Chinese Science Bulletin

Molecular motions and conformational transition between different conformational states of HIV-1 gp120 envelope glycoprotein

LIU ShuQun[1], FU YunXin[1,2] & LIU CiQuan[3†]

[1] Laboratory for Conservation and Utilization of Bio-resources, Yunnan University, Kunming 650091, China;
[2] Human Genetics Center, the University of Texas Health Science Center, Houston, Texas 77030, USA;
[3] Modern Biology Center, Yunnan University, Kunming 650091, China

The HIV-1 gp120 exterior envelope glycoprotein undergoes a series of conformational rearrangements while sequentially interacting with the receptor CD4 and coreceptor CCR5 or CXCR4 on the surface of host cells to initiate virus entry. Both the crystal structures of the HIV-1 gp120 core bound by the CD4 and antigen 17b and the SIV gp120 core pre-bound by CD4 are known. Despite the wealth of knowledge on these static snapshots of molecular conformations, the details of molecular motions involved in conformational transition that are crucial to intervention remain elusive. We presented comprehensive comparative analyses of the dynamics behaviors of the gp120 in its CD4-complexed, CD4-free and CD4-unliganded states based on the homology models with modeled V3 and V4 loops by means of CONCOORD computer simulation to generate ensembles of feasible protein structures that were subsequently analysed by essential dynamics analyses to identify preferred concerted motions. The revealed collective fluctuations are dominated by complex modes of combinational motions of the rotation/twisting, flexing/closure, and shortness/elongation between or within the inner, outer, and bridging-sheet domains, and these modes are related to the CD4 association and HIV neutralization avoidance. Further essential subspace overlap analyses were performed to quantitatively distinguish the preference for conformational transitions between the three states, revealing that the unliganded gp120 has a greater potential to translate its conformation into the conformational state adopted by the CD4-complexed gp120 than by the CD4-free gp120, whereas the CD4-free gp120 has a greater potential to translate its conformation into the unliganded state than the CD4-complexed gp120 does. These dynamics data of gp120 in its different conformations are helpful in understanding the relationship between the molecular motion/conformational transition and the function of gp120, and in gp120-structure-based subunit vaccine design.

HIV-1 gp120, comparative modeling, conformational transition, essential dynamics, essential subspace overlap

Over 40 million people are currently infected with the human immunodeficiency virus type 1 (HIV-1), the major cause of the acquired immune deficiency syndrome (AIDS)[1,2], which has brought about catastrophic consequences to human society. The development of a preventive vaccine, which optimally should elicit both virus-neutralizing antibodies and cellular immune response, is of high priority and urgency[3,4].

Entry of HIV-1 into target cells requires transformation of the protective envelope into a fusion-competent state. Infection is initiated by the selective interaction between the viral exterior envelope glycoprotein, gp120,

Received September 20, 2006; accepted September 21, 2007
doi: 10.1007/s11434-007-0478-4
†Corresponding author (email: shuqunliu@ynu.edu.cn)
Supported by the Yunnan University (Grant No. 2004Q013B), Yunnan Province (Grant No. 2006C008M), and partially supported by open fund from the Laboratory for Conservation and Utilization of Bio-resources, Yunnan University, and Innovation Group Project from Yunnan University

and receptors on the surface of the target cell, CD4, and obligatory chemokine receptors (CCR5 or CXCR4)[5−8]. The envelope glycoproteins precursor gp160 polypeptide chain is cleaved during transport process into the gp120 and gp41, which are organized into trimetric complexes on the viron surface[9,10]. Accumulating biochemical and structural evidences indicate that the initial binding of the CD4 to gp120 triggers conformational alterations in the HIV envelope that subsequently promote recognition of the coreceptors and ultimately lead to conformational changes of gp41 and membrane fusion[8,10]. This complex process involves a series of structural rearrangements in which conformational dynamics of the HIV envelope glycoproteins plays a crucial role.

Recently, the X-ray crystallographic structures of the core gp120 in complex with two amino-terminal domain (D1D2) of the CD4 and the antigen binding fragment (Fab) of the human neutralizing antibody, 17b, have been determined in a series of X-ray crystallographic studies[11,12]. These structures demonstrate that the gp120 core is composed of the inner and outer domains and a bridging sheet. Components of both the domains and the bridging sheet contribute to the CD4 binding, and several structural elements of this structure appear to depend on association with CD4 for their stabilization. Although the crystal structures of the CD4-bound HIV-1 gp120 core is incomplete due to deletions of residues from both the N- and C-termini; Gly-Ala-Gly tripeptide substitutions for 67 V1/V2 loop residues and 32 V3 loop residues; and no electronic density map for the V4 loop, they nevertheless provide tremendous structural insights for analyzing envelope antigenicity and infection mechanism[11−13]. Another advance is the more recent determination of the crystallographic structure of the SIVmac 32H gp120 in a CD4-unliganded conformation that reveals its prefusion state, before interaction with the CD4[14]. This unliganded SIV gp120 core is also composed of the inner and outer domains, and contains neither the variable loops V1/V2 nor the V3, but the V4 loop was resolved, which adopts an open conformation that extends away from the body of the outer domain. The relatively high overall sequence conservation between SIV and HIV and their shared receptor and coreceptors indicate that the crystal structure of the SIV gp120 does present the conformation in the CD4-unliganded state rather than one of the continuum of flexible conformation selected by the crystallization condition[14]. A rough comparison of the structures of the unliganded SIV gp120 and the CD4-bound HIV gp120 highlights the similarities and differences between the unliganded and CD4-bound states. The conformational changes involve the reconfiguration and translocation of structural components in the inner domain and the bridging sheet, while the outer domain is relatively unchanged. The unexpectedly extensive conformational rearrangement upon receptor binding is in agreement with the evidence from the thermodynamics studies suggesting that the binding of CD4 or antibodies to the gp120 induces substantial structural rearrangements primarily in the core structure[15,16].

Although the crystallographic data provide important snapshots of the core gp120 structures in different functional states, a more complete understanding of the function involved in gp120-CD4 interaction requires the detailed comparative analyses of the molecular motions of gp120 and the preference for conformational transition between different conformational states of gp120 because these analyses are helpful in a full understanding of the protein function.

The conformational changes involve range from very subtle, local changes to global conformational changes involving motions of significant amplitude for large parts of a protein. Dynamics plays an important role not only in the functional native state of many proteins, but also in the mechanism by which a protein reaches its functional conformation, e.g. the protein folding process is a highly dynamic process. In the case of the HIV-1 gp120, the dynamics is also crucial in the process of conformational transition from the CD4-unliganded state to CD4-bound state as unusually large structural difference between them. Although there is currently no experimental technique that allows monitoring protein conformational changes at atomic resolution, it has been found that further insights into the dynamic structure/function relationship can be gained through an analysis of the computational simulated dynamics[17]. For example, previous studies on the dynamic properties of gp120 in its CD4-free and CD4-complexed forms using 10 ns molecular dynamics (MD) simulations[18] have demonstrated that CD4 binding reduced conformational flexibility in certain loop regions of the outer domain. Other MD studies indicated that the CD4 binding partially locked the bridging sheet in gp120 but left the

β2-β3 strands flexible[19,20]. A shortcoming of the MD approach, however, is that currently available computing power permits only short simulations to be run for biological macromolecules, typically of the order of tens of nanoseconds. This time scale may insufficiently explore the configurational space available to the protein, and furthermore, is a few orders of magnitude smaller than that on which most biological processes take place. An alternative approach to MD is to generate an ensemble of structures randomly without using Newton's equations of motion. CONCOORD[21] takes this approach as a fast way to generate ensembles that explore configurational space more fully. Previous studies[21−26] have shown that ensembles derived from the CONCOORD are most useful in identifying global motions that a protein is able to perform around an equilibrium position, and can be extrapolated to predict the character of conformational transition.

We have applied comparative modeling technique to construct the three-dimensional structural models of gp120 in its different conformational states with modeled V3 and V4 loops, i.e. the gp120-CD4 complex, gp120 alone but in the CD4-bound state (CD4-free gp120), and gp120 in the CD4-unliganded state (unliganded gp120). These homology models were used as starting structures for CONCOORD to generate respective ensembles. The aims of this study were, (1) to characterize the molecular motions of gp120 in the three conformational states at global level to determine the hierarchy of tertiary structures that exist in gp120; (2) to identify the preference for conformational transition between the three conformational states of gp120. The results in this article are able to extend our understanding of the interrelationships between gp120 structures/dynamics and function, and will significantly assist with HIV inhibitor design.

1 Materials and methods

1.1 Generation of starting structures

The starting sequence of the HIV-1 HXBc2 isolate gp160 precursor (Swiss-Prot accession number P04578) was obtained from the Swiss-Prot protein sequence database[27], the sequence corresponding to the transmembrane glycoprotein gp41, 52 and 19 residues from the N- and C-termini of the gp120, and the V1/V2 loop (Gly-Ala-Gly substitutes for 67 V1/V2 loop residues) were removed. The final gp120 primary sequence consists of residues from 83 to 492, which includes sequences for the V3 and V4 loops. Structural templates of the gp120 core and the V3 loop were obtained from the PDB protein structure database[28]. PDB entries 1G9M (chain G)[11] and 2BF1[14] were used as templates to model the CD4-bound and unliganded gp120 cores, respectively. PDB entry 1CE4[29] was used as template to model the V3 loop.

The sequence alignment between the V3 loop (1CE4), the gp120 core (1G9M, chain G) templates and the gp120 target sequence was performed using the MODELLER software package[30]. The alignment was shown in Figure 1. 20 gp120 models in the CD4-bound state were generated based on the template structures of the gp120 core and the V3 loop using an in-house written script "get-model" for the MODELLER. The V4 loop coordinates were generated and optimized by a loop modeling sub-routine within the "get-model" with a 3D_INTERPOLATION algorithm and a thorough optimization protocol (a thorough molecular dynamics/simulated annealing procedure and the MD_LEVEL was defined as refine_3), respectively. Finally, these 20 models were clustered and an optimized average model was generated by a MODELLER script "cluster". The final optimized average gp120 model is in the CD4-bound state without including the CD4 molecule, which was referred to as the CD4-free (or CD4-bound) gp120 in the rest of the text.

The gp120 model in complex with CD4 was obtained as following steps: (1) the CD4-free gp120 model was superimposed onto the crystal structure of the complex (1G9M) to obtain the appropriate orientation with respect to the CD4; (2) the D2 domain of CD4 was removed but that the CD4-free gp120 and CD4 D1 domain taken from the crystal complex were preserved, which we referred to as the gp120-CD4 complex in the context. The monomer gp120 within the gp120-CD4 complex was referred to as the CD4-complexed gp120 in the rest of the text.

The homology model of the unliganded gp120 with the V3 loop was obtained as described for the CD4-free model with the exception that the structure of the SIV gp120 core (2BF1) was used as the template. The final optimized average model in the CD4-unliganded conformational state was referred to as the unliganded gp120 in the rest of the text.

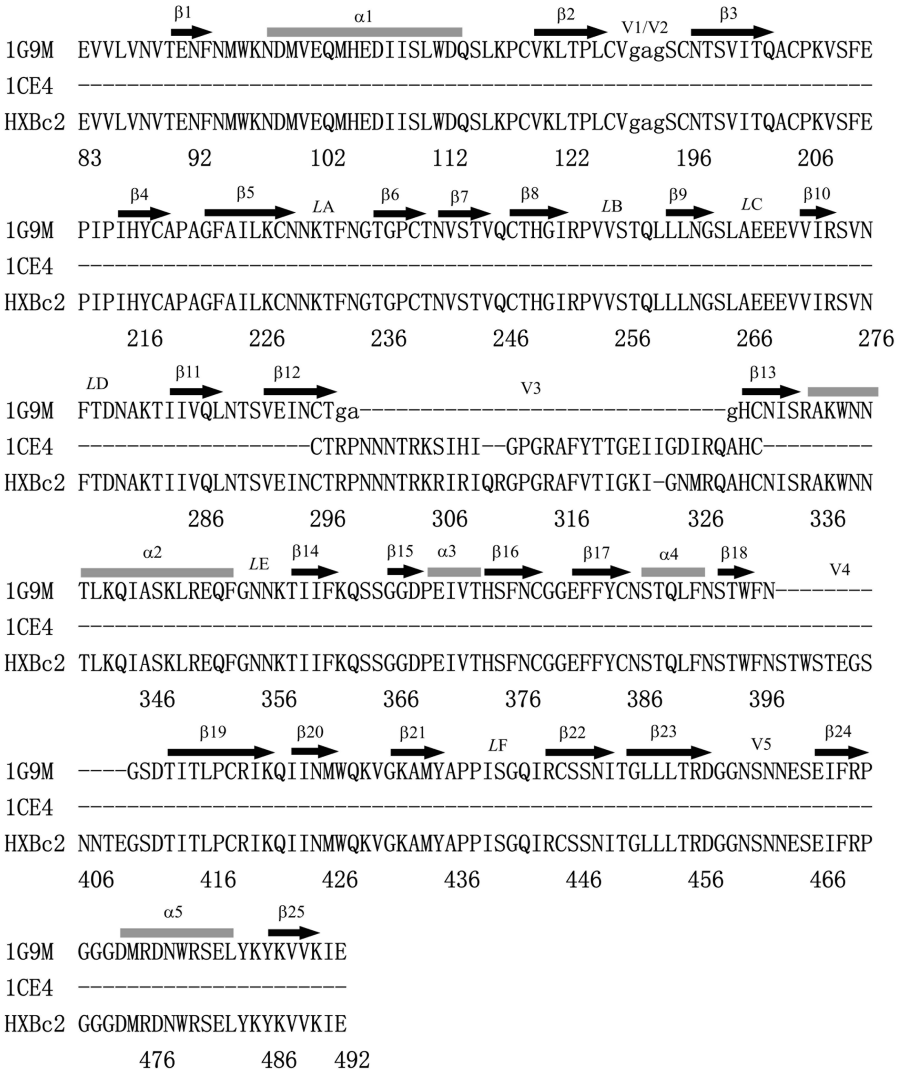

Figure 1 Multiple Sequence alignment between templates of the gp120 core (PDB entry 1G9M, Chain G) and V3 loop (PDB entry 1CE4), and the target sequence of the HXBc2 isolate (Swiss-Prot accession number P04578). The secondary structure elements are assigned according to the X-ray core structure with arrows for β-strands and rods for α-helices. The "gag" sequence in the V1/V2 and V3 loops of the gp120 core is the consequence of the truncation.

1.2 Generation of structural ensembles

The models of the gp120-CD4 complex, CD4-free and unliganded gp120s were used as starting structures for the CONCOORD program[21] to generate ensembles. CONCOORD works at two steps. The first step is to derive a table of upper and lower distance limits for the various interatomic distances (covalent, ionic, and hydrogen bond etc.), defining a region of configurational space that the protein can access. For example, the distance limit for the atomic pair connected by covalent bond is only minute (0.002 nm), whereas in the case of the atomic pairs involving in the salt bridge, the distance perturbation can be relatively larger (0.075 nm). The second step is to generate an ensemble of structures that fulfil the interatomic distance bounds for all pairs of atoms. The ensemble generated in this way can be thought of as a set of uncorrelated structural snapshots with the temporal order of frames scrambled. However, in many ways it can be treated and analyzed in exactly the same way as a trajectory of snapshots produced by an MD simulation[21−26].

CONCOORD was run 3 times, once for each of the models generated by MODELLER, and 500 conformations were finally generated for the gp120-CD4 complex, CD4-free gp120, and unliganded gp120, respectively.

1.3 Essential dynamics analysis

Essential dynamics (ED) method[31] is a powerful tool for filtering large-scale concerted motions from an en-

semble of structures and is related to principal component analysis (PCA) and quasiharmonic analyses[31−34]. It is based on the diagonalization of the covariance matrix built from the atomic fluctuations in a trajectory from which the overall translations and rotations have been removed. Resulting eigenvectors are directions in conformational space that represent collective motion and the corresponding eigenvalues define the mean square fluctuation (MSF) of the motion along these vectors. The required covariance matrix, eigenvectors, and corresponding eigenvalues for the ED analyses of the CONCOORD ensembles for the gp120-CD4 complex, CD4-free and unliganded gp120s were obtained using the GROMACS package[35,36]. Only the backbone atoms were included in the ED analyses in this article because it has been shown[31,37,38] that this simplification best detects the large-scale concerted motions in proteins.

The DYNAMITE program[23] was utilized to produce the "porcupine plot" that was able to give a graphical representation of the motion held in an eigenvector v_i. For instance, to visualize eigenvector 1, a cone is drawn for each residue starting from the C_α, projecting in the direction of component of the first eigenvector that corresponds to that residue. The length of the cone represents the motional amplitude. Given the eigenvectors of the ED decomposition, a script was generated using the DYNAMITE, allowing the molecular graphics program VMD[39] to automatically plot these cones on the protein.

1.4 Essential subspace overlap analyses

If eigenvectors obtained from the ED analysis are seen as the vectors that span a complex $3N\times3N$ (with N = the number of atoms of the protein) dimensional space, the few "essential" eigenvectors with the largest eigenvalues span a subspace, the essential subspace, in which all the large concerted motions occur. It is assumed that also the true conformational space of the most proteins contains a low-dimensional subspace in which the most positional fluctuations known to play a crucial role in the biological function of proteins[24,25,33,40]. The essential subspace derived from the simulation is an approximation of that subspace[21]. Furthermore, the essential subspace overlap between different simulated systems can provide a valid approach to assess their dynamic[21,31,32,34,41,42] or functional similarity[25]. To do so, a common approach is to select a subset of eigenvectors for each ensemble, and to calculate their cumulative mean square inner product (CMSIP) using the expression[43]:

$$\psi = \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n}(v_i^{\mathrm{A}}\cdot v_j^{\mathrm{B}})^2 , \quad (1)$$

where the v_i^{A} and v_j^{B} are the ith and jth eigenvectors of two different sets of ensembles, respectively, the $(v_i^{\mathrm{A}}\cdot v_j^{\mathrm{B}})^2$ is the square inner product. The subspace overlap ψ ranges from 0, when the eigenvector subsets are completely dissimilar, to 1 (or 100%) when they are identical. The number of eigenvectors used, n, is typically chosen so as to represent a significant proportion of fluctuations in the simulation.

We performed the comparative analyses of the essential subspace overlap between CONCOORD ensembles of the CD4-complexed, CD4-free, and unliganded gp120s to quantify the extent to which gp120 in different conformational states explore the same conformational space, and hence quantitatively distinguish the preference for conformational transition between the three states

2 Results

2.1 Descriptions of models of the gp120-CD4 complex, CD4-free and unliganded gp120s

A ribbon diagram of the CD4-free gp120 is shown in Figure 2(a). The structure closely resembles its template, the gp120 core from 1G9M[11], with the exception of the modeled V3 and V4 loops. The root mean square deviation (RMSD) between the shared backbone atoms of them is 0.36 Å. All the structural components constitute two major domains, the inner domain and the outer domain, with some excursions emanating from this body. The inner domain includes the N- and C-termini, a two-helix, a small two-stranded bundle, and a six-stranded β-sandwich at its termini-proximal end and a projection at the distal end from which the V1/V2 stem (β2-β3 hairpin) emanates. The outer domain is a stacked double barrel that lies alongside the inner domain. The V3 loop lies beneath the distal end of the outer domain and acts as a connection between the β12 and β13. The V4 loop extends away from the right side of the outer domain and adopts an external coiled conformation. The V1/V2 stem and β20-β21 constitute an antiparallel four-stranded bridging sheet, which stands below the distal ends of both the inner and outer domains.

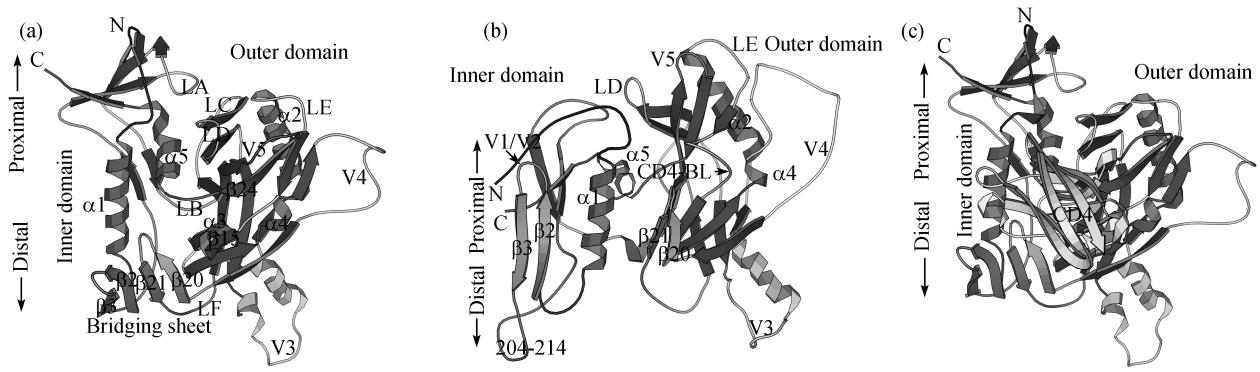

Figure 2 Structures of the gp120 homology models. (a) Ribbon diagram of the CD4-free gp120; (b) ribbon diagram of the unliganded gp120; (c) ribbon diagram of the gp120 in complex with CD4.

Figure 2(b) shows the model of the unliganded HIV-1 HXBc2 gp120, which resembles its template, the unliganded SIV gp120 core[14]. The common backbone atom RMSD between the model and the template is 0.82 Å. The unliganded gp120 model also has two major domains with inner domain comprising the N- and C-termini, a three-stranded β-sheet at its termini-proximal end, V1/V2 stem (β2-β3 hairpin), α-helix 1 (α1), and a short α-helix (α5) at the inner/outer domain junction. The unresolved connecting segment (residues 218—228) between the β3 from the V1/V2 stem and the β5 from the three stranded β-sheet in the crystal structure of the SIV gp120 corresponds to the residues 204—214 in the unliganded HIV-1 gp120. This segment was generated and optimized using the MODELLER loop modeling sub-routine. It presents a disordered loop conformation that extends away from the V1/V2 stem base of the unliganded gp120 inner domain (Figure 2(b)). The outer domain has, with some local exceptions, globally the same structural organization as that in the CD4-free gp120. The V3 and V4 loops form protruding excursions on the body of the outer domain as they do in the CD4-free gp120. The major differences in the outer domain between the CD4-free and unliganded gp120 are: (1) the lengths in loops, for instance, the V4, V5, *L*E are slightly longer in the unliganded form; (2) The connection between strands of the β14 and β16 adopts a extended (loop) conformation in the unliganded gp120. This segment can be regarded as the "CD4-binding-loop" (CD4-BL) because the GGDPE motif can move when CD4 associates and presents a extended strand (β15) and an α-helix (α3) in the CD4-bound conformation, with the β15 forming main-main hydrogen bonding with CD4's C″ strand. The bridging sheet presented in the CD4-free gp120 is absent in the unliganded gp120. Although the two β-hairpins, the β2-β3 and β20-β21 that can form bridging sheet is ordered, a space of 22—29 Å intervenes between them in the unliganded gp120.

The gp120-CD4 complex model is shown in Figure 2(c), the structural organization of the CD4-complexed gp120 is identical to that of CD4-free gp120 with the exception of the CD4 D1 domain docked onto it.

In the CD4-free gp120 model, the receptor CD4 binding sites are composed of the β20-β21 ribbon, V1/V2 stem, loops *L*D, *L*E, V5, β15-α3 excursion, and β24-α5 connection, which constitute an unusually large CD4 binding cavity (Figure 2(a)). The CD4 Phe 43 binding pocket is located at the center of the large cavity. Especially it is situated at the intersection between the inner domain, the outer domain, and the bridging sheet, deeply buried, extending into the hydrophobic interior of gp120. In the unliganded form, the α1, α5, β20-β21 ribbon, and the CD4-BL create a long, narrow channel at the intersection surfaces of the inner and outer domains. Apparently, the binding cavity for the CD4 and binding pocket for the CD4 Phe 43 are absent in the unliganded state and can not form unless induced by CD4 binding.

2.2 Essential dynamics analyses

ED analyses of the CONCOORD ensembles of the three gp120 models revealed that only a few eigenvectors possessed significant eigenvalues, which are shown in Figure 3 with sorted decreasing value. The traces of the covariance matrices after diagonalization, calculated over the ensembles of the CD4-complexed, CD4-free, and unliganded gp120 which represent the total mean square fluctuation (MSF), are 59.9, 78.2, and 79.6 nm^2, respectively. The first 4 eigenvectors contributing to the total MSF are 66.0%, 70.2%, and 70.8%, and the first 10

eigenvectors contributing to the total MSF are 84.5%, 87.8%, and 84.7% for the ensembles of the CD4-complexed, CD4-free, and unliganded gp120, respectively. These results indicate that the first 10 eigenvectors, especially the first 4 eigenvectors, emerge with appreciable freedom.

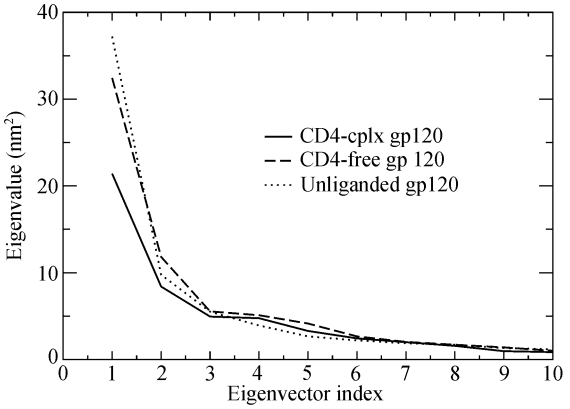

Figure 3 Eigenvalues obtained from essential dynamics analyses of CONCOORD ensembles for the CD4-complexed, CD4-free, and unliganded gp120s. Only the 10 largest eigenvalues are shown out of 3114. The CD4-complexed gp120 is abbreviated to CD4-cplx gp120.

2.3 Large-scale concerted motions

Figure 4 shows, in porcupine representation, the large-scale collective motions of the gp120-CD4 complex and the CD4-free and unliganded gp120s. For gp120 within the gp120-CD4 complex (Figure 4(a)), eigenvector 1 with the largest eigenvalue describes a rotation/twisting motion between the inner domain and the outer domains. Specially, the N-, C-termini and the termini-proximal β-bundle in the inner domain, and the V3 and V4 loops in the outer domain show the largest fluctuations as they have the longest "quill". The V1/V2 stem in the bridging sheet and the loops *L*C, V5, and *L*E in the outer domain show moderate fluctuations. The α1 and α5 in the inner domain and the structural components in the outer domain such as the β20-β21 ribbon and α3-β15 contacting with or close to the CD4 show the smallest fluctuations as they have the shortest "quill". The motion of the CD4 within the gp120-CD4 complex is similar to that of the gp120 inner domain, i.e. the CD4 and bridging sheet, together with the inner domain perform a twist around an axis running from the outer domain to the inner domain with respect to the outer domain, despite almost uniformly moderate fluctuations occurring across the entire CD4 molecule (Figure 4(a)). The second largest collective fluctuation of the gp120-CD4 complex described by the eigenvector 2 is illustrated in Figure 4(b). This mode can be described as a rotation/twisting motion of the structural region composed of the inner domain, bridging sheet, and V3 loop relative to the region composed of the entire CD4 and large parts of the outer domain with the exception of the V3 loop. The motion of the gp120-CD4 complex described by eigenvector 3 is illustrated in Figure 4(c). As shown in Figure 4(c), the gp120 outer domain and bridging sheet together with the CD4 are dominated by a vortex that rotates around an axis running from the center of the outer domain to CD4, another vortex formed by the N-, C-termini and V3 loop rotates about the axis in the opposite direction. The fourth ranked motion is dominated by elongations of the inner domain and outer domain along respective major axes, which is accompanied by thinnings of the two domains (data not shown).

For the CD4-free gp120 ensemble, the first eigenvector with the largest atomic displacement can be described as a twist of the inner domain relative to the outer domain (Figure 4(d)), i.e. the two domains rotate in opposite directions around an axis connecting the centers of the two domains. This mode resembles the twist motion described by eigenvector 1 for the CD4-complexed gp120, whereas the rotating directions of the inner domain and outer domain are reverse to the directions of the inner and outer domains of the CD4-complexed gp120. It should be pointed out that the opposite rotation direction does not demonstrate the substantial difference between the motional modes of these two structures, which is probably caused by the reversed expression sign for the eigenvector. Figure 4(d) emphasizes the increased amplitude of conformational freedom of the CD4-binding cavity (composed of bridging sheet, loops *L*D, *L*E, V5, β15-α3 and β24-α5) when CD4 is removed. The second most significant predicted motion of the CD4-free gp120 corresponds to a twisting mode of the two domains around an axis running roughly through the centers of them, with a large part of the outer domain (especially the proximal end and V4 loop) forming a vortex that rotates in opposite direction relative to a vortex formed by the structural components of the N-, C-termini, bridging sheet, and V3 loop (Figure 4(e)), which is similar to the second ranked motion of the CD4-complexed gp120. Eigenvector 3 describes a relatively larger vortex formed by the bridging sheet and the entire outer domain that rotates around an axis running through the Phe 43 pocket, behaving itself a flexing

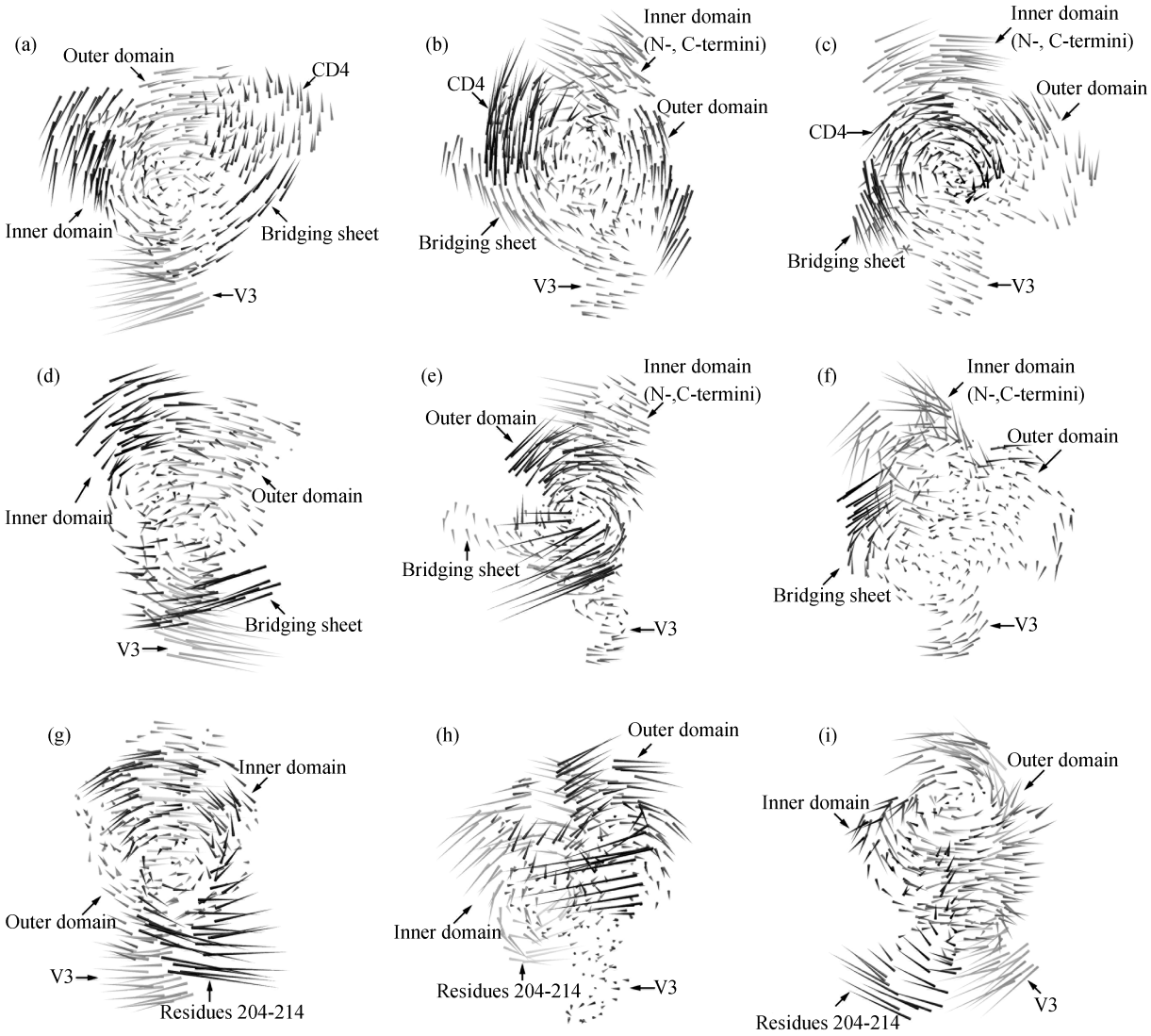

Figure 4 Porcupine plots of the principal modes of conformational variability calculated from CONCOORD ensembles for the gp120-CD4 complex, CD4-free, and unliganded gp120s. (a) The most significant motion for the gp120-CD4 complex is dominated by a twist of the structural region composed of the CD4, the inner domain and bridging sheet of gp120 relative to the outer domain of gp120. The view is from the inner domain to the outer domain. (b) The second ranked motion for the gp120-CD4 complex is dominated by a twist of the structural region composed of the CD4 and a large part of the outer domain relative to the region composed of the inner domain, bridging sheet, and V3 loop. The view is from the outer to inner domains. (c) The third ranked mode for the gp120-CD4 complex describes a twist of the region composed of the N-, C-termini and V3 loop relative to the region composed of the CD4, bridging sheet and a large part of the outer domain. This view is looking "down" from the CD4 molecule. (d) The most significant motion for the CD4-free gp120 is a twist of the inner domain (including the bridging sheet) relative to the outer domain. The view is from the inner to outer domains. (e) The second ranked motion for the CD4-free gp120 is dominated by a twist of a large part of the outer domain (without V3 loop) relative to the region composed of the inner domain, bridging sheet, and V3 loop. The view is from the outer to inner domains. (f) The third ranked mode for the CD4-free gp120 describes a flexing of the N-, C-termini of the inner domain relative to the clockwise vortex formed by the bridging sheet and outer domain. This view is from the CD4 binding cavity. (g) The most significant motion for the unliganded gp120 is dominated by a twist of the inner domain with respect to the outer domain. This view is from the inner to outer domains. (h) The second ranked motion for the unliganded gp120 is dominated by a twist of a large part of the outer domain relative to the inner domain, with the β20-β21 and V3 in the outer domain exhibiting the smallest fluctuations. The view is from the outer to inner domains. (i) The third ranked mode for the unliganded gp120 describes a relative flexing of the region composed of residues 204—214 and V3 loop relative to the anticlockwise vortex formed by large parts of the inner and outer domains. This view is from the inner to outer domain.

relative to the N-, C-termini of the inner domain (Figure 4(f)). The fourth eigenvector characterizes shortnesses of the inner and outer domains along respective major axes, resulting in condensations of the two domains (data not shown).

For the unliganded gp120 ensemble, the motion characterized by eigenvector 1 is a twist between the inner and outer domains. As shown by the porcupine plot

(Figure 4(g)), the two domains form a pair of opposite spin vortices that rotate around an axis connecting the centers of the two domains. The two domains rotate in the same direction as the twisting motion described by eigenvector 1 for the CD4-complexed gp120, so that the unliganded and CD4-complexed gp120 share the common most significant motional modes. Eigenvector 2 of the unliganded ensemble describes a twisting motion between the inner and outer domains by which two vortices rotating in the opposite directions are formed (Figure 4(h)). This collective motional mode is similar to the second largest modes obtained from the CD4-complexed and CD4-free ensembles, despite the smallest fluctuations in the β20-β21 ribbon, *L*F, and V3 loop of the unliganded gp120. We also note that the long, narrow channel at the interfaces of the inner and outer domains undergoes a slightly closing motion. Unlike the third ranked motions for the CD4-complexed and CD4-free gp120s, the obtained eigenvector 3 for the unliganded gp120 is dominated by a single anticlockwise vortex formed by large parts of the inner and outer domains, which rotates around an axis running from the outer domain to the inner domain. Other parts of the unliganded gp120 such as residues 204—214 in the inner domain and the V3 loop in the outer domain undergoes flexing motions in the opposite direction relative to the single vortex (Figure 4(i)). Eigenvector 4 describes elongation motions of the inner and outer domains, resulting in respective thinnings of the two domains (data not shown).

2.4 Conformational space overlap and conformational transition

Comparisons of motions of the three gp120 models in different conformational states were made by calculations of the essential subspace overlap between the eigenvectors derived from the diagonalization of the covariance matrices because the essential subspace is of particular interest. Generally, more than 80% of the observed structural fluctuations occur in a subspace spanned by the first 10 eigenvectors[21]. Figure 5 shows the cumulative mean square inner product (CMSIP) between the first 10 eigenvectors spanning an essential subspace from one CONCOORD simulation and all the eigenvectors from another CONCOORD simulation. For the CD4-complexed gp120 and CD4-free gp120, Figure 5(a) shows that the curves are steep, indicating a high degree of overlap is concentrated in the initial part. For example, 80% of the overlap with the first 10 eigenvectors obtained from the CD4-complexed gp120 simulation is reached within the first 14 eigenvectors obtained from the CD4-free gp120 simulation (CD4-cplx vs. CD4-free), and 80% of the overlap with the first 10 eigenvectors from the CD4-free simulation is reached within the first 12 eigenvectors from the CD4-complexed gp120 simulation (CD4-free vs. CD4-cplx), indicating that all essential fluctuations occurring in the

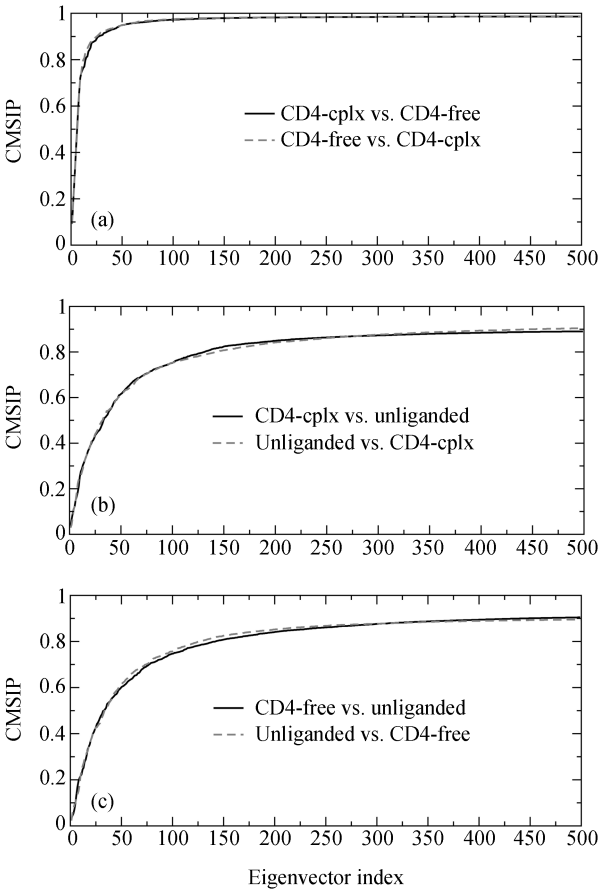

Figure 5 Cumulative mean square inner products (CMSIP) between the first 10 eigenvectors with the largest eigenvalues obtained from CONCOORD simulations and all the eigenvectors obtained from different CONCOORD runs. (a) The solid line represents the overlap of the first 10 eigenvectors from the CD4-complexed gp120 simulation against all the eigenvectors from the CD4-free gp120 simulation (CD4-cplx vs. CD4-free). The dashed line represents the overlap of the first 10 eigenvectors from the CD4-free simulation against all the eigenvectors from the CD4-complexed simulation (CD4-free vs. CD4-cplx). (b) The solid line represents the overlap of the first 10 CD4-complexed eigenvectors against all the unliganded gp120 eigenvectors (CD4-cplx vs. unliganded). The dashed line represents the overlap of the first 10 unliganded eigenvectors against all the CD4-complexed eigenvectors (unliganded vs. CD4-cplx). (c) The solid line is the overlap of the first 10 CD4-free eigenvectors against all the unliganded gp120 eigenvectors (CD4-free vs. unliganded). The dashed line is the result of the reversed overlap (unliganded vs. CD4-free).

CD4-free gp120 also occur in the CD4-complexed gp120. Furthermore, all eigenvectors from the CD4-free simulation can rebuild 98.6% of the first 10 eigenvectors from the CD4-complexed simulation, and *vice versa*, indicating that the two gp120 models have almost the identical motional freedom.

For the CD4-complexed g120 and unliganded gp120 (Figure 5(b)), the overlap curves are less steep, indicating that the overlaps converge slowly compared to Figure 5(a). The unliganded simulation requires its first 130 eigenvectors to rebuild 80% of the first 10 eigenvectors from the CD4-complexed simulation (CD4-cplx vs. unliganded), and the CD4-complexed simulation requires its first 140 eigenvectors to rebuild 80% of the first 10 eigenvectors from the unliganded gp120 simulation (unliganded vs. CD4-cplx). Furthermore, all eigenvectors of one set are able to rebuild 90.0% of the first 10 eigenvectors of the other set. These results indicate that there are clear differences in motional freedom between the CD4-complexed and unliganded gp120s. The probable reason for this is that the large conformational differences between them (Figure 2). However, the large overlaps (90%) between all eigenvectors of one set and the essential subspaces spanning the first 10 eigenvectors of the other set indicate the conformational space of one set can rebuild the essential subspace of the other set, implying that the unliganded gp120 has the potential to translate its conformation into the CD4-complexed state, and vise versa. Furthermore, the unliganded gp120 is likely to possess the larger potential to experience a transition towards the CD4-complexed (CD4-bound) conformation than the CD4-complexed gp120 towards the unliganded gp120 because the unliganded gp120 require less quantity of eigenvectors (130) than the CD4-complexed gp120 (140) for rebuilding 80% of the essential subspace of the other set.

For the CD4-free and unliganded gp120s (Figure 5(c)), the convergences of the overlaps are similar to those for the CD4-complexed and unliganded gp120s (Figure 5(b)). The first 140 eigenvectors from the unliganded simulation will rebuild 80% of the first 10 eigenvectors from the CD4-free simulation (CD4-free vs. unliganded). Under the reversed condition, the required number of the eigenvectors is 125 (unliganded vs. CD4-free). Despite the large difference in motional freedoms between the CD4-free and unliganded gp120s, it seems that the CD4-free gp120 has a greater transition potential towards the unliganded state, while the unliganded gp120 has a relatively weak transition potential towards the CD4-free form. The reasons for this is that the CD4-free gp120 requires its first 125 eigenvectors to reproduce 80% of the essential subspace obtained from the unliganded simulation, whereas the unliganded gp120 requires more eigenvectors (about 140) to reproduce 80% of the essential subspace obtained from the CD4-free simulation.

The overlaps of the essential subspaces obtained from CONCOORD simulations for the CD4-complexed, CD4-free, and unliganded gp120s have been evaluated in a more quantitative way in order to ascertain the preference for conformational transition between these three states. For all three simulation systems studied, the first 4 eigenvectors and the first 10 eigenvectors contribute more than 65% and 80% to the total fluctuations (for details, see Essential dynamics analyses in the Results section), respectively, indicating that the subspaces spanned by the first 10 eigenvectors, especially by the first 4 eigenvectors, cover the majority of the motional modes, namely the "essential" fluctuations. Therefore, the cumulative mean square inner products between the first 4 eigenvectors of one conformational state and the first 10 eigenvectors of another conformational state were calculated (Table 1), which provides a more strictly quantitative measure for the preference for conformational transition. Table 1 shows that the fist 4 eigenvectors obtained from the CD4-complexed simulation are to a large extent (83.8%) presented in the subspace spanned by the first 10 eigenvectors obtained from the CD4-free simulation (CD4-complex vs. CD4-free). Under the reversed condition, the first 4 eigenvectors from the CD4-free simulation are reproduced 85% by the first 10 eigenvectors from the CD4-complexed simulation (CD4-free vs. CD4-complex). Despite the subtle difference in the overlaps between the CD4-complex vs. CD4-free and CD4-free vs. CD4-complex, we could infer from which that the CD4-complexed gp120 has a larger potential to translate its conformation into the CD4-free state, while the CD4-free has a smaller potential to translate its conformational state into the CD4-complexed state. Similarly, the unliganded gp120 has larger potential to translate its conformation into the CD4-complexed form because its first 10 eigenvectors can reproduce 26.9% of the essential subspace spanned by the first 4 eigenvectors obtained from the CD4-complexed gp120 simulation, but that the value for the reversed overlap is only 22.0%, indicating a relatively

Table 1 Cumulative mean square inner products between the first 4 eigenvectors of one set and the first 10 eigenvectors of another set

Eigenvectors 1−4	Eigenvectors 1−10		
	CD4-free	unliganded	CD4-complexed
CD4-free	1.00	0.180	0.850
Unliganded	0.259	1.00	0.220
CD4-complexed	0.838	0.269	1.00

weak transition potential for the CD4-complexed gp120 towards the unliganded gp120. For the CD4-free gp120 and unliganded gp120, the preference for conformational transition between them is that CD4-free gp120 has larger transition potential (25.9%) towards the unliganded state, while the potential for the reversed transition is smaller (18.0%). Of particular interest is that the overlap (25.9%) of the first 10 eigenvectors from the CD4-free simulation against the first 4 eigenvectors from the unliganded simulation is larger than the overlap (22.0%) of the first 10 eigenvectors from the CD4-complexed simulation against the first 4 eigenvectors from the unliganded simulation, indicating that the CD4-free gp120 has a greater transition potential towards the unliganded form than the CD4-complexed gp120.

3 Discussion

The prerequisite for our analyses of the gp120 dynamics is to obtain in-depth knowledge of the structures in both the CD4-bound and unliganded states. The 100% sequence identity between the target sequence of the HXBc2 and its template 1G9M results in an accurate core structural model of the liganded gp120 (CD4-complexed or CD4-free gp120) with approximately accurate V3 and V4 loop structures modeled on this core as more than 98% residues in this model fall into the favoured regions of the Ramachandran plot obtained from the PROCHECK program[44]. The direct experimental data of the unliganded structure of the HIV-1 gp120 is not available at this state. However, given the high sequence identity between the SIV and HIV (identity 35%, similarity 70%), the unliganded HIV gp120 is likely to assume a similar conformation. It was noted[45] that it was possible to construct medium-accuracy unliganded HIV-1 gp120 models with ~90% of the main-chain atoms being modeled with 1.5 Å RMS error as the sequence similarity between the target and the template is more than 70%. In our unliganded gp120 model with the modeled V3 loop, there are more than 90% residues that have appropriate dihedral angel distributions. Therefore, most of the differences between our HIV-1 homology models of the unliganded gp120 and the CD4-complexed/CD4-free gp120 should reflect the conformational alterations induced by receptor binding. Thus, our gp120 models in the three conformational forms (CD4-complexed, CD4-free, and unliganded) are suitable for exploring the molecular motion and conformational flexibility in gp120.

To study the mobility and flexibility of a protein, simulation techniques such as the MD or CONCOORD are required to generate ensemble of structure in which each frame contains an accessible conformation. Although the MD structures are in principle free to access all physically obtainable conformation, the relatively short timescales that this approach can currently access; the incremental search protocol that is used to obtain individual members of the ensemble; and the Brownian and viscous effects of the solvent environment, may limit the full exploration of the conformational space available to the protein[22]. On the contrary, the CONCOORD structures are derived from constraints that circumscribe a region of the conformational space that can be accessed. This space is thoroughly explored in a CONCOORD ensemble as a result of the highly randomizing protocol by which each member of the ensemble is generated[21]. Thus, the ensembles derived from the CONCOORD are most useful in identifying global large-scale motions that a protein is able to perform around an equilibrium position. Also, these ensembles can be extrapolated to predict the character of conformational changes, and further to predict the preference for conformational transition between different conformational states of a protein[21−23,26].

For all the three simulated systems, the ED analyses of CONCOORD ensembles reveal that the first 4 eigenvectors account for a large proportion of the total positional fluctuation. The modes of the large-scale collective protein fluctuations are likely involved in the mechanism responsible for gp120-ligand (CD4, antibodies, or small molecular inhibitors) association and neutralization avoidance. Although the ED technique yields individual collective motions, all the modes inter-

play with each other in a complex manner, and hence not necessarily all individual modes will correspond to a specific functional task. However, for clarity, we will discuss the individual modes of the concerted motion one by one and relate them to putative functions.

The most significant motions of gp120 in the CD4-complexed, CD4-free, and unliganded forms described by eigenvector 1 are all the twist of the inner domain relative to the outer domain. However, the variation of twisting directions among the three conformational forms will result in different functional effects. For the gp120-CD4 complex (Figure 4(a)), the gp120 inner domain and bridging sheet together with the CD4 D1 domain form an anticlockwise vortex, which rotates with respect to a clockwise vortex formed by the gp120 outer domain. The effect of such twist will grip CD4 molecule tightly because the bridge sheet and components (in particular the *L*D, V5, and *L*E) that are located at the proximal end of the outer domain move towards the CD4. For the CD-free gp120 (Figure 4(d)), the clockwise vortex formed by the inner domain and bridging sheet rotates relative to the anticlockwise vortex formed by the outer domain, resulting in a twisting mode in the opposite direction compared to that of the CD4-complexed gp120. This mode will enlarge the CD4 cavity because the bridge sheet and components located at the proximal end of the outer domain move away from the cavity. We considered that, when the effect of motional directions of the eigenvectors is excluded, the removal of CD4 may lead to an increased closing or opening of the CD4-binding cavity, resulting in jumping out of energetic barrier for CD4-free gp120 to translate its conformation towards unliganded form. For the unliganded gp120, the twisting motion is similar to that for the CD4-complexed gp120 despite the large structural difference between them, i.e., the anticlockwise vortex formed by the inner domain rotates with respect to the clockwise vortex formed by the inner domain (Figure 4(g)). Although the twisting mode in the unliganded gp120 does not directly open or close the long, narrow channel, the common motional mode for the CD4-complexed and unliganded gp120s may endow the unliganded gp120 with transition potential towards the CD4-complexed state, or towards states that can favourably associate with the CD4.

All the second eigenvectors from the three simulation systems are dominated by a similar rotation/twisting motion of the outer domain with respect to the inner domain (Figure 4(b), (e) and (h)), although the bridging sheet and V3 loop in the CD4-complexed and CD4-free gp120s rotate in concert with the inner domain. For the CD4-complexed and CD4-free gp120s, this motion will enlarge the CD4-binding cavity and may be related to CD4 release, whereas for the unliganded gp120, this motion will slightly close the long, narrow channel. Of particular interest is that the V3 loop and *L*F in the CD4-complexed gp120 emerge with the largest motional amplitude among the three gp120 forms, which is probably caused by CD4 binding so that we speculate that the high mobility of the *L*F and V3 would be of great benefit to association with the coreceptor CCR5 or CXCR4.

The rotations of the two vortices described by eigenvectors 3 obtained from the CD4-complexed g120 simulation will close the narrow part of CD4 cavity that is located at the interfaces (composed of the α5, *L*C, *L*D, β9−β11, see Figure 2(a)) of the proximal ends of the inner and outer domains but open the Phe 43 pocket that is located at the intersection of the outer domain, the inner domain, and the bridging sheet. This mode may be related to the orientation and insertion of the CD4 Phe 43 into its pocket (Figure 4(c)). However, the two vortices of the CD4-free gp120 rotate in the opposite directions (Figure 4(f)) compared to those of the CD4-complexed gp120, which leads to the reversed effects on the narrow part of CD4 cavity and Phe 43 pocket, i.e. enlarges the former and reduces the latter. In general, the motions described by eigenvector 3 for the CD4-complexed gp120 govern the opening or closing of the CD4 Ph43-binding pocket given the direction effect is not considered. The third eigenvector of the unliganded gp120 describes a flexing motion of the outer domain and inner domain relative to the V3 loop and residues 204−214 (Figure 4(i)). This kind of flexing will influence reorientations of the V1/V2 stem, the V3 loop, and the connecting segment (residues 204−214) between the V1/V2 stem and three-stranded bundle at the termini-proximal end of the inner domain, resulting in the approach of the V1/V2 stem to the β20-β21 ribbon to drive the formation of the bridging sheet.

The fourth eigenvectors of the CD4-complexed, CD4-free, and unliganded gp120s describe the elongations or shortnesses of the inner and the outer domains along respective major axes. These modes will lead to the thinnings or condensations of the two domains and may contribute to conformational transition between

these three states.

It is worth pointing out that the large-scale collective motions described by the first 4 eigenvectors for the three gp120 forms do not represent explicit rotation/twisting or breathing (open/close) modes between the two domains as have been observed for the T4 lysozyme[40], hyaluronate lyase[24], and cyclin-dependent kinase 2[22]. Although gp120 in the CD4-bound and unliganded states can be divided into two major domains (the inner domain and the outer domain) based on their structural organizations[11,14], ED and DYNAMITE analyses demonstrate that some of the structural components such as the bridging sheet, V1/V2 stem, and V3 loop can move concertedly either with the inner domain or with the outer domain, demonstrating perplexing combinations of variable motional modes (rotation/twisting, open/close, or closure/flexing) of one mobile unit (or domain) with respect to the other mobile unit (or domain). We considered that these perplexing collective motions can be related to the dynamics mechanism involving in receptor association. On the other hand, these modes may also provide one probable explanation for the failure of gp120 to induce potent neutralizing antibodies against the CD4 and coreceptor binding sites.

The quantitative insights into the preference for conformational transition were obtained by comprehensive comparative analyses of the essential subspace overlaps (cumulative mean square inner products) between the three simulated models (Table 1). In the case of gp120 models in the same starting conformation such as CD4-free and CD4-complexed gp120s, we found that the essential subspace overlaps between them were greater than 80%, whereas in the case of gp120 models in the different starting conformation such as CD4-free/CD4-complexed and unliganded gp120s, the overlaps were less than 30%. The large structural differences between the CD4-free/CD4-complexed and unliganded forms are responsible for the small essential subspace overlap values, indicating the apparent differences in motional freedoms between the CD4-bound and CD4-unliganded states. As explained in the previous section, CONCOORD generates structures by a random search method that searches for solution in a predefined coordinate space with all the interactions implemented in the form of distance constraints[21], and therefore only short-range interactions within the protein make a serious contribution. We infer that the overestimation of the contribution of the non-covalent or short-range interaction in CONCOORD is also responsible for the small essential subspace overlap between the unliganded gp120 and the CD4-complexed/CD4-free gp120. To evaluate to what extent the three gp120 forms can sample the common configurational spaces under a more physically realistic condition, essential subspace overlap analyses based on the long-time MD simulations should be performed, which is beyond the scope of this paper. Nevertheless, here we mainly focus on the variations of the values for the similar essential subspace overlaps, i.e. the overlaps of models in the same starting conformational state against another model in the different starting conformation. Although the differences between similar overlapping values are small, they are significant for reasons that CONCOORD hardly suffers from a sampling problem when at least a few hundred structures are generated. Therefore the essential subspace overlapping value is a sensitive indicator for evaluating the dynamics similarity and conformational transition preference for a protein in different conformational states[21,26]. For example, for similar essential subspace overlaps, the CD4-free gp120 against the unliganded gp120 and the CD4-complexed gp120 against the unliganded gp120, the overlapping values are 0.259 and 0.220, respectively (Table 1), the former is greater than the latter despite the small overlapping difference, from which we can infer that the CD4-free gp120 has a larger transition potential towards the unliganded form than the CD4-complexed gp120. This result is in agreement with the MD data[19,20] and the extrapolation based on chemical intuition[11], suggesting that the removal of CD4 from the gp120-CD4 complex would destabilize the CD4-bound conformation and cause conformational diffusion towards the unliganded state. It has been generally accepted that the unliganded HIV-1 gp120 is structurally flexible, but that the CD4's association stabilizes the CD4-bound form which can be recognized by the second receptor, CCR5 or CXCR4[11,14]. In some cases, the HIV-1 gp120 glycoproteins exhibit the ability to bind chemokine receptor in the absence of CD4[47], which implies that the unliganded gp120 is likely to fluctuate spontaneously towards the CD4-bound state[14]. Our essential subspace overlap analyses shows that the overlap value of the unliganded gp120 against the CD4-free gp120 (18.0%) is smaller than the overlap value of the unliganded gp120 against the CD4-complexed gp120 (26.9%), reflecting that the unliganded

gp120 has a greater potential to fluctuate towards the CD4-complexed form than towards the CD4-free form. This provides an explanation for the CD4-independent infection phenomenon.

4 Conclusion

Our homology models of the HIV-1 gp120 in the CD4-complexed, CD4-free, and unliganded forms represent three conformational states, which are the excited or pre-fusogenic state, the excited state with the removal of the CD4, and the relaxed ground state, respectively. The CONCOORD computer simulations and subsequent ED analyses of the ensembles revealed that the large-scale concerted motions are dominated by perplexing combined rotations of the vortices formed between or within the inner domain, outer domain, bridging-sheet, and V3 loop. These modes will open/close the CD4 cavity and reposition certain structural components involved in ligand association. The comparative analyses of essential subspace overlaps quantitatively distinguish the preference for conformational transition between the three states. These results are in agreement with chemical intuition: (1) the removal of CD4 causes destabilization in the CD4-free gp120 which in turn enhances the fluctuations towards the unliganded state; (2) despite the large differences in conformation between the unliganded and CD4-free/CD4-complexed gp120s, the unliganded gp120 prefers the transition towards the CD4-complexed state to the transition towards the CD4-free state, providing a plausible explanation for the CD4 independent entry pathway. We have also demonstrated the sophisticated plasticity of gp120, from its rigid core to a high-flexibility exterior coating that provides recognition specificity without compromising the capability to avoid neutralization.

The authors thank two anonymous reviewers for their valuable comments and High Performance Computer Center in Yunnan University for computational support.

1 Barre-Sinoussi F, Chermann J C, Rey F, et al. Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immunodeficiency syndrome (AIDS). Science, 1983, 220: 868－871

2 Gallo R C, Salahuddin S Z, Popovic M, et al. Frequent detection and isolation of cytopathic retroviruses (HTLV-III) from patients with AIDS and at risk for AIDS. Science, 1984, 224: 500－503

3 Heeney J L, Hahn B H. Vaccines and immunology: Elucidating immunity to HIV-1 and current prospects for AIDS vaccine development. AIDS, 2000, 14(Suppl): s125－s127

4 Klein E, Ho R. Challenges in the development of an effective HIV vaccine: Current approaches and future directions. Clin Ther, 2000, 22: 295－314

5 Dalgleish A G, Beverley P C, Clapham P R, et al. The CD4 (T4) antigen is an essential component of the receptor for the AIDS retrovirus. Nature, 1984, 312: 763－767

6 Feng Y, Broder C C, Kennedy P E, et al. HIV-1 entry cofactor: Functional cDNA cloning of a seven-transmembrane, G protein-coupled receptor. Science, 1996, 272: 872－877

7 Trkola A, Dragic T, Arthos J, et al. CD4-dependent, antibody-sensitive interactions between HIV-1 and its co-receptor CCR-5. Nature, 1996, 384: 184－187

8 Wu L, Gerard N P, Wyatt R, et al. CD4-induced interaction of primary HIV-1 gp120 glycoproteins with the chemokine receptor CCR-5. Nature, 1996, 384: 179－183

9 Veronese F D, de Vico A L, Copeland T D, et al. Characterization of gp41 as the transmembrane protein coded by the HTLV-III/LAV envelope gene. Science, 1985, 229: 1402－1405

10 Trkola A, Purtscher M, Muster T, et al. Human monoclonal antibody 2G12 defines a distinctive neutralization epitope on the gp120 glycoprotein of human immunodeficiency virus type 1. J Virol, 1996, 70: 1100－1108

11 Kwong P D, Wyatt R, Robinson J, et al. Structure of an HIV gp120 envelope glycoprotein in complex with the CD4 receptor and a neutralizing human antibody. Nature, 1998, 393: 648－659

12 Kwong P D, Wyatt R, Majeed S, et al. Structures of HIV-1 gp120 envelope glycoproteins from laboratory-adapted and primary isolates. Struct Fold Des, 2000, 8: 1329－1339

13 Wyatt R, Kwong P D, Desjardins E, et al. The antigenic structure of the HIV gp120 envelope glycoprotein. Nature, 1998, 393: 705－711

14 Chen B, Vogan E M, Gong H, et al. Structure of an unliganded simian immunodeficiency virus gp120 core. Nature, 2005, 433: 834－841

15 Kwong P D, Doyle M L, Casper D J, et al. HIV-1 evades antibody-mediated neutralization through conformational masking of receptor-binding sites. Nature, 2002, 420: 678－682

16 Myszka D G, Sweet R W, Hensley P, et al. Energetics of the HIV gp120-CD4 binding reaction. Proc Natl Acad Sci USA, 2000, 97: 9026－9031

17 Berendsen H J C, Hayward S. Collective protein dynamics in relation to function. Curr Opin Struct Biol, 2000, 10: 165－169

18 Hsu S T, Bonvin A M. Atomic insight into the CD4 binding-induced conformational changes in HIV-1 gp120. Proteins, 2004, 55: 582－593

19 Pan Y, Ma B, Nussinov R. CD4 binding partially locks the bridging sheet in gp120 but leaves the beta2/3 strands flexible. J Mol Biol, 2005, 350: 514－527

20 Pan Y, Ma B, Keskin O, et al. Characterization of the conformational state and flexibility of HIV-1 glycoprotein gp120 core domain. J Biol Chem, 2004, 279: 30523－30530

21 de Groot B L, van Aalten D M F, Scheek R M, et al. Prediction of protein conformational freedom from distance constraints. Proteins, 1997, 29: 240－251

22 Barrett C P, Noble M E. Molecular motions of human cy-

clin-dependent kinase 2. J Biol Chem, 2005, 280: 13993－14005
23 Barrett C P, Hall B A, Noble M E. Dynamite: A simple way to gain insight into protein motions. Acta Crystallogr D Biol Crystallogr, 2004, 60: 2280－2287
24 Mello L V, de Groot B L, Li S, et al. Structure and flexibility of *Streptococcus agalactiae* hyaluronate lyase complex with its substrate. Insights into the mechanism of processive degradation of hyaluronan. J Biol Chem, 2002, 277: 36678－36688
25 Vreede J, van der Horst M A, Hellingwerf K J, et al. PAS domains. Common structure and common flexibility. J Biol Chem, 2003, 278: 18434－18439
26 Liu S Q, Liu C Q, Fu Y X, Molecular motions in HIV-1 gp120 mutants reveal their preferences for different conformations. J Mol Graphics Modell, 2007, 26: 306－318
27 Bairoch A, Apweiler R, Wu C H, et al. The Universal Protein Resource (UniProt). Nucl Acids Res, 2005, 33: 154－159
28 Deshpande N, Addess K J, Bluhm W F, et al. Describes the capabilities of the PDB Beta site. Nucl Acids Res, 2005, 33: 233－237
29 Vranken W F, Budesinsky M, Fant F, et al. The complete consensus V3 loop peptide of the envelope protein Gp120 of HIV-1 shows pronounced helical character in solution. FEBS Lett, 1995, 374: 117－121
30 Sali A, Blundell T L. Comparative protein modelling by satisfaction of spatial restraints. J Mol Biol, 1993, 234: 779－815
31 Amadei A, Linssen A B M, Berendsen H J C. Essential dynamics of proteins. Proteins, 1993, 17: 412－425
32 Levy R, Srinivasan A, Olson W, et al. Quasiharmonic method for studying very low frequency modes in proteins. Biopolymers, 1984, 23: 1099－1112
33 Garcia A E. Large-amplitude nonlinear motions in proteins. Phys Rev Lett, 1992, 68: 2696－2699
34 Hayward S, Go N. Collective variable description of native protein dynamics. Annu Rev Phys Chem, 1995, 46: 223－250
35 Berendsen H J C, van der Spoel D, van Drunen R. GROMACS: A message-passing parallel molecular dynamics implementation. Comp Phys Comm, 1995, 91: 43－56
36 Lindahl E, Hess B, van der Spoel D. GROMACS 3.0: A package for molecular simulation and trajectory analysis. J Mol Mod, 2001, 7: 306－317
37 van Aalten D M F, Findlay J B C, Amadei A, et al. Essential dynamics of the cellular retinol binding protein: evidence for ligand induced conformational changes. Prot Eng, 1995, 8: 1129－1136
38 van Aalten D M F, de Groot B L, Berendsen H J C, et al. A comparison of techniques for calculating protein essential dynamics. J Comp Chem, 1997, 18: 169－181
39 Humphrey W, Dalke A, Schulten K. VMD——Visual molecular dynamics. J Mol Graph, 1996, 14: 33－38; 27－28
40 de Groot B L, Hayward S, van Aalten D M, et al. Domain motions in bacteriophage T4 lysozyme: A comparison between molecular dynamics and crystallographic data. Proteins, 1998, 31: 116－127
41 Merlino A, Vitagliano L, Ceruso M A, et al. Dynamic properties of the N-terminal swapped dimer of ribonuclease A. Biophys J, 2004, 86: 2383－2391
42 Roccatano D, Daidone I, Ceruso M A, et al. Selective excitation of native fluctuations during thermal unfolding simulations: Horse Heart Cytochrome C as a case study. Biophys J, 2003, 84: 1876－1883
43 Hess B. Similarities between principal components of protein dynamics and random diffusion. Phys Rev E, 2000, 62: 8438－8448
44 Laskowski R A, MacArthur M, Moss D S, et al. PROCHECK: A program to check the stereochemical quality of protein structures. J Appl Crystallogr, 1993, 26: 283－291
45 Baker D, Sali A. Protein structure prediction and structural genomics. Science, 2001, 294: 93－96
46 Liu S Q, Liu S X, Fu Y X. Dynamic domains and geometrical properties of HIV-1 gp120 during conformational changes induced by CD4 binding. J Mol Mod, 2007, 13: 411－424
47 Kolchinsky P, Mirzabekov T, Farzan M, et al. Adaptation of a CCR5-using, primary human immunodeficiency virus type 1 isolate for CD4-independent replication. J Virol, 1999, 73: 8120－8126