side	ND2	ASP	22	Main	O	24.68
ASN	5	Side	ND2	SER	23	Main	O	36.76
ASN	6	Main	N	SER	23	Main	O	80.32
ASN	6	Side	ND2	SER	23	Main	O	25.67
THR	8	Side	OG1	LYS	25	Main	N	44.16
ARG	9	Side	NH1	SER	9	Side	OG	39.36
ARG	9	Side	NH2	SER	9	Side	OG	32.17
ARG	9	Side	NH1	ASP	10	Side	OD2	22.48
ARG	9	Side	NH2	ASP	10	Side	OD2	19.38
ARG	9	Side	NH1	ASN	11	Main	O	14.19
ARG	9	Side	NE	ASN	11	Side	OD1	50.65
ARG	9	Side	NH2	ASN	11	Side	OD1	48.75
LYS	10	Main	N	TYS	12	Side	OS3	37.26
LYS	10	Side	NZ	TYS	12	Side	OS2	27.27
LYS	10	Side	NZ	TYS	12	Side	OS4	25.97
LYS	10	Side	NZ	TYS	12	Side	OS3	16.98
LYS	10	Side	NZ	ASP	20	Main	O	96.90
LYS	10	Side	NZ	ASP	193	Side	OD1	97.00
LYS	10	Side	NZ	ASP	193	Side	OD2	60.24
LYS	10	Side	NZ	GLU	268	Side	OE1	32.17
LYS	10	Side	NZ	GLU	268	Side	OE2	14.99
ARG	11	Main	O	MET	1	Main	N	15.18
ARG	11	Side	NE	MET	1	Main	O	54.85
ARG	11	Side	NH1	MET	1	Main	O	53.15
ARG	11	Side	NH1	ASP	10	Side	OD2	76.82
ARG	11	Side	NH1	ASP	10	Side	OD1	60.24
ARG	11	Side	NH2	ASP	10	Side	OD1	56.34
ARG	11	Side	NE	ASP	10	Side	OD2	28.57
ARG	11	Side	NH2	ASP	10	Side	OD2	26.87
ARG	11	Side	NE	ASP	10	Side	OD1	20.88
ARG	11	Side	NH2	ASN	11	Main	O	14.49
ARG	11	Side	NH2	TYR	190	Main	O	38.46
ARG	11	Side	NH1	TYR	190	Main	O	12.89
VAL	12	Main	O	MET	1	Main	N	20.08
SER	13	Side	OG	MET	1	Main	N	97.20
SER	13	Side	OG	GLU	2	Main	O	17.08
LEU	14	Main	N	ASP	187	Side	OD2	47.95

LEU	14	Main	N	ASP	187	Side	OD1	46.15
GLY	15	Main	O	ARG	183	Side	NH1	14.79
ARG	18	Side	NH1	THR	117	Side	OG1	92.71
ARG	18	Side	NH1	THR	117	Side	OG1	38.76
ARG	18	Side	NH1	ASP	171	Side	OD2	75.42
ARG	18	Main	O	ARG	188	Side	NH2	99.20
ARG	18	Side	NH1	HIS	203	Side	NE2	99.00
ARG	18	Side	NH2	HIS	203	Side	NE2	97.00
ARG	18	Side	NH2	TYR	255	Side	OH	90.61
ARG	18	Side	NH2	TYR	255	Side	OH	68.73
ARG	18	Main	N	GLU	288	Side	OE2	35.56
ARG	18	Side	NH2	GLU	288	Side	OE1	97.30
ARG	18	Side	NE	GLU	288	Side	OE2	61.24
ARG	18	Side	NH2	GLU	288	Side	OE2	45.35
ARG	18	Side	NE	GLU	288	Side	OE1	30.37
TRP	20	Side	NE1	ARG	188	Side	NH2	74.53
TRP	20	Side	NE1	TYR	190	Side	OH	52.75
TRP	20	Side	NE1	TYR	190	Side	OH	13.29
TRP	20	Main	O	HIS	281	Side	ND1	78.02
TYR	21	Side	OH	ARG	30	Main	N	24.08
TYR	21	Side	OH	ASN	278	Side	ND2	42.46
TYR	21	Side	OH	ASN	278	Side	ND2	22.68
THR	23	Side	OG1	MET	1	Main	N	77.42
GLY	24	Main	O	CYS	28	Main	N	81.52
GLY	24	Main	N	CYS	28	Main	O	58.44
GLN	25	Side	OE1	SER	9	Main	N	13.69
GLN	25	Side	NE2	SER	9	Side	OG	15.18
ARG	31	Side	NH2	TYS	12	Side	OS4	47.05
ARG	31	Side	NH1	TYS	12	Side	OH	40.46
ARG	31	Side	NH1	TYS	12	Side	OS4	29.67
ARG	31	Side	NH2	TYS	12	Side	OS2	25.17
ARG	31	Side	NH1	TYS	12	Side	OS2	12.49
ARG	31	Side	NH2	TYS	12	Side	OS3	11.49
ARG	31	Side	NH2	ASP	20	Side	OD2	16.08
ARG	31	Side	NH2	SER	23	Main	O	11.79
LYS	32	Side	NZ	TYS	12	Side	OS3	11.49
LYS	32	Side	NZ	ASP	20	Side	OD2	53.35
LYS	32	Side	NZ	ASP	20	Side	OD1	51.75
HIS	34	Side	ND1	TYS	21	Main	O	29.57

Complex 13:

V3 loop Residue	Number	Group	Atom	CXCR4 Residue	Number	Group	Atom	Occupancy (%)
CYS	1	Main	N	GLU	31	Side	OE2	37.56
CYS	1	Main	N	GLU	31	Side	OE1	37.06
THR	2	Main	N	PRO	27	Main	O	33.37
THR	8	Main	N	MET	24	Side	SD	19.28
THR	8	Main	O	LYS	25	Main	N	92.41
ARG	9	Side	NH2	TYS	12	Side	OS3	24.58
ARG	9	Side	NH2	TYS	12	Side	OS4	16.08
ARG	9	Side	NH1	TYS	12	Side	OS4	12.59
ARG	9	Side	NH2	TYS	12	Side	OS2	10.19
ARG	9	Side	NH1	GLU	14	Side	OE2	55.44
ARG	9	Side	NH1	GLU	14	Side	OE1	12.09
ARG	9	Side	NH1	GLU	15	Main	O	35.96
ARG	9	Side	NH2	ASP	20	Side	OD1	94.71
ARG	9	Side	NH2	ASP	20	Side	OD2	92.41
ARG	9	Side	NH1	ASP	20	Side	OD1	50.85
ARG	9	Side	NH1	ASP	20	Side	OD2	28.67
ARG	9	Side	NH1	TYS	21	Main	O	25.87
ARG	9	Side	NH1	ASP	22	Main	O	18.38
ARG	9	Side	NH2	ASP	193	Side	OD1	99.70
ARG	9	Side	NE	ASP	193	Side	OD2	88.21
ARG	9	Side	NH2	ASP	193	Side	OD2	67.83
ARG	9	Side	NE	ASP	193	Side	OD1	40.76
LYS	10	Side	NZ	GLU	2	Side	OE1	9.79
LYS	10	Side	NZ	GLU	2	Side	OE2	9.69
LYS	10	Side	NZ	CYS	28	Main	O	94.51
ARG	11	Side	NH1	TYS	7	Main	O	17.08
ARG	11	Side	NH1	ASP	10	Side	OD2	76.72
ARG	11	Side	NH1	ASP	10	Side	OD1	57.14
ARG	11	Side	NH2	ASP	10	Side	OD1	40.46
ARG	11	Side	NH2	ASP	10	Side	OD2	30.87
ARG	11	Side	NH2	TYR	190	Main	O	71.63
ARG	11	Side	NH2	PRO	191	Main	O	47.95
SER	13	Side	OG	ARG	188	Main	O	32.97
LEU	14	Main	O	ARG	188	Side	NH2	99.70
LEU	14	Main	O	ARG	188	Side	NE	42.16
LEU	14	Main	N	GLN	200	Side	OE1	66.43
GLY	17	Main	O	ASN	37	Side	ND2	39.86
GLY	17	Main	O	SER	285	Side	OG	93.11
GLY	17	Main	N	GLU	288	Side	OE1	94.81
GLY	17	Main	N	GLU	288	Side	OE2	79.92
ARG	18	Side	NH2	ASP	181	Side	OD2	25.57
ARG	18	Side	NH1	ASP	181	Side	OD2	13.09
VAL	19	Main	N	HIS	281	Side	NE2	62.24
VAL	19	Main	O	HIS	281	Side	ND1	19.08
VAL	19	Main	N	HIS	281	Side	ND1	15.98

TYR	21	Main	O	GLU	2	Main	N	17.08
TYR	21	Side	OH	GLU	277	Main	O	46.65
TYR	21	Side	OH	ASN	278	Side	OD1	18.48
TYR	21	Side	OH	ASN	278	Side	ND2	15.58
TYR	21	Side	OH	HIS	281	Side	ND1	51.85
THR	22	Side	OG1	MET	1	Main	N	75.92
THR	23	Main	N	MET	1	Main	O	16.78
THR	23	Side	OG1	GLU	2	Side	OE2	51.75
THR	23	Side	OG1	GLU	2	Side	OE1	49.45
GLN	25	Side	NE2	ILE	6	Main	O	23.68
GLN	25	Main	N	TYS	7	Main	O	28.57
GLN	25	Side	NE2	THR	8	Side	OG1	13.49
GLN	25	Main	O	SER	9	Main	N	66.53
GLN	25	Main	O	ASP	10	Main	N	53.75
VAL	27	Main	N	ASP	10	Main	O	13.79
VAL	27	Main	N	ASN	11	Side	OD1	38.26
VAL	27	Main	O	ASN	11	Side	ND2	24.58
ARG	31	Side	NH1	GLU	14	Side	OE1	85.51
ARG	31	Side	NH2	GLU	14	Side	OE2	81.82
ARG	31	Side	NH2	GLU	14	Side	OE1	67.83
ARG	31	Side	NH1	GLU	14	Side	OE2	46.65
ARG	31	Side	NH2	SER	23	Main	O	96.30
ARG	31	Side	NE	SER	23	Main	O	49.85
LYS	32	Side	NZ	ASP	22	Side	OD2	79.02
LYS	32	Side	NZ	ASP	22	Side	OD1	78.42
LYS	32	Side	NZ	SER	23	Side	OG	32.77

Complex 14:

V3 loop Residue	Number	Group	Atom	CXCR4 Residue	Number	Group	Atom	Occupancy (%)
ARG	3	Side	NH1	ASP	22	Side	OD2	91.11
ARG	3	Side	NH2	ASP	22	Side	OD1	87.61
ARG	3	Side	NH2	ASP	22	Side	OD2	74.93
ARG	3	Side	NH1	ASP	22	Side	OD1	18.48
ASN	5	Side	ND2	ASP	20	Main	O	90.71
ASN	5	Side	ND2	ASP	22	Side	OD2	84.32
ASN	5	Side	ND2	ASP	22	Side	OD1	79.02
ASN	6	Main	N	ASP	22	Side	OD1	42.16
ASN	7	Side	ND2	ASP	20	Side	OD1	55.34
ASN	7	Side	ND2	ASP	20	Side	OD2	37.06
THR	8	Main	N	ASP	20	Side	OD1	53.55
THR	8	Main	N	ASP	20	Side	OD2	32.97
THR	8	Side	OG1	ASP	20	Main	O	11.59
THR	8	Side	OG1	TYS	21	Main	O	22.98
ARG	9	Side	NH1	SER	5	Side	OG	21.38
ARG	9	Side	NH2	PRO	191	Main	O	95.30
ARG	9	Side	NH1	PRO	191	Main	O	74.83
LYS	10	Side	NZ	GLY	19	Main	O	88.81
LYS	10	Side	NZ	TYS	21	Main	O	97.90
LYS	10	Side	NZ	SER	23	Main	O	50.45
LYS	10	Side	NZ	ASP	193	Side	OD2	95.90
LYS	10	Side	NZ	ASP	193	Side	OD1	23.98
ARG	11	Side	NH1	GLU	2	Side	OE2	70.43
ARG	11	Side	NH1	GLU	2	Side	OE1	64.54
ARG	11	Side	NH2	GLU	2	Side	OE2	62.64
ARG	11	Side	NH2	GLU	2	Side	OE1	58.94
ARG	11	Side	NH1	TYS	7	Side	OS4	81.32
VAL	12	Main	N	TYS	7	Side	OS2	95.90
VAL	12	Main	N	TYS	7	Side	OS4	10.29
SER	13	Main	O	ARG	30	Side	NH2	38.16
SER	13	Side	OG	ARG	30	Side	NH2	19.58
LEU	14	Main	O	ARG	30	Side	NH2	54.65
PRO	16	Main	O	ASN	37	Side	ND2	38.26
GLY	17	Main	N	ASP	97	Side	OD1	91.21
ARG	18	Side	NH1	HIS	113	Side	NE2	98.40
ARG	18	Side	NH2	TYR	116	Side	OH	11.79
ARG	18	Side	NH2	HIS	203	Side	NE2	97.10
ARG	18	Side	NH1	HIS	203	Side	NE2	78.62
ARG	18	Side	NH2	TYR	255	Side	OH	89.21
ARG	18	Side	NH2	TYR	255	Side	OH	68.73
ARG	18	Side	NE	GLU	288	Side	OE2	91.21
ARG	18	Side	NH2	GLU	288	Side	OE1	83.02
ARG	18	Side	NH2	GLU	288	Side	OE2	63.54
ARG	18	Side	NE	GLU	288	Side	OE1	62.24
TRP	20	Side	NE1	TYR	190	Side	OH	22.88

TRP	20	Side	NE1	TYR	190	Side	OH	21.18
TRP	20	Side	NE1	GLN	200	Side	OE1	25.07
TRP	20	Side	NE1	GLN	200	Side	NE2	18.38
TRP	20	Main	O	HIS	281	Side	ND1	67.33
TRP	20	Main	N	SER	285	Side	OG	19.78
TYR	21	Side	OH	ASN	33	Side	ND2	10.49
TYR	21	Side	OH	ASN	37	Side	ND2	18.48
THR	22	Main	O	LYS	25	Side	NZ	37.76
THR	22	Side	OG1	LYS	25	Side	NZ	79.02
THR	23	Side	OG1	CYS	28	Main	N	35.76
THR	23	Side	OG1	CYS	28	Main	O	26.17
GLY	24	Main	N	LYS	25	Main	O	46.05
GLY	24	Main	N	GLU	26	Main	O	13.79
GLN	25	Side	NE2	THR	8	Side	OG1	10.49
ARG	31	Side	NH1	GLU	14	Side	OE1	31.67
ARG	31	Side	NH1	GLU	14	Side	OE2	29.67
ARG	31	Side	NH1	SER	18	Side	OG	41.96
ARG	31	Side	NH1	SER	18	Side	OG	31.17
ARG	31	Side	NH2	TYS	21	Side	OS3	47.95
ARG	31	Side	NH1	TYS	21	Side	OS3	47.35
ARG	31	Side	NH2	TYS	21	Side	OS2	43.26
ARG	31	Side	NH2	TYS	21	Side	OS4	37.86
ARG	31	Side	NH1	TYS	21	Side	OS2	35.16
ARG	31	Side	NH1	TYS	21	Side	OS4	34.07
LYS	32	Side	NZ	TYS	21	Side	OS2	29.77
LYS	32	Side	NZ	TYS	21	Side	OS3	26.87
LYS	32	Side	NZ	TYS	21	Side	OS4	23.28

Complex 15:

V3 loop Residue	Number	Group	Atom	CXCR4 Residue	Number	Group	Atom	Occupancy (%)
CYS	1	Main	N	ASP	20	Main	O	58.84
CYS	1	Main	N	ASP	20	Side	OD2	81.52
CYS	1	Main	N	ASP	20	Side	OD1	41.26
CYS	1	Main	N	TYS	21	Main	O	58.04
CYS	1	Main	N	GLU	268	Side	OE1	43.76
CYS	1	Main	N	GLU	268	Side	OE2	15.88
ASN	5	Side	ND2	TYS	12	Side	OS2	39.56
ASN	6	Main	N	TYS	12	Side	OS4	45.45
ASN	6	Main	N	TYS	12	Side	OS2	21.18
ASN	7	Main	O	ASN	11	Main	N	22.88
ASN	7	Main	O	ASN	11	Side	ND2	18.38
ASN	7	Main	O	TYS	12	Main	N	36.76
ASN	7	Side	ND2	TYS	12	Main	O	22.08
ASN	7	Side	ND2	GLU	14	Side	OE1	17.68
ASN	7	Side	ND2	GLU	14	Side	OE2	14.69
THR	8	Main	O	SER	9	Main	N	45.75
THR	8	Side	OG1	ASN	11	Side	OD1	22.98
THR	8	Side	OG1	ASN	11	Side	ND2	18.98
ARG	9	Side	NE	TYS	12	Side	OS4	90.61
ARG	9	Side	NH2	TYS	12	Side	OS4	23.68
ARG	9	Side	NE	TYS	12	Side	OS3	14.89
ARG	9	Side	NH2	ASP	20	Side	OD1	22.98
ARG	9	Side	NH2	TYS	21	Main	O	17.38
ARG	9	Side	NH1	TYS	21	Main	O	13.39
ARG	9	Side	NH1	ASP	22	Main	O	30.27
ARG	9	Side	NH2	ASP	22	Main	O	13.29
LYS	10	Side	NZ	ILE	6	Main	O	25.37
LYS	10	Side	NZ	THR	8	Side	OG1	29.17
LYS	10	Side	NZ	SER	9	Side	OG	15.28
ARG	11	Main	O	MET	1	Main	N	28.97
ARG	11	Side	NH1	TYS	7	Main	O	15.88
ARG	11	Side	NH2	ASP	10	Side	OD1	99.30
ARG	11	Side	NH1	ASP	10	Side	OD2	99.30
ARG	11	Side	NH1	ASP	10	Side	OD1	54.65
ARG	11	Side	NH2	ASP	10	Side	OD2	41.26
ARG	11	Side	NH2	PRO	191	Main	O	46.75
ARG	11	Side	NE	PRO	191	Main	O	40.86
SER	13	Side	OG	MET	1	Main	N	62.04
PRO	16	Main	O	ASN	37	Side	ND2	10.79
PRO	16	Main	O	ARG	183	Side	NH1	84.12
PRO	16	Main	O	ARG	183	Side	NH2	52.35
GLY	17	Main	N	GLU	288	Side	OE1	78.02
ARG	18	Side	NH2	TYR	45	Side	OH	98.30
ARG	18	Side	NH1	TYR	45	Side	OH	43.46
ARG	18	Side	NH1	TRP	94	Side	NE1	50.85

ARG	18	Side	NH2	TYR	116	Side	OH	12.09
ARG	18	Main	O	ARG	188	Side	NE	35.86
ARG	18	Main	O	ARG	188	Side	NH2	32.67
ARG	18	Main	N	GLU	288	Side	OE1	98.60
ARG	18	Side	NH2	GLU	288	Main	O	28.87
ARG	18	Side	NH2	GLU	288	Side	OE2	97.80
ARG	18	Side	NE	GLU	288	Side	OE2	94.01
ARG	18	Side	NE	GLU	288	Side	OE1	44.46
ARG	18	Side	NH2	GLU	288	Side	OE1	13.09
VAL	19	Main	O	ARG	188	Side	NH2	34.97
VAL	19	Main	O	TYR	190	Side	OH	89.51
TRP	20	Side	NE1	ARG	188	Side	NH2	86.51
TRP	20	Side	NE1	ARG	188	Side	NH2	42.26
TRP	20	Side	NE1	HIS	203	Side	NE2	23.68
TYR	21	Side	OH	MET	1	Main	N	39.66
TYR	21	Side	OH	ASP	187	Side	OD2	37.06
TYR	21	Side	OH	ASP	187	Side	OD1	15.98
TYR	21	Side	OH	ARG	188	Main	O	20.98
THR	22	Side	OG1	GLU	277	Side	OE2	60.54
THR	23	Main	O	LYS	25	Side	NZ	98.50
THR	23	Side	OG1	ASP	193	Side	OD2	30.97
THR	23	Side	OG1	ASP	193	Side	OD1	19.28
GLY	24	Main	O	CYS	28	Main	N	27.47
ILE	26	Main	N	LYS	25	Main	O	57.04
CYS	35	Main	N	SER	23	Main	O	37.56

Complex 16:

V3 loop Residue	Number	Group	Atom	CXCR4 Residue	Number	Group	Atom	Occupancy (%)
ARG	3	Side	NE	SER	23	Side	OG	10.19
ARG	3	Side	NE	SER	23	Side	OG	9.79
ASN	6	Side	ND2	ASN	11	Side	OD1	20.08
ASN	7	Side	ND2	ASP	22	Main	O	10.09
ASN	7	Side	ND2	SER	23	Main	O	21.88
ASN	7	Side	OD1	LYS	25	Main	N	25.67
THR	8	Main	O	LYS	25	Main	N	24.18
ARG	9	Side	NH1	GLU	14	Side	OE1	62.74
ARG	9	Side	NH1	GLU	14	Side	OE2	53.75
ARG	9	Side	NH2	ASP	20	Main	O	10.59
ARG	9	Side	NH1	ASP	193	Side	OD2	82.32
ARG	9	Side	NH2	ASP	193	Side	OD2	61.74
ARG	9	Side	NH1	ASP	193	Side	OD1	56.84
ARG	9	Side	NH2	ASP	193	Side	OD1	11.89
ARG	9	Side	NH2	LEU	266	Main	O	17.38
LYS	10	Side	NZ	CYS	28	Main	O	34.07
ARG	11	Side	NH1	TYS	12	Side	OS2	43.16
ARG	11	Side	NH2	TYS	12	Side	OS2	36.56
ARG	11	Side	NH1	TYS	12	Side	OS3	30.57
ARG	11	Side	NH1	TYS	12	Side	OS4	23.48
ARG	11	Side	NH2	TYS	12	Side	OS4	21.58
ARG	11	Side	NH2	TYS	12	Side	OS3	12.59
ARG	11	Side	NH1	GLU	14	Side	OE2	68.23
ARG	11	Side	NE	GLU	14	Side	OE2	40.76
ARG	11	Side	NH2	PRO	191	Main	O	58.54
ARG	11	Side	NE	ASP	193	Side	OD1	23.78
ARG	11	Side	NH1	ASP	193	Side	OD1	22.98
SER	13	Side	OG	ARG	188	Main	O	59.54
LEU	14	Main	O	ARG	188	Side	NH2	99.30
LEU	14	Main	O	ARG	188	Side	NE	57.34
LEU	14	Main	N	GLN	200	Side	OE1	28.77
GLY	17	Main	O	SER	285	Side	OG	93.51
GLY	17	Main	N	GLU	288	Side	OE1	96.20
GLY	17	Main	N	GLU	288	Side	OE2	88.51
ARG	18	Side	NH2	ARG	30	Side	NH2	33.47
ARG	18	Side	NH1	ARG	30	Side	NH2	31.77
ARG	18	Side	NH1	ARG	30	Side	NH2	11.79
ARG	18	Side	NH2	ASP	181	Side	OD2	86.51
ARG	18	Side	NH1	ASP	181	Side	OD2	26.47
ARG	18	Side	NH1	ARG	183	Side	NH2	17.08
ARG	18	Side	NH1	ARG	183	Side	NH2	16.38
TYR	21	Main	O	MET	1	Main	N	61.44
TYR	21	Main	N	MET	1	Side	SD	20.18
TYR	21	Side	OH	GLU	277	Main	O	45.35
TYR	21	Side	OH	ASN	278	Side	OD1	15.98

THR	22	Side	OG1	MET	1	Main	N	96.10
THR	22	Side	OG1	MET	1	Main	O	20.48
GLN	25	Main	N	THR	8	Side	OG1	16.18
GLN	25	Side	OE1	SER	9	Main	N	13.19
GLN	25	Main	O	ASP	10	Main	N	57.24
GLN	25	Main	N	ASP	10	Side	OD2	10.69
GLN	25	Side	OE1	ASP	10	Main	N	19.18
GLN	25	Main	O	ASN	11	Main	N	70.63
GLN	25	Side	OE1	ASN	11	Main	N	22.08
GLN	25	Side	NE2	ASN	11	Side	OD1	37.86
VAL	27	Main	N	TYS	12	Main	O	70.03
VAL	27	Main	N	GLU	14	Side	OE1	16.38
VAL	27	Main	N	GLU	14	Side	OE2	14.99
ASP	29	Side	OD1	SER	18	Side	OG	11.59
ASP	29	Main	N	SER	23	Side	OG	71.33
ILE	30	Main	N	ASP	20	Side	OD1	12.19
ARG	31	Side	NH1	ASP	20	Main	O	14.69
ARG	31	Side	NH1	ASP	20	Side	OD1	16.78
ARG	31	Side	NH1	ASP	20	Side	OD2	15.38
ARG	31	Side	NH2	TYS	21	Main	O	16.58
ARG	31	Side	NH1	GLU	268	Side	OE1	73.13
ARG	31	Side	NH2	GLU	268	Side	OE1	72.23
ARG	31	Side	NH2	GLU	268	Side	OE2	70.53
ARG	31	Side	NH1	GLU	268	Side	OE2	41.66
LYS	32	Side	NZ	ASP	22	Side	OD1	55.54
LYS	32	Side	NZ	ASP	22	Side	OD2	37.56

Complex 17:

V3 loop Residue	Number	Group	Atom	CXCR4 Residue	Number	Group	Atom	Occupancy (%)
ARG	3	Side	NH2	TYS	21	Main	O	27.47
ARG	3	Side	NH1	TYS	21	Main	O	14.79
ARG	3	Side	NH2	ASP	22	Main	O	40.36
ARG	3	Side	NH2	SER	23	Main	O	10.49
ARG	9	Side	NH2	TYS	12	Side	OS3	19.28
ARG	9	Side	NH2	TYS	12	Side	OS2	13.09
ARG	9	Side	NH2	ASP	193	Side	OD1	22.28
ARG	9	Side	NE	ASP	193	Side	OD1	11.99
LYS	10	Side	NZ	GLU	2	Side	OE1	41.66
LYS	10	Side	NZ	GLU	2	Side	OE2	40.36
LYS	10	Main	O	LYS	25	Side	NZ	68.73
ARG	11	Side	NH2	TYS	12	Side	OS2	90.81
ARG	11	Side	NH2	TYS	12	Side	OS4	44.76
ARG	11	Side	NH2	TYS	12	Side	OS3	19.28
ARG	11	Side	NH2	ASP	193	Side	OD1	93.91
SER	13	Side	OG	ASP	187	Side	OD2	99.20
SER	13	Side	OG	ASP	187	Side	OD1	60.94
SER	13	Side	OG	PHE	189	Main	N	72.23
SER	13	Side	OG	PHE	189	Main	N	68.63
SER	13	Side	OG	TYR	190	Main	N	87.41
LEU	14	Main	O	ARG	188	Side	NE	91.81
LEU	14	Main	O	ARG	188	Side	NH2	88.91
GLY	17	Main	N	TYR	45	Side	OH	16.68
GLY	17	Main	O	SER	285	Side	OG	80.62
GLY	17	Main	N	GLU	288	Side	OE1	95.70
GLY	17	Main	N	GLU	288	Side	OE2	84.92
ARG	18	Side	NH1	ASP	181	Side	OD1	91.51
ARG	18	Side	NH2	ASP	181	Side	OD1	85.71
ARG	18	Side	NH2	HIS	281	Side	NE2	47.75
ARG	18	Side	NE	HIS	281	Side	NE2	46.75
ARG	18	Side	NH2	HIS	281	Side	ND1	29.27
TRP	20	Side	NE1	ARG	183	Side	NH2	27.47
TRP	20	Side	NE1	ARG	183	Side	NH2	25.57
TRP	20	Side	NE1	ARG	183	Side	NH1	9.59
TYR	21	Main	N	ASP	181	Side	OD2	19.18
TYR	21	Main	N	ASP	181	Side	OD1	14.99
TYR	21	Side	OH	GLU	277	Side	OE2	46.05
TYR	21	Side	OH	GLU	277	Side	OE1	34.97
TYR	21	Side	OH	ASN	278	Side	ND2	12.59
TYR	21	Side	OH	HIS	281	Side	ND1	10.79
THR	22	Side	OG1	TYR	190	Main	O	91.91
THR	23	Main	O	MET	1	Main	N	94.91
THR	23	Main	N	MET	1	Main	O	39.96
THR	23	Side	OG1	GLU	2	Side	OE2	47.95
THR	23	Side	OG1	GLU	2	Side	OE1	46.25

GLN	25	Side	NE2	SER	9	Side	OG	37.66
GLN	25	Main	O	ASN	11	Side	ND2	24.38
GLN	25	Side	OE1	ASN	11	Side	ND2	34.67
GLN	25	Side	NE2	ASN	11	Side	OD1	19.48
GLN	25	Main	O	TYS	12	Main	N	57.84
ILE	26	Main	O	ASN	11	Side	ND2	11.89
VAL	27	Main	N	ASN	11	Side	OD1	10.89
VAL	27	Main	N	TYS	12	Main	O	52.95
GLY	28	Main	N	GLU	14	Side	OE2	17.28
GLY	28	Main	N	GLU	14	Side	OE1	11.09
ARG	31	Side	NH1	GLU	14	Side	OE1	34.07
ARG	31	Side	NH1	GLU	14	Side	OE2	22.28
ARG	31	Side	NH2	GLU	14	Side	OE1	20.08
ARG	31	Side	NH2	SER	18	Side	OG	11.59
ARG	31	Side	NH2	SER	18	Side	OG	11.59
ARG	31	Side	NH2	ASP	20	Side	OD1	87.21
ARG	31	Side	NH1	ASP	20	Side	OD2	79.12
ARG	31	Side	NH2	ASP	20	Side	OD2	58.34
ARG	31	Side	NH1	ASP	20	Side	OD1	35.06
LYS	32	Side	NZ	GLU	268	Side	OE1	77.42
LYS	32	Side	NZ	GLU	268	Side	OE2	62.34

Supporting Table 4: The table consists of three individual panels divided by thick black frames. The left panel presents the average interaction energy per V3 loop residue in descending order, by summing up the interaction energies of the specific residue with all of its possible CXCR4 interacting residues. The middle and right panels present the (%) residue propensity per V3 loop position for all 35-residue CXCR4 and CXCR/CCR5 V3 loops, respectively, deposited in the Los Alamos National Laboratory database (http://www.hiv.lanl.gov). All values have been computed by analysis of 1000 snapshots, extracted from the 20-ns simulation of complex 1 at 20-ps intervals.

	Interaction Free Energies per V3-loop residue (kcal/mol)			Residue (%) Propensity for CXCR4 V3 loops				Residue (%) Propensity for CXCR4/CCR5 V3 loops			
Residue	Total	Polar	Non-Polar	Residue 1:	Residue 2:	Residue 3:	Residue 4:	Residue 1:	Residue 2:	Residue 3:	Residue 4:
Arg18	-124.03	-111.61	-12.42	Arg(60)	Gln(31)	Lys(4)	Gly(2)	Arg(57)	Gln(35)	Gly(2)	Lys(2)
Lys10	-57.74	-42.30	-15.44	Lys(63)	Arg(15)	Thr(12)	Gln(7)	Lys(68)	Arg(15)	Gln(9)	Thr(6)
Arg3	-42.21	-37.32	-4.89	Arg(100)	Ile(0)	Met(0)	Ser(0)	Arg(100)	Ile(0)	Met(0)	Ser(0)
Arg31	-37.87	-29.54	-8.33	Arg(96)	Lys(2)	Gly(1)	Ile(1)	Arg(98)	Ile(1)	Lys(1)	Val(0)
Arg11	-37.58	-15.91	-21.67	Ser(42)	Arg(30)	Gly(24)	Asp(1)	Ser(47)	Gly(27)	Arg(22)	Asp(1)
Arg9	-32.58	-16.99	-15.59	Arg(88)	Ile(4)	Lys(3)	Ala(2)	Arg(89)	Ala(3)	Ile(3)	Lys(3)
Trp20	-28.95	-8.94	-20.01	Phe(59)	Trp(10)	Val(9)	Leu(8)	Phe(60)	Trp(10)	Leu(9)	Val(8)
Leu14	-27.42	-7.00	-20.42	Ile(65)	Leu(18)	Met(11)	Ala(1)	Ile(70)	Leu(16)	Met(10)	Ala(1)
Lys32	-26.28	-24.74	-1.54	Gln(50)	Lys(31)	Arg(18)	Ile(0)	Gln(55)	Lys(28)	Arg(15)	His(1)
Tyr21	-19.79	-2.61	-17.18	Tyr(82)	Phe(8)	His(6)	Ser(2)	Tyr(81)	Phe(10)	His(7)	Val(0)
Pro16	-13.58	-0.73	-12.85	Pro(91)	Gln(4)	Gly(2)	Leu(1)	Pro(94)	Gln(3)	Gly(2)	Ala(0)
Gly15	-12.67	-6.48	-6.20	Gly(95)	Ile(2)	Ser(2)	Ala(0)	Gly(97)	Ile(2)	Ser(1)	Ala(0)
Ser13	-12.49	-6.72	-5.77	Arg(29)	His(27)	Thr(15)	Ser(14)	His(35)	Arg(28)	Thr(12)	Ser(10)
Asn5	-12.48	-6.14	-6.34	Asn(60)	Gly(17)	Ser(11)	Tyr(6)	Asn(62)	Gly(20)	Ser(11)	Tyr(5)
Gly17	-12.11	-6.12	-6.00	Gly(96)	Pro(1)	Gln(1)	Arg(1)	Gly(97)	Pro(1)	Arg(1)	Ala(0)
Gln25	-11.96	-1.05	-10.91	Asp(21)	Glu(19)	Gln(18)	Arg(12)	Asp(22)	Glu(18)	Gln(14)	Lys(13)
Gly24	-11.49	-4.43	-7.07	Gly(66)	Glu(8)	Arg(7)	Thr(6)	Gly(68)	Glu(8)	Thr(7)	Lys(4)
Val12	-10.31	0.50	-10.82	Ile(56)	Val(29)	Thr(5)	Leu(4)	Ile(59)	Val(29)	Thr(6)	Phe(2)
Thr8	-10.22	-1.96	-8.26	Thr(84)	Ile(9)	Lys(3)	Met(1)	Thr(92)	Ile(4)	Lys(1)	Arg(1)
Val19	-9.52	-0.98	-8.53	Ala(49)	Val(28)	Thr(15)	Arg(2)	Ala(50)	Val(23)	Thr(21)	Tyr(2)
Thr22	-8.33	-1.94	-6.4	Thr(43)	Ala(40)	Lys(7)	Arg(5)	Ala(48)	Thr(42)	Arg(3)	His(2)
Thr23	-8.32	-0.19	-8.13	Thr(87)	Ala(6)	Met(2)	His(1)	Thr(86)	Ala(7)	Arg(2)	Ser(2)
Asn6	-7.40	-3.53	-3.87	Asn(90)	Lys(4)	Tyr(2)	Asp(1)	Asn(95)	Asp(1)	Glu(1)	Lys(1)
Val27	-6.38	-0.22	-6.15	Ile(71)	Val(13)	Thr(11)	Glu(1)	Ile(80)	Val(12)	Thr(7)	Met(1)
Ile26	-5.90	-0.33	-5.57	Ile(95)	Val(3)	Gly(1)	Lys(1)	Ile(95)	Val(3)	Gly(1)	Lys(0)
Asn7	-4.90	-2.14	-2.76	Asn(84)	Lys(5)	Tyr(3)	Thr(2)	Asn(93)	Lys(2)	Tyr(2)	Ile(1)
His34	-2.99	-0.71	-2.28	His(65)	Tyr(34)	Arg(0)	Ile(0)	His(72)	Tyr(27)	Val(0)	His(0)
Cys1	-0.71	-0.68	-0.04	Cys(100)	-	-	-	Cys(100)	-	-	-
Pro4	-0.67	-0.07	-0.61	Pro(100)	Ile(0)	Met(0)	Ser(0)	Pro(100)	Ile(0)	Met(0)	Ser(0)
Ile30	-0.44	-0.15	-0.29	Ile(96)	Met(1)	Asn(1)	Thr(1)	Ile(97)	Asn(1)	Thr(1)	Val(1)
Gly28	-0.32	0.06	-0.38	Gly(99)	Lys(1)	Thr(0)	Glu(0)	Gly(99)	Lys(1)	Thr(0)	Met(0)
Thr2	-0.19	-0.03	-0.16	Thr(79)	Ile(9)	Met(3)	Ser(3)	Thr(74)	Ile(12)	Met(4)	Ser(4)
Ala33	-0.16	0.02	-0.18	Ala(98)	Pro(1)	Arg(0)	Ile(0)	Ala(98)	Pro(1)	Val(1)	His(0)
Cys35	0.98	1.03	-0.06	Cys(100)	-	-	-	Cys(100)	-	-	-
Asp29	1.04	1.25	-0.21	Asp(83)	Asn(11)	Tyr(2)	Gly(1)	Asp(85)	Asn(13)	Gly(1)	Met(0)

Left panel: The interaction energies are decomposed into polar and non-polar components. Middle and right panels: The four most probable V3 loop residues per position (according to the descending order of the left panel) are presented, and the corresponding residue propensity is

shown in parenthesis. If a residue in the middle and right panels, matches the corresponding residue of the left panel (V3 loop of this study), it is highlighted in grey background. If a residue in the middle and right panels, possesses similar physicochemical properties to the corresponding residue of left panel (V3 loop of this study), it is highlighted in light green background.

Supporting Coordinates

The MD coordinates, extracted every 2 are provided as Supporting Information in PDB format. The structures are aligned with regard to the backbone of the CXCR4 intramembrane helical region. The correspondence of PDB files and time in the MD simulation is as follows:

PDB file	Time (ns)
v3loop_cxcr4.1.pdb	2
v3loop_cxcr4.2.pdb	4
v3loop_cxcr4.3.pdb	6
v3loop_cxcr4.4.pdb	8
v3loop_cxcr4.5.pdb	10
v3loop_cxcr4.6.pdb	12
v3loop_cxcr4.7.pdb	14
v3loop_cxcr4.8.pdb	16
v3loop_cxcr4.9.pdb	18
v3loop_cxcr4.10.pdb	20

Supporting Video

A video demonstrating the simulation trajectory of the V3 loop : CXCR4 complex structure, and depicting the gradual stabilization and preservation of the key salt bridges is provided as Supporting Video.

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, the corresponding heavy atom distances are denoted in Å, and the participating V3 loop and CXCR4 residue moieties are shown in licorice. The Cα atoms of V3 loop and CXCR4 salt bridge participating residues are annotated. Hydrogen atoms are omitted for clarity and the V3 loop disulfide bridge is shown in fat transparent licorice representation.

Supporting References

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

2. Neria, E., S. Fischer, and M. Karplus. 1996. Simulation of Activation Free Energies in Molecular Systems. *J. Chem. Phys.* 105:1902-1921.

3. Ryckaert J.P., G.*Ciccotti, and H.J.C. Berendsen.* Numerical integration of the Cartesian Equations of Motion of a System with Constraints: Molecular Dynamics of n-Alkanes. *J. Comput. Phys.* 23:327-341.

4. MacKerell, A. D. Jr, D. Bashford, M. Bellott, R. L. Dunbrack Jr, J. D. Evanseck, M. J. Field, S. Fischer, J. Gao, H. Guo, , S. Ha, D. Joseph-McCarthy, L. Kuchnir, , K. Kuczera, F. T. K. Lau, C. Mattos, S. Michnick, T. Ngo, D. T. Nguyen, B. Prodhom, W. E. Reiher III, B. Roux, M. Schlenkrich, , J. C. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiorkiewicz-Kuczera, D. Yin, and M. Karplus. 1998. All-atom empirical potential for molecular modeling and dynamics Studies of proteins. *J. Phys. Chem. B* 102:3586-3616.

5. Buck, M., S. Bouguet-Bonnet, R. W. Pastor, and A. D. MacKerell. 2005. Importance of the CMAP correction to the CHARMM22 protein force field: Dynamics of Hen Lysozyme. *Biophys. J.* 90, L36-38.

6. Vanommeslaeghe K., E. Hatcher, C. Acharya, S. Kundu, S. Zhong, J. Shim, E. Darian, O. Guvench, P. Lopes, I. Vorobyov, and A. D. MacKerell Jr. 2010. CHARMM General Force Field (CGenFF): "A force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields", *J. Comput. Chem.* 31:671-690.

7. Seeber, M., M. Cecchini, F. Rao, G. Settanni, and A. Caflisch. 2007. WORDOM: A program for efficient analysis of molecular dynamics simulations. *Bioinformatics* 23, 2625-2627.

8. Brooks, B. R., C. L. Brooks, III, A. D. Mackerell Jr., L. Nilsson, R. J. Petrella, B. Roux, Y. Won, G. Archontis, C. Bartels, S. Boresch, , A. Caflisch, L. Caves, Q. Cui, A. R. Dinner, M. Feig, S. Fischer, J. Gao, M. Hodoscek, W. Im, K. Kuczera, T. Lazaridis, J. Ma, V. Ovchinnikov, E. Paci, R. W. Pastor, C. B. Post, J. Z. Pu, M. Schaefer, B. Tidor, R. M. Venable, H. L. Woodcock, X. Wu, W. Yang, D. M. York, and M. J. Karplus. 2009. CHARMM: the biomolecular simulation program. *J. Comput. Chem.* 30:1545–1614.

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

10. Wu, B., E. Y. T. Chien, C. D. Mol, G. Fenalti, W. Liu, V. Katritch, R. Abagyan, A. Brooun, P. Wells, F. C. Bi, D. J. Hamel, P. Kuhn, T. M. Handel, V. Cherezov, and R. C. Stevens. 2010. Structures of the CXCR4 chemokine GPCR with small-molecule and cyclic peptide antagonist. *Science* 330: 1066-1071.