BIOCHEMISTRY
including biophysical chemistry & molecular biology

Article

pubs.acs.org/biochemistry

A Model of Peptide Triazole Entry Inhibitor Binding to HIV-1 gp120 and the Mechanism of Bridging Sheet Disruption

Ali Emileh,*[,†,‡] Ferit Tuzer,[‡] Herman Yeh,[∥] Muddegowda Umashankara,[‡] Diogo R. M. Moreira,[‡] Judith M. LaLonde,[§] Carole A. Bewley,[∥] Cameron F. Abrams,*[,†] and Irwin M. Chaiken*[,‡]

[†]Chemical and Biological Engineering, Drexel University, Philadelphia 19104, Pennsylvania, United States

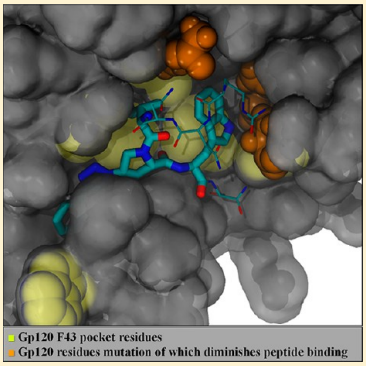
[‡]Department of Biochemistry and Molecular Biology, Drexel University College of Medicine, Philadelphia, Pennsylvania 19129, United States

[§]Department of Chemistry, Bryn Mawr College, Bryn Mawr, Pennsylvania 19010, United States

[∥]Laboratory of Bioorganic Chemistry, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, Maryland 20892, United States

Supporting Information

ABSTRACT: Peptide triazole (PT) entry inhibitors prevent HIV-1 infection by blocking the binding of viral gp120 to both the HIV-1 receptor and the coreceptor on target cells. Here, we used all-atom explicit solvent molecular dynamics (MD) to propose a model for the encounter complex of the peptide triazoles with gp120. Saturation transfer difference nuclear magnetic resonance (STD NMR) and single-site mutagenesis experiments were performed to test the simulation results. We found that docking of the peptide to a conserved patch of residues lining the "F43 pocket" of gp120 in a bridging sheet naïve gp120 conformation of the glycoprotein led to a stable complex. This pose prevents formation of the bridging sheet minidomain, which is required for receptor−coreceptor binding, providing a mechanistic basis for dual-site antagonism of this class of inhibitors. Burial of the peptide triazole at the gp120 inner domain−outer domain interface significantly contributed to complex stability and rationalizes the significant contribution of hydrophobic triazole groups to peptide potency. Both the simulation model and STD NMR experiments suggest that the I-X-W [where X is (2*S*,4*S*)-4-(4-phenyl-1*H*-1,2,3-triazol-1-yl)pyrrolidine] tripartite hydrophobic motif in the peptide is the major contributor of contacts at the gp120−PT interface. Because the model predicts that the peptide Trp side chain hydrogen bonding with gp120 S375 contributes to the stability of the PT−gp120 complex, we tested this prediction through analysis of peptide binding to gp120 mutant S375A. The results showed that a peptide triazole KR21 inhibits S375A with 20-fold less potency than WT, consistent with predictions of the model. Overall, the PT−gp120 model provides a starting point for both the rational design of higher-affinity peptide triazoles and the development of structure-minimized entry inhibitors that can trap gp120 into an inactive conformation and prevent infection.



The envelope glycoprotein spikes on the HIV surface are the receptor recognition machinery of the virus, responsible for the initiation of host cell entry.[1,2] Each envelope spike is a trimer of noncovalent dimers of transmembrane gp41 and the highly glycosylated gp120.[2] Viral entry occurs as a sequence of interactions with cell surface receptors: gp120 binds the CD4 receptor resulting in a conformational change that decrypts an epitope on gp120 that binds the coreceptor, a member of the chemokine receptor family, usually CCR5 or CXCR4.[2,3] Binding to the coreceptor results in exposure of the gp41 fusion peptide, its insertion into the target membrane, and subsequent formation of the "six-helix bundle" of gp41 that likely drives fusion of the target cell and viral membranes, allowing the viral capsid to enter the cell.[4,5]

The multistep process of viral entry provides a series of targets for intervention and blocking infection before the virus establishes a foothold in the host.[6] Enfuvirtide (T20) is a subcutaneously injected 36-amino acid peptide that binds to the transiently exposed HR1 of gp41 and prevents formation of the six-helix bundle, preventing viral entry.[7] It was approved for clinical use in 2003, although its low bioavailability and poor pharmacokinetics have spurred research aimed at its improvement.[8] Maraviroc, an orally administered small molecule that binds to CCR5 and inhibits gp120 recognition, was approved by the Food and Drug Administration in 2007.[7] The emergence of X4-tropic or R5X4-dualtropic viruses utilizing CXCR4 is the most common pathway of resistance to maraviroc.[6,9] The success of combination therapy has shown the value of using multiple agents to target sequentially diverse HIV.[8] Prominent examples of targeting gp120 through inhibiting its interactions with CD4 or CCR5/CXCR4 include NBD-556 and DMJ-II-228, which inhibit gp120−CD4 interaction, and BMS378806 and

Received: September 28, 2012
Revised: March 7, 2013
Published: March 7, 2013

   dx.doi.org/10.1021/bi400166b | *Biochemistry* 2013, 52, 2245−2261

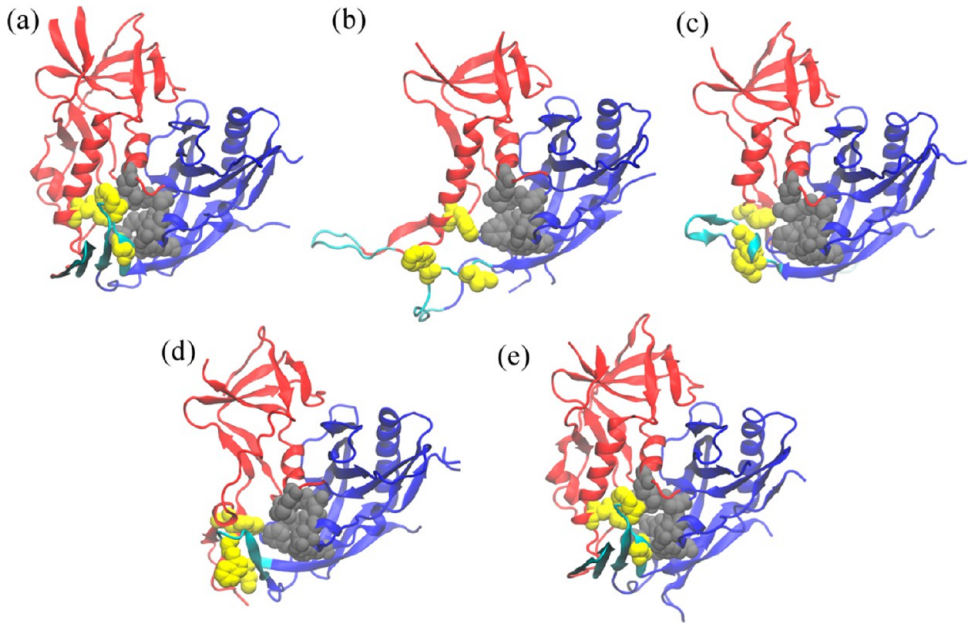


Figure 1. Conformations of the gp120 core that have been crystallographically resolved to date: (a) unliganded, (b) F105-bound, (c) b13-bound, (d) b12-bound, and (e) CD4/48d-bound. Structures d and e were resolved using highly engineered and disulfide-stitched gp120 sequences. The sequence used in panel e contained the pocket filling S375W and T257S mutations. The inner domain is colored red and the outer domain blue, and components of the bridging sheet are colored cyan. Colored gray are the outer domain residues of the F43 pocket (V255, T257, E370, F382, Y384, M475, S256bb, S375bb, F376bb, and N377bb where the bb suffix indicates backbone atoms), which are conformationally invariant between different structures. Colored yellow are the inner domain (W112) or β20 and β21 (W427, I424, and N425bb) members of the pocket.

BMS488043, which prevent viral entry through an unknown mechanism involving conformational changes in the envelope spike.[10−12] The former is an agonist of coreceptor binding,[13] although very recent work has shown that this undesirable phenotype can be eliminated by structural modification.[14] The BMS class has a very small barrier to resistance as one or two mutations in gp120 result in 40−500-fold resistance to the inhibitor.[7]

12p1 is a small peptide (with the sequence RINNIPW-SEAMM) that allosterically inhibits interactions of gp120 with both the receptor and the coreceptor.[15,16] Replacing the middle proline with azidoproline and derivatizing the latter with various alkynes through click chemistry led to the identification of triazole variants with an approximately 400-fold increase in potency, generating a new class of peptide triazole (PT) entry inhibitors with low nanomolar affinity for gp120.[17,18] HNG-156, which contains a ferrocenyl triazole substitution, binds to gp120 with an equilibrium dissociation constant (K_D) of 7 nM, compared to the K_D of 5400 nM for 12p1.[19] It antagonizes interactions of gp120 with both CD4 and the coreceptor and inhibits HIV-1_{BaL} entry at a nanomolar 50% effective concentrations (EC_{50}). This class of entry inhibitors is broadly active against various HIV clades and nontoxic and can be synergistically combined with other entry inhibitors.[20,21] Recently, it was found that some variants of the peptide can disrupt the viral membrane, leading to cell-free HIV virolysis and inactivation.[22] gp120 shedding and virolysis were shown to be dependent on the specific interactions of the peptides with gp120.[22] Clearly, such peptide triazoles hold great promise for both therapeutic and microbicidal applications. Understanding the binding mode of peptide triazoles in the context of gp120 structure would help greatly in further developing peptide designs with improved potency and peptidomimetic character.

The crystal structure of deglycosylated HIV gp120 lacking the V1−V3 variable loops (gp120 core) in complex with soluble CD4 and a coreceptor surrogate antibody (mAb 48d) shows gp120 in its canonical activated state (Figure 1a).[23] In the simplest interpretation, gp120 in this conformation can be divided into the inner domain, the outer domain, and, at their interface, a four-strand β-sheet (composed of β2, β3, β20, and β21) called the bridging sheet. The latter is where the coreceptor binds.[23,24] At the intersection of the inner and outer domains with the bridging sheet lies a deep cavity called the "F43 pocket". Residues lining this cavity are hydrophobic, well-conserved, and thought to be important for gp120 function.[24] Of 23 crystal structures of various HIV gp120 cores available in the Protein Data Bank (PDB), only a handful are resolved in a conformation significantly different than the activated state.[25] These include the mAb F105-bound (Figure 1b),[26] the mAb b13-bound (Figure 1c),[26] and the mAb b12-bound (Figure 1d)[27] conformations, all of which show a structure without a bridging sheet. Very recent studies suggest that the core gp120 has a propensity to more frequently sample the activated state with a folded bridging sheet (Figure 1e),[25] while the same study suggests that the full-length gp120 monomer and the native state spike-bound gp120 are flexible, conformationally diverse, and not in an activated state.[25] The latter is in line with previous findings suggesting conformational diversity as a mechanism of viral evasion.[28,29] It has been proposed that peptide triazoles inhibit interactions of gp120 with CD4 and the coreceptor surrogate, mAb 17b, by trapping the flexible ground state viral protein in one or more inactive conformations[15,16,30] and possibly preventing bridging sheet formation.[16]

Currently, there are no crystal structures of a peptide triazole−gp120 complex, hampering efforts to optimize the peptide potency and develop non-natural peptidomimetics through structure-based design. The only direct study aimed at

dx.doi.org/10.1021/bi400166b | *Biochemistry* 2013, 52, 2245−2261

identifying binding determinants on the peptide employed saturation transfer difference nuclear magnetic resonance (STD NMR) to investigate binding of the parent 12p1 peptide to gp120. Those results showed that the bulky hydrophobic side chain of the peptide tryptophan binds to gp120,[31] consistent with previous synthetic variation studies of the peptide sequence.[15−17] However, neither a mechanistic explanation of the interaction nor a model of the complex was suggested. Binding of peptide triazoles to full-length gp120 is characterized by a smaller negative entropy change ($-T\Delta S \cong 6.3$ kcal/mol)[30] compared to that of CD4−gp120 binding ($-T\Delta S$ = 44.2 kcal/mol),[28] suggesting that gp120 maintains substantial conformational flexibility when bound to the peptide. This might be one of the reasons the gp120−peptide complex has resisted crystallization efforts to date. In this work, we set out to build atomistic models of gp120−peptide triazole complexes. We hypothesized that the peptide binds to a large hydrophobic cavity on the glycoprotein that is exposed in bridging sheet naïve conformations of gp120, like the mAb F105-bound conformation. We used all-atom, explicit-solvent molecular dynamics to populate various conformations of an encounter complex consistent with this hypothesis. The most stable and energetically favorable complex showed a specific orientation of the peptide with its tryptophan inside the F43 pocket and its isoleucine plugging a shallow cavity at the junction of the inner and outer domains. This remarkably stable pose was reached after backbone movements of the β20−β21 loop that buried the large phenyl-triazole appendage on the derivatized proline. Independently, we used STD NMR to study the peptide residues that come into contact with gp120 using a less potent version of the peptide. These results are in good agreement with the simulation model and show that the hydrophobic core of the peptide provides the major contact surface for gp120 interaction. In addition, the peptide−protein interface from the model encompasses all the residues that were found in a parallel study to be important for binding of the peptide to gp120.[32] Finally, we tested the predictive utility of the model by modifying a PT−gp120 contact that was found to be important in the simulations and verified this prediction experimentally. Altogether, these findings provide a working model for the mechanism of dual inhibition of peptide triazoles and can help both in the design of peptides with improved potencies and as a guide for efforts to produce a stabilized complex for future crystallographic analysis.

MATERIALS AND METHODS

Peptide Synthesis. Peptide triazoles (Table 1) were synthesized via manual solid phase synthesis using Fmoc chemistry on a Rink amide resin at 0.25 mmol scale, purified using reverse phase HPLC, and validated by matrix-assisted laser desorption ionization time-of-flight mass spectrometry (MALDI-TOF MS) as detailed elsewhere.[30]

Protein Production. Monoclonal antibody 17b was obtained from Strategic BioSolutions. Soluble CD4 (sCD4) and full-length HIV-1_{YU2} gp120 were produced and purified as described previously.[30] The S375A mutation was produced and purified as reported previously.[32] Briefly, DNA for transient transfection was purified using the Qiagen MaxiPrep kit (Qiagen) and transfected into confluent HEK 293F cells according to the manufacturer's protocol (Invitrogen). Five days later, cells were harvested and spun down, and the supernatant was filtered through 0.2 μm filters. Purification was performed using nickel affinity chromatography as described previously,[33] followed by size-exclusion chromatography on a Hiload 26/60 Superdex 200 column (GE Healthcare) to separate gp120 aggregates from monomers. The size and purity of the final product were verified by separation via 7.5% acrylamide nonreducing sodium dodecyl sulfate−polyacrylamide gel electrophoresis and staining with Coomassie blue.

Surface plasmon resonance (SPR) kinetic interaction analysis experiments were performed on a Biacore 3000 (GE) machine. gp120 (WT and S375A) was immobilized on a CM5 chip using standard NHS/amine conjugation. Anti IL5α receptor mAb 2B6R was immobilized to 2000 RUs on one chip flow cell for use as a control surface. For direct binding experiments with the KR21 peptide or sCD4 to gp120, analytes were injected at a rate of 100 μL/min for 2.5 min, followed by a 5 min dissociation. Surfaces were regenerated using 10 mM HCl injection for 3 s. All experiments were conducted in filtered and degassed PBS with 0.005% Tween 20 and 0.1% sodium azide (pH 7.2) at 25 °C. For KR21/sCD4 competition experiments, serial dilutions of the peptide with a constant concentration of sCD4 (50 nM) were injected over the gp120 surface. Data were fit to the Langmuir 1:1 binding equation to estimate steady state affinity in Biaevaluation version 4.1 (direct binding) or to a four-parameter sigmoidal (competition) in Origin version 7.0 (OriginLab).

Saturation Transfer Difference (STD) NMR. All NMR data were recorded at 298 K on a Bruker Avance 600 NMR instrument equipped with a cryogenically cooled *z*-shielded gradient probe. Series of one-dimensional STD spectra were acquired with selective irradiation at −1.5, 11, and 12 ppm for on-resonance spectra and 30 ppm for off-resonance "reference" spectra, using a train of 50 ms Gaussian-shaped radiofrequency pulses separated by 1 ms delays and an optimized power level of 55 db. Water suppression was achieved using excitation sculpting.[34] ^{1}H and ^{13}C chemical shift assignments were made by interpreting COSY, HOHAHA, and NOESY spectra of the free peptides in solution. Samples were prepared in 20 mM sodium phosphate buffer containing 50 mM NaCl and 1 mM EDTA at final concentrations of 15 μM gp120 and 1.2 mM peptide. The pH of each sample was adjusted to 6.85 with 0.5 mM NaOH or HCl (not corrected for ^{2}H). The NMR data were processed and analyzed with Topspin version 2.1.6; in particular, each on-resonance spectrum was subtracted from its corresponding off-resonance spectrum to give a difference spectrum that was phase corrected using Topspin version 2.1.6. Relative STD enhancements were determined by integrating peak intensities and normalizing to peak(s) showing the largest enhancement (here Trp indole protons). To ensure that artifacts were not introduced with protein irradiation, enhancements for spectra recorded with the three different on-resonance carrier frequencies described above were compared. Though no STD signals were observed in control spectra using the same pulse sequence in the absence of protein, slight differences (<5%) in enhancements were observed. We therefore chose to integrate the upfield signals in spectra irradiated at 12 ppm and the aromatic signals in spectra irradiated at −1.5 ppm.

Molecular Simulations. Molecular dynamics (MD) simulations were performed using NAMD version 2.7[35] and the CHARRM22 force field.[36] CHARMM parameters for the triazole appendage on the peptide were estimated by similarity to those available in the CHARMM general force field,[37] using ParamChem online tool version 0.9.19 (https://www.paramchem.org/)[38,39] and, if necessary, manual addition of missing parameters (Supporting Information). Initial coordinates of the HIV-1

Table 1. Sequence and Functional Properties of Parent 12p1 and Peptide Triazole Variants[a]

designation	sequence[d]	triazole derivative		chirality Trp C_α	chirality Pro C_γ	$IC_{50(ELISA)}$ (nM)[b] sCD4	$IC_{50(ELISA)}$ (nM)[b] mAb 17b	$IC_{50(SPR)}$ (nM)[c] sCD4	$IC_{50(SPR)}$ (nM)[c] mAb 17b
UM101	Cit^1-N^2-N^3-I^4-X^5-W^6-S^7	phenyl	*S*	*S*	230	670	–	–	
UM25	Cit^1-N^2-N^3-I^4-X^5-W^6-S^7	*p*-aminophenyl	*S*	*S*	1100	18000	–	–	
UM25RS	Cit^1-N^2-N^3-I^4-X^5-W^6-S^7	*p*-aminophenyl	*S*	*R*	$>1 \times 10^5$	$>1 \times 10^5$	–	–	
UM25RR	Cit^1-N^2-N^3-I^4-X^5-W^6-S^7	*p*-aminophenyl	*R*	*R*	$>1 \times 10^5$	$>1 \times 10^5$	–	–	
UM25SR	Cit^1-N^2-N^3-I^4-X^5-W^6-S^7	*p*-aminophenyl	*R*	*S*	$>1 \times 10^5$	$>1 \times 10^5$	–	–	
12p1	R^1-I^2-N^3-N^4-I^5-P^6-W^7-S^8-E^9-A^{10}-M^{11}-M^{12}	–		*S*	–	–	–	1100	1600
HNG156	R^1-I^2-N^3-N^4-I^5-X^6-W^7-S^8-E^9-A^{10}-M^{11}-M^{12}	ferrocenyl		*S*	*S*	–	–	112	152
UM24	Cit^1-N^2-N^3-I^4-X^5-W^6-S^7	ferrocenyl	*S*	*S*	–	–	118	138	
KR21	R^1-I^2-N^3-N^4-I^5-X^6-W^7-βA^{8e}-$BtLys^{9f}$-G^{10}	ferrocenyl		*S*	*S*	16.4	80	85	110

[a]Cit in the peptide sequence denotes citrulline. [b]IC_{50} determined using an ELISA as outlined in Materials and Methods. [c]From refs 19 and 30. The IC_{50} was determined using SPR, and values are shown here for comparison. [d]X in the sequence denotes derivatized azidoproline. [e]βA denotes β-alanine. [f]BtLys denotes N^e-biotinyl-L-lysine.

YU2 gp120 core in complex with the antigen-binding fragment (Fab) of antibody F105 were extracted from PDB entry 3HI1.[26] Missing residues (V4 loop and the base of the V3 loop) were modeled in silico as unstructured loops. Hereafter, we will refer to this system as the "F105-bound" state. All structural manipulations and image rendering were conducted in VMD.[40]

For a typical system setup for MD runs, the polypeptide was solvated in a TIP3P[41] water box with a distance between the polypeptide molecule and the periodic box boundary of at least 20 Å. Neutralizing Na^+ and Cl^- ions were added to a total NaCl concentration of 0.025 M. A 2 fs time step was used in all integrations. The system temperature was kept constant at 310 K by coupling all non-hydrogen atoms to a Langevin thermostat with a friction coefficient of 5 ps^{-1}. Nonbonded interactions were cut off above 9 Å and smoothed to zero beginning from 8 Å. PME long-range electrostatics with a grid spacing of 1 Å was used, and all bonds involving hydrogens were constrained using RATTLE.[42] Equilibration runs were performed at a constant volume (*NVT* ensemble).

We used the F105-bound state directly from the crystal structure (PDB entry 3HI1) as the target for docking studies. UM101 or UM25 was minimized for 20 ps and equilibrated for 20 ns to sample conformations for initial complex pose generation. The minimized conformation of UM25 was manually oriented close to the F43 pocket of the F105-bound gp120 from the crystal structure (PDB entry 3HI1). In this latter state, the F43 pocket is open and exposed, with only the outer domain wall of the pocket formed such that the peptide tryptophan is located close to the F43 pocket and the peptide isoleucine is at the inner domain−outer domain interface. Twenty conformations were randomly selected from the UM25 equilibration run and aligned over this initial conformation using the heavy atoms of these two peptide residues. For UM101 initial poses, the same initial conformations were used, replacing the para amino group on the triazole with hydrogen. Water and neutralizing ions were added to these initial conformations (as explained above), and the system was minimized for 40 ps, keeping all protein non-hydrogen atoms fixed to relieve protein−peptide clashes. An additional 20 ps minimization was performed allowing for movement of the protein atoms. After all runs had been manually inspected to check for any remaining clashes, runs with no overlaps were submitted for 15 ns of free MD equilibration. This resulted in MD simulation for 165 and 225 ns for UM101 and UM25 complexes, respectively.

For each peptide, the MD-generated systems were sampled every 100 ps and pooled. After the conformations had been oriented over the gp120 outer domain, coordinates of the peptide heavy atoms were extracted and clustered using the following procedure. Each peptide conformation represents a point in a 3*N*-dimensional space of Cartesian coordinates, where *N* is the number of atoms considered for analysis. We first performed standard *k*-means clustering on the set of *f* data points extracted from the trajectory, for $2 \leq k \leq 50$. The task was then to find *k* for which the set of cluster centers most efficiently described the complete data set, which we termed *k**. We opted to use similarity between the ordered eigenvectors of the global covariance matrix $C_{\alpha\beta}$ and those of the covariance matrix computed from among the *k*-means $C^k_{\alpha\beta}$. The rationale was that, because covariance matrix eigenvectors defined directions in Cartesian space along which the data were most variable, the mean values (cluster centers) whose covariance matrix eigenvectors best recapitulated these most highly variable directions of the global data set were themselves the most efficient "coarse-graining" of the data set. We used the following rule: *k** is that *k* for which the dot-products of the $k - 1$ eigenvalue-descending ordered eigenvectors of the covariance matrix computed using the *k*-means with their respective eigenvectors of the global covariance matrix are all greater than 0.5, and such that this number of "near-parallel" eigenvectors of the global and *k*-mean covariance matrices is at least equal to the number of global eigenvectors that account for at least 80% of the variance in the global data. The amount of global variance for which each principal component (PC) accounts is equal to the variance of that PC divided by the total data set variance.[43] Because the *k*-means algorithm is sensitive to the choice of the initial cluster seeds,[44] the whole loop was repeated 50 times, and the result that gave the most well-packed clusters (as measured by the sum of Euclidean distances of cluster members from the cluster center) was used as the global optimum. All data points were transformed to standard *z*-scores before analysis. All analyses were conducted in MATLAB (MathWorks). The protein−peptide interaction energy was measured in NAMD as the nonbonded (van der Waals and electrostatic) part of the CHARMM energy between the two molecules, in the absence of water and with a dielectric constant of 80. We here report the root-mean-square deviation (rmsd) of peptide atoms averaged for all member conformations of a cluster as a metric of cluster uniformity and stability. The average nonbonded interaction energy (hereafter termed the interaction energy or simply energy) and average peptide

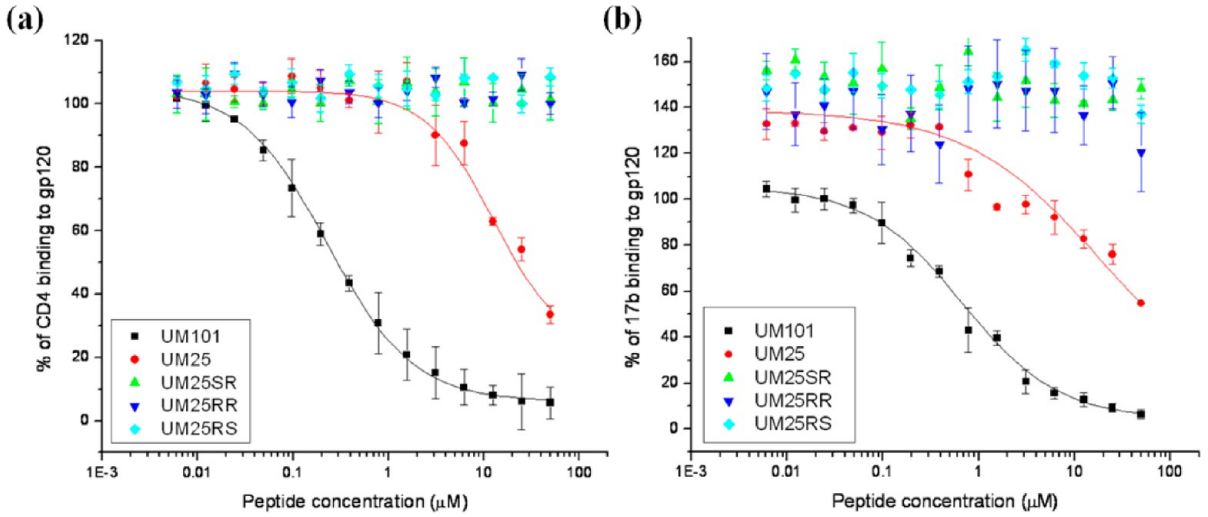


Figure 2. Effect of peptide triazoles on the interaction of HIV-1_{YU2} gp120 with (a) sCD4 and (b) 17b. IC_{50} values were determined by fitting the UM101 and UM25 data to a logistic function in Origin (OriginLab) and are listed in Table 1.

rmsd of each cluster were used as the principal scoring metrics. The internal energy of the protein or the peptide was not included in the scoring as the former would have swamped and smoothed out any differences in the interaction energy, and both were found to be fairly constant throughout all simulations.

We submitted the F105-bound gp120 conformation from PDB entry 3HI1, along with an initial peptide conformation, to the SwissDock server (http://swissdock.vital-it.ch/).[45] SwissDock is a free service that automates pocket detection, pose generation, and scoring process. We used this service as an unbiased method for generation of PT−gp120 encounter complexes. The scoring was based on a combination of contributions from CHARMM energy terms and surface desolvation terms.[46] The top highest-ranking results from the web server were selected, and the top five that were more than 1 Å different in peptide rmsd were submitted for explicit-solvent all-atom MD. Results were clustered and analyzed the same way as the two production sets explained above.

RESULTS

Choice and Characterization of Peptide Triazoles Used in Modeling and NMR Studies. Most of our past work on PTs has used peptides containing a ferrocenyl group on the triazole appendage (Table 1). Although this was found to lead to the most potent peptides, a phenyl moiety on the triazole leads to almost equally potent peptides.[19] We chose to use phenyl-containing triazoles for our modeling and NMR studies as we can use available CHARMM parameters for the phenyl and our most diverse structure−activity studies on the triazole were conducted on a phenyl scaffold.[17] To facilitate STD NMR studies of interactions of peptide triazole with gp120, it was necessary to select peptide constructs that bound gp120 with sufficiently low affinity, because the efficiency of saturation transfer to the ligand depends on its off-rate from the protein.[47] Thus, we introduced a para amino group on the phenyl triazole that, on the basis of past work, was expected to decrease the peptide affinity for gp120[17] and increase solubility.

We used a competition ELISA to measure the effects of the peptide triazole on interactions of gp120 with sCD4 and 17b. The results are shown in Figure 2. Mean inhibitory concentrations (IC_{50}) were calculated after fitting the data to a logistic function in Origin (OriginLab) and are reported in Table 1. Replacing the ferrocenyl triazole of HNG156 with phenyl triazole in UM101 results in a decrease in 17b inhibitory potency and has a similar but smaller effect on sCD4 inhibitory potency (Table 1).[17,19] Changing the phenyl triazole of UM101 to *p*-aminophenyl triazole in UM25 has a much stronger negative effect on the peptide potency and almost reverts the sCD4 IC_{50} to that of the parent peptide 12p1 with no triazole. 17b inhibitory potency is almost lost, and an approximate IC_{50} value of 18 μM can be estimated from the dose response in Figure 2. These results are in line with previous findings for the correlation of the triazole derivative hydrophobicity and peptide affinity, with the more hydrophobic and bulky derivatives producing more potent inhibitors.[17]

Previous work has shown that the stereochemistries of the chiral centers on peptide Pro-C_γ and Trp-C_α are very important for activity: when D-Trp is used along with a ferrocenyl appendage on the triazole, the peptide is almost inactive.[30] Moreover, while the *S* configuration of the triazole on Pro-C_γ is active, the *R* configuration has activity similar to that of the parent peptide with no triazole [12p1 (see Table 1)].[17] Here we aimed to further elucidate these findings through a systematic investigation of the stereochemistry around the two chiral centers and correlate that with NMR data and structural models. To simultaneously investigate the role of stereochemistry in both stereocenters of the hydrophobic X-W core in peptide activity, we compared the UM25 inhibitory potency to those of other diastereomers in which the Pro-C_γ and Trp-C_α atoms possessed the *R* or *S* configuration. The original configuration with *S*-Pro-C_γ and *S*-Trp-C_α (UM25) is active, but any change in the configuration of either chiral center results in an inactive peptide (Figure 2 and Table 1), consistent with prior work on similar but longer peptides containing a ferrocenyl triazole.[17,30] Considering these results, we focused our modeling work on only UM101 and UM25, both of which have the *S* configuration in the chiral centers denoted above.

Simulation of a Peptide Triazole−gp120 Core Encounter Complex. *Target gp120 Conformation for Docking Studies.* Previous work has shown that the parent peptide 12p1 does not inhibit binding of sCD4 to a mutant gp120 (S375W)[15] that has a propensity to sample the activated,

dx.doi.org/10.1021/bi400166b | *Biochemistry* 2013, 52, 2245−2261

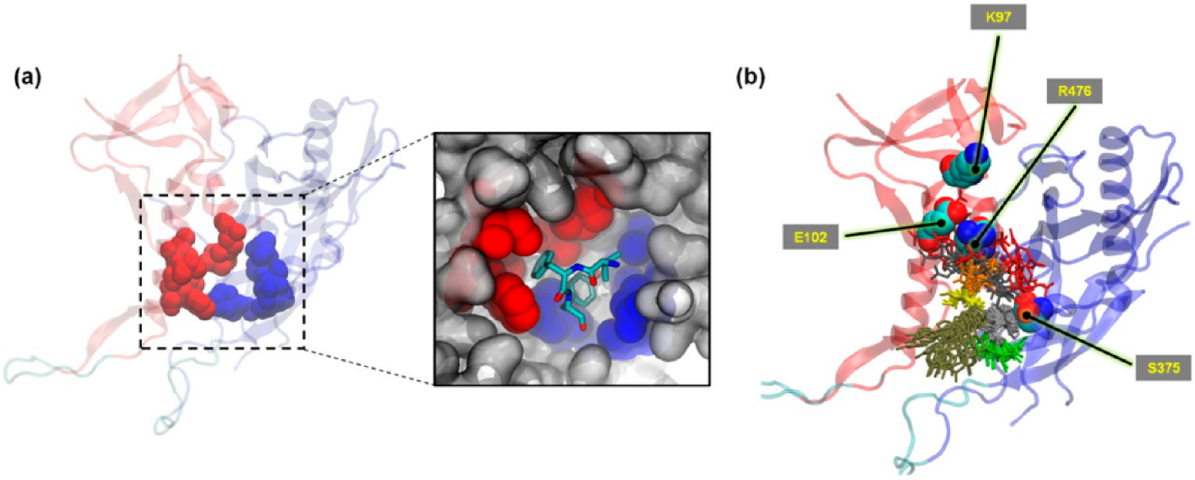


Figure 3. (a) F105-bound gp120 with the Fab removed. Red spheres represent the residues that make up the small pocket at the interface of the inner and outer domains and blue spheres some of the outer domain residues that are part of the (unformed) F43 cavity. The inset shows the placement of the tip of F105 CDR H3 loop residues V100a, F100b, and Y100c (in cyan stick representation) at the gp120 interface. F100b inserts into the small pocket at the interface of the inner and outer domains (red spheres), and Y100c is buried deep in the (unformed) F43 pocket (blue spheres). (b) Initial conformations of the UM25–gp120 encounter complex before the start of minimization and equilibration steps. Peptide residues are colored on the basis of residue number, with red depicting the N-terminal (citrulline) and light green depicting the C-terminal (serine) residues. The gp120 backbone color-coding is that of Figure 1. gp120 residues for which mutation was found to affect binding of the parent peptide 12p1[15]are shown as spheres and labeled along with S375.

CD4-bound state,[48] while it inhibits binding of the F105 antibody to a mutant gp120 (I423P) that is inefficient at assuming the activated state.[15,48] Additionally, on the basis of surface plasmon resonance (SPR) results, it was suggested that 12p1 inhibits binding of sCD4 to gp120 through an allosteric mechanism.[15] The observations described above were used to propose that 12p1 traps gp120 in an inactive conformation that does not favor the CD4-bound state and lacks the bridging sheet. This is consistent with follow-up calorimetric studies that showed that binding of both 12p1 and peptide triazoles results in significantly less structuring of gp120 compared to binding of sCD4.[17,30] Because a major determinant of gp120 structuring upon CD4 binding and activation is formation of the bridging sheet[49] and consistent with the aforementioned studies, we hypothesized that the gp120 conformation targeted by peptide triazoles is a conformation that has an unstructured bridging sheet.

There are three gp120 structures resolved to date that lack the bridging sheet. These are gp120 complexes with mAbs F105, b12, and b13[26,27] (Figure 1b–d, respectively). We used the freely available program Fpocket (http://fpocket.sourceforge.net) to detect druggable pockets on these structures,[50,51] where "druggable" is any pocket with a drug score of >0.5.[51] The algorithm failed to find any druggable pockets on the b12- and b13-bound conformations, and indeed, it failed to find any pocket at all on the b12-bound conformation. This is consistent with the fact that these structures were resolved with a highly engineered gp120 sequence containing multiple pocket-filling mutations and disulfide stitches that were introduced to stabilize the complexes for crystallization.[26] In contrast, the Fpocket algorithm detected a large druggable pocket in the F105-bound structure overlapping the domain where the F43 pocket is located in the CD4-bound state (Figure S1 of the Supporting Information). Additionally, the F105-bound conformation is the only bridging sheet naïve conformation of gp120 that has been crystallographically resolved with a WT sequence.[26] Finally, the F105-bound conformation is unique in lacking a secondary structure on both loops that make up the bridging sheet, consistent with recent findings on the solution structure of these loops.[52]Taking these points together, we selected this conformation as the target structure for further studies.

Role of the Peptide Hydrophobic Core in Binding. The I-X-W hydrophobic sequence in the middle of the peptide triazoles is crucial for their activity[30] and requires a specific X-W stereochemistry.[17,30] A survey of 103 high-resolution peptide–protein complexes in a prior study has shown that only a few amino acids of a peptide sequence typically contribute most of the interaction energy with its partner, commensurate with the "hot spot" hypothesis for general protein–protein interactions.[53,54] More interestingly, it was found that phenylalanine, leucine, tryptophan, tyrosine, and isoleucine are by far the most frequently observed amino acids in peptide–protein interface hot spots.[54] On the basis of these observations, and considering the previous work on truncated peptides,[30] we hypothesized that the I-X-W hydrophobic core acts as the main interacting motif of the peptide.

Initial Pose Generation for Simulations. We combined two points (use of the F105-bound conformation and the possible role of the peptide hydrophobic core in gp120 binding) to formulate a strategy for generating initial peptide–gp120 poses. Examination of the F105–gp120 interface in the crystal structure revealed that the tip of the complementarity-determining region of the third variable loop in the heavy chain (CDR H3), comprised mainly of three residues, V100a, F100b, and Y100c, inserted deeply into the (yet unformed) F43 pocket (Figure 3). Previously, it was reported that a pocket-filling S375W mutation in gp120 completely abrogated binding of the parent 12p1 peptide.[15] Considering these observations, we further hypothesized that the peptide tryptophan could bind inside the F43 pocket, like Y100c of F105, and that the peptide isoleucine could insert into the small cavity formed by I108, I109, W112, F210, V255, and M475 of gp120, similar to F100b of F105 (Figure 3a). Interestingly, in the activated state of

dx.doi.org/10.1021/bi400166b | *Biochemistry* 2013, 52, 2245–2261

gp120, this latter pocket (hereafter termed the "W427" pocket) is occupied by W427 on the β20–β21 loop, which is part of the bridging sheet. This conserved residue is necessary for CD4–gp120 interaction,[24] and its mutation to valine can improve F105 recognition,[55] leading one to speculate that in an unactivated conformation, this W427 pocket is unoccupied. Taken together, we postulated that the components of the peptide triazole hydrophobic core that are common between PTs and the parent 12p1 bind to these gp120 pockets (F43 and W427 pockets), partially mimicking the F105 CDR H3 binding in the F105-bound conformation of gp120. Thus, we used the Ile and Trp of the peptide to orient all MD-sampled peptide conformations over gp120 for initial pose generation. It is worth noting that such peptide orientation points the hydrophilic N-terminal sequence of the peptide toward the residues the mutation of which was shown to affect binding of the parent peptide 12p1 (K97, R476, and E102). These points are depicted in Figure 3b.

Simulation Results. We tested the hypotheses described above by applying the MD-based sampling scheme outlined in Materials and Methods to a set of initial poses for UM101 and UM25 peptides. Figure 3b shows the initial poses of the MD runs of the UM25–gp120 complex. As an unbiased control, we also used the SwissDock server to produce energy-minimized peptide–gp120 docked complexes that we then subjected to the same all-atom explicit solvent MD scheme explained above (code-named UM101SD and UM25SD, further discussed below). After the solvent and neutralizing ions had been removed, a PCA-based *k*-means clustering algorithm was used to group similar conformations together. The average peptide–gp120 interaction energy of each cluster was used as the main scoring criterion. The average rmsd of peptide heavy atoms within each cluster was used as a measure of cluster uniformity and pose stability, postulating that a low-energy favorable encounter complex should result in a structurally stable peptide and a cluster of similar conformations. Note that the rmsd can be calculated against any reference conformation and the choice of the reference conformation will not have any significant effect on the relative difference between clusters. Figure 4 shows the mean interaction energy of each cluster plotted against the mean cluster rmsd. For UM101, there is a distinct collection of clusters that clearly are separated from the other clusters, showing a significant difference in potential energy and also lower intracluster fluctuations (as measured by the intracluster rmsd). The three best scoring clusters were on average more than 10 kcal/mol lower in energy scale than the next cluster. For UM25, there was no such obvious separation of clusters using similar metrics, although among the top three best scoring clusters (8, 18, and 22), cluster 22 has the lowest mean rmsd, almost half that of the lowest-energy cluster (cluster 8). As cluster 22 demonstrated the lowest energy and highest stability (bottom left of the plot in Figure 4), we chose cluster 22 as the top ranking cluster for UM25.

To compare these results with those of another docking method and to probe the effect of initiating our MD runs from a limited set of initial poses, we submitted UM101 and UM25 to the SwissDock server.[45] We chose SwissDock over other freely available docking software primarily because it supports inclusion of non-natural amino acids out of the box and without the need for user-supplied force filed parameters, provides an all-atom model of the complex, and is suitable for relatively large ligands with many (>10) rotatable bonds. The results of MD runs and clustering for UM101 and UM25, starting from these optimized SwissDock poses, are represented in Figure 4 as UM101SD and UM25SD. The initial poses, based on a priori knowledge of gp120 conformation and past experimental knowledge of 12p1–gp120 interactions, result in clusters that are more homogeneous and have lower interaction energies (compare empty and filled symbols in Figure 4). Interestingly, after MD equilibration, the UM101SD pose with the lowest global interaction energy resulted in a conformation in which the peptide hydrophobic core was roughly positioned similarly to our initial UM101–UM25 docking poses, although UM101SD was initiated from a different conformation generated by the SwissDock server (Figure S2 of the Supporting Information).

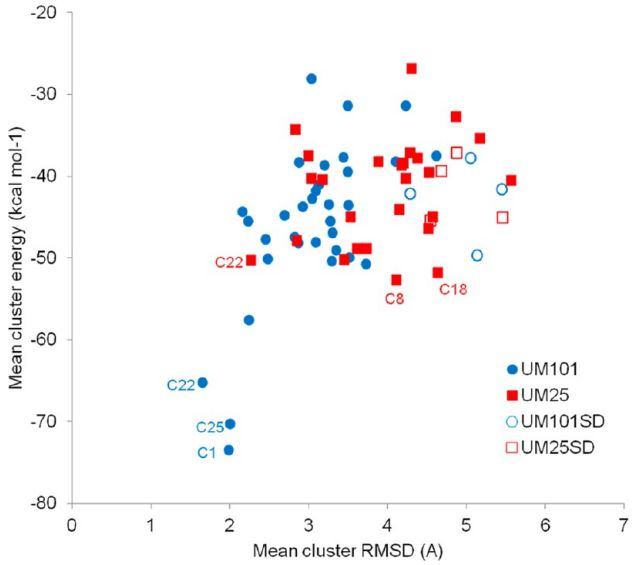


Figure 4. Average scoring energy (CHARMM intermolecular interaction energy) of each cluster plotted vs average cluster variation (average rmsd over all cluster members) for all the runs after clustering of MD-generated conformations. UM101SD and UM25SD were runs initiated from best results out of SwissDock. The top three clusters from UM101 and UM25 are colored blue and red, respectively.

We further examined specific gp120–peptide triazole contacts for the top ranking UM101–UM25 clusters (Figure 5).

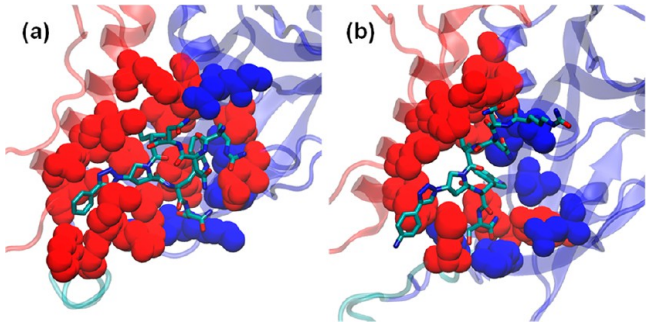


Figure 5. Minimal energy representatives from the top scoring clusters for (a) UM101 (cluster 1) and (b) UM25 (cluster 22). Heavy atoms of gp120 residues that contribute more than 2% of the total interaction energy for each cluster are shown as red spheres. Heavy atoms of residues that act as a donor or acceptor of hydrogen bonds with occupancy values of >1% are shown as blue spheres. For residues that appear in both groups, those with a hydrogen bond occupancy of <10% are colored red. Peptide heavy atoms are depicted in sticks. Backbones of the gp120 inner and outer domains are color-coded as in Figure 1.

dx.doi.org/10.1021/bi400166b | *Biochemistry* 2013, 52, 2245–2261

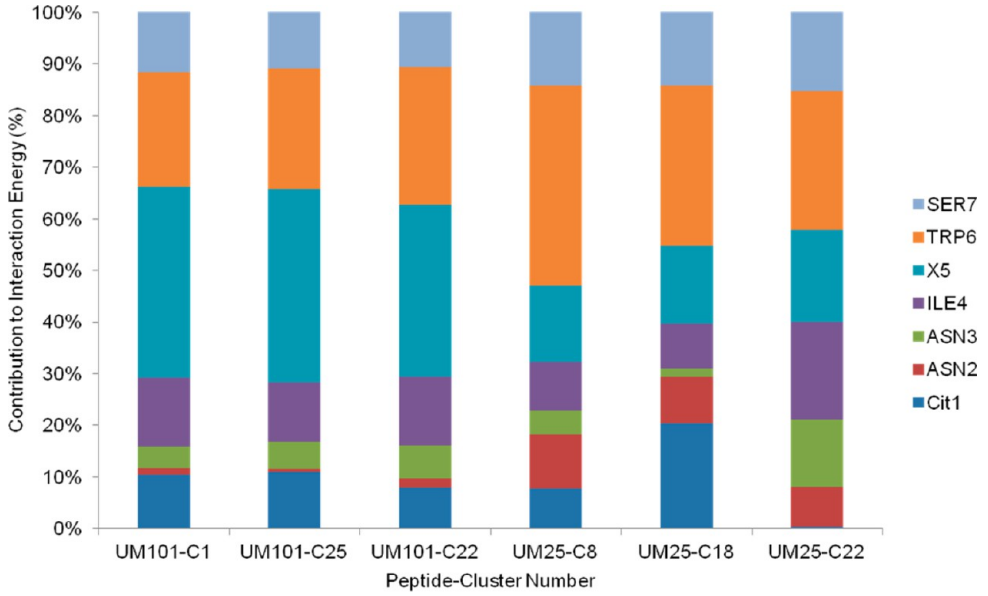


Figure 6. Contribution of each peptide residue to peptide−protein interaction energy in the three top ranking clusters of UM101 and UM25 from the simulations. C1, C25, etc., denote cluster numbers.

We measured the intermolecular interaction energy for all members of a cluster and determined the fraction of that energy contributed by each gp120 residue (Table S1 of the Supporting Information). We also counted the intermolecular hydrogen bonds (with a distance cutoff of 3 Å and angular cutoff of 30°) and determined the fraction of conformations in each cluster where a gp120 residue participates in hydrogen bonds. This fraction was used as the intermolecular hydrogen bond occupancy for that residue (Table S2 of the Supporting Information). These data are summarized in Figure 5 (and further detailed in Figures S3 and S4 of the Supporting Information). Here, we have colored gp120 residues that contribute more than 2% of the intermolecular interaction energy for each cluster red. gp120 residues participating in intermolecular hydrogen bonds are colored blue, and for residues belonging to both groups, if the hydrogen bond occupancy was <10%, we colored the residues red.

The major structural difference between the top ranking clusters of UM101 and UM25 is the burial of the triazole moiety at the interface of the α1 C-terminus and β20−β21 loop in the former. This contributes on average 26−27 kcal/mol to the potential energy of interaction between the peptide and gp120. Measuring the fraction of potential energy contributed by each peptide residue, and averaging the results for the top three clusters, shows that the X-W cluster contributed 60 and 50% of the interaction energy in UM101 and UM25, respectively (Figure 6). This is consistent with previous results that showed these two residues to be critical for any functionality in the peptide.[30] Although both peptides are similar in the overall contribution of this hydrophobic core to the interaction energy, the relative role of the two residues is drastically different. While in UM101 the triazole makes up 36% of the energy versus 24% for the tryptophan, in UM25 it is the tryptophan that contributes more (32%) than the triazole (16%). Considering the depicted models of the complex (Figure 5), such results reflect the observation that in UM101 a large hydrophobic moiety is desolvated while in UM25 it stays solvated though still in contact with the protein. Interestingly, UM25 has an IC_{50} for inhibition of gp120−sCD4 interaction similar to that of 12p1, which does not contain a triazole appendage (Table 1). The role of hydrophobic group desolvation in improving the affinity of ligands is well-documented[56,57] and is consistent with previous results that show a direct correlation between the hydrophobic solvent accessible surface area (SASA) of different triazole derivatives and the affinity of peptide triazoles.[17] We also observed that UM101 is ∼5 times more potent than UM25 in inhibiting binding of gp120 to sCD4 (Table 1). Burial of the more hydrophobic triazole at the interface of the inner and outer domains provides an explanation for the significantly improved potency of UM101 versus that of UM25.

Correlation of the Model of the Peptide−gp120 Complex with Experimental Data and Prediction of Important Contacts. We used the past work on the interaction of the parent 12p1 with gp120 to guide the initial pose generation for a peptide triazole−/gp120 encounter complex.[28,58] More specifically, we took into account the effect that the S375W and I423P mutations have on 12p1 inhibition of interactions of gp120 with different ligands, along with information about the conformation targeted by the peptide gleaned from antibody binding data and isothermal titration calorimetry. More recent work in our group, through analysis of binding of peptide triazole to a panel of gp120 alanine mutants, demonstrated that D474 and T257 had a significant effect on PT−gp120 interactions, with smaller effects observed for E370, M475, V255, and R476.[32] All of these residues are among those that contribute to van der Waals and electrostatic interactions with the peptide (Table S1 of the Supporting Information) or hydrogen bonds (Table S2 of the Supporting Information) in the simulations. Because those studies were conducted using a longer, high-affinity peptide with a ferrocenyl triazole (KR21), we consider only our UM101 simulations for comparison. We show the top UM101 pose (cluster 1) along with the residues mentioned above in Figure 7.

Focusing first on the hydrophilic residues surrounding the F43 pocket, we found E370 or D474 side chains serve as acceptors of hydrogen bonds from the charged N-terminal backbone of both UM101 and UM25 peptides (see Table S2 of

dx.doi.org/10.1021/bi400166b | *Biochemistry* 2013, 52, 2245−2261

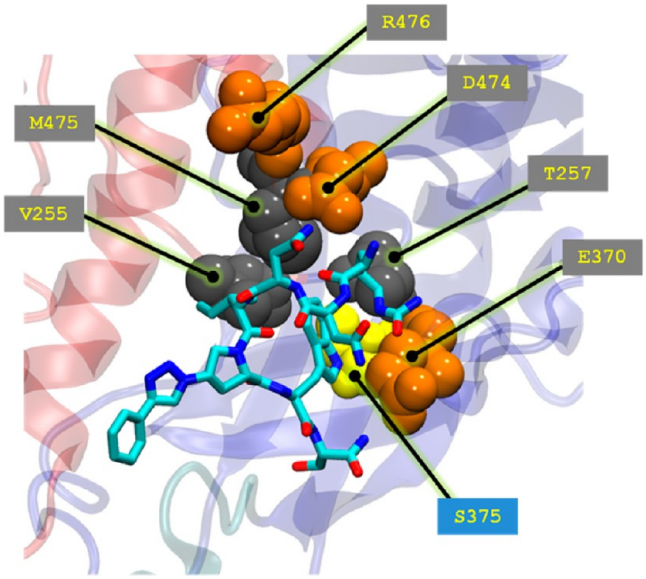


Figure 7. Experimentally verified gp120 residues important for peptide binding (spheres) in the context of the UM101 minimal energy conformation of cluster 1 (Figure 5a). Hydrophilic residues surrounding the putative F43 pocket are colored orange; residues lining the inside of the pocket are colored gray, and gp120 S375 (discussed later in the text) is colored yellow. The gp120 backbone coloring is that of Figure 1.

the Supporting Information). In contrast, persistent interactions were observed with R476 only in the UM25 docked pose. In addition, such a positioning of the peptide N-terminus puts the two asparagines at interaction distances from a series of hydrophilic gp120 residues on the CD4 binding loop (E370 and D368), the α5 helix (D474 and D476), and the α1 helix (E102). Indeed, the top two UM101 clusters are different only in how the N-terminus is placed relative to these gp120 domains (not shown). In cluster 1, the N-terminus and N3 of the peptide interact with D474 [hydrogen bond occupancy of ~80% (Table S2 of the Supporting Information)], while the citrulline (Cit1) side chain and N2 interact with E370 and D368 [vdW and electrostatic (Table S1 of the Supporting Information)]. In cluster 25 (the second ranking cluster), their placements are exchanged [D370−Cit1 hydrogen bond occupancy of ~42%, N3−D474 occupancy of ~5% (data not shown)]. In other words, the hydrophilic N-terminus of the peptide might be able to use the ring of hydrophilic residues surrounding the F43 pocket for hydrogen bonding interactions with the side chains of the two peptide asparagines or salt bridge formation with the charged N-terminus of the peptide. This is also consistent with previous studies that have shown that hydrophilic mutations toward the N-terminus of the peptide yield compounds that still maintain activity.[30] For example, mutation of the citrulline side chain to arginine or glutamic acid barely has any effect on peptide binding. Even complete deletions at these positions are tolerated without a loss of function. Side chain variability is much less tolerated at the N-terminus and quickly vanishes as one moves closer to the hydrophobic core. Although peptides with hydrophilic mutations at the N-terminal asparagine maintained some activity, variations in the second asparagine led to an inactive peptide.[30] These observations are consistent with our model, in which the peptide N-terminus can be localized on either side of the entrance to the (putative unformed) F43 pocket (D474 or E370) and allow formation of hydrogen bonding networks through interactions of the two asparagine side chains with the remaining hydrophilic residues on gp120.

gp120 residues T257, M475, and V255 line the inside of the F43 cavity in the CD4-bound conformation and are completely buried in that state, while they are more exposed in the nonactivated conformations such as the F105-bound conformation. In the latter, more "open" inactive conformations, these residues form the outer domain wall of the yet unformed F43 pocket. Combined, these residues contribute ~12% of the total potential energy of interaction between UM101 and gp120 cluster 1. V255 and M475 contribute to the gp120 W427 pocket, which is plugged by the peptide isoleucine in the docked model. While these residues seem to interact with both peptide Ile and Trp, the critical T257 seems to solely interact with the important peptide tryptophan. Together, these data show that our model is consistent with recent experimental data for the PT−gp120 interface.

The gp120 residues that were experimentally determined to affect peptide−gp120 interactions are all among the contacts that contribute either to vdW−electrostatic interaction (Table S1 of the Supporting Information) or hydrogen bonds (Table S2 of the Supporting Information) between the peptide and gp120 in the simulations. The only exception is R476, which only mildly affects the PT−gp120 interaction.[32] This residue was found to be more important for the parent, lower-affinity peptide 12p1.[15] Interestingly, in our simulations, this residue contributed to the complex stability of only the lower-affinity peptide UM25. In addition to these contacts, we observed other gp120 residues that may be important for interactions with PTs but were not included in the studies of Tuzer et al.[32] The functionally critical peptide tryptophan donates a persistent hydrogen bond to gp120 S375, while its phenyl ring rests against the methyl group of T257 (the role of this latter contact will be discussed further below). The orientation of the peptide tryptophan shown in UM101 cluster 1 stacks the indole ring against the highly conserved bulky F382 inside the pocket, commensurate with the importance of such $\pi-\pi$ interactions in ligand recognition[56] and mimicking the interaction of F105 Y100c with gp120 F382 (Figure 3a). W112, F382, I423, M426, I109, L116, and N425 all contribute significantly to the intermolecular interaction energy of the UM101−gp120 complex (Table S1 of the Supporting Information). Overall, these observations suggest that the proposed model of the peptide triazole−gp120 interaction is consistent with the current data for the role of different gp120 residues in modulating peptide interactions.

Characterization of Binding of Peptide Triazole to gp120 by NMR. We used STD NMR to independently characterize the binding of peptide triazoles to soluble gp120. In this method, saturation of protein resonances results in the transfer of magnetization from the receptor to the bound ligand in a distance-dependent manner. We introduced a *p*-aminophenyl group in place of the phenyl in UM101 to increase the off-rate of the peptide−gp120 interaction for optimal results in STD NMR experiments[47] and to improve peptide solubility. In a difference spectrum, signals corresponding to protons closest to the protein in the bound state appear to be the strongest signals. Together with full chemical shift assignments (Table S3 of the Supporting Information) made via $^{1}H-^{1}H$ and $^{1}H-^{13}C$ two-dimensional NMR experiments, an atomic level description of a binding epitope can be obtained.

dx.doi.org/10.1021/bi400166b | *Biochemistry* 2013, 52, 2245−2261

As seen in Figure 8, strong STD enhancements (red spectrum) are seen for numerous protons of UM25. These include

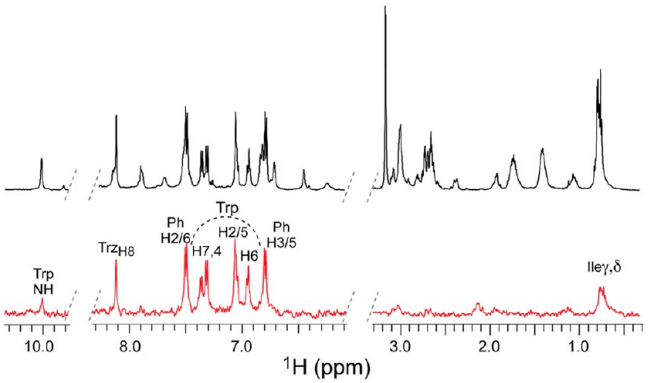


Figure 8. STD NMR spectrum of UM25. Reference and difference spectra (see Materials and Methods) are colored black and red, respectively (see Materials and Methods). Protons exhibiting the largest enhancements are labeled (see Table 2 for quantification). Protons of Trp are located under the curved, dotted line. Trz and Ph denote the triazole and *p*-aminophenyl rings, respectively (see Figure 9).

protons of the indole ring of Trp, and the *p*-aminophenyl triazole group, as well as γ- and δ-methyl groups of isoleucine. When compared to the reference spectrum, the strongest signals in the difference spectrum correspond to H-2, H-4, H-6, and H-7 of tryptophan (Figure 9). The percent enhancement

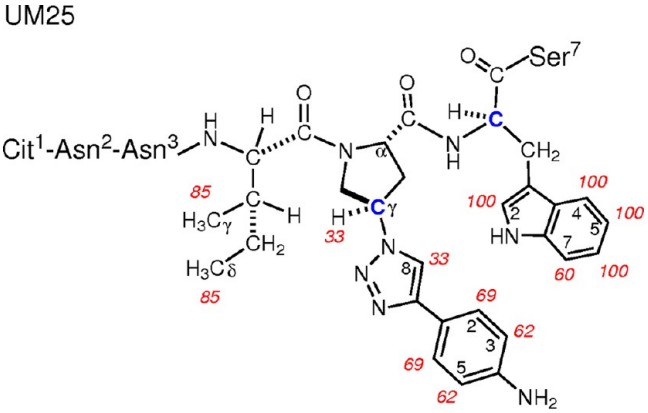


Figure 9. Chemical structure of UM25 showing protons involved in gp120 binding. Percent enhancements observed in the STD NMR spectrum are labeled in red (Table 2). The two stereocenters studied for UM25 are indicated as blue carbons.

for each of the remaining signals was normalized to these protons, and the percent enhancements for each are shown in Figure 9 and Table 2. In addition to Trp, the side chains of Ile and Pro-Hγ strongly contribute to the binding to gp120, with STD enhancements ranging from 60 to 85%. Apart from protons located within these three residues, there are no other large enhancements observed in the difference spectra. Together, these data suggest that these three residues in large part govern binding of UM25 to gp120. This is consistent with suggestions from the simulation model that the hydrophobic core of peptide triazoles is the major binding partner to gp120, with the I-X-W motif contributing ~60% of the interaction energy in UM25.

Binding of UM25 to gp120 Is Stereospecific. Previously, we demonstrated that the *S* configurations on Trp-C_α and Pro-C_γ

Table 2. Percent Enhancements for Binding of UM25 to gp120 Observed by STD NMR

residue	proton	enhancement (%)
Ile[4]	Meγ	~85[a]
	Meδ	~85[a]
X[5]	Hγ	33
	H-2/6	69[b]
	H-3/5	62[c]
	H-8$_{Trz}$	33
Trp[6]	H-2/5	100[c]
	H-4	100
	H-6	100
	H-7	60

[a]Average of two methyl signals. [b]Average of H-2 and H-6. [c]Average of H-3 and H-5.

were important for the inhibitory function of peptide triazoles.[17,30] In this work, we changed the stereochemistry of these two chiral centers and observed that changing the Trp-C_α configuration from *S* to *R* renders the peptide almost inactive, while changing the Pro-C_γ at the triazole derivative attachment point from *S* to *R* reverts the activity back to the parent peptide without the triazole (Table 1 and Figure 2). To directly compare binding interactions among the four diastereomeric analogues of UM25 to gp120, we recorded STD NMR spectra on each of these peptides in complex with gp120. All samples were prepared and spectra recorded in an identical manner (Materials and Methods). For comparison, the four difference spectra are shown as stacked plots in Figure 10. The STD

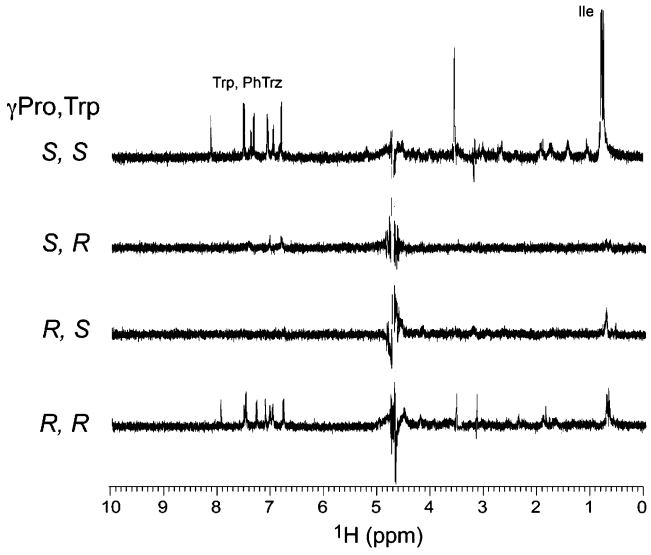


Figure 10. STD NMR spectra of the four diastereomers of UM25 in complex with gp120. The stereochemistry of the two chiral centers is indicated at the left. PhTrz denotes the *p*-aminophenyl triazole.

NMR spectra show that the UM25 analogue bearing D-Trp (*S* configuration of the side chain) in the wild-type background and the analogue bearing *R*-Pro-Hγ in the presence of L-Trp do not bind to gp120. Interestingly, for the UM25 analogue in which both stereocenters have been inverted, to *R*-Pro-Hγ and D-Trp, binding to gp120 is observed. However, the signal-to-noise ratio for this spectrum is less than half that measured for wild-type UM25, and while enhancements are seen for protons

dx.doi.org/10.1021/bi400166b | *Biochemistry* 2013, 52, 2245−2261

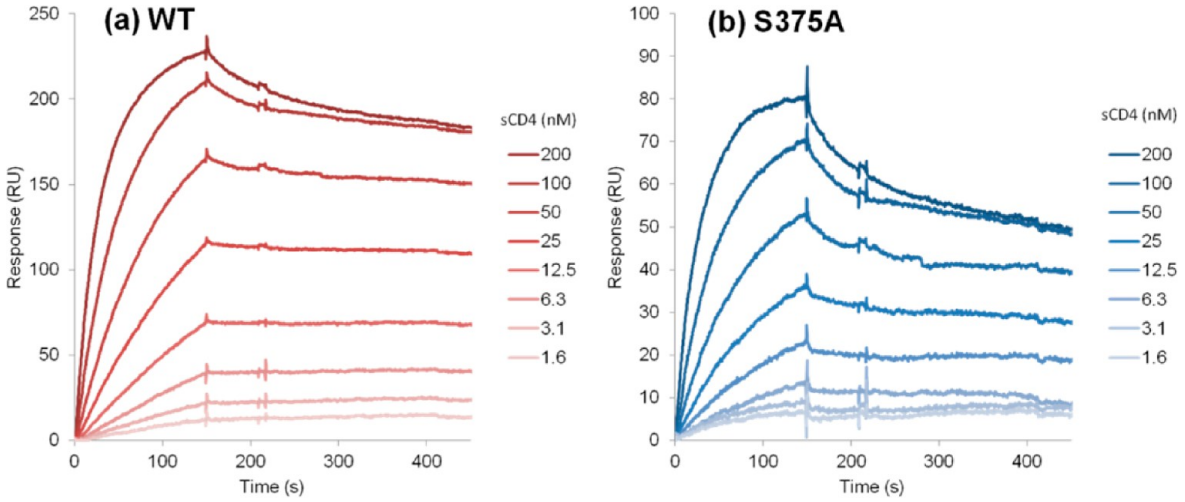


Figure 11. Binding of sCD4 to immobilized (a) WT and (b) S375A gp120. A lighter color of the SPR signal denotes a lower sCD4 concentration. Fitting of the data to a steady state affinity model yielded K_D values of 37 nM for WT and 41 nM for S375A gp120. Data are averages of duplicate determinations.

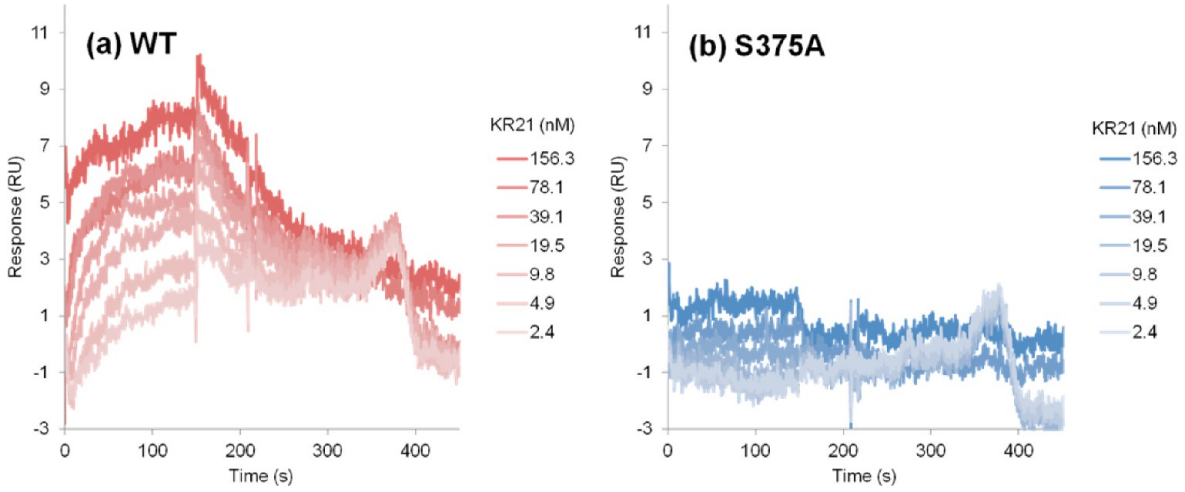


Figure 12. Direct binding of KR21 peptide triazole to immobilized (a) WT and (b) S375A gp120. A darker color of the SPR signal denotes a higher peptide concentration. Fitting of the data for WT to a steady state affinity model yielded a K_D of 9 nM. S375A sensorgrams did not approach saturation and hence could not be fit. Data are averages of triplicate determinations.

in the Pro-Hγ and Trp residues, negligible enhancements are observed for Ile. Thus, in addition to an apparent reduction in binding affinity, these data suggest that the entire tripartite arrangement presented by Ile, *S*-Pro-Hγ, and Trp is required for functional binding to gp120.

Experimental Test of Model Predictions. While the MD simulation model is consistent with recent mutational studies of PT–gp120 interaction[32] and the STD NMR studies presented above, a more direct verification of the model validity would be testing the predictions it provides. Considering the critical role the peptide tryptophan plays in its activity,[16,17] we examined its contacts with gp120 in the model more closely. Interestingly, a persistent hydrogen bond between the Trp side chain and gp120 S375 is observed in the UM101 model (as well as the UM25 model) (Figure 5 and Table 2). S375 is among the residues that line the inside of the F43 cavity (Figures 1 and 7) and, as such, can be considered a hydrogen bonding partner in a mostly hydrophobic cavity. Hydrogen bonds in hydrophobic pockets have been shown to be major contributors to stabilization of protein–ligand interactions.[51,59,60] These observations led us to postulate that the hydrogen bond between the peptide Trp side chain and gp120 S375 is important for peptide activity and that breaking this bond should lead to decreased peptide triazole activity. To test this hypothesis, we probed the interaction of a biotinylated peptide triazole [KR21 (Table 1)] with the gp120 S375A mutant.

We first verified the functionality of the S375A gp120 mutant by measuring its affinity for sCD4 (Figure 11). Fitting of the sensorgram data to a steady state 1:1 affinity model (Materials and Methods) yielded K_D values of 37 nM for WT and 41 nM for gp120 S375A. Hence, S375A maintains correct gp120 folding and activity. We then probed the affinity of KR21 for this mutant by a direct binding SPR assay wherein serial dilutions of the peptide were injected over immobilized WT and S375A (Figure 12). Steady state fits for WT sensograms yielded a K_D of 9 nM for KR21–gp120 interactions; however, binding of KR21 to S375A was severely disrupted, and we could not fit the data. A reduction in KR21 affinity for S375A, combined with the unaltered affinity of the mutant for sCD4, leads one to speculate that the latter would be more resistant to

dx.doi.org/10.1021/bi400166b | *Biochemistry* 2013, 52, 2245–2261

KR21 inhibition of the gp120−sCD4 interaction. Measuring the inhibitory potency of KR21 for PT−gp120 interaction provides an alternative means of quantifying the effect of the S375A mutation on PT−gp120 interaction. We mixed a nonsaturating amount of sCD4 (50 nM) with serial dilutions of KR21 and injected the mixture over immobilized gp120. The resulting SPR sensograms are shown in Figure S5 of the Supporting Information. Steady state response values at 290 s (end of the association phase), normalized to the response in the absence of the peptide, were used to generate the dose−response data in Figure 13. Fitting of the results to a four-parameter

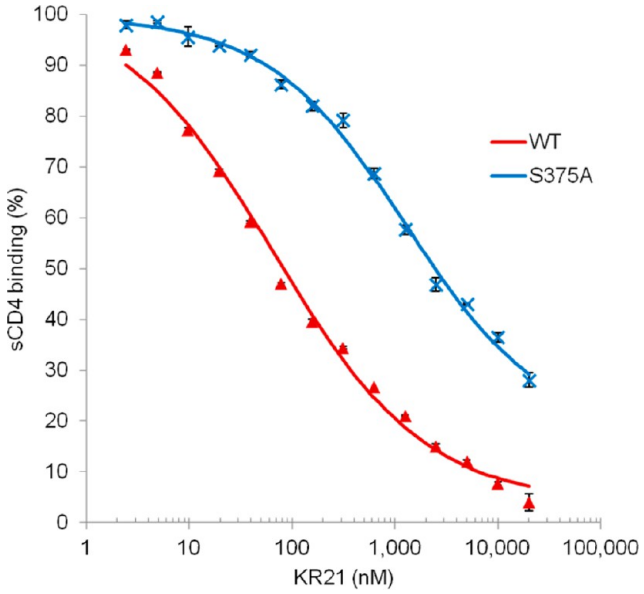


Figure 13. KR21 inhibition of interaction of WT (red) and S375A (blue) gp120 with sCD4. Data points were extracted from sensograms of Figure S5 of the Supporting Information and fit to a four-parameter sigmoidal equation resulting in IC_{50} values of 65 nM for WT and 1283 nM for S375A.

logistic curve yielded IC_{50} values of 65 and 1283 nM for WT and S375A, respectively. Consistent with the direct binding data, the competition results show a large-scale reduction (20-fold) in the sCD4−S375A inhibitory potency of KR21 compared to that of WT gp120. This is consistent with the simulation-based hypothesis that the peptide tryptophan side chain makes a hydrogen bond with the gp120 S375 inside the F43 pocket, and that disruption of this bond will significantly affect PT−gp120 interaction.

DISCUSSION

The ability of peptide triazoles to inactivate HIV-1 by specifically targeting a highly conserved site on Env gp120,[32] combined with the inability to date to obtain a crystallographic structure of the complex, provided the impetus in this study to simulate the complex using molecular dynamics. Prior work has suggested that peptide triazoles stabilize gp120 in a conformation that is not as structured as the activated form of the protein and lacks an ordered bridging sheet.[21,22] Here, we used the bridging sheet disordered F105-bound gp120 conformation as the simulation target and employed an explicit-solvent MD-based scheme to sample structures of the peptide−protein encounter complex. Clusters of conformation with minimal average energy and intracluster rmsd were selected for further analysis. Measuring the peptide−gp120 interaction energy in these clusters showed that the I-X-W hydrophobic cluster is the hot spot for peptide−gp120 interactions, contributing 50−60% of the intermolecular nonbonded energy. These were consistent with previous reports of the importance of the N-N-I-X-W motif for peptide activity.[15,30] In addition, STD NMR analysis confirmed the importance of the I-X-W motif in the peptide for gp120 binding and showed that binding of this tripartite hydrophobic domain is the sole contributor to STD NMR signals of the peptide. In simulations, a specific geometry in the peptide hydrophobic core was required to insert into the hydrophobic cavities present in the F105-bound gp120: the minimal energy poses from simulations showed that a change in stereochemistry at Trp-C_α and Pro-C_γ chiral centers led to clashes with the protein surface or forced the triazole moiety away from the protein surface. Consistent with the simulation results, our NMR experiments also showed no signal for binding of peptide diastereomers with *R*-Trp-C_α and *R*-Pro-C_γ. We measured the contribution to the interaction energies from different gp120 residues and found that those that were experimentally determined to be important for peptide triazole binding were also major contributors to the interaction energy or participated in persistent hydrogen bonds with the peptide. Furthermore, we identified additional gp120 residues the mutation of which would be predicted to reduce gp120 affinity. Among those, we tested the effect of mutating gp120 S375 on binding of a peptide triazole variant (KR21). Consistent with our prediction, this mutation decreased the sCD4 IC_{50} of KR21 by 20-fold. Together, these results provide a self-consistent model of the peptide triazole−gp120 encounter complex, which can be used both to design a stabilized covalent conjugate for further structural studies and to test ideas for the rational design of improved entry inhibitors.

We used the F105-bound gp120 conformation as our target for PT−gp120 complex simulations, based on previous studies suggesting that peptide triazoles bind to a bridging sheet naïve gp120 conformation.[23,25,28,29] We used the Fpocket algorithm[25] to look for suitable pockets on available bridging sheet naïve gp120 crystal structures. Previous studies have shown that peptides mostly bind to the largest pockets on protein surfaces,[54] and the F105-bound structure was the only conformation that displayed a large druggable cavity. In addition, in contrast to the b12- and b13-bound gp120 conformations that were resolved with highly engineered sequences, the F105−gp120 crystal structure was resolved with a WT sequence and did not contain any engineered mutations.[28,30] Interestingly, recent hydrogen−deuterium exchange studies of the full-length gp120 conformations in solution showed that the $\beta2-\beta3$ and $\beta20-\beta21$ domains that fold into the bridging sheet in the activated state are largely disordered in unliganded gp120.[52] This is a unique property of the gp120 conformation in complex with F105, in contrast to the b12- and b13-bound structures. Finally, recent simulation studies have demonstrated the flexibility of the $\beta20-\beta21$ domain in the F105-bound conformation[25] and the corresponding amenability of the F43 pocket to expansion.[61] Despite these findings, we used the SwissDock server to generate possible initial PT−gp120 docked poses; however, all of the top 20 poses were on the gp120 silent face, and none were close to the functionally relevant gp120 F43 pocket. Taken together, these results all support our choice of the F105-bound state as the target for docking studies.

We clustered the PT−gp120 complexes from MD simulations based on peptide conformations and selected the most homogeneous clusters with the lowest values of the intermolecular part of CHARMM interaction energy as the best scoring clusters (Figure 4). Simulation of the encounter complex of both UM101 and UM25 showed a similar binding pose for the peptide hydrophobic core and slightly different trajectories for the peptide N-termini (Figure 5). The common hydrophobic core contacts were reflected in gp120 residues contributing significantly to the interaction energy (Table S1 of the Supporting Information) or intermolecular hydrogen bonds (Table S2 of the Supporting Information). Our simulations point to the I-X-W motif as the most important contributor of interaction energy between gp120 and PT (Figure 6). Similarly, our STD NMR results point to the hydrophobic I-X-W motif as the only peptide−gp120 interface (Figure 9). Both observations are consistent with previous studies that established the N-N-I-X-W sequence as the minimal sequence required for peptide activity.[26]

We used peptide triazoles with phenyl (UM101) and *p*-aminophenyl triazoles (UM25) on the central proline for our simulation and STD NMR studies. The latter facilitated our STD NMR studies by providing better solubility and faster off-rates, while the high affinity of the former for gp120 provided us with a peptide that better represented the high-potency members of this class of molecules. We also synthesized stereoisomers of UM25 with different configurations around the tryptophan C_α and azidoproline C_γ (Figure 9). Consistent with our ELISA results on their lack of inhibitory potency (Figure 2), we did not observe any significant STD NMR signal for any of the stereoisomers (Figure 10). Our simulations suggest that a precise geometry around the tripartite hydrophobic core of PTs is necessary for their function, consistent with previous studies of longer PT sequences.[23] Hence, this model provides a working hypothesis for these effects by proposing that a change in stereochemistry at the chiral centers mentioned above will lead to clashes with the protein surface (for the tryptophan side chain) or force the triazole moiety away from the protein and into the solvent (for triazole).

While the hydrophobic core conformation in the encounter complex for both UM101 and UM25 is similar, more variability is observed in the N-terminus. Past work in our group has shown that the N-terminal hydrophilic extension on the peptides can be modified to include positively or negatively charged or neutral residues without much difference in the affinity of the peptide for gp120.[30] Such findings are consistent with our simulations. It has been suggested that electrostatic interactions serve to guide peptide−protein interactions to an initial, metastable encounter complex that is then stabilized through binding of a small number of hot spot residues.[54,62] Therefore, we think that hydrogen-bonding, hydrophilic N-terminal extension residues of the peptide bind to the ring of charged residues surrounding the F43 pocket [like E370 and D474 (Figure 7 and Tables S1 and S2 of the Supporting Information)] and localize it close to this flexible pocket, and then in a second step, the hydrophobic hot spot residues bind to the (possibly transiently exposed) F43 pocket, latching the peptide onto gp120. Launching a 200 ns MD simulation from minimal energy cluster 1 of UM101, we observed that the hydrophobic core of the peptide is solidly bound to gp120 (peptide core heavy atom rmsd of ≅2 Å), while the hydrophilic N-terminus is flexible (data not shown) and transiently samples the two hydrophilic domains surrounding the F43 cavity (centered on D474 and E370, both of which were found to be important PT−gp120 contacts).[32]

The most significant difference in the peptide pose of the top scoring complexes for UM101 and UM25 was at the triazole appendage. While this moiety was fully buried (FSASA = 0.13) in the simulated structure of the UM101−gp120 complex (Figure 5a) and contributed the largest fraction of intermolecular interaction energy, it stayed well-solvated (FSASA = 0.62) in the case of UM25, and its contribution to the interaction energy was reduced, surpassed by that of the peptide tryptophan (Figure 6). Previous studies established that hydrophobic substitutions on the triazole enhance the peptide potency,[17] consistent with the generally accepted positive contribution of hydrophobic group desolvation to increasing drug potency.[56,57] These results are further corroborated by our STD NMR results that show difference signals for the triazole moiety (Figure 9), although the signal is smaller than the tryptophan signals. Though we did not observe complete, persistent burial of the triazole in UM25 (which was used in NMR studies here) simulations, the triazole in this peptide stayed close to the surface of gp120. Such proximity can give rise to saturation transfer difference signals.[63] Interestingly, the UM25 inhibitory potency of gp120−sCD4 interaction is similar to that of 12p1, which lacks the triazole moiety and hence will lack the positive contribution its burial provides for the intermolecular interaction energy. Complete burial of the phenyl triazole of the high-affinity UM101 peptide at the interface of the β20−β21 loop and the α1 C-terminus and lack of such persistent desolvation for the lower-affinity UM25 *p*-aminophenyl triazole provide a structural explanation for the improved potency of the former peptide.

The strongest STD NMR signals arise from the hydrogens on the tryptophan indole, indicating that this residue is in intimate contact with gp120. This is consistent with the simulation results that show that the peptide tryptophan is stabilized inside the (yet unformed) gp120 F43 cavity, particularly through a side chain hydrogen bond with gp120 S375. These two independent observations are consistent with past results that showed that mutation of this tryptophan abrogates peptide activity.[15,16]

The next set of strong signals arises from Ile4 of the peptide, which in our model resides inside a shallow cavity at the interface of the inner and outer domains, which we called the "W427" pocket (Figures 3 and 5). As such, the peptide tryptophan fills a cavity that is occupied by F105 F100b. On the basis of our model, we think introduction of more hydrophobic residues (like Phe) at this position on the peptide can improve its inhibitory potency.

The triazole appendage makes up the last bit of the I-X-W "tripartite" pharmacophore, with STD enhancements as high as 80% of those observed for the tryptophan protons. We find it surprising that moving from the triazole ring toward the *p*-aminophenyl appendage, the magnitude of the STD NMR signal seems to decrease. This may indicate burial of this moiety inside a transient or flexible pocket. The interface between two flexible domains may be a plausible option for such a pocket.

Our data on gp120 contacts at the PT−gp120 interface correlate with recent studies in which gp120 residues important for PT binding were investigated through alanine scanning mutagenesis of some gp120 mutants.[32] Residues that were shown to be important for binding of PT to gp120 (D474, T257, E370, M475, V255, and R476)[32] were also found in our simulations to contribute to either the intermolecular

interaction energy or hydrogen bond formation in the PT−gp120 complex for both UM101 and UM25 simulations, although R476 was implicated only in simulations of the low-affinity UM25 peptide. This latter residue was found to have an impact on interactions of the parent low-affinity 12p1 peptide with gp120[15] but had a smaller impact on the interactions of the higher-affinity variants (like KR21) than D474 or T257. Because the mutational studies were performed with a high-affinity peptide (KR21), we focused on the UM101−gp120 pose generated from the simulations.

In addition to being consistent with these observations, our model suggests additional contacts that were not studied by Tuzer et al.:[32] W112, F382, I423, M426, I109, L116, and N425 were gp120 residues that made the largest contribution to interaction energy with UM101. It would be interesting to test the effect of mutations at these sites on PT−gp120 interactions. Therefore, our model not only is consistent with recent experimental data on the PT−gp120 interaction but also can be used to make specific predictions about how one might further test the validity of the model using additional mutations.

As an example, and to further test the predictive utility of our model, we focused on gp120 S375. The side chain of this residue in our model accepts a persistent hydrogen bond from the critical peptide tryptophan side chain (Figure 7 and Table S2 of the Supporting Information). We hypothesized that this buried hydrogen bond contributes significantly to the stability of the PT−gp120 complex. To test this, we probed the interactions of a PT variant with a C-terminal biotin (KR21) with the gp120 mutant S375A using SPR. Direct binding experiments (Figure 12) showed that the affinity of KR21 for gp120 is significantly reduced. To quantify the extent of this effect, we tested the KR21 inhibitory potency against the interaction of S375A with sCD4. Although the affinity of S375A for sCD4 is almost unchanged compared to that of WT, the KR21 inhibitory potency for this mutant is 20-fold reduced (Figure 13). This was consistent with our model-based prediction and provides direct evidence of the utility of this model in structure-based studies of PT−gp120 interactions.

Among all the clades tested (clades A−D), only a gp120 sequence (90CM243) from the circulating recombinant form AE (CRF01_AE) was fully resistant to peptide triazole inhibition of sCD4 binding (up to 4 μM peptide triazoles);[20] 99% of sequences in this strain of HIV-1 have a histidine at position 375.[64] Interestingly, the S375W gp120 mutant was found to be resistant to the parent 12p1 peptide,[15] while we observed that S375A is less sensitive to PTs (Figure 13). We think the lower potency of KR21 for S375A argues for the role of a direct contact effect in this case and possibly a combination of direct contact and conformational alteration for S375W. The significantly larger side chain of S375W would completely obstruct the critical peptide tryptophan from accessing contacts inside the F43 pocket. Considering the similarity of the aromatic tryptophan and histidine side chains,[64] we think complete blockade of all peptide tryptophan contacts inside the F43 pocket (Figure 7) might explain the resistance of the 90CM243 strain to peptide triazoles. Therefore, we think our model can provide a structure-based hypothesis for the observed resistance of a gp120 strain. Furthermore, we predict that all clade AE sequences will be resistant to peptide triazoles, as long as they have a bulky side chain like a histidine at position 375.

Further studies will try to overcome the limitations of our current model. We do not know the conformation of gp120 that is targeted by the peptide. Using the "right" target conformation immensely impacts the success of docking models.[65] In using MD to drive conformational sampling of both the peptide and the protein, our aim was to start from a potentially "wrong" conformation and arrive at a local minimum on some complex free energy surface that would provide a better picture of the peptide−gp120 interactions. Previous work in our lab has generated ensembles of bridging sheet naïve gp120 conformations[66] that can be used in ensemble docking methods.[65] Our use of similarity-derived CHARMM parameters for the triazole moiety might have negatively impacted the quality of the results. Although surveys of PDB structures have found that ligand strain energies of several kT are common,[67] supporting some acceptable margin of error in dihedral parameters used in this work, we recognize that improving the quality of these parameters would be important for improving the current model. Finally, we used a simple nonbonded interaction energy score for ranking the MD-generated poses. Although other, more advanced scoring schemes can be used, the ambiguity in the target conformation of gp120 argues against investing in such expensive methods because, using such a simple scoring scheme, we were able to distinguish poses that are consistent with independently derived experimental results. In addition to the single mutation studied here (S375A), testing additional predictions of the current model experimentally (i.e., introduction of phenylalanine in place of Ile4 on the peptide to improve its affinity or mutating W112 or I109 on gp120 to probe the effect on peptide binding) can provide more direct cues for refining the model using more rigorous simulation methods.

■ ASSOCIATED CONTENT

Supporting Information

Tables S1−S3, Figures S1−S5, and CHARMM force field parameters for triazole. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*A.E.: Department of Biochemistry and Molecular Biology, Drexel University College of Medicine, 11102 New College Building, 245 N. 15th St., Philadelphia, PA 19102; phone, (215) 762-3535; e-mail, ali.emileh@drexel.edu. C.F.A.: Chemical and Biological Engineering, Drexel University, 3141 Chestnut St., Philadelphia, PA 19104; phone, (215) 895-2231; e-mail, cfa22@drexel.edu. I.M.C.: Department of Biochemistry and Molecular Biology, Drexel University College of Medicine, 11102 New College Building, 245 N. 15th St., Philadelphia, PA 19102; phone, (215) 762-4197; e-mail, ichaiken@drexelmed.edu.

Author Contributions

A.E. and F.T. contributed equally to this work.

Funding

We acknowledge the National Institutes of Health (NIH) for support through Grants 5R21AI093248-02 and 5R21AI091513-02 (A.E., C.F.A., and I.M.C.) and Grant 5P01GM056550 (A.E., F.T., M.U., J.M.L., and I.M.C.) and the National Science Foundation (NSF) for computational support through TeraGrid/XSEDE allocation MCB070073N (C.F.A.). Funding for D.R.M.M. was provided through a CAPES-Fullbright scholarship. This work was supported in part by the Intramural AIDS Targeted Antiviral Program, Office of the Director, NIH (C.A.B.), and the NIH Intramural Research Program (National Institute of Diabetes

and Digestive and Kidney Diseases). Part of the computational work was performed on the DRACO GPU cluster in the Department of Physics at Drexel University (NSF Grant AST-0959884).

Notes

The authors declare no competing financial interest.

■ ABBREVIATIONS

HIV, human immunodeficiency virus; mAb, monoclonal antibody; STD NMR, saturation transfer difference nuclear magnetic resonance; HPLC, high-performance liquid chromatography; MALDI-TOF MS, matrix-assisted laser desorption ionization time-of-flight mass spectrometry; ELISA, enzyme-linked immunosorbent assay; PBS, phosphate-buffered saline; PBST, phosphate-buffered saline with Tween; BSA, bovine serum albumin; SA, streptavidin; HRP, horseradish peroxidase; OPD, *o*-phenylenediamine dihydrochloride; OD, optical density; COSY, correlation spectroscopy; HOHAHA, homonuclear Hartmann−Hahn spectroscopy; NOESY, nuclear Overhauser effect spectroscopy; EDTA, ethylenediaminetetraacetic acid; NaOH, sodium hydroxide; HCl, hydrochloric acid; MD, molecular dynamics; Fab, antigen-binding fragment; PDB, Protein Data Bank; PC, principal component; PCA, principal component analysis; rmsd, root-mean-square distance or deviation; WT, wild-type; IC_{50}, half-maximal inhibitory concentration; EC_{50}, half-maximal effective concentration; SASA, solvent accessible surface area; RU, response unit.

■ REFERENCES

(1) Poignard, P., Saphire, E. O., Parren, P. W., and Burton, D. R. (2001) gp120: Biologic aspects of structural features. *Annu. Rev. Immunol. 19*, 253−274.

(2) Wyatt, R., and Sodroski, J. (1998) The HIV-1 envelope glycoproteins: Fusogens, antigens, and immunogens. *Science 280*, 1884−1888.

(3) Doms, R. W. (2000) Beyond receptor expression: The influence of receptor conformation, density, and affinity in HIV-1 infection. *Virology 276*, 229−237.

(4) Gallo, S. A., Finnegan, C. M., Viard, M., Raviv, Y., Dimitrov, A., Rawat, S. S., Puri, A., Durell, S., and Blumenthal, R. (2003) The HIV Env-mediated fusion reaction. *Biochim. Biophys. Acta 1614*, 36−50.

(5) Harrison, S. C. (2008) Viral membrane fusion. *Nat. Struct. Mol. Biol. 15*, 690−698.

(6) Didigu, C. A., and Doms, R. W. (2012) Novel approaches to inhibit HIV entry. *Viruses 4*, 309−324.

(7) Tilton, J. C., and Doms, R. W. (2010) Entry inhibitors in the treatment of HIV-1 infection. *Antiviral Res. 85*, 91−100.

(8) Berkhout, B., Eggink, D., and Sanders, R. W. (2012) Is there a future for antiviral fusion inhibitors? *Curr. Opin. Virol. 2*, 50−59.

(9) Moore, J. P., and Kuritzkes, D. R. (2009) A piece de resistance: How HIV-1 escapes small molecule CCR5 inhibitors. *Curr. Opin. HIV AIDS 4*, 118−124.

(10) Si, Z., Madani, N., Cox, J. M., Chruma, J. J., Klein, J. C., Schon, A., Phan, N., Wang, L., Biorn, A. C., Cocklin, S., Chaiken, I., Freire, E., Smith, A. B., III, and Sodroski, J. G. (2004) Small-molecule inhibitors of HIV-1 entry block receptor-induced conformational changes in the viral envelope glycoproteins. *Proc. Natl. Acad. Sci. U.S.A. 101*, 5036−5041.

(11) Zhao, Q., Ma, L., Jiang, S., Lu, H., Liu, S., He, Y., Strick, N., Neamati, N., and Debnath, A. K. (2005) Identification of N-phenyl-N′-(2,2,6,6-tetramethyl-piperidin-4-yl)-oxalamides as a new class of HIV-1 entry inhibitors that prevent gp120 binding to CD4. *Virology 339*, 213−225.

(12) Lin, P. F., Blair, W., Wang, T., Spicer, T., Guo, Q., Zhou, N. N., Gong, Y. F., Wang, H. G. H., Rose, R., Yamanaka, G., Robinson, B., Li, C. B., Fridell, R., Deminie, C., Demers, G., Yang, Z., Zadjura, L., Meanwell, N., and Colonno, R. (2003) A small molecule HIV-1 inhibitor that targets the HIV-1 envelope and inhibits CD4 receptor binding. *Proc. Natl. Acad. Sci. U.S.A. 100*, 11013−11018.

(13) Schon, A., Madani, N., Klein, J. C., Hubicki, A., Ng, D., Yang, X., Smith, A. B., III, Sodroski, J., and Freire, E. (2006) Thermodynamics of binding of a low-molecular-weight CD4 mimetic to HIV-1 gp120. *Biochemistry 45*, 10973−10980.

(14) LaLonde, J. M., Kwon, Y. D., Jones, D. M., Sun, A. W., Courter, J. R., Soeta, T., Kobayashi, T., Princiotto, A. M., Wu, X., Schon, A., Freire, E., Kwong, P. D., Mascola, J. R., Sodroski, J., Madani, N., and Smith, A. B., III (2012) Structure-based design, synthesis, and characterization of dual hotspot small-molecule HIV-1 entry inhibitors. *J. Med. Chem. 55*, 4382−4396.

(15) Biorn, A. C., Cocklin, S., Madani, N., Si, Z., Ivanovic, T., Samanen, J., Van Ryk, D. I., Pantophlet, R., Burton, D. R., Freire, E., Sodroski, J., and Chaiken, I. M. (2004) Mode of action for linear peptide inhibitors of HIV-1 gp120 interactions. *Biochemistry 43*, 1928−1938.

(16) Ferrer, M., and Harrison, S. C. (1999) Peptide ligands to human immunodeficiency virus type 1 gp120 identified from phage display libraries. *J. Virol. 73*, 5795−5802.

(17) Gopi, H., Umashankara, M., Pirrone, V., LaLonde, J., Madani, N., Tuzer, F., Baxter, S., Zentner, I., Cocklin, S., Jawanda, N., Miller, S. R., Schon, A., Klein, J. C., Freire, E., Krebs, F. C., Smith, A. B., Sodroski, J., and Chaiken, I. (2008) Structural determinants for affinity enhancement of a dual antagonist peptide entry inhibitor of human immunodeficiency virus type-1. *J. Med. Chem. 51*, 2638−2647.

(18) Gopi, H. N., Tirupula, K. C., Baxter, S., Ajith, S., and Chaiken, I. M. (2006) Click chemistry on azidoproline: High-affinity dual antagonist for HIV-1 envelope glycoprotein gp120. *ChemMedChem 1*, 54−57.

(19) Gopi, H., Cocklin, S., Pirrone, V., McFadden, K., Tuzer, F., Zentner, I., Ajith, S., Baxter, S., Jawanda, N., Krebs, F. C., and Chaiken, I. M. (2009) Introducing metallocene into a triazole peptide conjugate reduces its off-rate and enhances its affinity and antiviral potency for HIV-1 gp120. *J. Mol. Recognit. 22*, 169−174.

(20) Cocklin, S., Gopi, H., Querido, B., Nimmagadda, M., Kuriakose, S., Cicala, C., Ajith, S., Baxter, S., Arthos, J., Martin-Garcia, J., and Chaiken, I. M. (2007) Broad-spectrum anti-human immunodeficiency virus (HIV) potential of a peptide HIV type 1 entry inhibitor. *J. Virol. 81*, 3645−3648.

(21) McFadden, K., Fletcher, P., Rossi, F., Kantharaju, Umashankara, M., Pirrone, V., Rajagopal, S., Gopi, H., Krebs, F. C., Martin-Garcia, J., Shattock, R. J., and Chaiken, I. (2012) Antiviral breadth and combination potential of peptide triazole HIV-1 entry inhibitors. *Antimicrob. Agents Chemother. 56*, 1073−1080.

(22) Bastian, A. R., Kantharaju, McFadden, K., Duffy, C., Rajagopal, S., Contarino, M. R., Papazoglou, E., and Chaiken, I. (2011) Cell-free HIV-1 virucidal action by modified peptide triazole inhibitors of Env gp120. *ChemMedChem 6*, 1318 , 1335−1339.

(23) Pancera, M., Majeed, S., Ban, Y. E., Chen, L., Huang, C. C., Kong, L., Kwon, Y. D., Stuckey, J., Zhou, T., Robinson, J. E., Schief, W. R., Sodroski, J., Wyatt, R., and Kwong, P. D. (2010) 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. 107*, 1166−1171.

(24) Kwong, P. D., Wyatt, R., Robinson, J., Sweet, R. W., Sodroski, J., and Hendrickson, W. A. (1998) Structure of an HIV gp120 envelope glycoprotein in complex with the CD4 receptor and a neutralizing human antibody. *Nature 393*, 648−659.

(25) Kwon, Y. D., Finzi, A., Wu, X., Dogo-Isonagie, C., Lee, L. K., Moore, L. R., Schmidt, S. D., Stuckey, J., Yang, Y., Zhou, T., Zhu, J., Vicic, D. A., Debnath, A. K., Shapiro, L., Bewley, C. A., Mascola, J. R., Sodroski, J. G., and Kwong, P. D. (2012) Unliganded HIV-1 gp120 core structures assume the CD4-bound conformation with regulation by quaternary interactions and variable loops. *Proc. Natl. Acad. Sci. U.S.A. 109*, 5663−5668.

(26) Chen, L., Kwon, Y. D., Zhou, T., Wu, X., O'Dell, S., Cavacini, L., Hessell, A. J., Pancera, M., Tang, M., Xu, L., Yang, Z. Y., Zhang, M. Y.,

Arthos, J., Burton, D. R., Dimitrov, D. S., Nabel, G. J., Posner, M. R., Sodroski, J., Wyatt, R., Mascola, J. R., and Kwong, P. D. (2009) Structural basis of immune evasion at the site of CD4 attachment on HIV-1 gp120. *Science 326*, 1123−1127.
(27) Zhou, T., Xu, L., Dey, B., Hessell, A. J., Van Ryk, D., Xiang, S. H., Yang, X., Zhang, M. Y., Zwick, M. B., Arthos, J., Burton, D. R., Dimitrov, D. S., Sodroski, J., Wyatt, R., Nabel, G. J., and Kwong, P. D. (2007) Structural definition of a conserved neutralization epitope on HIV-1 gp120. *Nature 445*, 732−737.
(28) Kwong, P. D., Doyle, M. L., Casper, D. J., Cicala, C., Leavitt, S. A., Majeed, S., Steenbeke, T. D., Venturi, M., Chaiken, I., Fung, M., Katinger, H., Parren, P. W., Robinson, J., Van Ryk, D., Wang, L., Burton, D. R., Freire, E., Wyatt, R., Sodroski, J., Hendrickson, W. A., and Arthos, J. (2002) HIV-1 evades antibody-mediated neutralization through conformational masking of receptor-binding sites. *Nature 420*, 678−682.
(29) Myszka, D. G., Sweet, R. W., Hensley, P., Brigham-Burke, M., Kwong, P. D., Hendrickson, W. A., Wyatt, R., Sodroski, J., and Doyle, M. L. (2000) Energetics of the HIV gp120-CD4 binding reaction. *Proc. Natl. Acad. Sci. U.S.A. 97*, 9026−9031.
(30) Umashankara, M., McFadden, K., Zentner, I., Schon, A., Rajagopal, S., Tuzer, F., Kuriakose, S. A., Contarino, M., Lalonde, J., Freire, E., and Chaiken, I. (2010) The active core in a triazole peptide dual-site antagonist of HIV-1 gp120. *ChemMedChem 5*, 1871−1879.
(31) Celigoy, J., Ramirez, B., Tao, L., Rong, L., Yan, L., Feng, Y. R., Quinnan, G. V., Broder, C. C., and Caffrey, M. (2011) Probing the HIV gp120 envelope glycoprotein conformation by NMR. *J. Biol. Chem. 286*, 23975−23981.
(32) Tuzer, F., Madani, N., Kamanna, K., Zentner, I., Lalonde, J., Holmes, A., Upton, E., Rajagopal, S., McFadden, K., Contarino, M., Sodroski, J., and Chaiken, I. (2013) HIV-1 ENV gp120 structural determinants for peptide triazole dual receptor site antagonism. *Proteins 81*, 271−290.
(33) Gift, S. K., Zentner, I. J., Schon, A., McFadden, K., Umashankara, M., Rajagopal, S., Contarino, M., Duffy, C., Courter, J. R., Zhang, M. Y., Gershoni, J. M., Cocklin, S., Dimitrov, D. S., Smith, A. B., Freire, E., and Chaiken, I. M. (2011) Conformational and Structural Features of HIV-1 gp120 Underlying the Dual Receptor Antagonism by Cross-Reactive Neutralizing Antibody m18. *Biochemistry 50*, 2756−2768.
(34) Hwang, T. L., and Shaka, A. J. (1995) Water Suppression That Works: Excitation Sculpting Using Arbitrary Wave-Forms and Pulsed-Field Gradients. *J. Magn. Reson., Ser. A 112*, 275−279.
(35) Phillips, J. C., Braun, R., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E., Chipot, C., Skeel, R. D., Kale, L., and Schulten, K. (2005) Scalable molecular dynamics with NAMD. *J. Comput. Chem. 26*, 1781−1802.
(36) MacKerell, A. D., Bashford, D., Bellott, M., Dunbrack, R. L., Evanseck, J. D., Field, M. J., Fischer, S., Gao, J., Guo, H., Ha, S., Joseph-McCarthy, D., Kuchnir, L., Kuczera, K., Lau, F. T. K., Mattos, C., Michnick, S., Ngo, T., Nguyen, D. T., Prodhom, B., Reiher, W. E., Roux, B., Schlenkrich, M., Smith, J. C., Stote, R., Straub, J., Watanabe, M., Wiorkiewicz-Kuczera, J., Yin, D., and Karplus, M. (1998) All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J. Phys. Chem. B 102*, 3586−3616.
(37) Vanommeslaeghe, K., Hatcher, E., Acharya, C., Kundu, S., Zhong, S., Shim, J., Darian, E., Guvench, O., Lopes, P., Vorobyov, I., and MacKerell, A. D. (2010) CHARMM General Force Field: A Force Field for Drug-Like Molecules Compatible with the CHARMM All-Atom Additive Biological Force Fields. *J. Comput. Chem. 31*, 671−690.
(38) Vanommeslaeghe, K., and Mackerell, A. D., Jr. (2012) Automation of the CHARMM General Force Field (CGenFF) I: Bond Perception and Atom Typing. *J. Chem. Inf. Model. 52*, 3144−3154.
(39) Vanommeslaeghe, K., Raman, E. P., and Mackerell, A. D., Jr. (2012) Automation of the CHARMM General Force Field (CGenFF) II: Assignment of Bonded Parameters and Partial Atomic Charges. *J. Chem. Inf. Model. 52*, 3155−3168.
(40) Humphrey, W., Dalke, A., and Schulten, K. (1996) VMD: Visual molecular dynamics. *J. Mol. Graphics 14*, 27−38 , 33−38.
(41) Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W., and Klein, M. L. (1983) Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys. 79*, 926−935.
(42) Andersen, H. C. (1983) Rattle: A Velocity Version of the Shake Algorithm for Molecular-Dynamics Calculations. *J. Comput. Phys. 52*, 24−34.
(43) Jackson, J. E. (1991) *A user's guide to principal components*, Wiley, New York.
(44) Rencher, A. C. (2002) *Methods of multivariate analysis*, 2nd ed., Wiley, New York.
(45) Grosdidier, A., Zoete, V., and Michielin, O. (2011) SwissDock, a protein-small molecule docking web service based on EADock DSS. *Nucleic Acids Res. 39*, W270−W277.
(46) Grosdidier, A., Zoete, V., and Michielin, O. (2007) EADock: Docking of small molecules into protein active sites with a multiobjective evolutionary optimization. *Proteins 67*, 1010−1025.
(47) Meyer, B., and Peters, T. (2003) NMR spectroscopy techniques for screening and identifying ligand binding to protein receptors. *Angew. Chem., Int. Ed. 42*, 864−890.
(48) Xiang, S. H., Kwong, P. D., Gupta, R., Rizzuto, C. D., Casper, D. J., Wyatt, R., Wang, L. P., Hendrickson, W. A., Doyle, M. L., and Sodroski, J. (2002) 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. 76*, 9888−9899.
(49) Myszka, D. G., Morton, T. A., Doyle, M. L., and Chaiken, I. M. (1997) Kinetic analysis of a protein antigen-antibody interaction limited by mass transport on an optical biosensor. *Biophys. Chem. 64*, 127−137.
(50) Le Guilloux, V., Schmidtke, P., and Tuffery, P. (2009) Fpocket: An open source platform for ligand pocket detection. *BMC Bioinf. 10*, 168.
(51) Schmidtke, P., and Barril, X. (2010) Understanding and predicting druggability. A high-throughput method for detection of drug binding sites. *J. Med. Chem. 53*, 5858−5867.
(52) Guttman, M., Kahn, M., Garcia, N. K., Hu, S. L., and Lee, K. K. (2012) Solution Structure, Conformational Dynamics, and CD4-Induced Activation in Full-Length, Glycosylated, Monomeric HIV gp120. *J. Virol. 86*, 8750−8764.
(53) Clackson, T., and Wells, J. A. (1995) A Hot-Spot of Binding-Energy in a Hormone-Receptor Interface. *Science 267*, 383−386.
(54) London, N., Movshovitz-Attias, D., and Schueler-Furman, O. (2010) The Structural Basis of Peptide-Protein Binding Strategies. *Structure 18*, 188−199.
(55) Wyatt, R., Thali, M., Tilley, S., Pinter, A., Posner, M., Ho, D., Robinson, J., and Sodroski, J. (1992) Relationship of the human immunodeficiency virus type 1 gp120 third variable loop to a component of the CD4 binding site in the fourth conserved region. *J. Virol. 66*, 6997−7004.
(56) Bissantz, C., Kuhn, B., and Stahl, M. (2010) A medicinal chemist's guide to molecular interactions. *J. Med. Chem. 53*, 5061−5084.
(57) Freire, E. (2008) Do enthalpy and entropy distinguish first in class from best in class? *Drug Discovery Today 13*, 869−874.
(58) Madani, N., Schon, A., Princiotto, A. M., LaLonde, J. M., Courter, J. R., Soeta, T., Ng, D., Wang, L. P., Brower, E. T., Xiang, S. H., Do Kwon, Y., Huang, C. C., Wyatt, R., Kwong, P. D., Freire, E., Smith, A. B., and Sodroski, J. (2008) Small-Molecule CD4 Mimics Interact with a Highly Conserved Pocket on HIV-1 gp120. *Structure 16*, 1689−1701.
(59) Gao, J., Bosco, D. A., Powers, E. T., and Kelly, J. W. (2009) Localized thermodynamic coupling between hydrogen bonding and microenvironment polarity substantially stabilizes proteins. *Nat. Struct. Mol. Biol. 16*, 684−690.
(60) Schmidtke, P., Luque, F. J., Murray, J. B., and Barril, X. (2011) Shielded Hydrogen Bonds as Structural Determinants of Binding

Kinetics: Application in Drug Design. *J. Am. Chem. Soc. 133*, 18903−18910.
(61) Korkut, A., and Hendrickson, W. A. (2012) Structural Plasticity and Conformational Transitions of HIV Envelope Glycoprotein gp120. *PLoS One 7*, e52170.
(62) Dagliyan, O., Proctor, E. A., D'Auria, K. M., Ding, F., and Dokholyan, N. V. (2011) Structural and Dynamic Determinants of Protein-Peptide Recognition. *Structure 19*, 1837−1845.
(63) Jayalakshmi, V., and Krishna, N. R. (2005) Determination of the conformation of trimethoprim in the binding pocket of bovine dihydrofolate reductase from a STD-NMR intensity-restrained CORCEMA-ST optimization. *J. Am. Chem. Soc. 127*, 14080−14084.
(64) Schader, S. M., Colby-Germinario, S. P., Quashie, P. K., Oliveira, M., Ibanescu, R. I., Moisi, D., Mesplede, T., and Wainberg, M. A. (2012) HIV gp120 H375 is unique to HIV-1 subtype CRF01_AE and confers strong resistance to the entry inhibitor BMS-599793, a candidate microbicide drug. *Antimicrob. Agents Chemother. 56*, 4257−4267.
(65) Lill, M. A. (2011) Efficient Incorporation of Protein Flexibility and Dynamics into Molecular Docking Simulations. *Biochemistry 50*, 6157−6169.
(66) Abrams, C. F., and Vanden-Eijnden, E. (2010) Large-scale conformational sampling of proteins using temperature-accelerated molecular dynamics. *Proc. Natl. Acad. Sci. U.S.A. 107*, 4961−4966.
(67) Mobley, D. L., and Dill, K. A. (2009) Binding of Small-Molecule Ligands to Proteins: "What You See" Is Not Always "What You Get". *Structure 17*, 489−498.