Available online at www.sciencedirect.com
ScienceDirect
Journal of Molecular Graphics and Modelling 26 (2007) 306–318
Journal of Molecular Graphics and Modelling
www.elsevier.com/locate/JMGM

Molecular motions in HIV-1 gp120 mutants reveal their preferences for different conformations

Shu-Qun Liu [a], Ci-Qun Liu [a], Yun-Xin Fu [a,b,*]

[a] *Laboratory for Conservation and Utilization of Bio-resources, Yunnan University, Kunming 650091, PR China*
[b] *Human Genetics Center, The University of Texas Health Science Center, Houston, TX 77030, USA*

Received 31 August 2006; accepted 10 December 2006
Available online 15 December 2006

Abstract

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. We built the homology models of gp120 with loops V3 and V4 in the CD4-complex, CD4-free and CD4-unliganded states, and models of the 375 S/W and 423 I/P mutants in the CD4-free and unliganded states, respectively. CONCOORD was utilized for generating ensembles of the seven gp120 models that were analyzed by essential dynamics analyses to identify their preferred concerted motions. The revealed large-scale concerted motions are related to either the receptor association/release or the conformational transition between different conformational states. Essential subspace overlap analyses were performed to quantitatively distinguish the preference for conformational transitions between states of the gp120 mutants and further to ascertain what kind of conformational state that the mutants prefer to adopt. Results indicate that the 375 S/W mutant, in which the tryptophan indole group is predicted to occupy the phe43 pocket in the gp120 interior, favors a conformation close to the CD4-bound state. However, the other mutant 423 I/P inclines to prevent the formation of bridging sheet and stabilize the conformation in the unliganded state. Our theoretical analyses are in agreement with experimentally determined mutation effects, and can be extended to a new approach to design or screen mutants that have effects on conformation/function of a protein.

Keywords: gp120 mutants; Conformational transition; Essential dynamics; Essential subspace overlap; Mutation effect

1. Introduction

The human immunodeficiency viruses (HIV-1 and HIV-2) and simian immunodeficiency viruses (SIV) are the etiologic agents of acquired immunodeficiency syndrome (AIDS) in their respective human and simian hosts [1–4]. Entry of the HIV-1 into target cells is initiated by the selective interaction between the viral exterior envelope glycoprotein, gp120, and the receptor on the surface of the host cell, CD4, as well as the obligatory chemokine receptors CCR5 or CXCR4 [5–7]. Accumulating biochemical and structural evidences indicate that the initial binding of the CD4 to the gp120 triggers conformational alterations in the HIV envelope that subsequently promote recognition of the coreceptor and ultimately lead to conformational changes of gp41 and membrane fusion [8–10]. This complex mechanism involves a series of structural rearrangements in which the conformational dynamics of the HIV envelope glycoproteins plays crucial role.

Both the crystal structures of the HIV-1 gp120 core [11,12] bound by the CD4 and antigen binding fragment (Fab) of the human neutralizing antibody 17b, and the SIV gp120 core [13] are known. They are conformations corresponding to post- and pre-bound by the CD4, respectively. Although both of the cores contained two domains (the inner and outer domains), the rough structural comparison revealed unexpectedly extensive conformational rearrangements upon receptor binding [13], which was in line with thermodynamics evidence suggesting that the binding of CD4 or antibodies to gp120 induced substantial structural rearrangements of the gp120 [14,15].

Although the crystallographic studies have provided two static snapshots of gp120 cores in the CD4-bound and CD4-unliganded states, it has been generally accepted that the

* Corresponding author at: Laboratory for Conservation and Utilization of Bio-resources, Yunnan University, Kunming 650091, PR China.
Tel.: +86 871 5035257; fax: +86 871 5035257.
E-mail address: shuqunliu@ynu.edu.cn (Y.-X. Fu).

1093-3263/$ – see front matter
doi:10.1016/j.jmgm.2006.12.007

conformational flexibility of HIV-1 envelope glycoproteins is important for their functions in mediating virus entry and evading humoral immune response [13–16]. The comparative analyses of molecular motions and dynamics data for gp120 in different conformational states are helpful in a full understanding of the protein function. In particular, the understanding of ranges of conformations available to the gp120 in different functional states and further distinguishing the preference for conformational transition between states are of importance for the appreciation of their functional roles and for guiding attempts at intervention. Xiang et al. [17] have identified two groups of HIV-1 gp120 mutants using the ligand-binding phenotypes to reach a conclusion that the gp120 glycoprotein can assume at least two distinct conformations. Concretely, for a mutant 375 S/W, both the enthalpy change and the entropy penalty observed when the wild-type gp120 bound to the CD4 were significantly reduced, and it was recognized better than the wild-type gp120 by the CD4 and CD4-induced (CD4i) antibodies, whereas the affinity to antibodies elicited against the CD4-binding site (CD4BS) was almost totally abolished, indicating the 375 S/W mutant favors a conformation near the CD4-bound state. Another mutant, the 423 I/P, did not bind the CD4 or CD4i antibodies, even though recognition by the CD4BS antibodies was efficient, indicating that the 423 I/P mutant favors a conformation differing from the CD4-bound state, i.e. the CD4-unliganded or ground state.

The conformational changes involved range from very subtle, local changes to global conformational changes involving motions of significant amplitude for a 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 crucial in the process of conformational transition from the CD4-unliganded state to the CD4-bound state as unusually large structural differences exist between them. Although there is currently no experimental technique that allows monitoring of protein conformational changes at atomic resolution, it has been found that further insights into the dynamic structure/function relationship can be gained through analysis of computational simulated dynamics [18]. For example, previous studies on the dynamics properties of gp120 in the CD4-free and gp120–CD4-complex states using 10 ns molecular dynamics (MD) simulations [19] have demonstrated that the CD4 association reduced conformational flexibility in certain loop regions of the outer domain. Other MD studies indicated that the CD4 can merely lock the β20–β21 hairpin of the bridging sheet in gp120 but leave the β2–β3 hairpin flexible [20,21]. A shortcoming of the MD approach, however, is that the 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 conformational space available to a protein, and is a few orders of magnitude smaller than that on which most biological processes take place. Thus the MD-derived ensembles are most useful in predicting both the detailed behavior and local conformational changes of a protein. An alternative approach to MD is to generate an ensemble of structures randomly without using Newton's equations of motion. The CONCOORD [22] takes this approach as a fast way to generate ensembles that explore the conformational space more fully. Previous studies [22–26] have shown that ensembles derived from CONCOORD are most useful in identifying global motions that a protein is able to perform and predicting the character of conformational changes.

We have performed the comparative modeling technique to generate the three-dimensional structural models of gp120 in three 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). We also introduced point mutation into the wild-type gp120 models to generate the 375 S/W and 423 I/P mutants in the CD4-free and unliganded states, respectively. These seven homology models were used as starting structures for CONCOORD to generate respective ensembles. One of our aims was to characterize the molecular motions of gp120 in different conformational states at global level to determine the hierarchy of tertiary structures that exists in gp120. The other aims were to investigate the variations of molecular motions caused by point mutations, and further identify the preference for conformation transition between gp120 mutants and finally, ascertain which conformational states that the 375 S/W and 423 I/P mutants prefer to adopt. Our results show that the 375 S/W mutant sampled a conformational space that is very close to the CD4-bound state, while the 423 I/P mutant preferred a conformation near the unliganded state. The consistency of our theoretical results with experimentally determined mutation effects not only compensates the mutagenesis experiments and MD studies, but also extends a new approach to design or screen mutants that have effects on conformation/function of a protein.

2. Materials and methods

2.1. Generation of the starting structures

The starting sequence of the HIV-1 HXBc2 isolate gp160 precursor (Swiss-Prot accession no. P04578) was obtained from the Swiss-Prot protein sequence database [27]. The sequence for the transmembrane glycoprotein gp41, 52 and 19 residues from the N- and C-termini and the V1/V2 loop (Gly-Ala-Gly substitutes for 67 V1/V2 loop residues) of the gp120 were removed. The final gp120 primary sequence consists of residues from 83 to 492, including sequences of the loops V3 and V4. Structural templates of the gp120 core and V3 loop were obtained from the PDB protein structure database [28], out of them PDB entries 1G9M (chain G) [11] and 2BF1 [13] were used as templates to model the CD4-bound and unliganded gp120 cores, respectively, and 1CE4 [29] for the V3 loop.

To obtain homology models of the gp120 with V3 and V4 loops in the CD4-bound conformational state, the sequence