Published in final edited form as:

Biochemistry. 2012 October 2; 51(39): 7783–7793. doi:10.1021/bi300878d.

Spontaneous Rearrangement of the β 20/ β 21 Strands in Simulations of Unliganded HIV-1 Glycoprotein, gp120

Indira Shrivastava^{*}, Kaylee Wendel[#], and Judith M. LaLonde[#]

Indira Shrivastava: ihs2@pitt.edu; Judith M. LaLonde: jlalonde@brynmawr.edu

^{*}Department of Systems and Computational Biology, School of Medicine, University of Pittsburgh, 3083 Biomedical Science Tower 3, 3501 Fifth Avenue, Pittsburgh PA 15213

[#]Chemistry Department, Bryn Mawr College, 101 N. Merion Avenue, Bryn Mawr, PA 19010

Abstract

Binding of the viral spike drives cell entry and infection by HIV-1 to the cellular CD4 and chemokine receptors with associated conformational change of the viral glycoprotein envelope, gp120. Crystal structures of the CD4-gp120-antibody ternary complex reveal a large internal gp120 cavity formed by three domains, the inner domain, outer domain and bridging sheet domain, and is capped by CD4 residue Phe43. Several structures of gp120 envelope in complex with various antibodies indicated that the bridging sheet adopts varied conformations. Here, we examine bridging sheet dynamics using a crystal structure of gp120 bound to the F105 antibody exhibiting an open bridging sheet conformation and with an added V3 loop. The two strands of the bridging sheet β 2/ β 3 and β 20/ β 21 are dissociated from each other and are directed away from the inner and outer domains. Analysis of Molecular Dynamics (MD) trajectories indicates that the β 2/ β 3 and β 20/ β 21 strands rapidly rearrange to interact with the V3 loop and the inner and outer domains, respectively. Residues N425 on β 20 leads to the conformational rearrangement of the β 20/ β 21 strands by interacting with W112 on the inner domain and F382 on the outer domain. An accompanying shift is observed in the inner domain as helix α 1 exhibits a loss in helicity and pivots away from helix α 5. The two simulations provide a framework for understanding the conformational diversity of the bridging sheet and the propensity of the β 20/ β 21 strand to refold between the inner and outer domains of gp120, in the absence of a bound ligand.

Keywords

HIV; gp120; CD4 binding; Conformational Change; Bridging Sheet; CD4-Antibody binding; Chemokine Receptor; Molecular Dynamics; protein-ligand interactions; HIV entry inhibitor

Infection of HIV-1 begins with a series of dynamic binding events between the trimeric glycoprotein envelope spike and the host cell CD4 and chemokine receptors.^{1–7} The envelope trimer (gp160) is composed of three gp120 glycoproteins and three transmembrane gp41 proteins.^{8–10} The first dynamic event occurs via binding of gp120 to the host T-cell CD4 receptor^{3, 11} followed by extensive restructuring of gp120.^{12, 13} This conformational change results in the exposure of the chemokine binding site on gp120, thus permitting binding to either of the chemokine receptors, CCR5 or CXCR4.^{14–17} Upon chemokine receptor binding a second conformational change occurs in gp41 to form the fusion peptide

CORRESPONDING AUTHOR FOOTNOTE Indira Shrivastava, ihs2@pitt.edu, 412 383 5806 (Voice), 412 648 3163 (Fax).

Supporting Information Available

Supplementary Figures S1–S9 and the MD movies are available free of charge via the Internet at <http://pubs.acs.org>.

that inserts in the host cell membrane, leading to viral entry.^{6, 18–20} The CD4 induced gp120 conformational change has been characterized thermodynamically, showing a highly favorable binding enthalpy balanced with a highly unfavorable molecular ordering.^{21, 22} This thermodynamic signature resembles protein folding, rather than binding, and reflects the large molecular ordering of gp120 upon CD4 binding.²¹ A similar thermodynamic signature is exhibited by soluble CD4 (sCD4) binding to both full-length gp120 (gp120_{full}) and a core gp120 (gp120_{core}) containing truncations in gp120 variable loops.^{22–24} Furthermore, the large entropic penalty associated with CD4 binding was substantially reduced for a cavity filling mutant (S375W) indicating that this mutant gp120 is stabilized in a CD4 bound-like state.^{25, 26}

There are several gp120_{core}-structures bound with CD4 receptor,^{27–31} antibodies,^{32–35} mini-protein^{36, 37} and small molecule ligands^{38, 39} detailing gp120-ligand interactions. The CD4-gp120_{core} crystal structures reveal the CD4 induced formation of a large internal gp120 cavity formed by the inner, outer and bridging sheet domains (the CD4 bound conformation of gp120 is denoted as gp120_{CD4}) (Figure 1A).^{27, 28} The highly conserved CD4 Phe43 side-chain binds at the top of the gp120 “Phe43 cavity”. The structure of the biphenyl conjugated scyllatoxin derived mini-protein (2I5Y)³⁶ reveals the depth of the cavity and biphenyl interactions with gp120 aromatic residues lining the Phe43 cavity from each of the three domains, namely, W112 (inner domain) F382 and Y384 (outer domain), and W427 (bridging sheet). In gp120_{CD4}, the bridging sheet domain is composed of a four-stranded antiparallel β -sheet (β_2 , β_3 , β_{20} , β_{21}). The β_2/β_3 strands stem from the inner domain, while the β_{20}/β_{21} strands arise from the outer domain. Thus, the bridging sheet spans the interface of both the inner and outer domain and forms a third of the surface of the Phe43 cavity. As yet, structural information of the pre-structured, unbound form of trimeric gp120 has been elusive. However, the structure of the unbound form of SIV gp120 has been solved revealing a structurally invariant outer domain with a markedly different conformation and arrangement of the inner domain and bridging sheet domains compared to HIV-1 gp120_{CD4} structures.⁴⁰ In fact the structure of gp120_{CD4} containing portions of the gp41-interacting region, displays a three-layered domain architecture adjacent to the invariant outer domain.³¹ This topological layering of the gp41-interacting elements, the inner domain and bridging sheet domain is postulated to form a “shape-changing spacer” that forms the basis of glycoprotein conformational mobility.^{31, 41} The plasticity of monomeric gp120_{core} in complex with several monoclonal antibodies, b12 (2NY7),²⁹ b13 (3IDX), and F105 (3HI1),³² and the N-terminal of the CCR5 receptor (2QAD)⁴² also reveal the structural variability of the inner domain, bridging sheet and variable loops.⁴³ The gp120 bound N-terminal of the CCR5 receptor (2QAD), which is the structure of tyrosine-sulfated 412d antibody complexed with HIV-1 YU2 gp120 and CD4, reveals the structure of the V3 loop situated adjacent to the bridging sheet. This highly variable, glycosylated, 35 residue V3 loop determines HIV-1 trimer preference for binding to either the CCR5 or CXCR4 chemokine receptors.^{44–46} Furthermore, Yokoyama et al. found that the V3 loop acts as an electrostatic modulator that impacts the fluctuation and conformation of CD4, coreceptor and antibody binding loops.⁴⁷ Other disulfide bond linked gp120_{core} variants complexed to b12 and b13 display an altered conformation compared to the CD4-bound state with shifts in the V1/V2 stem and alternations in the position and orientation of the bridging sheet β_2/β_3 and β_{20}/β_{21} strands. Furthermore, the b12-gp120_{core} complex displays an unstructured helix α_1 in the inner domain. This is in contrast to the F105-gp120_{core} (gp120_{F105}) complex (Figure 1B), which shows a structured helix α_1 , but with narrowing of the interface between the inner domain and outer domain. This narrowing is accompanied by the displacement and disassociation of the 30 residues of the four-stranded anti-parallel β -sheet bridging sheet into distinct conformations for the β_2/β_3 and β_{20}/β_{21} strands. The two strands extend away from the inner-outer domain interface resulting in a shallower and more open “Phe43 cavity” bound by the F105 antibody Val-Phe-Tyr motif (Figure 1B). Thus, these antibody-

stabilized complexes represent gp120 conformations that may be sampled along the transition path between the unliganded and the CD4 bound form. Recently, Kwon et al. have solved the structure of the unliganded gp120 extended core (gp120_{coreE}) from clades B, C and E and demonstrate that while the glycoprotein trimer spike rarely assumes a CD4-bound like conformation, gp120_{coreE} spontaneously adopts the CD4 bound conformation in the absence of ligand when not restricted by the presence of the V1, V2 and V3 variable loops.³⁸ Therefore, ligand binding is not a prerequisite for the creation of the gp120_{CD4} conformation, suggesting that this conformation may indeed be sampled in solution in the context of the gp120_{core}.

The conformational dynamics of gp120 has been analyzed using various computational models reiterating the bridging sheet flexibility seen in gp120-antibody crystal structures. Floppy Inclusion and Rigid Substructure Topography analysis of several gp120 structures bound with various ligands demonstrated that the inner domain and bridging sheet domains are more flexible while the outer domain is more rigid.⁴⁸ Several simulations of gp120 bound with CD4, mini-protein or small molecule using the coarse-grained Gaussian network model also concluded that the outer domain was more rigid and exhibited less flexibility, than the inner domain and the bridging sheet.⁴⁹ Application of temperature accelerated MD predicted a counter rotation between the inner and outer domains and a disruption of the bridging sheet in the unbound form of gp120.⁵⁰ Additional, MD studies by Hsu et al. indicate that there are concerted loop motions in the vestibule of the CD4 cavity, stabilization of the bridging sheet and a coalescing of the bridging sheet and V3 loop to form the co-receptor binding site.^{51, 52} MD studies demonstrated that the bridging sheet refolds for a S375A mutant,⁵³ while the bridging sheet did not refold for either a S375W cavity mutant⁵⁴ or double mutant, S375W/T257.⁵³ Furthermore, Liu et al. showed that a gp120 homology model containing the S375W mutation favors the CD4 bound conformation, while the I423P mutation prefers an unliganded conformation.⁵⁵ Non-equilibrium steered MD showed that the bridging sheet strands $\beta 2/\beta 3$ had more flexibility than the $\beta 20/\beta 21$ strands, which preferred interactions with the inner domain.⁵⁶ MD simulation of gp120 showed that the binding of the small molecule, NBD-556, within the Phe43 cavity enhances the dynamics of gp120 with increased mobility, especially in the outer domain, as compared to the CD4 bound or mini-protein bound complex.⁵⁷ More recently, Yokoyama et al. describe MD simulations of gp120 in the CD4 bound form containing the V3 loop, concluding that the V3 loop acts as an electrostatic modulator that impacts the fluctuation and conformation of CD4, coreceptor and antibody binding loops.⁴⁷

The thermodynamic, crystallographic and dynamic studies all suggest that gp120 is in equilibrium among a number of conformational states, which when bound by ligand (CD4, antibody, mini-protein or small molecule) can drive the equilibrium toward a conformational state that favors binding to the chemokine co-receptor. However, the molecular mechanism of the transition between these conformational states remains undetermined. Nonetheless, as revealed by the CD4-gp120 structure containing portions of the gp41-interacting region³¹ and the gp120_{F105} structure (3HI1),³² the inner domain and bridging sheet are most likely to facilitate/mediate conformational changes compared to the outer domain, which has thus far been observed in a conserved conformation. We postulate that the extended conformations of the $\beta 2/\beta 3$ and $\beta 20/\beta 21$ as observed in the gp120_{F105} structure represent one potential conformation of the metastable unliganded gp120 monomer. Moreover, the observation that the unliganded gp120_{coreE} adopts the gp120_{CD4} conformation without induction by ligand binding suggested further investigation of inter-domain interactions that may contribute to gp120 conformational transitions. In this study, we use MD simulations to explore the conformational rearrangement of the bridging sheet as observed in the unliganded gp120_{F105} structure containing the V3 loop. The two 100 ns simulations reported herein, demonstrate that the $\beta 2/\beta 3$ and $\beta 20/\beta 21$ strands and the V3 loop rearrange spontaneously, leading to

reorganization of the $\beta 20/\beta 21$ strand wedged between the inner domain and outer domain accompanied by a shift in the position of helix $\alpha 1$ on the inner domain.

Materials and Methods

For this study, we used the crystal structure of YU2 gp120-core complexed to the F105-antibody (3HI1),³² The conformation of the third variable loop (V3) was revealed in the crystal structure of the gp120 complexed to the tyrosine-sulfated 412d antibody (2QAD).⁴³ Superposition of the outer domain of these two structures indicated that the V3 loop could be grafted onto the V3 stem in the outer domain of 3HI1 without disrupting the extended position of the bridging sheet strands $\beta 20/\beta 21$. Using the homology modeling module in MOE⁵⁸ the V3 loop from 2QAD was grafted onto the corresponding YU2 sequence in context of the 3HI1 structure. This loop was added to provide completeness to the monomer structure and to assess the role of adjacent V3 loop on bridging sheet dynamics. The missing residues corresponding to the V4 loop in 3HI1 were added from the 1G9N²⁷ crystal structure and the resulting structure was energy minimized. Hydrogen atoms were added and the tautomeric states and orientations of Asn, Gln and His residues were determined with Molprobit (<http://molprobit.biochem.duke.edu/>).^{59, 60} The construct was then minimized using the OPLS-AA force field⁶¹ in MOE⁵⁸ with the following procedure: all hydrogen atoms were minimized to a root mean square (rms) gradient of 0.01, holding the heavy atoms fixed. A stepwise minimization followed for all atoms, using a quadratic force constant ($100 \text{ kcal mol}^{-1} \text{ \AA}^{-1}$) to tether the atoms to their starting geometries; for each subsequent minimization, the force constant was reduced by a half until $0.25 \text{ kcal mol}^{-1} \text{ \AA}^{-1}$. This was followed by a final cycle of unrestrained minimization. Water molecules, Asn-linked acetyl-D-glucosamine, 2-(acetylamino)-2-deoxy-a-d-glucopyranose and small molecules were removed prior to MD calculations.

Molecular Dynamics Simulations

The model of gp120, including the bridging sheet and the V3 loop, in explicit solvent, was initially energy minimized, using the steepest descent algorithm, in GROMACS, (<http://www.gromacs.org>) software.⁶² This energy minimized model included explicit SPC⁶³ water molecules resulting in a simulation model of ~55000 atoms. After energy minimization, the system was equilibrated for 1 nanosecond (ns), during which the backbone atoms were restrained by a harmonic potential of $1000 \text{ kJ mol}^{-1} \text{ nm}^{-1}$, while the side-chain atoms and the water molecules were allowed to relax. This was then followed by a production run of 100 ns, during which all restraints were removed. The simulations were all performed under constant number of particles, pressure and temperature conditions with temperature set at 310K. A weak temperature coupling using Berendsen's bath with a time constant of 0.1 picosecond (ps) was used, while a pressure coupling of 1 atm/bar with a time constant of 1ps was applied, also using the Berendsen's algorithm.⁶⁴ The LINCS algorithm⁶⁵ was used to constrain bonds involving hydrogen atoms and the integration time step was set to 2 femto seconds. The electrostatic interactions were calculated using Ewald particle mesh method.⁶⁶ Two 100 ns simulation trajectories were generated from this energy minimized and equilibrated system each started off with a different set of random initial velocities. In the following discussion, these will be denoted as MD-1 and MD-2. The simulation trajectories were all generated on the supercomputing resources available at the Pittsburgh Center for Simulation and Modeling (<http://www.sam.pitt.edu>).

Results and Discussion

gp120 Exhibits a Conformational Transition After 40ns of Simulation

The unliganded monomer gp120 core extracted from the 3HI1 complex³² was used for simulations. The V3 loop was grafted from the gp120-412d antibody structure (2QAD)⁴³ to provide completeness to the gp120 monomer structure (gp120_{F105-V3}) and to assess its impact on bridging sheet dynamics Figure 1 (see materials and methods). Two 100 ns simulations were conducted for gp120_{F105-V3} and are denoted MD-1 and MD-2. The root mean square deviations (rmsd), of the C- α atoms, typically provide a measure of the stability of the protein, as a function of time during the course of simulation. In both MD-1 and MD-2, the rmsd, initially increases up to almost 0.8 nm (Figure 2A). The high rmsd in both the simulations indicates substantial conformational change from the starting structure. In MD-1, an abrupt transition is observed near 25 ns, where the rmsd jumps from ~0.5 nm to ~0.8 nm, while in the second simulation, MD-2, the rmsd steadily increases up to 0.8 nm during the first 10 ns of the simulation. Superposition of the 100 ns averaged structures of MD-1 and MD-2 onto their starting structures (Figure 2B and C), clearly shows that the change in conformation is primarily due to motion of the bridging sheet β 2/ β 3 and β 20/ β 21 strands and the V3 loop, in both MD-1 (Figure 2B) and MD-2 (Figure 2C). In both simulations, the β 2/ β 3 strands and V3 loops move towards each other, forming non-covalent interactions. This finding is interesting given that the V3 loop determines gp120 usage of either the CCR5 or CXCR4 co-receptor⁴⁴⁻⁴⁶ and further suggests that V3 loop sequence variability may effect V3 structure and mobility as previously described.^{47, 67} The β 20/ β 21 strand, on the other hand, moves towards W112 on inner domain helix α 1, closing off the wall of the Phe43 binding cavity between the inner and outer domains. A correlation coefficient of 0.832 in root mean square fluctuations (from the respective 100ns average) (Supplementary Figure S1), suggests remarkably similarity between the two simulations. Moreover, the 100 ns averaged structure, corresponding to the folded structure, in both MD-1 and MD-2, are also remarkably similar, with a low rmsd of 0.45 nm (Supplementary Figure S2). In both the simulations, the topology and structure of gp120 core outer domain remains similar to the initial structure, while both the β 2/ β 3 and β 20/ β 21 strands and V3 loop dramatically change. Interestingly, helix α 1 in the inner domain has also shifted from the starting position in MD-1 and MD-2 and has partially lost helicity (Figure 2B and C). The transition point, observed within 30 ns in both MD-1 and MD-2 (Figure 2A), is triggered by the movement of β 20/ β 21 towards the inner domain helix α 1 and the formation of a stabilizing interaction between W112 and W427 (MD-1) or W112 and Y435 (MD-2) followed by the concerted motions of β 2/ β 3 and V3 loops towards (Figure 2B and C) each other.

Events Leading to Bridging Sheet Rearrangement

To assess the hierarchy of folding events, we monitored the time-evolution of the minimum distances between any two atoms on residues W427 or Y435 on β 20/ β 21 and W112 on inner domain helix α 1 (Figure 3A and B). These residues were considered since they form part of an aromatic cluster (W112, F210, F382, W427, Y435) that coalesces between the inner, outer and bridging sheet upon formation of the Phe43 cavity as observed in gp120_{CD4} structures. In MD-1, the minimum distance between W112 and W427 initially increases (Figure 3A, black line), however, within 10 ns this distance drops to ~0.5 nm for up to 40 ns and then finally to ~0.25 nm for the remainder of the simulation, suggesting a close interaction between W112 and W427. In MD-2 (Figure 3A, red line), residues W112 and W427 are farther apart (0.75 nm) during the first 40 ns, then later increases to ~1.0 nm, as a result of the W427 aromatic ring rotating away from W112. Closer inspection of W112 interactions with β 20 indicates that in MD-2, the helix α 1- β 20 interaction is stabilized instead by W112- Y435 interaction (0.25 nm) (Figure 3B, red line) maintaining the cavity

like structure. In MD-1, W427 approaches W112 in a T-stacking aromatic-aromatic interaction but remains on the outer flank of W112 and the Phe43 cavity (Supplementary material, WT1_MD_movie). In the MD-2 simulation, after both W427 and Y435 form T-stacking interactions with W112, W427 then slides over the surface of W112 to settle over the top of the cavity within van der Waals interaction distance of C- β atom of S375 on the outer domain (Supplementary material, WT2_MD_movie). Thus, monitoring interactions of W112 with W427 and Y435 indicates that the β 20/ β 21 rearranges rapidly (within 40 ns) to form close associations with inner and outer domain residues, enclosing the Phe43 cavity.

A study by Xiang et al. concluded that the V3 loop and bridging sheets assist in stabilizing the unliganded conformation of the gp120 trimer.⁶⁸ Kwon et al. further posit that gp120_{coreE} exhibits an altered conformational equilibrium as a consequence of the V3 loop truncations.³⁸ To assess the possible role of the V3 loop in assisting bridging sheet dynamics, in the absence of a bound ligand, the interaction distance was monitored between Q203 on β 2/ β 3 and Y318 on the V3 loop (Figure 3C). These two residues were selected because they are observed to be in close contact (less than 0.5 nm) at the end of the simulation (Supplementary Table S1 and movies). In MD-1 (Figure 3C, black line) β 2/ β 3 and V3 initially shift apart, then at approximately 30 ns, β 2/ β 3 and V3 move within 0.5 nm, and remain associated for the duration of the simulation. In contrast, in MD-2, association between the β 2/ β 3 and the V3 loop is seen to be formed almost instantaneously, with the two residues approaching within ~0.25 nm of each other. Analysis of the movies (Supplementary Material, Movies) and the interaction plots for residue pairs W112:W427 (Figure 3A) and W112:Y435 (Figure 3B) and Q203:Y318 (Figure 3C) reveal that in both MD-1 and MD-2 the β 2/ β 3 and β 20/ β 21 strands and V3 loop reorganize within the same time frame. The crystal structures of the unliganded gp120_{coreE}³⁸ indicate that residues in the stem of the truncated V3 loop stabilize the gp120_{CD4} conformation by forming hydrogen bonds with the main-chain atoms of the β 20/ β 21 strands. In the case of gp120_{F105-V3}, MD-1 and MD-2 trajectories reveal only transiently formed hydrogen bonds between V3 loop and β 20/ β 21 strands during simulation. Thus, the rearrangement of β 20/ β 21 in the absence of ligand in less than 40 ns is not impeded by V3 loop interaction. Nonetheless, the V3 loop closely associates with the β 2/ β 3 throughout the simulations, likely preventing β 2/ β 3 interaction with β 20/ β 21. Thus, the β 20/ β 21 strand spontaneously rearranges to enclose the Phe43 cavity via close interactions with the inner and outer domain residues.

Role of Conserved N425 in Driving Bridging Sheet Rearrangement

We next assessed interaction distances for all residues pairs between the β 20/ β 21 strand and residues from inner domain and the outer domain within the vicinity of the Phe43 cavity. A strong interaction, is defined as an interaction in which the minimum distance between any two atoms of the residue pair is less than or equal to 0.4nm, and the distance is maintained over a period of at least 20ns or more. The residue pairs with the strongest interactions are listed in Table 1. Both N425 and K421 have the strongest interactions with the outer domain throughout the simulation. The inter-residue distances (N425 to W112 and N425 to F328) were monitored for MD-1 and MD-2 (Figure 4A–C). Further, inspection of the MD-1 and MD-2 trajectories reveals that N425 assists in the rearrangement of β 20/ β 21 (Supplementary Material, Movies). As shown in the 100 ns snapshots for MD-1 and MD-2 (Figure 4D and F) N425 rearranges to interact with F382 in the outer domain over the course of the simulation. The transition begins within the first 4 ns of the simulation, with the side-chain of N425 reversing its orientation from solvent exposed to interacting with W112 on the inner domain. In the MD-1 simulation (Figure 4A, black line), N425 remains in contact with W112 (0.25 nm) for ~35 ns and then moves across the Phe43 cavity and interacts with F382 (Figure 4A, red line and 4D) and S375 in the outer domain for the remainder of the simulation. In MD-2, after brief contact (~5 ns) with W112, N425 shifts to interact with

F382 for the duration of the simulation (Figure 4B red line and 4F). The N425-F382 interaction is not observed in the CD4 bound form of gp120 (for example in 1G9N); instead N425 hydrogen bonds to E370 anchoring the bridging sheet to the outer domain in gp120_{CD4} (Figure 4E and Supplementary Figure S3). Interestingly, Residue N425 is 75% conserved among HIV strains and is most often substituted by arginine. Furthermore, a gp120 triple mutation, I423M/N425K/G431E, exhibits both reduced soluble CD4 and antibody 17b binding²⁶ adding support to the role of N425 strand in mediating bridging sheet reorganization. Therefore, N425 appears to initiate $\beta 20/\beta 21$ rearrangement by first interacting W112 on the inner domain and then F382 on the outer domain. This rearrangement then positions W427 and Y435 in close proximity of W112, early on during the simulation, stabilizing $\beta 20/\beta 21$ interaction with the inner and outer domains.

The pivotal role of N425 led us to examine adjacent $\beta 20$ residues (K421, N422, I423, I424 and M426) as these are all in close proximity to the outer domain in gp120_{CD4}. In the gp120_{F105} crystal structure K421 is within interacting distance of F382 and Y384 in the outer domain (Table 1). In both MD-1 and MD-2, K421-Y384 interaction is maintained (Table 2 and Supplementary Figure S4). The K421-F382 interaction is maintained in MD-2, while in MD-1 it dissipates at ~40 ns. In MD1, the key bridging-sheet/inner-domain interactions are all maintained over 60% of the simulation period, with an average standard deviation of 0.1nm (Table 2). In MD2, except for W112-W427, the remaining bridging-sheet/inner-domain key interactions are all maintained over 97% of the simulation time (Table 2) with an average standard deviation of 0.05 nm (Table 2).

Clustering analysis, revealed two major clusters in MD1, and MD2 (Supplementary Figure S5). The clustering was performed using the Jarvis-Patrick method, which uses the nearest neighbor approach.⁶⁹ In this method, the similarity is determined using a distance metric and the neighbors of a given structure are the closest structures within a cutoff (0.1nm). In MD1, the major structural difference between the two clusters is in the bridging-sheet conformation (Supplementary Figure S5, A–B) whereas in MD2 a less dramatic difference is observed between the orientations of the $\beta 20/\beta 21$ strands (Supplementary Figure S5, C–D). The key bridging-sheet/inner-domain interactions excluding the W112-W427 interaction are typically weaker in cluster 2 of MD1 (~0.4 nm) than in cluster 2 of MD2 (~0.22 nm) (Table 2).

The next residue on $\beta 20$, Q422, forms a hydrogen bond with E381 and is solvent exposed in the gp120_{F105} crystal structure, however, this interaction is not maintained and Q422 remains solvent exposed throughout the MD-1 and MD-2 simulations. The behavior of neighboring $\beta 20$ residues, I423 and I424, varied between the two simulations. In MD-1, I423 has close interactions with F382, while in MD-2 I424 maintains interaction with F382 suggesting both I423 and I424 have the capacity to stabilize interactions with the outer domain. Residue M426, rather than interacting with the outer domain, remains closely associated with W427 while interacting with W112 as well. To validate the importance of $\beta 20$ residues in driving bridging sheet rearrangement, gp120_{F105-V3} containing the I423P mutation was simulated for 100 ns. Xiang et al. introduced the I423P mutant to constrain the conformation of the bridging sheet and found that the mutant eliminated both CD4 and CD4i antibody binding.²⁶ Examination of the I423P trajectory demonstrates that I423P interacts with F328 and W112 blocking both the interaction of N425 with F382 and W112 as well as the refolding of $\beta 20/\beta 22$ between the inner and outer domains (Supplementary Figure S6) consistent with observations reported by Liu et al.⁵⁵ Thus, we posit that the $\beta 20/\beta 21$ strands initiate gp120 conformational change that leads to a CD4 binding competent conformation and that there is a natural propensity of the $\beta 20/\beta 21$ to reside between the inner and outer domains in the preformed Phe43 cavity. The finding that $\beta 20/\beta 21$ reorients between the inner and outer domain is generally consistent with the conformation exhibited by $\beta 20/\beta 21$

in the only crystal of unliganded gp120, from SIV (Supplementary Figure S7) and suggests that $\beta 20/\beta 21$ has a tendency to adopt this conformation. The results of the simulations are further supported by experiment: mutant gp120 envelope proteins containing K421D, N422L, I423S, or I424S on $\beta 20$ reduced binding to both soluble CD4 and CCR5;⁶⁸ and deletion of both the $\beta 20/\beta 21$ strands enhances both F105 and b12 antibody binding to gp120 suggesting a reduction of steric hindrance with its removal.⁷⁰ As the initial distance between N425, and W112 on helix $\alpha 1$ is 11.2 Å, a longer-range electrostatic interaction as noted by Kwong et al. may be triggering the rearrangement.²⁸

Inner Domain Conformational Changes

The inner domain has been described as containing three structurally mobile layers, comprised of helix $\alpha 0$, helix $\alpha 1$ and helix $\alpha 5$ ³¹ (we note that helix $\alpha 0$ is not observed in the gp120_{F105} crystal structure). Interestingly, helix $\alpha 1$ is partially unfolded in the b13-gp120 and completely unfolded in the b12-gp120 cocrystal structures.^{29, 32} When comparing the position of helix $\alpha 1$ and helix $\alpha 5$ in gp120_{F105} versus gp120_{CD4}, the two helices are further apart in the gp120_{CD4} implying that the helix $\alpha 1$ and $\alpha 5$ shift relative to each other in order to accommodate the formation of the bridging sheet and the Phe43 cavity. Furthermore, previous computational analysis posited that H105 on helix $\alpha 1$ and M475 on the tip of helix $\alpha 5$ behave like a hinge region, stabilizing interactions with the bridging sheet.⁴⁹ Hence, we monitored the distances between W427 on $\beta 20/\beta 21$ and helices $\alpha 1$ (W112) and $\alpha 5$ (M475). The distance plot between W427-M475 indicates that W427 moves from 2.0 nm to within 0.5 nm of M475, in less than 40 ns in both MD-1 and MD-2 (Figure 5A–C, red line). Interaction between W427 and H105 on helix $\alpha 1$ is weaker for both MD-1 and MD-2 (Figure 5A–C, black line). The distances between M475 and H105 or W112 as an indicator of helix $\alpha 1$ and $\alpha 5$ interaction (Figure 5D–F) indicates that in MD-1, M475 maintains close interaction with W112 (~0.25 nm), and no interaction with H105. On the other hand, in MD-2, M475 maintains steady interactions with H105 (~0.5 – 0.6 nm) while the distances between M475 and W112 fluctuates. We further analyzed the interaction of residue W479 on helix $\alpha 5$ with residues, M104, H105, I108 and I109, on helix $\alpha 1$ to assess the relative motions of these two layers of the inner domain. As shown in Figure 6A–C for MD-1 and MD-2, the strongest interaction occurs between W479 and I109 (blue line), with either H105 (MD-1, red line (Figure 6A), or I108 (MD-2, green line (Figure 6B) providing additional stabilization between helix $\alpha 1$ and helix $\alpha 5$. Interestingly, these residues are associated much more closely in simulations of small ligand bound (NBD) or CD4 bound gp120_{CD4} (GCD2) compared to the unliganded gp120_{CD4} (GPO3)⁴⁹ (Supplementary Figure S8).

The average distances between N-terminus and C-terminus of helix $\alpha 1$ with the C-terminus and N-terminus of helix $\alpha 5$ are ~0.75 nm and ~1.57 nm respectively. Closer inspection indicates that a kink develops in helix $\alpha 1$ at M104 and H105 in MD-1 and MD-2, respectively as early as 40 ns (Figure 6D and F). Thus, the C-terminal portion of helix $\alpha 1$ pivots away from the outer-domain while the N-terminal portion remains more closely associated with the inner domain β -sandwich and helix $\alpha 5$ (Figure 6D and F). In fact, the structure of helix $\alpha 1$ as exhibited in the average structure of MD-2 resembles that observed in the b13-gp120 complex Figure 6E. As expected, helix $\alpha 5$ in MD-1 and MD-2 maintains the same position as observed in gp120_{CD4} (rmsd 0.19 and 1.6 nm) and gp120_{F105} (rmsd 0.25 and 0.25 nm) crystal structures. Thus, the tight interaction between M475 and W479 on helix $\alpha 5$ and H105, I108 and I109 on helix $\alpha 1$ anchors the N-terminus region of the helix $\alpha 1$ while the C-terminus portion of helix $\alpha 1$ pivots away from the outer domain permitting the $\beta 20/\beta 21$ strands to wedge between the inner and outer domains. Since a hydrogen bond is observed between E102 and R476 in gp120_{CD4} crystal structure and the distance between E102 and R476 or R480 side-chains is less than 0.4 nm in gp120_{F105-V3}, we also examined the electrostatic interactions between E102 (helix $\alpha 1$ and R476 or R480 (helix $\alpha 5$). This

hydrogen bond or salt-bridge does not form between E102 and R476 or R480 side-chains in either MD-1 or MD-2. Comparison of 100 ns simulations of GPO3, NBD, and GCD2 indicates that stabilizing E102 and R476 or R480 side-chains interactions are maintained (Supplementary Figure S9). Thus, in the unliganded form, the majority of the interactions between helix $\alpha 1$ and helix $\alpha 5$ are hydrophobic in nature, although weak electrostatic interactions also form between E102 and R476 or R480.

Conclusions

In this study we hypothesized that the extended $\beta 2/\beta 3$ and $\beta 20/\beta 21$ arrangement as exhibited in the antibody F105-gp120 complex represents one potential conformation of the metastable unliganded gp120 monomer for study of gp120 conformational change. The two 100 ns MD simulations demonstrate remarkably similar results with a conformational transition occurring within the first 30 ns. The transition commences with $\beta 20/\beta 21$ strand reorienting from an extended conformation exhibited in the gp120_{F105} crystal structure to a folded conformation lying between the inner and outer domains, reminiscent of its position observed in gp120_{CD4}. Within the first 4 ns, residue N425 on $\beta 20$ reorients by interacting with W112 before settling against the F382 in the outer domain positioning $\beta 20/\beta 21$ residues W427 or Y435 to interact with W112 on helix $\alpha 1$. The conformational transition was not observed when simulating a $\beta 20/\beta 21$ constraining mutant, I423P. The $\beta 2/\beta 3$ and the V3 loop also reorganize by folding over the $\beta 20/\beta 21$ in the simulated unliganded form of gp120_{F105-V3}. The rearrangement of the $\beta 20/\beta 21$ was accompanied by unwinding of helix $\alpha 1$ which also pivots about the N-terminus. The $\beta 20/\beta 21$ once refolded remains between the inner and outer domain for the entire 100 ns simulation period, even in the absence of a ligand. Thus, the folding of $\beta 20/\beta 21$ encloses one face of the pre-formed Phe43 cavity and is stabilized by interactions maintained between residues on both inner domain and outer domain.

Previous dogma held that CD4 or small molecule binding induces gp120 conformational transition from an unstructured form to the gp120_{CD4} form. However, the recent crystal structure of the unliganded gp120³⁸ indicates that gp120_{coreE} spontaneously adopts the CD4 bound conformation when not restricted by the presence of the V1, V2 and V3 variable loops. In our studies the presence of the V3 loop did not constrain the conformational rearrangement of the $\beta 20/\beta 21$ strands. However, the V3 loop formed strong interactions with the $\beta 2/\beta 3$ preventing the formation of the four-stranded anti-parallel bridging sheet observed in gp120_{CD4}. The propensity for V3 loop reorganization to interact with the $\beta 2/\beta 3$ strand during MD simulation is consistent with V3 determinants of co-receptor usage^{45, 46} and the sensitivity of V3 sequence in eliciting humoral immunity.⁷¹ The previously reported charge dependent mobility of the V3 loop as observed in MD simulation of gp120_{CD4}⁴⁷ is also consistent with the mobility of the loop observed in MD simulations of gp120_{F105}. We conclude from the simulations reported herein that the $\beta 20/\beta 21$, in unliganded gp120, has a natural propensity to adopt a conformation between the inner and outer domain. This conclusion is consistent with the position of the $\beta 20/\beta 21$, as observed in the SIV- structure⁴⁰ and as suggested by mutant data reported by Finzi et al.⁴¹ and others.^{26, 68, 70} Experimentally, we suggest probing the binding affinity of small molecules inhibitors⁷² that display varied thermodynamic signatures in full-length gp120 containing the N425 mutant to further explore the role of $\beta 20/\beta 21$ in mediating gp120 conformational equilibrium. Moreover, the insights provided herein should be helpful in developing strategies to entrap unproductive conformations of gp120 in its unliganded form as means of preventing HIV-1 entry.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The authors appreciate the useful suggestions and discussions with Arne Schon, Navid Madani and Ivet Bahar. Support from NIH grant 1U54GM087519-01A1 to Ivet Bahar's lab is gratefully acknowledged by IHS. This work was supported by a grant from the National Institutes of Health GM56550 (JL).

FUNDING:

Support from NIH grant 1U54GM087519-01A1 to Ivet Bahar's lab is gratefully acknowledged by IHS. IHS acknowledges NIH for support (R01LM007994-06). This work was supported by a grant from the National Institutes of Health GM56550 (JL).

ABBREVIATIONS

MD	Molecular Dynamics
HIV-1	Human immunodeficiency virus, 1
RMSF	root mean square fluctuations
rmsd	root mean square deviations
gp120_{CD4}	CD4 bound conformation of gp120
gp120_{coreE}	gp120 extended core
gp120_{core}	truncated core gp120
gp120_{F105}	antibody F105 bound gp120 _{core} complex
gp120_{F105-V3}	gp120 _{F105} , antibody F105 bound gp120 _{core} complex with a grafted V3 loop.

References

1. Barre-Sinoussi F, Chermann JC, Rey F, Nugeyre MT, Chamaret S, Gruest J, Dautet C, Axler-Blin C, Vezinet-Brun F, Rouzioux C, Rozenbaum W, Montagnier L. Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS). *Science*. 1983; 220:868–871. [PubMed: 6189183]
2. Gallo RC, Salahuddin SZ, Popovic M, Shearer GM, Kaplan M, Haynes BF, Palker TJ, Redfield R, Oleske J, Safai B, et al. Frequent detection and isolation of cytopathic retroviruses (HTLV-III) from patients with AIDS and at risk for AIDS. *Science*. 1984; 224:500–503. [PubMed: 6200936]
3. Dalgleish AG, Beverley PC, Clapham PR, Crawford DH, Greaves MF, Weiss RA. The CD4 (T4) antigen is an essential component of the receptor for the AIDS retrovirus. *Nature*. 1984; 312:763–767. [PubMed: 6096719]
4. McDougal JS, Kennedy MS, Sligh JM, Cort SP, Mawle A, Nicholson JK. Binding of HTLV-III/LAV to T4+ T cells by a complex of the 110K viral protein and the T4 molecule. *Science*. 1986; 231
5. Klatzmann D, Champagne E, Chamaret S, Gruest J, Guetard D, Hercend T, Gluckman JC, Montagnier L. T-lymphocyte T4 molecule behaves as the receptor for human retrovirus LAV. *Nature*. 1984; 314:767–768. [PubMed: 6083454]
6. Wyatt R, Sodroski J. The HIV-1 envelope glycoproteins fusogens antigens and immunogens. *Science*. 1998; 280:1884–1888. [PubMed: 9632381]
7. Ugolini S, Moulard M, Mondor I, Barois N, Demandolx D, Hoxie J, Brelot A, Alizon M, Davoust J, Sattentau QJ. HIV-1 gp120 induces an association between CD4 and the chemokine receptor CXCR4. *J Immunol*. 1997; 159:3000–3008. [PubMed: 9300725]

8. Blacklow SC, Lu M, Kim PS. A trimeric subdomain of the simian immunodeficiency virus envelope glycoprotein. *Biochemistry*. 1995; 34:14955–14962. [PubMed: 7578108]
9. Lu M, Blacklow SC, Kim PS. A trimeric structural domain of the HIV-1 transmembrane glycoprotein. *Nat Struct Biol*. 1995; 2:1075–1082. [PubMed: 8846219]
10. Kowalski M, Potz J, Basiripour L, Dorfman T, Goh WC, Terwilliger E, Dayton A, Rosen C, Haseltine W, Sodroski J. Functional regions of the envelope glycoprotein of human immunodeficiency virus type 1. *Science*. 1987; 237:1351–1355. [PubMed: 3629244]
11. Dragic T, Litwin V, Allaway GP, Martin SR, Huang Y, Nagashima KA, Cayanan C, Maddon PJ, Koup RA, Moore JP, Paxton WA. HIV-1 entry into CD4+ cells is mediated by the chemokine receptor CC-CKR-5. *Nature*. 1996; 381:667–673. [PubMed: 8649512]
12. Sattentau QJ, Moore JP. Conformational changes induced in the human immunodeficiency virus envelope glycoprotein by soluble CD4 binding. *J Exp Med*. 1991; 174:407–415. [PubMed: 1713252]
13. Sattentau QJ, Moore JP, Vignaux F, Traincard F, Poignard P. Conformational changes induced in the envelope glycoproteins of the human and simian immunodeficiency viruses by soluble receptor binding. *J Virol*. 1993; 67:7383–7393. [PubMed: 7693970]
14. Wu L, Gerard NP, Wyatt R, Choe H, Parolin C, Ruffing N, Borsetti A, Cardoso AA, Desjardin E, Newman W, Gerard C, Sodroski J. CD4-induced interaction of primary HIV-1 gp120 glycoproteins with the chemokine receptor CCR-5. *Nature*. 1996; 384
15. Choe H, Farzan M, Sun Y, Sullivan N, Rollins B, Ponath PD, Wu L, Mackay CR, LaRosa G, Newman W, Gerard N, Gerard C, Sodroski J. The beta-chemokine receptors CCR3 and CCR5 facilitate infection by primary HIV-1 isolates. *Cell*. 1996; 85:1135–1148. [PubMed: 8674119]
16. Deng H, Liu R, Ellmeier W, Choe S, Unutmaz D, Burkhart M, Di Marzio P, Marmon S, Sutton RE, Hill CM, Davis CB, Peiper SC, Schall TJ, Littman DR, Landau NR. Identification of a major co-receptor for primary isolates of HIV-1. *Nature*. 1996; 381:661–666. [PubMed: 8649511]
17. Doranz BJ, Rucker J, Yi Y, Smyth RJ, Samson M, Peiper SC, Parmentier M, Collman RG, Doms RW. A dual-tropic primary HIV-1 isolate that uses fusin and the beta-chemokine receptors CKR-5, CKR-3, and CKR-2b as fusion cofactors. *Cell*. 1996; 85:1149–1158. [PubMed: 8674120]
18. Bosch ML, Earl PL, Fargnoli K, Picciafuoco S, Giombini F, Wong-Staal F, Franchini G. Identification of the fusion peptide of primate immunodeficiency viruses. *Science*. 1989; 244:694–697. [PubMed: 2541505]
19. Brasseur R, Cornet B, Burny A, Vandenbranden M, Ruyschaert JM. Mode of insertion into a lipid membrane of the N-terminal HIV gp41 peptide segment. *AIDS Res Hum Retroviruses*. 1988; 4:83–90. [PubMed: 3259143]
20. Helseth E, Olshevsky U, Gabuzda D, Ardman B, Haseltine W, Sodroski J. Changes in the transmembrane region of the human immunodeficiency virus type 1 gp41 envelope glycoprotein affect membrane fusion. *J Virol*. 1990; 64:6314–6318. [PubMed: 2243396]
21. Myszka DG, Sweet RW, Hensley P, Brigham-Burke M, Kwong PD, Hendrickson WA, Wyatt R, Sodroski J, Doyle ML. Energetics of the HIV gp120-CD4 binding reaction. *Proc Natl Acad Sci U S A*. 2000; 97:9026–9031. [PubMed: 10922058]
22. Dey B, Pancera M, Svehla K, Shu Y, Xiang SH, Vainshtein J, Li Y, Sodroski J, Kwong PD, Mascola JR, Wyatt R. Characterization of human immunodeficiency virus type 1 monomeric and trimeric gp120 glycoproteins stabilized in the CD4-bound state: antigenicity, biophysics, and immunogenicity. *J Virol*. 2007; 81:5579–5593. [PubMed: 17360741]
23. Kwong PD, Wyatt R, Desjardins E, Robinson J, Culp JS, Hellmig BD, Sweet RW, Sodroski J, Hendrickson WA. Probability analysis of variational crystallization and its application to gp120, the exterior envelope glycoprotein of type 1 human immunodeficiency virus (HIV-1). *J Biol Chem*. 1999; 274:4115–4123. [PubMed: 9933605]
24. Schon A, Madani N, Klein JC, Hubicki A, Ng D, Yang X, Smith AB 3rd, Sodroski J, Freire E. Thermodynamics of binding of a low-molecular-weight CD4 mimetic to HIV-1 gp120. *Biochemistry*. 2006; 45:10973–10980. [PubMed: 16953583]
25. Dey B, Svehla K, Xu L, Wycuff D, Zhou T, Voss G, Phogat A, Chakrabarti BK, Li Y, Shaw G, Kwong PD, Nabel GJ, Mascola JR, Wyatt RT. Structure-based stabilization of HIV-1 gp120

- enhances humoral immune responses to the induced co-receptor binding site. *PLoS Pathog.* 2009; 5:e1000445. [PubMed: 19478876]
26. Xiang SH, Kwong PD, Gupta R, Rizzuto CD, Casper DJ, Wyatt R, Wang L, Hendrickson WA, Doyle ML, Sodroski J. Mutagenic stabilization and/or disruption of a CD4-bound state reveals distinct conformations of the human immunodeficiency virus type 1 gp120 envelope glycoprotein. *J. Virol.* 2002; 76:9888–9899. [PubMed: 12208966]
 27. Kwong PD, Wyatt R, Majeed S, Robinson J, Sweet RW, Sodroski J, Hendrickson WA. Structures of HIV-1 gp120 envelope glycoproteins from laboratory-adapted and primary isolates. *Structure. Fold. Des.* 2000; 8:1329–1339. [PubMed: 11188697]
 28. Kwong PD, Wyatt R, Robinson J, Sweet RW, Sodroski J, Hendrickson WA. Structure of an HIV gp120 envelope glycoprotein in complex with the CD4 receptor and a neutralizing human antibody. *Nature.* 1998; 393:648–659. [PubMed: 9641677]
 29. Zhou T, Xu L, Dey B, Hessel AJ, Van Ryk D, Xiang SH, Yang X, Zhang MY, Zwick MB, Arthos J, Burton DR, Dimitrov DS, Sodroski J, Wyatt R, Nabel GJ, Kwong PD. Structural definition of a conserved neutralization epitope on HIV-1 gp120. *Nature.* 2007; 445:732–737. [PubMed: 17301785]
 30. Diskin R, Marcovecchio PM, Bjorkman PJ. Structure of a clade C HIV-1 gp120 bound to CD4 and CD4-induced antibody reveals anti-CD4 polyreactivity. *Nat Struct Mol Biol.* 2010; 17:608–613. [PubMed: 20357769]
 31. Pancera M, Majeed S, Ban YE, Chen L, Huang CC, Kong L, Kwon YD, Stuckey J, Zhou T, Robinson JE, Schief WR, Sodroski J, Wyatt R, Kwong PD. Structure of HIV-1 gp120 with gp41-interactive region reveals layered envelope architecture and basis of conformational mobility. *Proc Natl Acad Sci U S A.* 2010; 107:1166–1171. [PubMed: 20080564]
 32. Chen L, Kwon YD, Zhou T, Wu X, O'Dell S, Cavacini L, Hessel AJ, Pancera M, Tang M, Xu L, Yang Z-Y, Zhang M-Y, Arthos J, Burton DR, Dimitrov DS, Nabel GJ, Posner MR, Sodroski J, Wyatt R, Mascola JR, Kwong PD. Structural Basis of Immune Evasion at the Site of CD4 Attachment on HIV-1 gp120. *Science.* 2009; 326:1123–1127. [PubMed: 19965434]
 33. Huang C-C, Lam SN, Acharya P, Tang M, Xiang S-H, Shahzad-ul Hussan S, Stanfield RL, Robinson S, Sodroski J, Wilson IA, Wyatt R, Bewley CA, Kwong PD. Structures of the CCR5 N Terminus and of a Tyrosine-Sulfated Antibody with HIV-1 gp120 and CD4. *Science.* 2007; 319:1930–1934. [PubMed: 17901336]
 34. Zhou T, Georgiev I, Wu X, Yang ZY, Dai K, Finzi A, Kwon YD, Scheid JF, Shi W, Xu L, Yang Y, Zhu J, Nussenzweig MC, Sodroski J, Shapiro L, Nabel GJ, Mascola JR, Kwong PD. Structural basis for broad and potent neutralization of HIV-1 by antibody VRC01. *Science.* 2010; 329:811–817. [PubMed: 20616231]
 35. Wu X, Zhou T, Zhu J, Zhang B, Georgiev I, Wang C, Chen X, Longo NS, Louder M, McKee K, O'Dell S, Perfetto S, Schmidt SD, Shi W, Wu L, Yang Y, Yang ZY, Yang Z, Zhang Z, Bonsignori M, Crump JA, Kapiga SH, Sam NE, Haynes BF, Simek M, Burton DR, Koff WC, Doria-Rose NA, Connors M, Mullikin JC, Nabel GJ, Roederer M, Shapiro L, Kwong PD, Mascola JR. Focused evolution of HIV-1 neutralizing antibodies revealed by structures and deep sequencing. *Science.* 2011; 333:1593–1602. [PubMed: 21835983]
 36. Huang CC, Stricher F, Martin L, Decker JM, Majeed S, Barthe P, Hendrickson WA, Robinson J, Roumestand C, Sodroski J, Wyatt R, Shaw GM, Vita C, Kwong PD. Scorpion-toxin mimics of CD4 in complex with human immunodeficiency virus gp120 crystal structures, molecular mimicry, and neutralization breadth. *Structure.* 2005; 13:755–768. [PubMed: 15893666]
 37. Stricher F, Huang C-c, Descours A, Duquesnoy S, Combes O, Decker JM, Do Kwon Y, Lusso P, Shaw GM, Vita C, Kwong PD, Martin L. Combinatorial Optimization of a CD4-Mimetic Miniprotein and Cocrystal Structures with HIV-1 gp120 Envelope Glycoprotein. *Journal of Molecular Biology.* 2008; 382:510–524. [PubMed: 18619974]
 38. Kwon YD, Finzi A, Wu X, Dogo-isonagie C, Lee LK, Moore LR, Schmidt SR, Stuckey J, Yang Y, Zhou T, Zhu J, Viceda DA, Debnath AK, Shapiro L, Bewley CA, Mascola JR, Sodroski J, Kwong PD. Unliganded HIV-1 gp120 Structures Reveal Propensity of Core to Assume the CD4-Bound Conformation with Regulation by Quaternary Interactions and Variable Loops. *Proc Natl Acad Sci U S A.* 2012

39. LaLonde JM, Kwon YD, Jones DM, Sun AW, Courter JR, Soeta T, Kobayashi T, Princiotta AM, Wu X, Schön A, Freire E, Kwong PD, Mascola JR, Sodroski J, Madani N, Smith AB III. Dual Hot Spot Structure-Based Design, Synthesis and Characterization of Small-Molecule HIV-1 Entry Inhibitors. *J. Med. Chem.* 2012; 55:4382–4396. [PubMed: 22497421]
40. Chen B, Vogan EM, Gong H, Skehel JJ, Wiley DC, Harrison SC. Structure of an unliganded simian immunodeficiency virus gp120 core. *Nature.* 2005; 433:834–841. [PubMed: 15729334]
41. Finzi A, Xiang SH, Pacheco B, Wang L, Haight J, Kassa A, Danek B, Pancera M, Kwong PD, Sodroski J. Topological layers in the HIV-1 gp120 inner domain regulate gp41 interaction and CD4-triggered conformational transitions. *Mol Cell.* 2010; 37:656–667. [PubMed: 20227370]
42. Huang CC, Tang M, Zhang MY, Majeed S, Montabana E, Stanfield RL, Dimitrov DS, Korber B, Sodroski J, Wilson IA, Wyatt R, Kwong PD. Structure of a V3-containing HIV-1 gp120 core. *Science.* 2005; 310:1025–1028. [PubMed: 16284180]
43. Huang C-C, Venturi M, Moore MJ, Phogat S, Zhang M-Y, Hendrickson WA, Robinson J, Sodroski J, Wyatt R, Choe H, Farzan M, Kwong PD. Structural basis of tyrosine sulfation and VH-gene usage in antibodies that recognize the HIV type 1 coreceptor-binding site on gp120. *Proc Natl Acad Sci U S A.* 2004; 101:2706–2711. [PubMed: 14981267]
44. Hartley O, Klasse PJ, Sattentau QJ, Moore JP. V3: HIV's switch-hitter. *AIDS Res Hum Retroviruses.* 2005; 21:171–189. [PubMed: 15725757]
45. Cann AJ, Churcher MJ, Boyd M, O'Brien W, Zhao JQ, Zack J, Chen IS. The region of the envelope gene of human immunodeficiency virus type 1 responsible for determination of cell tropism. *J Virol.* 1992; 66:305–309. [PubMed: 1727490]
46. Westervelt P, Trowbridge DB, Epstein LG, Blumberg BM, Li Y, Hahn BH, Shaw GM, Price RW, Ratner L. Macrophage tropism determinants of human immunodeficiency virus type 1 in vivo. *J Virol.* 1992; 66:2577–2582. [PubMed: 1548783]
47. Yokoyama M, Naganawa S, Yoshimura K, Matsushita S, Sato H. Structural Dynamics of HIV-1 Envelope Gp120 Outer Domain with V3 Loop. *PLoS One.* 2012; 7:e37530. Epub 2012 May 18. [PubMed: 22624045]
48. Tan H, Rader AJ. Identification of putative, stable binding regions through flexibility analysis of HIV-1 gp120. *Proteins.* 2009
49. Shrivastava I, LaLonde JM. Fluctuation dynamics analysis of gp120 envelope protein reveals a topologically based communication network. *Proteins: Structure, Function, and Bioinformatics.* 2010; 78:2935–2949.
50. Abrams CF, Vanden-Eijnden E. Large-scale conformational sampling of proteins using temperature-accelerated molecular dynamics. *Proc Natl Acad Sci U S A.* 2010; 107:4961–4966. [PubMed: 20194785]
51. Hsu S-TD, Bonvin AMJJ. Atomic insight into the CD4 binding-induced conformational changes in HIV-1 gp120. *Proteins: Struct. Funct. & Bioinformatics.* 2004; 55:582–593.
52. Hsu S-TD, Peter D, van Gunsteren WF, Bonvin AMJJ. Entropy Calculation of HIV-1 Env gp120, its Receptor CD4, and their Complex: An Analysis of Configurational Entropy Changes upon Complexation. *Biophysical Journal.* 2005; 88:14–24.
53. Da L-T, Quan J-M, Wu YD. Understanding of the Bridging Sheet Formation of HIV-1 Glycoprotein gp120. *J. Phys. Chem. B.* 2009; 113:14536–14543. [PubMed: 19813706]
54. Pan Y, Ma B, Keskin O, Nussinov R. Characterization of the conformational state and flexibility of HIV-1 glycoprotein gp120 core domain. *J Biol Chem.* 2004; 279:30523–30530. [PubMed: 15131118]
55. Liu SQ, Liu SX, Fu YX. Molecular motions of human HIV-1 gp120 envelope glycoproteins. *J Mol Model.* 2008; 14:857–870. [PubMed: 18594881]
56. Pan Y, Ma B, Nussinov R. CD4 binding partially locks the bridging sheet in gp120 but leaves the beta2/3 strands flexible. *J Mol Biol.* 2005; 350:514–527. [PubMed: 15946678]
57. Shrivastava I, Lalonde JM. Enhanced Dynamics of HIV gp120 Glycoprotein by Small Molecule Binding. *Biochemistry.* 2011; 50:4173–4183. [PubMed: 21488663]
58. Molecular Operating Environment (MOE). Montreal, QC, Canada: Chemical Computing Group Inc.: 1010 Sherbooke St. West, Suite #910; 2011. 2011.10; H3A 2R7

59. Lovell SC, Davis IW, Arendall WB 3rd, de Bakker PI, Word JM, Prisant MG, Richardson JS, Richardson DC. Structure validation by C α geometry: ϕ , ψ and C β deviation. *Proteins*. 2003; 50:437–450. [PubMed: 12557186]
60. Word J, Lovell S, Richardson J, Richardson D. Asparagine and glutamine: Using hydrogen atom contacts in the choice of side-chain amide orientation. *J. Mol. Biol.* 1999; 285:1735–1747. [PubMed: 9917408]
61. Jorgensen WL, Maxwell DS, Tirado-Rives J. Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids. *J. Am. Chem. Soc.* 1996; 117:11225–11236.
62. Lindahl EHB, van der Spoel D. GROMACS 30 A package for molecular simulation and trajectory analysis. *J Mol Mod.* 2001; 7:306–317. <http://www.gromacs.org/>.
63. Van der Spoel E, Lindahl B, Hess G, Mark Groenhof AE, Berendsen HJC. GROMACS Fast Flexible and Free. *J Comp Chem.* 2005; 26:1701–1718. [PubMed: 16211538]
64. Berendsen HJC, Grigera JR, Straatsma PR, Berendsen HJC, Grigera JR, Straatsma PR. *J. Phys. Chem.* 1987; 91:6269–6271.
65. Hess B, Bekker H, Berendsen HJC, Fraaije J. LINCS: A linear constraint solver for molecular simulations. *J. Comput. Chem.* 1997; 18:1463–1472.
66. Essmann U, Perera L, Berkowitz ML, Darden T, Lee H, Pedersen LG. A smooth particle mesh ewald method. *J. Chem. Phys.* 1995; 103:8577–8593.
67. Andrianov AM, Anishchenko IV, Tuzikov AV. Discovery of novel promising targets for anti-AIDS drug developments by computer modeling: application to the HIV-1 gp120 V3 loop. *J Chem Inf Model.* 2011; 51:2760–2767. Epub 2011 Sep 22. [PubMed: 21888425]
68. Xiang SH, Finzi A, Pacheco B, Alexander K, Yuan W, Rizzuto C, Huang CC, Kwong PD, Sodroski J. A V3 loop-dependent gp120 element disrupted by CD4 binding stabilizes the human immunodeficiency virus envelope glycoprotein trimer. *J Virol.* 2010; 84:3147–3161. [PubMed: 20089638]
69. Jarvis RA, Patrick EA. Clustering Using a Similarity Measure Based on Shared Near Neighbors. *IEEE Transactions on Computers.* 1973; C22:1025–1034.
70. Berkower I, Patel C, Ni Y, Virnik K, Xiang Z, Spadaccini A. Targeted deletion in the beta20–beta21 loop of HIV envelope glycoprotein gp120 exposes the CD4 binding site for antibody binding. *Virology.* 2008; 377:330–338. Epub 2008 Jun 2. [PubMed: 18519142]
71. Goudsmit J, Debouck C, Meloen RH, Smit L, Bakker M, Asher DM, Wolff AV, Gibbs CJ Jr, Gajdusek DC. Human immunodeficiency virus type 1 neutralization epitope with conserved architecture elicits early type-specific antibodies in experimentally infected chimpanzees. *Proc Natl Acad Sci U S A.* 1988; 85:4478–4482. [PubMed: 2454471]
72. Schön A, Madani N, Smith IAB, LaLonde JM, Freire E. Some Drug Properties are Dependent on Thermodynamic Signature. *Chem Biol Drug Des.* 2011; 77:161–165. [PubMed: 21288305]

Figure 1. Conformations of gp120

(A) The CD4 (magenta) bound conformation of gp120 (gp120_{CD4}). The CD4 residue Phe43 (magenta spheres) binds to a large internal cavity (the "Phe43 cavity") formed by the inner domain (red), outer domain (yellow) and the bridging sheet composed of the β2/β3 strands (green) and the β20/β21 strands (cyan). The crystal complex 1G9N also contains the chemokine antibody surrogate, 17b, which is not shown here. (B) The F105 antibody (gray and pink) bound conformation of gp120 (gp120_{F105}). The F105 Val-Phe-Tyr motif (gray spheres) binds to a shallower and more open "Phe43 cavity". The β2/β3 strands (green) and the β20/β21 strands (cyan) are extended away from the Phe43 cavity and disassociated. (C) The

gp120_{F105-V3} model used in MD simulations contains the V3 loop (dark blue) grafted onto gp120_{F105}.

Figure 2. gp120 Conformational Transition

(A) The Root Mean Square Deviation (rmsd) for backbone C- α atoms of MD-1 (black line) and MD-2 (red line) from the initial conformation. (B) Superimposed structures of MD-1 starting structure (blue) and 100 ns structure (red) and (C) MD-2 starting structure (blue) and 100 ns structure (red). In both the simulations, the outer-domain has moved less, as compared to inner-domain. The high rmsd at the conformational transition is due to restructuring of the $\beta 2/\beta 3$ and $\beta 20/\beta 21$ strands and the V3 loop.

Figure 3. Monitoring $\beta 2/\beta 3$ and $\beta 20/\beta 21$ strands and the V3 loop rearrangement

(A) The minimum distances between W112 (red spheres) and W427 (cyan spheres) in MD-1 and MD-2 black and red line, respectively. (B) The minimum distances in MD-1 and MD-2 black and red line, respectively for the residue pairs, W112 (red spheres) and Y435 (cyan spheres) (C) The minimum distances in MD-1 and MD-2 black and red line, respectively for the Q203 (green spheres) and Y318 (cyan spheres) residue pair.

Figure 4. The Role of N425 in $\beta 20/\beta 21$ Rearrangement

(A) The minimum distances between N425 and W112 (black) and N425 and F382 (red) in MD-1 and (C) MD-2. (B) Ribbon diagram of the starting structure, in which the residue pairs whose distances were measured are highlighted, W112 (red spheres), N425 (cyan spheres), and F382 (yellow spheres). (D) The MD-1 100 ns snapshot showing the association between N425 (cyan spheres) and W112 (red spheres) and F382 (yellow spheres). (E) In gp120_{CD4}, x-ray structure (pdb 1G9N) the corresponding locations of W112 (red spheres), N425 (cyan spheres), and F382 (yellow spheres) (F) The MD-2 100 ns snapshot, with the residues highlighted the same as in D.

Figure 5. Inner Domain Stabilization of $\beta 20$ Residue, W427

(A) For MD-1, the minimum distances between W427 and H105 (black line) and W427 and M475 (red line). (B) Position of the residues W427 (cyan sphere) and H105 (red sphere) is indicated in the starting structure, (C) MD-2, same as in (A). (D) For MD-1, the minimum distances between M475 and H105 (black line) and M475 and W112 (red line). (E) Position of the residues M475, H105 and W112 (red spheres) is indicated in the starting structure. (F) MD-2, same as in (D).

Figure 6. Shifts in Inner Domain Helix $\alpha 1$

(A) For MD-1, the minimum distances between W479 and M104 (black line), W479 and H105 (red line), W479 and I108 (green line) and W479 and I109 (blue line). (B) In the starting structure, the distances between residues W479 (red spheres) and residues M104, H105, I108 and I109 (ball and stick) are highlighted. (C) MD-2 same as in (A). (D) The MD-1, 100 ns snapshot showing the kink in that develops in helix $\alpha 1$ near H105. (E) Comparison of MD-2 and the antibody b13 bound gp120 conformation (gp120_{b13}) (pdb 2IDY) (pink ribbon). The loss of helicity at I108 is also apparent. (F) The MD-2, 100 ns snapshot showing the kink and loss of helicity in helix $\alpha 1$ around I108.

Table 1

Distance for Strongly Interacting Residues Pairs in gp120 Crystal Structures

β 20/ β 21	Inner Domain	Outer Domain	Distance (nm)* gp120 _{CD4} (1G9N)	Distance (nm)*
			gp120 _{F105-V3}	
K421		F382	0.309	0.190
K421		Y384	0.398	0.232
N425		V254	1.341	1.509
N425		F382	0.633	0.907
	W112	G380	0.994	0.553
W427	W112		0.324	1.023

* The minimum distance between any two atoms of the residue pairs,

Table 2

Analysis of Key interactions in MD simulations

Interactions [*]	Average ^{**}	Std-deviation	Percent [†]	Cluster1 [#]	Cluster2 ^{##}
MD1					
K421-F382	0.392	0.129	60.3	0.320	0.531
K421-Y384	0.296	0.042	100	0.202	0.308
N425-V254	0.337	0.171	85.4	0.379	0.389
N425-F382	0.311	0.104	92.2	0.267	0.422
W112-W427	0.373	0.149	74.3	0.422	0.366
W112-G380	0.843	0.122	0.05	0.657	0.829
MD2					
K421-F382	0.199	0.020	100	0.223	0.177
K421-Y384	0.291	0.040	100	0.223	0.278
N425-V254	0.299	0.115	97.7	0.293	0.201
N425-F382	0.269	0.061	98.9	0.275	0.253
W112-W427	0.737	0.165	3.29	0.791	0.649
W112-G380	0.969	0.122	0.45	1.064	0.927

^{*}The first 5 interactions are bridging-sheet/inner domain interactions, the sixth is inner-domain/outer domain interaction.

^{**}Minimum distance between any two atoms, averaged over 100ns simulations.

[†]Percentage time a given interaction (minimum distance between any two atoms, is $\leq 0.45\text{nm}$), over the 100ns period.

[#]Cluster1 corresponds to clusters centered around structure, at $t = 18600.0$ and 34100.00 ps respectively for MD1 and MD2

^{##}Cluster2 corresponds to clusters centered around structure, at $t = 65800.00$ and 86550.00 ps respectively for MD1 and MD2.