THE JOURNAL OF BIOLOGICAL CHEMISTRY Vol. 279, No. 29, Issue of July 16, pp. 30523–30530, 2004
Printed in U.S.A.

Characterization of the Conformational State and Flexibility of HIV-1 Glycoprotein gp120 Core Domain*

Received for publication, April 20, 2004, and in revised form, May 6, 2004
Published, JBC Papers in Press, May 6, 2004, DOI 10.1074/jbc.M404364200

Yongping Pan‡, Buyong Ma‡, Ozlem Keskin‡§, and Ruth Nussinov‡¶‖

From the ‡Basic Research Program, Science Applications International Corporation-Frederick, Inc., Laboratory of Experimental and Computational Biology, NCI-Frederick, National Institutes of Health, Frederick, Maryland 21702, the §Koc University, Center for Computational Biology and Bioinformatics and College of Engineering, Rumelifeneri Yolu, 34450 Sariyer Istanbul, Turkey, and the ¶Sackler Institute of Molecular Medicine, Department of Human Genetics and Molecular Medicine, Sackler School of Medicine, Tel Aviv University, Tel Aviv 69978, Israel

gp120 is key to the human immunodeficiency virus type 1 viral cell entry. Knowledge of the detailed conformational states of gp120 is crucial to intervention, yet the unbound form is still resistant to structural characterization probably because of its flexibility. Toward this goal, we performed molecular dynamics simulations on the wild type gp120 core domain extracted from its ternary crystal structure and on a modeled mutant, S375W, that experimentally has a significantly different phenotype from the wild type. Although the mutant retained a bound-like conformation, the wild type drifted to a different conformational state. The wild type β strands 2 and 3 of the bridging sheet were very mobile and partially unfolded, and the organization among the inner and outer domains and β strands 20 and 21 of the bridging sheet, near the mutation site, was more open than in the bound form, although the overall structure was maintained. These differences were apparently a result of the strengthening of the hydrophobic core in the mutant. This stabilization further explains the experimentally significantly different thermodynamic properties between the wild type and the mutant. Taken together, our results suggest that the free form, although different from the bound state, shares many of the bound structural features. The observed loss of freedom near the binding site, rather than the previously hypothesized more dramatic conformational transition from the unbound to the bound state, appears to be the major contributor to the large entropy cost for the CD4 binding to the wild type.

The human immunodeficiency virus type 1 virus continues to be the main cause for AIDS infection affecting millions of people every year. The persistence of the infection is believed to be related to the "wise" evolution of the structure of the envelope glycoprotein gp120 on the virion surface that can efficiently evade the immune system. Contributing elements include the heavy glycosylation on the surface (1), residue variation (2, 3), oligomerization (4, 5), and conformational alterations (6, 7). The structure of the gp120 core domain revealed in a co-crystal with CD4 and a Fab from the neutralizing antibody 17b (6), along with information derived from sequence alignments (8, 9) of different primary immunodeficiency viruses, shows a conserved core domain (see Fig. 1) with five long variable loops V1–V5 interspersed over the surface of the core domain. The molecule has been shown to be able to adopt multiple conformations along the pathway of the infection. Upon binding to CD4, the principal receptor in the course of the infection (10, 11), variable loops V1 and V2 move out of the way allowing transient exposure of the otherwise shielded conserved residues (12). CD4 binding then induces a conformational change that poses the necessary epitopes for its coreceptor association, usually CCR5 (6, 13–15), another major receptor of gp120 (16–20). Analysis of the crystal structure revealed a peculiar three-dimensional organization within the core domain with a deeply recessed CD4 binding pocket shaped by the inner domain, the outer domain, and the bridging sheet. As Kwong *et al.* have argued (6), the hydrophobic core that holds together the three subdomains and sits at the bottom of the recessed pocket might be sensitive to a change in the environment because of the unique organization among the interacting residues. The flexibility of the association between the inner and outer domains is thought to be responsible for the low immunogenicity due to the spread of the epitope residues over separate regions (21). This highly flexible nature of the gp120 has necessarily made it difficult to characterize the detailed conformational features of the free form. Yet, such a characterization of the gp120 would be extremely valuable to the understanding of its biology and intervention, particularly the inhibition of the gp120-CD4 association, which is of primary interest yet has met only limited success (22–24).

Despite the lack of direct structural determination, a large body of experimental work relating to the CD4-gp120 association, including thermodynamics (25), mutagenesis (26–32), and crystallography (21, 33–36), has provided considerable insight into the gp120 structure and flexibility. Of particular note is the thermodynamics data (see Table I) (25), which revealed exceptionally large enthalpy and entropy changes, suggesting that the unbound form of the gp120 may be very flexible and significantly different from the CD4-bound state. This was further supported by the fact that the CD4 molecule shows little conformational change upon ligand binding (6, 33–35), and multiple binding partners for gp120 consistently indicated similar thermodynamic behavior (7). The similar thermodynamic behavior for both the full protein and the core domain of gp120 (see Table I) going from the free-form to the CD4-complexed state, confirms that it is the core domain that is largely

* This work was supported in whole or in part with federal funds from the NCI, National Institutes of Health, under Contract Number NO1-CO-12400. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‖ Supported in part by the Center of Excellence in Geometric Computing and its Applications funded by the Israel Science Foundation administered by the Israel Academy of Sciences. To whom correspondence should be addressed: NCI-Frederick, Bldg. 469, Rm. 151, Frederick, MD 21702. Tel.: 301-846-5579; Fax: 301-846-5598; E-mail: ruthn@ncifcrf.gov.

Downloaded from http://www.jbc.org/ by guest on July 3, 2017

TABLE I
Experimental thermodynamic data for the CD4-gp120 binding
Experimental data were taken from Refs. 7 and 25.

	ΔG	ΔH	$-T\Delta S$	ΔC	K_d
	kcal/mol	*kcal/mol*	*kcal/mol*	*kcal/mol/K*	*nM*
WD61 full-length gp120	−11.8 ± 0.3	−63 ± 3	51.2 ± 3	−1.2 ± 0.2	5 ± 3
Core gp120	−9.5 ± 0.1	−62 ± 3	52.5 ± 3	−1.8 ± 0.4	190 ± 30
Wild-type YU2 gp120 + sCD4	−10.52 ± 0.03	−52.1 ± 0.2	41.6 ± 0.2		38 ± 2
S375W gp120 + sCD4	−11.62 ± 0.08	−35.5 ± 0.2	23.0 ± 0.2		6.4 ± 0.9

responsible for the observed enthalpy and entropy changes. Accordingly, the unusual entropic cost has been interpreted to be the outcome of a large interdomain flexibility, including a loose contact between the inner and outer domains and the unfolded bridging sheet in the unbound state (6, 21).

Interestingly, a mutant, S375W, designed to fill the space at the bottom of the CD4 binding pocket that is not occupied by the CD4, to alter the stability of gp120 and thus the nature for CD4 binding, displayed very different phenotypes than the wild type (7). Both the enthalpy change and the entropy penalty were significantly reduced for the CD4 binding (Table I). Furthermore, the S375W mutant binds CD4 tighter than the wild type, whereas the affinity to antibodies elicited against the CD4 binding site (CD4BS)[1] is almost totally abolished (7) rendering the possible changes in the organizations of the epitopes and the flexibility of the structure. The underlying factors that contribute to the different phenotypes are expected to reside in the structural and flexibility differences between the wild type and the mutant in free and bound forms.

Prompted by the available crystal structure of the wild type core domain in complex with CD4 and an antibody (6), we aimed to characterize the conformations of the gp120 core domain in its free form. We are interested in its structural flexibility, the difference between the free and bound forms, and the effect of the mutation. Toward this goal, we performed molecular dynamics simulations on both the monomers and the CD4-bound complexes for the wild type and the S375W mutant. The detailed structural information from simulations allow us to address several important issues, such as whether there are multiple conformational states in the free form, what the conformational changes are, and how large they are, and how the constellation of the epitope residues changes upon mutation so as to significantly alter the neutralizing interactions. Answers to these questions can not only shed light on the structural biology of the gp120 but further assist in the identification of specific conformations for efficient drug design. Despite the extensive loop truncations in the crystal structure, its biological functions are largely intact, and we expect that information derived from such a reduced model should allow us to capture the structural properties of the core domain. Our results show that whereas the S375W mutant sampled a conformational space that is very close to the CD4-bound state, the wild type preferred a conformation with a very flexible bridging sheet and a relaxed association between the inner and outer domains. Although the findings here are in very good agreement with experimental data in general they further suggest that only the bridging sheet, particularly the β2 and β3 portion, is very mobile, whereas the flexibility between the inner and outer domains is relatively limited in the unbound state.

MATERIALS AND METHODS

All structural manipulations and molecular dynamics simulations were performed using the CHARMM program (37) with the CHARMM 22 force field (38). The starting structures of the gp120 monomers and the gp120-CD4 complex were extracted from the ternary crystal structure (PDB 1g9m) (6). Those with the S375W mutations were modeled from the wild type crystal structure using CHARMM. The missing loop 4 was also modeled in by first building the corresponding 14-residue peptide (the central 12 residues were missing) and subjecting it to molecular dynamics simulation at a temperature of 300 K in the gas phase with the distance between the Cα atoms of the terminal residues restrained to the desired target, which corresponds to 24.3 Å in the crystal structure. The resulting conformations with these matched distances were selected and integrated into the core domain by matching the terminal residues of the Cα of the peptide with the corresponding atoms in the protein. A structure with no van der Waals overlap between loop 4 and the core domain was selected and minimized for 500 steps of the steepest descent algorithm. For the monomer simulations, the gp120 core domain was first solvated in a TIP3P (39) water box with dimensions of $\sim 88 \times 85 \times 71$ Å^3, with a minimum distance of 10 Å from any edge of the box to any protein atom, resulting in a system size of ~52,000 atoms. Minimizations were first performed for 500 steps with the steepest decent algorithm with the gp120 constrained and for additional 500 steps for the whole system. The systems were further pre-equilibrated for 20 ps with the *NVT* ensemble before the production simulations, which lasted for 3 ns each, with the *NPT* ensemble at 300 or 325 K. A time step of 2 fs and a force-switched nonbonded cutoff of 12 Å were used in the trajectory production, and structures were saved every 2 ps for analysis. For the complex simulations, the EEF1 implicit solvent model was used instead, and a total of 4-ns duration of the trajectories was generated with a time step of 1 fs and a group-based electrostatic calculation.

RESULTS

Structural Deviation from the Bound State—The monomer core domain consists of the inner domain, outer domain, and the bridging sheet with its β-strands 20 and 21 folded against the inner-outer domain interface and β-strands 2 and 3 protruding out (Fig. 1). Such a unique monomer organization taken from the ternary complex structure was expected to undergo conformational changes to restore its unbound free form. Tables II and III summarize the core domain RMSD from all four simulations for the monomers. Over the 3-ns period at 300 K, the structural change was moderate with a Cα RMSD for the inner and the outer domains together (Table II) of 1.9 and 1.7 Å for the wild type and the mutant, respectively. When a portion of the bridging sheet, β20,21, was added, the RMSD was essentially unchanged (Table II). However, when the whole domain was included, the RMSD climbed to 2.6 Å for the wild type and only slightly changed for the mutant (Table II), indicating a larger movement of the β2,3 motif in the wild type than in the mutant. The RMSDs, however, did not show dramatic structural changes within each subdomain (inner and outer domains, β20,21 and β2,3). The results from the 325 K simulations were very similar to those from the 300 K simulations except that the difference between the wild type and the mutant was much smaller.

The difference in the structural deviations of the bridging sheet between the wild type and the mutant was further highlighted when the structures were aligned against the inner and outer domains (Table III). At both 300 and 325 K, the RMSD difference between the wild type and the mutant is much larger for the β2,3 and β20,21, whereas it is similar for the inner and the outer domains. These results indicate that the S375W mutation appears to stabilize the bridging sheet from drifting

[1] The abbreviations used are: CD4BS, CD4 binding site; RMSD, root mean square distance.

Downloaded from http://www.jbc.org/ by guest on July 3, 2017

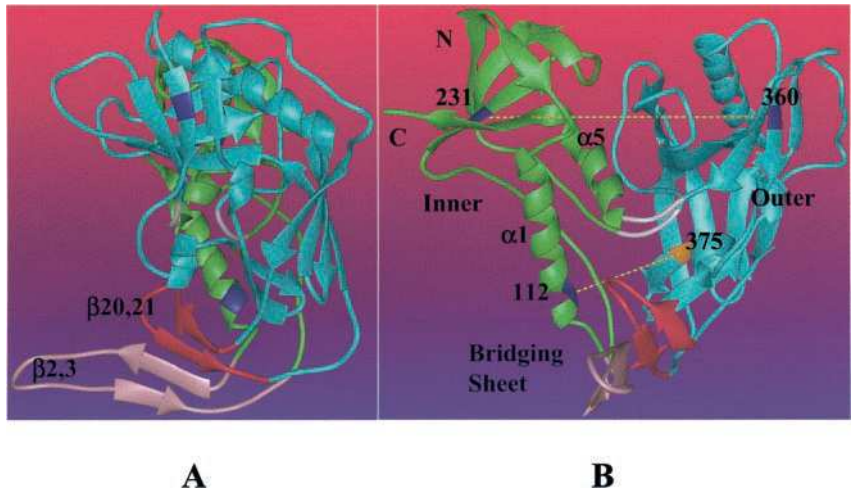


FIG. 1. **Ribbon representation of the gp120 core structure with subdomains color-coded, *green* (inner domain), *cyan* (outer domain), *red* (β-strands 20, 21 of the bridging sheet), and *pink* (β-strands 2,3 of the bridging sheet).** The orientations in *A* and *B* are perpendicular to each other to better view the geometrical relationship between the inner domain, the outer domain, and the protruding bridging sheet.

TABLE II
Cα RMSD for structural motifs of the gp120 with each motif (by itself) aligned
The three terminal residues and loop 4 were not included in the calculation, and averages were based on the last 2 ns of the trajectories.

	Inner	Outer	β20,21	β2,3	Inner + outer	Inner + outer + β20,21	Inner + outer + β2,3,20,21
300K							
WT	1.3 ± 0.1	1.8 ± 0.2	0.9 ± 0.1	1.9 ± 0.3	1.9 ± 0.1	2.0 ± 0.1	2.6 ± 0.3
S375W	1.1 ± 0.1	1.7 ± 0.1	0.8 ± 0.1	1.8 ± 0.3	1.7 ± 0.1	1.7 ± 0.1	1.9 ± 0.1
325K							
WT	1.5 ± 0.1	1.6 ± 0.1	1.1 ± 0.2	2.0 ± 0.4	1.8 ± 0.2	1.9 ± 0.2	2.3 ± 0.3
S375W	1.6 ± 0.2	1.7 ± 0.1	0.8 ± 0.1	2.0 ± 0.4	1.9 ± 0.1	1.9 ± 0.1	2.3 ± 0.2

TABLE III
Cα RMSD for structural motifs of gp120 with inner and outer domains aligned
The three terminal residues and loop 4 were not included in the calculation, and averages were based on the last 2 ns of the trajectories.

	Inner	Outer	β20,21	β2,3
300K				
WT	1.8 ± 0.3	2.1 ± 0.2	3.0 ± 0.6	6.4 ± 1.4
S375W	1.4 ± 0.2	1.8 ± 0.1	1.9 ± 0.5	4.0 ± 0.9
325K				
WT	1.7 ± 0.2	1.8 ± 0.1	3.1 ± 0.9	5.8 ± 1.3
S375W	1.9 ± 0.2	1.9 ± 0.1	2.4 ± 0.4	4.8 ± 1.0

away from the inner and outer domains, which in turn shows that the bridging sheet is the most mobile region of the core domain for the wild type.

Analysis of Interdomain Motions—Above, our results have indicated that the bridging sheet, especially the β2,3 motif, was the most mobile in the core domain of the wild type. To detect additional potential motions between the inner and outer domains, two selected distances between the inner and the outer domains were measured. One distance is between the Cαs of residue 231 and 360, and the other is between the Cα of residue 375 and the nearest heavy atom of residue 112 (Figs. 1*B*, and 2). Residues 231 and 360 are located at the centers of the well structured β sheets, and the distance between them was expected to be a good measure for the opening-closing motions between the inner and outer domains. Residues 112 and 375 were selected to monitor similar motions between the inner and outer domains. These residues are located at the narrower region of the heart-shaped inner domain-outer domain configuration (Fig. 1*B*). Because residue 375 was the mutation site in the S375W mutant, the mutational effect on the local conformations is also of interest.

The distance between residues 231 and 360 fluctuated between 39.5 and 40.5 Å for the wild type at 300 K (Fig. 2*A*). Compared with that in the crystal structure (39.5 Å), the expansion was very mild considering the thermal effect on the structure. The mutant behaved similarly. To increase the sampled conformational space, simulations were also performed at elevated temperature (325 K). However, no significant difference was observed for either the wild type or the mutant (Fig. 2*A*). Thus, it seems that the motions between the domains were quite limited by this measure.

The distance between residues 112 and 375 also increased for the wild type at 300 K (Fig. 2*B*). Starting from 9.4 Å in the bound state, it fluctuated between 10 and 12 Å after 1 ns of the simulation. Such a magnitude of fluctuation is more significant compared with the fluctuation for the distance between residues 231 and 360, indicating that this part of the structure might be more flexible and relatively more open in the free form. Interestingly, this distance for the mutant did not change as much and was consistently shorter than in the wild type (Fig. 2*B*). Such a difference between the wild type and the mutant was more significant when measured at 325 K (Fig. 2*B*). These results indicate that the effect of the S375W mutation was the stabilization of the bound-state association between the inner domain and the outer domain, consistent with the assertion based on the experimental observations (7).

To examine the conformational changes of the gp120 observed by molecular dynamics from a different perspective, we performed a normal mode analysis on both the monomer and the gp120-CD4 dimer with a method developed by Keskin *et al.* (40). This normal mode analysis method uses a coarse-grained model of the system and focuses on the lowest frequency modes.

Downloaded from http://www.jbc.org/ by guest on July 3, 2017

Results from the monomer analysis show that two of the three lowest frequency modes correspond to the β2,3 motif of the bridging sheet (Fig. 3, *A* and *B*), and the third lowest frequency mode involves the β4,5 motif (Fig. 3*C*). In addition, the magnitudes of two modes for β2,3 were significantly larger than the rest of the vibration modes. On the other hand, the magnitudes of these two motions were significantly reduced to a level that was comparable with the rest of the small amplitude motions in the gp120-CD4 dimer complex form (data not shown). This result agrees with our simulation, confirming that β2,3 is indeed the most flexible and might experience a large conformational change during the CD4 binding process.

S375W Mutation Reduces the Flexibility of gp120—The S375W mutation was designed to fill the partially occupied CD4 binding pocket (7) so that the stability or the conformation of gp120 would be altered. Because the mutation site is at the bottom of the CD4 binding pocket, and the pocket is surrounded by several hydrophobic/aromatic residues, this mutation was likely to perturb these interactions in its vicinity. Fig. 4 shows the constellations of the hydrophobic and aromatic residues near the mutation site in the average structures from the 300 K simulations (Fig. 4, *A* and *B*) and snapshots from 325 K simulations (Fig. 4, *C* and *D*). Compared with the wild type, the hydrophobic residues near the mutation site in the mutant shifted slightly toward each other within the pocket, resulting in a better packed hydrophobic cluster because of the presence of Trp-375. As a consequence, the distance between residues 112 and 375 was shortened (Figs. 2 and 4). Such a change in the constellation of the hydrophobic residues might have an effect on the stability of this region. To test whether the stability was affected, residuewise root mean square fluctuations were compared between the wild type and the mutant (Fig. 5). Overall, residues that were nearby the hydrophobic cluster displayed lower fluctuations in the mutant than in the wild type (Fig. 5, *B* and *C*), whereas residues far from this site either became more flexible or stayed the same in the mutant. Interestingly, β2,3 of the bridging sheet was also stabilized, possibly because of the lower mobility of the helix α1 or of the β20,21 of the bridging sheet. Other structural effects resulting from the stabilization include the better retained helical structure of helix α5, especially at the C-terminal (data not shown).

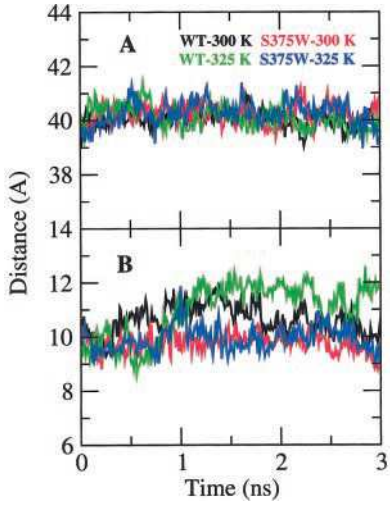


FIG. 2. **Minimum distances between the Cαs of residues 231 and 360 and between the Cα of residue 112 and heavy atoms of residue 375.** *Black* and *red* are for the wild type (*WT*) and the S375W mutant at 300 K, respectively, and *green* and *blue* are for the wild type and the S375W mutant at 325 K, respectively. The locations of these atoms are shown in Fig. 1*B*.

Of particular note are the residues that were most stabilized upon mutation (Fig. 5, *B* and *C* in *blue*) that were also within 4.5 Å of CD4 in the complex (Fig. 5*D*). Residues within 4.5 Å of CD4 include 124 (P), 126 (C), 196 (C), 279–281 (DNA), 283 (T), 365–368 (SGGD), 370–371 (EI), 425–430 (NMWQKV), 455–460 (TRDGGN), 469 (R), 472–475 (GGDM), and 477 (D) (Fig. 5*D*). Most of these residues are from the outer domain and the bridging sheet except for Met-475 and Asp-477. Only three residues from β2,3 of the bridging sheet are within 4.5 Å of CD4, two of which are Cys-126 and -196, which form the disulfide bond (1) and the other is Pro-124, which may also be important in maintaining the structure. β20,21 of the bridging sheet contains the longest continuous sequence of residues (NMWQKV) that were in contact with CD4. The only residue (Val-430) that made hydrophobic contact with CD4 also resides in this part of the bridging sheet. The rest of the residues were mostly charged or polar. Interestingly, three segments are rich in Gly (Fig. 5*D*, *red*). For each of the Gly-rich segments, there is an Asp immediately next to the GG motif. Because the interaction between gp120 and CD4 is quite polar with mostly polar and charged residues within the 4.5 Å (6), these Asp residues are important for the interactions. Changes in Asp-368 and Glu-370, for example, will disrupt interactions with

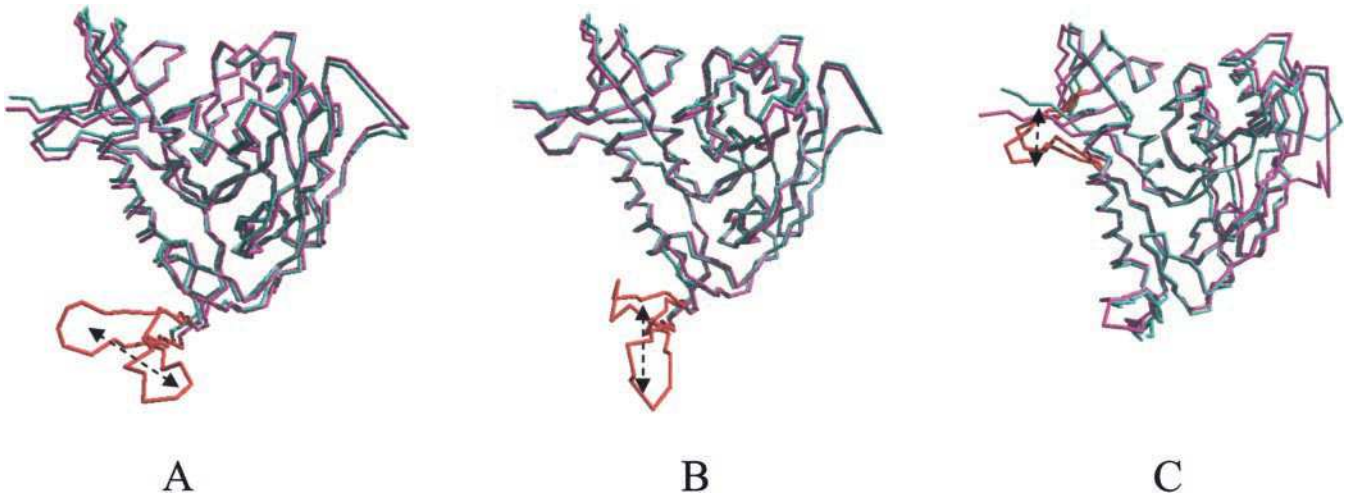


FIG. 3. **The three lowest motional modes generated from normal mode analysis.** The first two modes (*A* and *B*) correspond to the motions of the β strands 2 and 3 of the bridging sheet, and the third mode corresponds to β strands 4 and 5. The structural motifs involved in the modes were colored *red*, and the directions and the relative magnitudes of the motions are shown with *arrows*. Note that the third mode has much smaller magnitude of motions.

Downloaded from http://www.jbc.org/ by guest on July 3, 2017

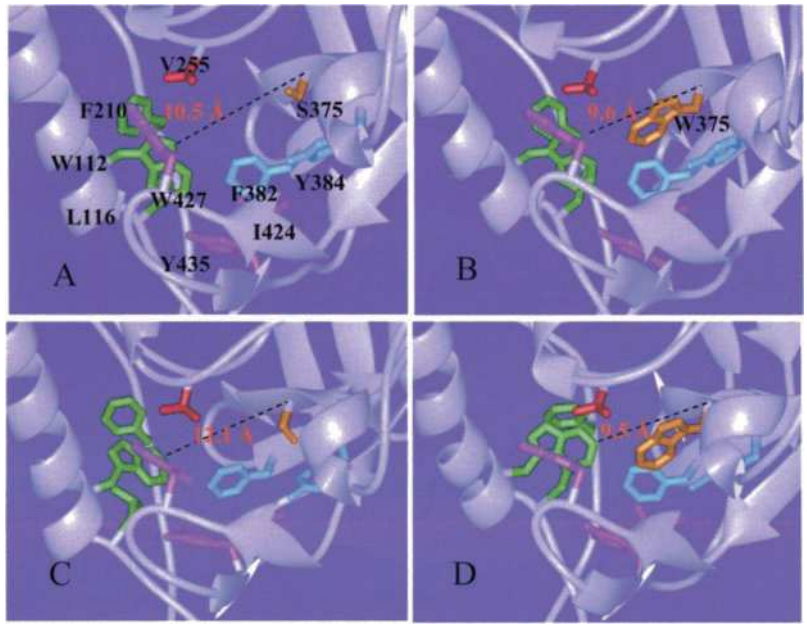


FIG. 4. **The average structures of the wild type (*A*) and mutant (*B*) from the last 2 ns of the trajectories at 300 K, and snapshots at 1.5 ns of the wild type (*C*) and mutant (*D*) from trajectories at 325 K.** For clarity, only a portion of the structure near the hydrophobic cluster was displayed. The hydrophobic residues near the S375W mutation site are shown to highlight the better hydrophobic interactions in the mutant for both cases of 300 K and 325 K simulations. Distances between residues 112 and 375 are shown.

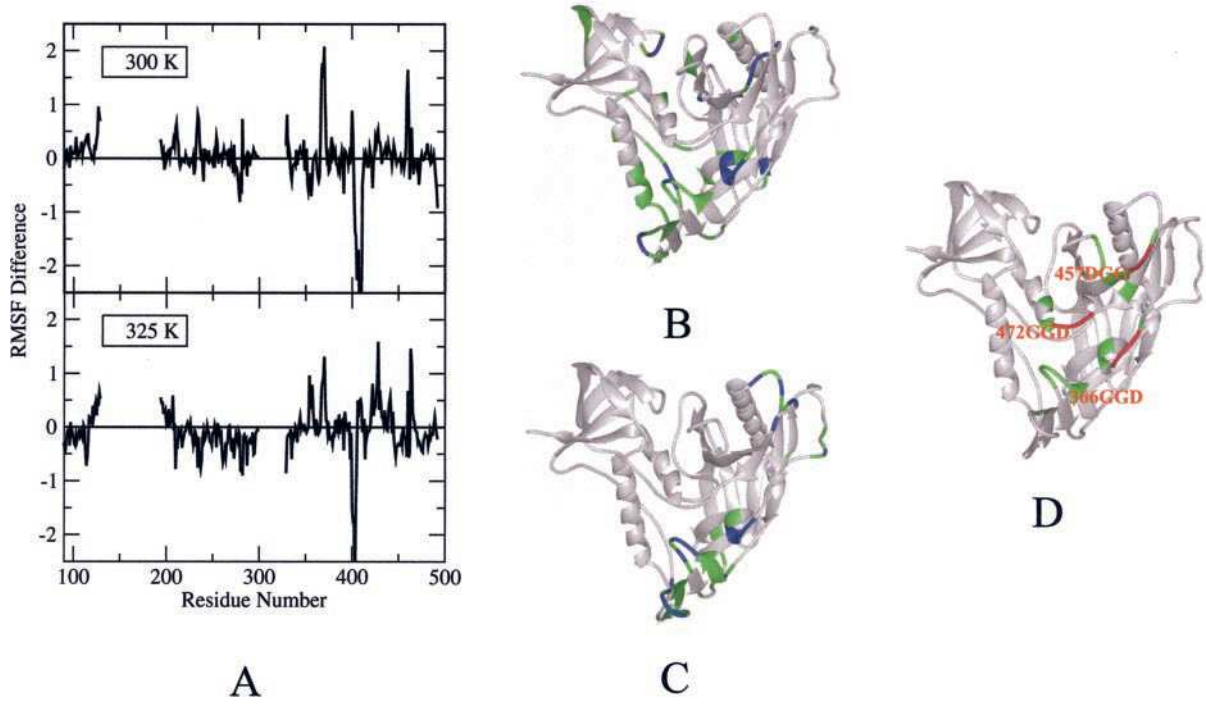


FIG. 5. *A*, differences of residue root mean square fluctuations between wild type and the S375W mutant from both 300 and 325 K simulations. A *positive number* indicates less fluctuation of the corresponding residue in the mutant than in the wild type. Stabilized residues are mapped onto the structure for the 300 K simulations (*B*) and 325 K simulations (*C*). A difference of root mean square fluctuations of more than 0.5 Å and between 0.2 and 0.5 Å are color-coded *blue* and *green*, respectively. Residues within 4.5 Å of CD4 based on the co-crystal structure are highlighted (*D*) and glycine-rich segments are colored *red* for comparison with the locations of the stabilized residues.

many CD4BS antibodies (41). However, to interact favorably with CD4, these residues need to be able to adjust their conformations. The Glys next to them make it possible. The stabilization of these segments upon mutation, however, will hinder the ability of these motifs to adjust their orientations with respect to CD4, resulting in the reduced enthalpy change as well as the entropy cost.

Change of Solvent-accessible Surface Areas upon Mutation—For the antibodies to bind, the protein needs to assume certain conformations that allow the access to the specific epitopes. The mutagenesis experiment shows that the S375W mutant binds CD4BS antibodies much weaker than the wild type and binds CD4 6-fold better (7). Thus, it is interesting to see whether the effect of the mutation is reflected in the solvent-accessible surface areas of the epitope residues. We calculated the solvent-accessible surface area difference between the wild type and the mutant for both the monomer and the gp120-CD4 complex (Fig. 6, *A–C*). For a direct comparison, the epitope residues for the CD4BS antibodies were mapped onto the structure (21) (Fig. 6*D*). Interestingly, we observed that most epitope residues for the CD4BS antibody binding became less accessible upon mutation in both the monomer and the complex simulations, with more dramatic changes in the complex simulations (Fig. 6). Mutagenic analysis indicates that many residues important for the binding of the CD4BS antibodies were not exposed on the surface of the

Downloaded from http://www.jbc.org/ by guest on July 3, 2017

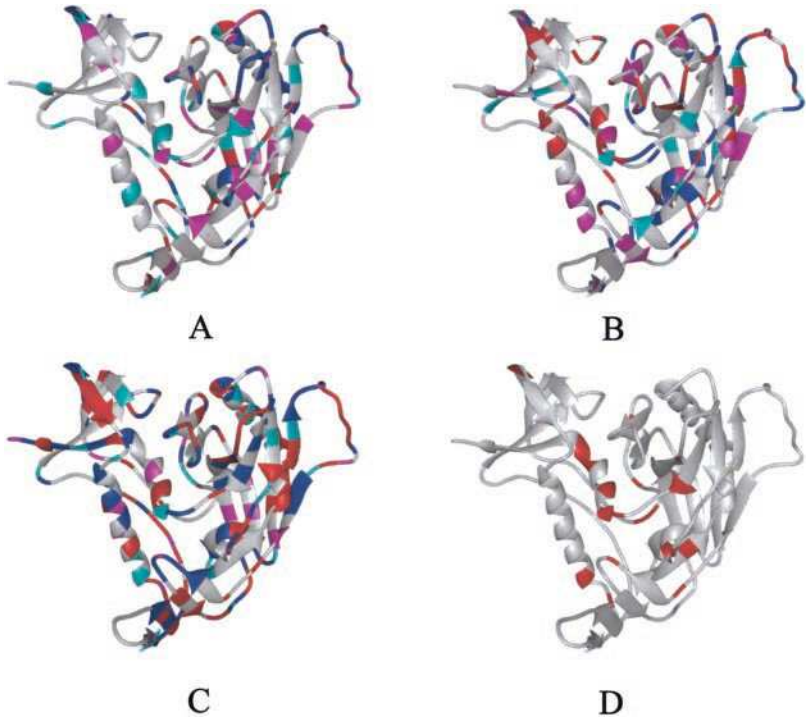


FIG. 6. Solvent-accessible surface area differences between the wild type and the mutant calculated from monomer simulations at 300 K (*A*) and 325 K (*B*), and complex simulations at 300 K (*C*) are shown. Residues becoming more buried upon mutation are colored *blue* (differ by at least 10 Å^2) or *cyan* (differ by 5–10 Å^2). Those becoming more exposed upon mutation are shown in *red* (differ by at least 10 Å^2) or *magenta* (differ by 5–10 Å^2). *D*, CD4BS antibody epitopes are shown (*magenta*) for comparison.

CD4-bound gp120 (21). This result shows that part of the mutational effect on antibody binding is the change of the accessibility of the epitope residues. Furthermore, the more dramatic effect on the surface accessibility of these epitope residues when in the complex form, indicates that CD4 binding is also partially responsible for the conformational change that reduced the accessibility of the CD4BS antibody epitopes. Because the mutation did not affect the binding affinity of CD4, this again confirms that CD4 might have interacted with a different set of residues or with different configurations of the residues.

The Source of Enthalpy and Entropy Differences—The interesting part of the binding free energy of gp120 and CD4 is the significant difference for the wild type and the mutant (Table I). The entropy cost for the mutant is only about half that of the wild type and the enthalpy gain is ~17 kcal/mol less favorable for the mutant. In addition to the large surface area burial and the conformational changes such as the folding of the bridging sheet as suggested earlier (6, 33), the reduction of the flexibility of residues near the mutation site was also responsible for the entropy changes. To compare the enthalpy changes, we calculated the interaction energy between gp120 and CD4, which is roughly proportional to the enthalpy change upon CD4 binding. Surprisingly, the calculated interaction energy was much smaller for the mutant than for the wild type after 1 ns of equilibration (Fig. 7*A*). The averages for the last 2 ns from the trajectories are −217 ± 16 kcal/mol and −175 ± 11 kcal/mol for the wild type and the mutant, respectively, which is ~40 kcal/mol more favorable for the wild type. Thus, our calculated free energy profile was qualitatively in very good agreement with the experimental result.

The calculated difference in interaction energies between the wild type and the mutant is very significant, because the only difference between the wild type and the mutant is the S375W mutation. Further, the mutated residue was not in contact distance with CD4 in either the wild type or the mutant and therefore did not directly contribute to the interaction energy. Therefore, its impact must be on the conformations. A simple measure of the distances between the Cα of residue Phe-43 of CD4, a residue that pointed to the hydrophobic pocket of the

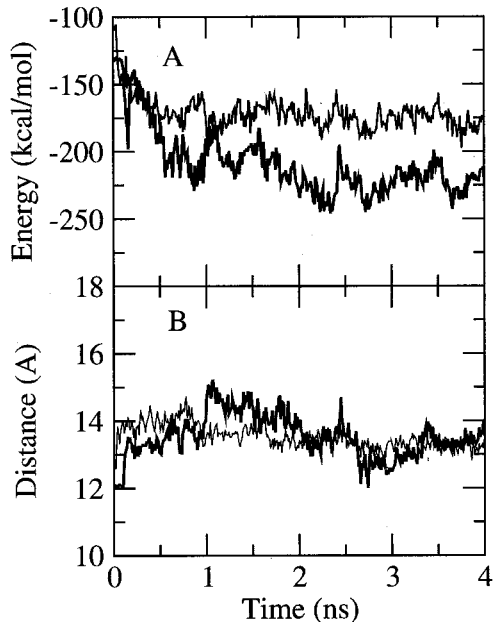


FIG. 7. *A*, interaction energies between the gp120 core domain and CD4 derived from complex simulations for wild type (*thick line*) and mutant (*thin line*). *B*, distance between Phe-43 of CD4 and S375W of gp120 for wild type (*thick line*) and mutant (*thin line*).

CD4 binding site, and the Cα of residue 375 of gp120 shows that this distance is quite stable in the mutant, but it fluctuated much more in the wild type (Fig. 7*B*). The interpretation of this result is that the low flexibility of the mutant at the interface of the inner domain, outer domain, and the bridging sheet, because of the stabilization effect, imposed some rigidity to the molecule. Because the CD4 molecule is known to be quite rigid, the low flexibility of the gp120 conformation will limit the freedom of the binding partners to adjust their conformations to yield optimal interactions, resulting in the less favorable interaction energy. For the same reason, the entropy cost is largely reduced.

Downloaded from http://www.jbc.org/ by guest on July 3, 2017

DISCUSSION

Unbound States of gp120 and Its Flexibility—Our simulation results have shown that the wild type and the mutant preferred different conformations. In particular, the mutant assumed a conformation that is more similar to the CD4-bound structure, whereas the wild type is relatively more open with respect to the inner and outer domains, with part of the bridging sheet (β2,3) partially unfolded. Such an observation is consistent with the previous speculations about the conformations of gp120 in its free form (7).

However, the putative open-state conformation observed in our simulations appears to be not as different from the bound state as was suggested previously (6). For example, it was proposed that the whole bridging sheet might be in an unfolded state in its free form, and the inner and outer domains may have larger flexibility with respect to each other (7). Our results show that the β2,3 part of the bridging sheet indeed experienced large conformational fluctuations and was partially unfolded. The β20,21 part of the bridging sheet, although also showing some mobility with respect to the inner/outer domains, was well structured and associated with the inner/outer domains throughout the trajectories. The inner and outer domains displayed only limited flexibility with respect to each other. The hydrophobic/aromatic cluster composed of residues Trp-112, Leu-116, and Phe-210 from the inner domain, Phe-382 and Tyr-384 from the outer domain, Ile-424, Trp-427, Tyr-435 from the bridging sheet, and Val-255 from the linker (Fig. 4), appeared to play an important role in the conformation stabilization. The hydroxyl group of Tyr-435 was also bonded to Lys-116 from helix α1, further enhancing the stability of the conformation. Breaking such a hydrophobic network needs sufficient energy to overcome the potentially high free energy barrier. In addition, although there is no other hydrophobic core along the inner/outer domain interface, there are extensive interactions between the two domains and they are connected by two fairly extended linkers (Fig. 1). Given these facts, one would not expect a much larger flexibility between the inner and outer domains than what is observed in the simulation. On the other hand, the β2,3 of the bridging sheet did not share a hydrophobic core with the rest of the molecule and was accordingly much more mobile. Thus, it is not unreasonable to believe that the unbound state still has a well structured inner domain, outer domain and part of the bridging sheet (β20,21) that are associated with each other, although we do not exclude the possibility that a more dramatic difference between the unbound and bound states may exist, given our limited simulation time. We further note that the flexibility of the core domain observed in this work may be changed in the context of the whole protein.

Conformational Features Versus Antibody Recognition—Mutagenesis studies show that structural perturbation of the bridging sheet affected the binding of gp120 to the CD4, CCR5, and the CD4i antibodies but not to the CD4BS antibodies (7). This indicates that the CD4BS antibodies recognize conformations of gp120 that are different from those recognized by the CD4 and CD4i antibodies. Consistently, most of the CD4BS epitopes have been mapped onto the inner and outer domain (21). The fact that the S375W mutant did not bind to any of the CD4BS antibodies indicates a conformational alteration in the inner and outer domain region. Our results suggest that these antibodies may prefer a conformation that is more open than that of the crystal structure particularly at the interface of the inner domain, outer domain, and the bridging sheet, because the mutant had a more contracted inner domain to outer domain distance. It is known that the antibodies are usually larger in size than CD4. Thus, the narrowed binding pocket induced by the S375W mutation explains perfectly the low binding affinity of this mutant to the CD4BS antibodies which were elicited against the potentially more open free form of gp120.

Implications for Drug Design—Anti-AIDS drugs are still in great need. However, there has been only limited success in the search for potent anti-AIDS agents even with the available structure of the gp120 in a complexed form. One of the reasons is that the prevailing conformation of gp120 in its free form may be very different from the CD4-bound state. Thus, targeting the bound-state structure that is not well populated is less likely to succeed. In addition, chemical agents targeting the CD4-bound form may act in a way similar to CD4 and thus possibly even trigger conformational changes needed for the chemokine receptor binding and consequently enhance the infection. It is known that several CD4 binding inhibitors are actually infection enhancers, and CD4 binding inhibitors are often known to be CCR5-binding activators (42). Thus, targeting the open form would be more efficient and possibly can avoid the CD4-binding inhibitor-infection-enhancing dilemma. With no available experimental structure of the gp120 in free form, a molecular-dynamics-generated open form structure provides another venue for the search for potent gp120 inhibitors. Encouragingly, our results suggest that the free form of gp120, although flexible as is true for many other proteins, assumes a well defined three-dimensional conformation that is targetable relative to the previously believed very mobile structure. In addition, multiple conformations of the backbone and side chains can be easily explored in this regard. Thus, with the use of our simulated models and the consideration of structural flexibility and multiple conformations, more potent and specific inhibitors can be expected from such structure-based efforts.

Conclusions—We have characterized the conformational features of the gp120 core domain in its free form via molecular dynamics simulations. Although the overall free-state structure is very similar to the bound-state, two specific regions have been identified to be conformationally different. First, when compared with the crystal CD4-bound form, β strands 2 and 3 of the bridging sheet in the free form were very flexible and partially unfolded. Second, the association among the inner domain, the outer domain, and β strands 20 and 21 of the bridging sheet was not as tight as in the bound form, resulting in a more expanded organization. Furthermore, analyzing the difference in stability and flexibility between the wild type and the mutant, we were able to show that the stabilization of the residues and the overall conformation in the S375W mutant account for an important part of the experimentally observed thermodynamic properties that differentiate between the wild type and the mutant. Overall, our results are in very good agreement with experiment (6, 7, 25), and explain the surprising experimental behavior of the S375W mutant. This mutant was designed with the goal of filling the space at the bottom of the CD4 binding pocket, which is not occupied by the CD4. Unexpectedly, the mutant bound CD4 tighter than the wild type did, and its affinity to antibodies elicited against the CD4BS was almost totally abolished (7). The conformational behavior shown here points toward the underlying factors that contribute to the altered phenotypes and highlights the differences between the CD4-bound and free states.

Hence, our results may further suggest the reason that drugs designed toward the gp120-bound state have had limited success. The conformational features of the free-state gp120 model obtained in our simulations should therefore be of great value in efforts toward new inhibitor development.

Acknowledgments—We thank Drs. D. Zanuy, K. Gunasekaran, Chung-Jung Tsai, and Hui-Hsu (Gavin) Tsai for many useful comments and suggestions.

Downloaded from http://www.jbc.org/ by guest on July 3, 2017

REFERENCES

1. Leonard, C. K., Spellman, M. W., Riddle, L., Harris, R. J., Thomas, J. N., and Gregory, T. J. (1990) *J. Biol. Chem.* **265,** 10373–10382
2. Wei, X., Decker, J. M., Wang, S., Hui, H., Kappes, J. C., Wu, X., Salazar-Gonzalez, J. F., Salazar, M. G., Kilby, J. M., Saag, M. S., Komarova, N. L., Nowak, M. A., Hahn, B. H., Kwong, P. D., and Shaw, G. M. (2003) *Nature* **422,** 307–312
3. Arendrup, M., Nielsen, C., Hansen, J. E., Pedersen, C., Mathiesen, L., and Nielsen, J. O. (1992) *J. Acquired Immune Defic. Syndr.* **5,** 303–307
4. Weiss, C. D., Levy, J. A., and White, J. M. (1990) *J. Virol.* **64,** 5674–5677
5. Kwong, P. D., Wyatt, R., Sattentau, Q. J., Sodroski, J., and Hendrickson, W. A. (2000) *J. Virol.* **74,** 1961–1972
6. Kwong, P. D., Wyatt, R., Robinson, J., Sweet, R. W., Sodroski, J., and Hendrickson, W. A. (1998) *Nature* **393,** 648–659
7. Xiang, S. H., Kwong, P. D., Gupta, R., Rizzuto, C. D., Casper, D. J., Wyatt, R., Wang, L., Hendrickson, W. A., Doyle, M. L., and Sodroski, J. (2002) *J. Virol.* **76,** 9888–9899
8. Starcich, B. R., Hahn, B. H., Shaw, G. M., McNeely, P. D., Modrow, S., Wolf, H., Parks, E. S., Parks, W. P., Josephs, S. F., Gallo, R. C., and Wong-Staal, F. (1986) *Cell* **45,** 637–648R. C.
9. Modrow, S., Hahn, B. H., Shaw, G. M., Gallo, R. C., Wong-Staal, F., and Wolf, H. (1987) *J. Virol.* **61,** 570–578
10. Dalgleish, A. G., Beverley, P. C., Clapham, P. R., Crawford, D. H., Greaves, M. F., and Weiss, R. A. (1984) *Nature* **312,** 763–767
11. Klatzmann, D., Champagne, E., Chamaret, S., Gruest, J., Guetard, D., Hercend, T., Gluckman, J. C., and Montagnier, L. (1984) *Nature* **312,** 767–768
12. Morikita, T., Maeda, Y., Fujii, S., Matsushita, S., Obaru, K., and Takatsuki, K. (1997) *AIDS Res. Hum. Retroviruses* **13,** 1291–1299
13. Rizzuto, C. D., Wyatt, R., Hernandez-Ramos, N., Sun, Y., Kwong, P. D., Hendrickson, W. A., and Sodroski, J. (1998) *Science* **280,** 1949–1953
14. Wu, L., Gerard, N. P., Wyatt, R., Choe, H., Parolin, C., Ruffing, N., Borsetti, A., Cardoso, A. A., Desjardin, E., Newman, W., Gerard, C., and Sodroski, J. (1996) *Nature* **384,** 179–183
15. Trkola, A., Dragic, T., Arthos, J., Binley, J. M., Olson, W. C., Allaway, G. P., Cheng-Mayer, C., Robinson, J., Maddon, P. J., and Moore, J. P. (1996) *Nature* **384,** 184–187
16. Alkhatib, G., Combadiere, C., Broder, C. C., Feng, Y., Kennedy, P. E., Murphy, P. M., and Berger, E. A. (1996) *Science* **272,** 1955–1958
17. Choe, H., Farzan, M., Sun, Y., Sullivan, N., Rollins, B., Ponath, P. D., Wu, L., Mackay, C. R., LaRosa, G., Newman, W., Gerard, N., Gerard, C., and Sodroski, J. (1996) *Cell* **85,** 1135–1148
18. Deng, H., Liu, R., Ellmeier, W., Choe, S., Unutmaz, D., Burkhart, M., Di Marzio, P., Marmon, S., Sutton, R. E., Hill, C. M., Davis, C. B., Peiper, S. C., Schall, T. J., Littman, D. R., and Landau, N. R. (1996) *Nature* **381,** 661–666
19. Doranz, B. J., Rucker, J., Yi, Y., Smyth, R. J., Samson, M., Peiper, S. C., Parmentier, M., Collman, R. G., and Doms, R. W. (1996) *Cell* **85,** 1149–1158
20. Dragic, T., Litwin, V., Allaway, G. P., Martin, S. R., Huang, Y., Nagashima, K. A., Cayanan, C., Maddon, P. J., Koup, R. A., Moore, J. P., and Paxton, W. A. (1996) *Nature* **381,** 667–673
21. Wyatt, R., Kwong, P. D., Desjardins, E., Sweet, R. W., Robinson, J., Hendrickson, W. A., and Sodroski, J. G. (1998) *Nature* **393,** 705–711
22. McKnight, A., and Weiss, R. A. (2003) *Proc. Natl. Acad. Sci. U. S. A.* **100,** 10581–10582
23. Lin, P. F., Blair, W., Wang, T., Spicer, T., Guo, Q., Zhou, N., Gong, Y. F., Wang, H. G., 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) *Proc. Natl. Acad. Sci. U. S. A.* **100,** 11013–11018
24. Boussard, C., Doyle, V. E., Mahmood, N., Klimkait, T., Pritchard, M., and Gilbert, I. H. (2002) *Eur. J. Med. Chem.* **37,** 883–890
25. 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) *Proc. Natl. Acad. Sci. U. S. A.* **97,** 9026–9031
26. Arthos, J., Deen, K. C., Chaikin, M. A., Fornwald, J. A., Sathe, G., Sattentau, Q. J., Clapham, P. R., Weiss, R. A., McDougal, J. S., Pietropaolo, C., Axel, R., Truneh, A., Maddon, P. J., and Sweet, R. W. (1989) *Cell* **57,** 469–481
27. Brodsky, M. H., Warton, M., Myers, R. M., and Littman, D. R. (1990) *J. Immunol.* **144,** 3078–3086
28. Cordonnier, A., Montagnier, L., and Emerman, M. (1989) *Nature* **340,** 571–574
29. Lasky, L. A., Nakamura, G., Smith, D. H., Fennie, C., Shimasaki, C., Patzer, E., Berman, P., Gregory, T., and Capon, D. J. (1987) *Cell* **50,** 975–985
30. Moebius, U., Clayton, L. K., Abraham, S., Harrison, S. C., and Reinherz, E. L. (1992) *J. Exp. Med.* **176,** 507–517
31. Olshevsky, U., Helseth, E., Furman, C., Li, J., Haseltine, W., and Sodroski, J. (1990) *J. Virol.* **64,** 5701–5707
32. Peterson, A., and Seed, B. (1988) *Cell* **54,** 65–72
33. Kwong, P. D., Wyatt, R., Majeed, S., Robinson, J., Sweet, R. W., Sodroski, J., and Hendrickson, W. A. (2000) *Structure Fold. Des.* **8,** 1329–1339
34. Ryu, S. E., Kwong, P. D., Truneh, A., Porter, T. G., Arthos, J., Rosenberg, M., Dai, X. P., Xuong, N. H., Axel, R., Sweet, R. W., and Hendrickson, W. A. (1990) *Nature* **348,** 419–426
35. Wang, J. H., Yan, Y. W., Garrett, T. P., Liu, J. H., Rodgers, D. W., Garlick, R. L., Tarr, G. E., Husain, Y., Reinherz, E. L., and Harrison, S. C. (1990) *Nature* **348,** 411–418
36. Wu, H., Kwong, P. D., and Hendrickson, W. A. (1997) *Nature* **387,** 527–530
37. Brooks, B. R., Bruccoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S., and Karplus, M. (1983) *J. Comput. Chem.* **4,** 187–217
38. MacKerell, A. D. Jr., Bashford, D., Bellott, M., Dunbrack, R. L. Jr., 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., III, Roux, B., Schlenkrich, M., Smith, J. C., Stote, R., Straub, J., Watanabe, M., Wiorkiewicz-Kuczera, J., Yin, D., and Karplus, M. (1998) *J. Phys. Chem. B.* **102,** 3586–3616
39. Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W., and Klein, M. L. (1983) *J. Chem. Phys.* **79,** 926–935
40. Keskin, O., Durell, S. R., Bahar, I., Jernigan, R. L., and Covell, D. G. (2002) *Biophys. J.* **83,** 663–680
41. Thali, M., Furman, C., Ho, D. D., Robinson, J., Tilley, S., Pinter, A., and Sodroski, J. (1992) *J. Virol.* **66,** 5635–5641
42. 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) *Biochemistry* **43,** 1928–1938

Downloaded from http://www.jbc.org/ by guest on July 3, 2017

Characterization of the Conformational State and Flexibility of HIV-1 Glycoprotein gp120 Core Domain

Yongping Pan, Buyong Ma, Ozlem Keskin and Ruth Nussinov

J. Biol. Chem. 2004, 279:30523-30530.
doi: 10.1074/jbc.M404364200 originally published online May 6, 2004

Access the most updated version of this article at doi: 10.1074/jbc.M404364200

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

This article cites 40 references, 14 of which can be accessed free at
http://www.jbc.org/content/279/29/30523.full.html#ref-list-1

Downloaded from http://www.jbc.org/ by guest on July 3, 2017