J Mol Model (2003) 9:329–336
DOI 10.1007/s00894-003-0154-9

ORIGINAL PAPER

Shuqun Liu · Shixiu Fan · Zhirong Sun

Structural and functional characterization of the human CCR5 receptor in complex with HIV gp120 envelope glycoprotein and CD4 receptor by molecular modeling studies

Received: 17 March 2003 / Accepted: 2 July 2003 / Published online: 29 August 2003

Abstract The entry of human immunodeficiency virus (HIV) into cells depends on a sequential interaction of the gp120 envelope glycoprotein with the cellular receptors CD4 and members of the chemokine receptor family. The CC chemokine receptor CCR5 is such a receptor for several chemokines and a major coreceptor for the entry of R5 HIV type-1 (HIV-1) into cells. Although many studies focus on the interaction of CCR5 with HIV-1, the corresponding interaction sites in CCR5 and gp120 have not been matched. Here we used an approach combining protein structure modeling, docking and molecular dynamics simulation to build a series of structural models of the CCR5 in complexes with gp120 and CD4. Interactions such as hydrogen bonds, salt bridges and van der Waals contacts between CCR5 and gp120 were investigated. Three snapshots of CCR5–gp120–CD4 models revealed that the initial interactions of CCR5 with gp120 are involved in the negatively charged N-terminus (Nt) region of CCR5 and positively charged bridging sheet region of gp120. Further interactions occurred between extracellular loop2 (ECL2) of CCR5 and the base of V3 loop regions of gp120. These interactions may induce the conformational changes in gp120 and lead to the final entry of HIV into the cell. These results not only strongly support the two-step gp120–CCR5 binding mechanism, but also rationalize extensive biological data about the role of CCR5 in HIV-1 gp120 binding and entry, and may guide efforts to design novel inhibitors.

S. Liu · Z. Sun (✉)
Institue of Bioinformatics,
Department of Biological Sciences and Biotechnology,
Tsinghua University,
100084 Beijing, People's Republic of China
e-mail: sunzhr@mail.tsinghua.edu.cn
Tel.: +86-10-62772237
Fax: +86-10-62772237

S. Fan
Cellular and Molecular Evolutionary Key Laboratory,
Kunming Institute of Zoology,
Chinese Academy of Sciences,
650223 Kunming, People's Republic of China

Keywords CCR5 · Gp120 · Interaction · Docking · Molecular dynamics simulation · Complex

Introduction

The entry of HIV[1]-1 into the target cell is mediated by CD4 as the primary receptor [1, 2], as well as by chemokine receptors such as CCR5, CXCR4, CCR3 and CCR2b as obligatory coreceptors [3, 4, 5, 6]. These chemokine receptors belong to the superfamily of G-protein-coupled receptors (GPCR) that possess seven transmembrane (TM) helices [7], among which CCR5 is of special importance for HIV pathogenesis because R5 virus strains are largely responsible for virus transmission and individuals who lack CCR5 due to a natural knock-out mutation in the CCR5 gene (ccr5 Δ32 allele) are highly resistant to HIV-1 infection [8, 9, 10]. All the biochemical and structural information available up to now permits us to propose a putative molecular mechanism of HIV-1 entry into cells: [11, 12, 13, 14, 15]. First, the HIV-1 glycoprotein gp120 recognizes the CD4 receptor on the cell surface, and binds to the most amino-terminal of the four immunoglobulin-like domains of CD4. Secondly, CD4 binding induces conformational changes in the gp120 glycoprotein, which forms or exposes the binding site for specific chemokine receptors. Thirdly, the binding of CCR5 to the gp120–CD4 complex triggers additional conformational changes in the envelope glycoprotein complex that ultimately lead to the fusion of the viral and target cell membranes. Among these steps, the interaction of CCR5 with the viral glycoprotein gp120 is critical for membrane fusion and virus entry because blockade of such binding can inhibit HIV-1 infection efficiently [15]. For this to occur, molecular events accompanying coreceptor-virus recognition and binding, and the precise mechanism for the engagement of CCR5 by the CD4-activated gp120 should be understood. Unfortunately, structural information of

[1] Human immunodeficiency virus

GPCR is difficult to obtain using experimental techniques because of enormous difficulties in the preparation of samples suitable for subsequent X-ray, NMR or electron microscopic analysis [16]. Thus, despite considerable efforts, no 3D structure of CCR5 has been solved to atomic resolution by experimental methods to date. However, a large amount of data is available for CCR5 from various experiments such as the chimeric chemokine receptors [17, 18, 18, 19, 20, 21], site-directed mutagenesis [22, 23, 24] and monoclonal antibody (mAb) competition binding experiments [25, 26], which have revealed the regions in CCR5 important for its interaction with HIV-1 and natural ligands. For instance, the Nt[2] extracelluar region plays a crucial role for high affinity binding of HIV-1 gp120 while ECL[3] is important for natural ligands (RANTES[4] and MIP[5]-1β) binding, besides, the extracellular loops are also important for inducing the conformational changes in gp120 that lead to membrane fusion. These available experimental data can also be used to help build molecular models of CCR5.

Recently, a structure was solved for a truncated core gp120 protein in a ternary complex with soluble CD4 and an antigen-binding fragment (Fab) of the neutralizing antibody 17b [13]. In this complex, gp120 is organized into an inner and an outer domain connected by a bridging sheet. The gp120 third variable (V3) loop has been implicated in chemokine receptor binding [11]. However, the use of the CCR5 chemokine receptor by diverse primate immunodeficiency viruses suggests the involvement of an additional, conserved gp120 region. The extended conformation of the bridging sheet constitutes key elements of the chemokine receptor binding site [15]. The core structure of CD4–gp120–17b reveals that these conserved gp120 regions form discontinuous structures important for the interaction with the coreceptors on the target cell.

Although the regions in CCR5 important for gp120 binding have been identified by various experiments, and crystallographic data have provided an important snapshot of the structures involved in the CD4–gp120–17b complex, the corresponding interaction sites in CCR5 and gp120 have not been matched yet because no structure of CD4–gp120–CCR5 has been resolved. Furthermore, a more complete understanding of molecular events accompanying coreceptor–virus recognition and binding, and of the precise mechanism for virus entry, requires building the structure of the CCR5 complexes with gp120 and CD4. In our latest paper, we built structural models of CCR5 through the approach of protein structure modeling and molecular dynamics simulation, the template selected for constructing the CCR5 structure is the crystal structures of bovine rhodopsin at 2.80 Å resolution. Although bovine rhodopsin has a low sequence homology with CCR5 (the identity is 20.6%), it belongs to the GPCR family and its crystal structure displays an explicit conformational feature of a bundle of seven TM α-helices shared by other GPCRs. Also, the sequence identity in transmembrane segment (TMS) between CCR5 and rhodopsin is about 30%. We found that in the CCR5 model with two disulfide linkages Cys20–269 and Cys101–178 formed, the Nt region could adopt two primary conformational states: it can either adsorb on the surface of the domain formed by extracellular loop1–3 (ECL1–3) or extend into the outer space of this domain (Liu S.Q. and Shi X.F. et al., submitted). In this paper, based on the CCR5 models and the crystal structure of gp120–CD4–17b [13], we try to build gp120–CD4–CCR5 models by docking and molecular dynamics, testing the two-step gp120–CCR5 binding mechanism proposed by us (Liu S.Q. and Shi X.F. et al., submitted) and other research groups [27] from the structural view.

Materials and methods

All computer simulations were performed on a Silicon Graphics Fuel workstation. Energy minimization and molecular dynamics were carried out with the commercial software package InsightII/Discover3 molecular simulation program version 2.98 (Accelrys Inc, San Diego, CA). The consistent valence force-field (CVFF) was used. To approximate solvation, calculations were carried out with a distance-dependent dielectric constant $1\times r$. The nonbonded part of the energy calculations was carried out employing a group-based summation method for van der Waals interactions with cut-off 13 Å, and for electrostatic interactions with cut-off 20 Å, spline width 1.0 Å and buffer width 1.0 Å, respectively. All hydrogen atoms were included in the calculations.

Docking gp120–CD4 complex onto models of CCR5

The high-resolution crystal structure of HIV gp120 envelope glycoprotein in ternary complex with the CD4 receptor and neutralizing human antibody 17b (PDB ID 1G9M) [13] was obtained from the PDB [28]. The antibody 17b was then removed from the ternary complex and only the complex of gp120–CD4 was retained as the ligand to dock to both CCR5 model B (PDB ID 1ND8), where the Nt region stays away from and locates at the top of the extracellular domain formed by ECL1–3, and model A (PDB ID 1NE0), where the Nt region adsorbs on the surface of the extracellular domain. The initial structures of gp120–CD4–CCR5 were generated by employing the simulated-annealing-docking (SA-Docking) method of InsightII/Affinity program, the total SA stage was set as 100 (100 fs per stage) and the system temperature was cooled from 1,000 K to 300 K in 10 ps. The extracellular domains of CCR5 were defined as the docking targets for gp120–CD4. The complex of gp120–CD4 was separately docked to two of the CCR5 models (models A and B) with different orientations of the Nt regions. Finally, two models of CCR5 in complex with gp120–CD4 were obtained. Further, the two structures were extensively energy minimized, and further high temperature (1,000 K) and long time (200 ps) molecular dynamics simulations were subsequently performed to search the possible interaction conformations of the coreceptor–gp120 complex more efficiently. Since the extracellular domains of CCR5 are the interface for gp120 binding, the structures of the transmembrane helices of CCR5 were fixed during the dynamics simulation, and the conformation of gp120–CD4 was also tethered throughout the calculation to maintain its original crystal structure. These procedures proved to be efficient in searching the interaction sites between CCR5 and gp120.

[2] N terminal/N terminus
[3] Extracellular loop
[4] Regulated on activation normal T cell expressed and secreted
[5] Macrophage inflammatory protein

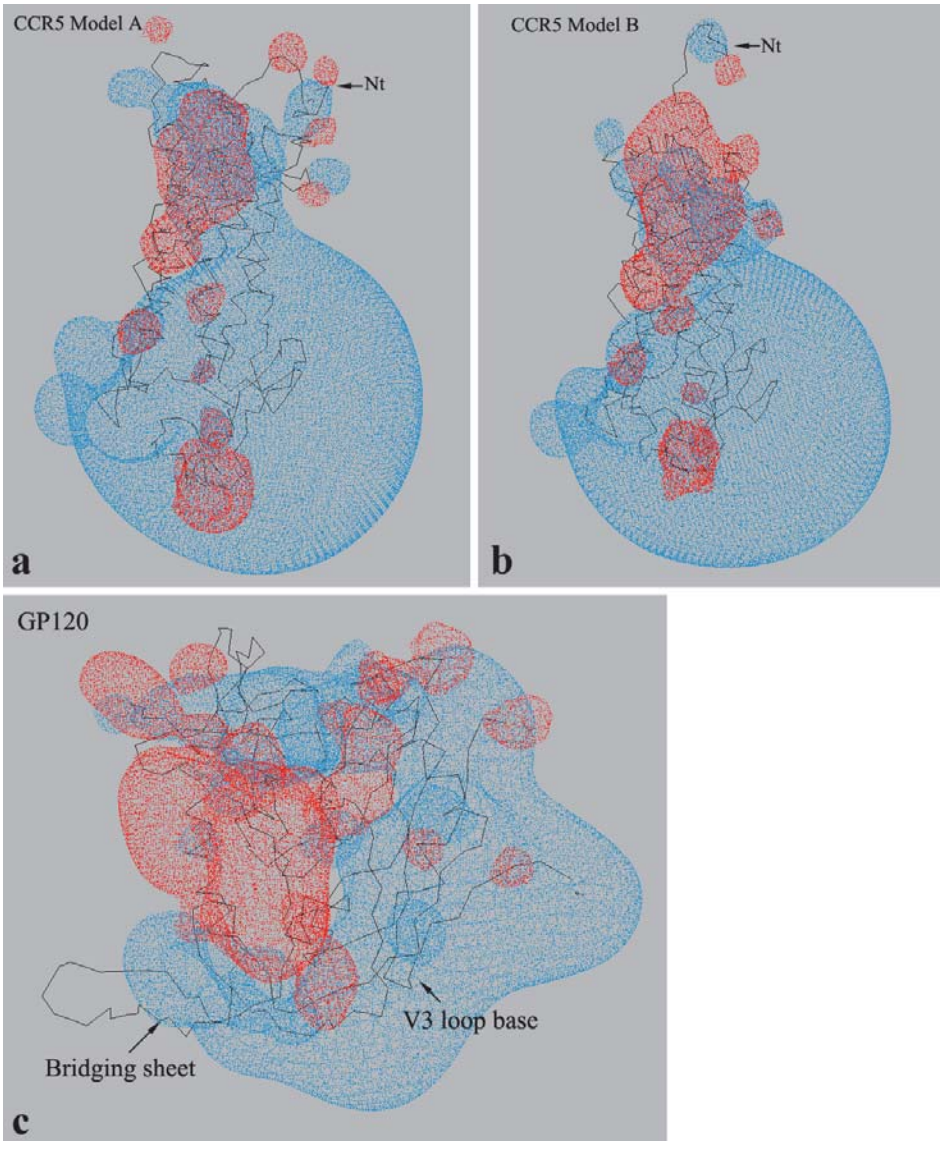

Fig. 1a–c The electrostatic potential map of the CCR5 and gp120. *Red* indicates −3 and *blue* +3.The extracellular domain of CCR5 dominated by strong negative potential (*red* color) is located towards the top of the diagram in a and b. **a** Electrostatic potential map of CCR5 model A with Nt adsorbing on the extracellular domain. **b** Electrostatic potential map of CCR5 model B with Nt protruding out from extracellular domain. **c** Electrostatic potential map of the gp120 crystal structure, the most negative region (*red*) is mainly located at the interfacial cavities for CD4 contact and binding, the regions of bridging sheet and theV3 loop base dominated by positive electrostatic potential is located at the diagram base (*blue*). These figures were generated by Swiss-pdbViewer Version 3.7 [29]

Result

Evaluation of electrostatic properties for CCR5 and gp120

CCR5 consists of 352 residues, out of which 37 are basic (R+K) and 17 are acidic (D+E). Thus, CCR5 is a basic protein and shows a strong net positive charge (+20) in neutral solution (all His are considered as neutral). Interestingly, the charged residues show polar distributions in the CCR5 molecule because there are seven negatively charged and seven positively charged groups in the extracellular domain, whereas 21 positively charged and six negatively charged groups are located in the cytoplasmic part. Electrostatic potential maps were calculated for the two models of CCR5 by Swiss-pdbViewer Version 3.7 [29], and are shown in Fig. 1a and b. The polar distributions of the charged groups can be seen on the electrostatic potential map: the red color indicating −3 dominates over the extracellular part with a few blue (indicating +3) balls caused mainly by K22, K26, R168 and K171 exposed on the red surface, this indicates the strong negative potential distributed in the extracellular domain. In contrast, a big blue ball surrounds the cytoplasmic part with only a few red balls scattered, indicating the strong positive potential there.

The core gp120 comprises 25 β-strands, 5 α-helices and 10 defined loop segments, and the polypeptide chain of gp120 is folded into two major domains: an inner domain and an outer domain, and a third element, the "bridging sheet" [13]. Out of 305 residues in the gp120 core, 33 are basic (R+K) and 25 are acidic (D+E), which indicates that the core of gp120 is a basic protein, too. An electrostatic potential map of core gp120 was also calculated and is shown in Fig. 1c, which reveals an uneven distribution of electrostatic potential on the surface of core gp120. The most negative center (red color) is located mainly at the interfacial cavities for CD4 contact and binding [13], while the regions of the bridging sheet, V3 loop base and other

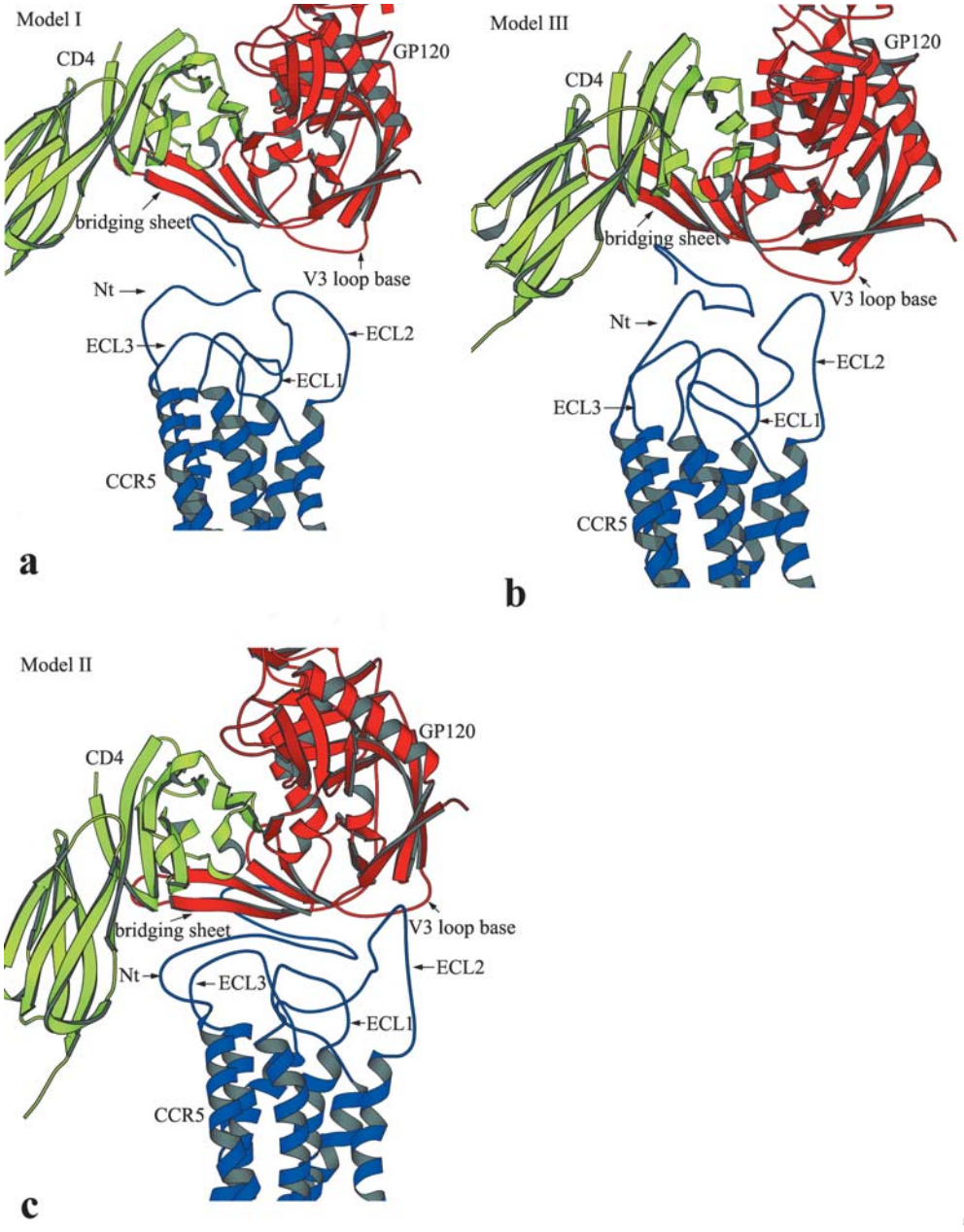

Fig. 2a–c Three structural models of CCR5 in complex with gp120–CD4, which represent snapshots of different intermediates of the complex that characterize the dynamic interaction process of CCR5 with gp120. The ribbon diagram shows gp120 in *red*, the two N-terminal domains of CD4 in *yellow*, and the CCR5 in *blue*. The Nt, ECL1, 2, and 3 in CCR5, as well as the bridging sheet and V3 loop base of gp120 are labeled. **a** Model I shows Nt as it begins to interact with bridging sheet, the ECLs of CCR5 make no interaction with gp120 and there is a big gap between them; it is a snapshot of the initial recognition of CCR5 with gp120. **b** Model III shows the ECL2 begins to interact with V3 loop base in addition to the interaction between Nt and bridging sheet; it is a snapshot of the transition state of the binding process. **c** Model II shows the full interaction between CCR5 and gp120, more residues within Nt interact with bridging sheet, ECL2 and ECL3 interact with V3 base loop and bridging sheet respectively; it is a snapshot a "fusion-active" state prior to virus entry. The interaction intensities between CCR5 and gp120 are model I<model III <model II. These figures were drawn by Molscript program [34]

parts were dominated by positive electrostatic potential (blue color), although a few small red balls (negative potential) are scattered there. In addition, the regions possessing of strong positive potential make up the binding sites for 17b [13]. According to the potential map of the gp120 core and the available experimental data [15], the negatively charged regions that are located at the bottom of gp120 (Fig. 1c) were deduced to be the interaction regions of gp120 with the extracellular domain of CCR5 for the following docking.

Structural models of gp120–CD4–CCR5

SA-docking separately generated two groups of structural models of gp120–CD4 in complex with CCR5 model A or B, and each group comprised 10 gp120–CD4–CCR5 complexes. Out of each group, one structure was selected according to two criteria, the interaction energy between coreceptor and virus gp120, and the available experimental data [13, 15, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26]. Concretely, for each group, five structures with the lower interaction energy were selected, then the interaction sites between CCR5 and gp120 were checked according to the available experimental data (for details, see the discussion below) and finally two complex models were obtained from each group. In both representative models, the gp120–CD4 is located at the top of the extracellular domain of CCR5. In gp120–CD4–CCR5 model I (Fig. 2a), the Nt region of CCR5 points at and has contacts with the bridging sheet region of gp120, and a

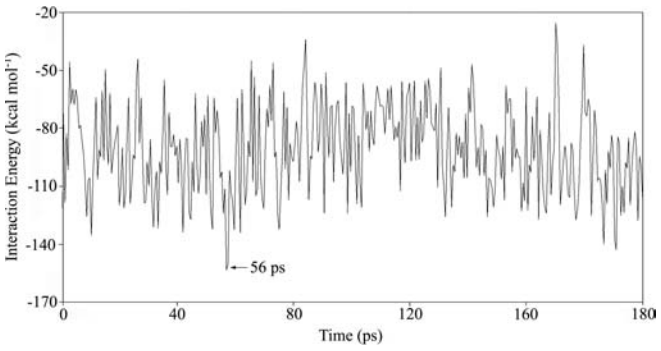

Fig. 3 Interaction energy between CCR5 and gp120 extracted from the 200 ps dynamics trajectory. Energy calculation was performed by using InsightII/Discover3. A series of structures were extracted from the 200-ps molecular dynamics simulation. The interaction energy was used as one of the criterions to selected possible model of the complex because a lower interaction energy may indicate a more favorable interaction between receptor and ligand. The conformer at simulation time 56 ps has the lowest interaction energy of about –152 kcal mol^{-1} between CCR5 and gp120 and is highlighted in this figure

large gap exists between the V3 loop base of gp120 and the tip of the ECLs of CCR5. However, in gp120–CD4–CCR5 model II (Fig. 2c), there is no large gap between gp120 and the extracellular domain of CCR5. The gp120–CD4 complex covers CCR5 just like a lip, the Nt and ECL3 of CCR5 are close to the bridging sheet of gp120, while the ECL2 is close to the V3 loop base.

Since the Nt and ECL2 are two of the most flexible parts among the extracellular segments of CCR5 (Liu S.Q. and Shi X.F. et al., submitted), their conformations should not remain invariable during the process of gp120 binding to CCR5. To simulate the dynamic binding process and investigate the general tendencies in behavior of the CCR5 extracellular domain interaction with core gp120, a 200-ps high temperature (1,000-K) molecular dynamics simulation was performed on the complex of gp120–CD4–CCR5. The initial structure for the molecular dynamics simulation was model II with the distance of the interfaces between CCR5 and gp120 enlarged to allow the CCR5 extracellular domain to search the possible interaction conformations more efficiently. 500 structures of the complex gp120–CD4–CCR5 were extracted from the molecular dynamics trajectory at a time interval of 400 fs. All structures were energy minimized, and the interaction energy between core gp120 and CCR5 was calculated for each structure (Fig. 3). These structures were clustered into about three conformational families based on the backbone structural similarity of Nt and ECL2. The representative structures for each family were analyzed according to available experimental data and interaction energy. One structure with the most favorable binding energy was finally selected as a possible model for the complex (Fig. 2b, model III). The structural organization of model III is similar to that of model II (Fig. 2c): the Nt region of CCR5 makes contact with bridging sheet of core gp120 and ECL2 direct at the V3 loop base. However, the difference between complex model II and III exists mainly in the structural orientation of Nt and ECL2 of CCR5, the Nt and ECL2 project from the extracellular domain to reach gp120 in model III, while in model II, the extracellular domain of CCR5 is comparatively packed, the Nt interact with gp120 without evident conformational extending. Thus, three models of CCR5 in complex with gp120–CD4 were obtained, they represent snapshots of different intermediates of coreceptor–gp120 that may characterize the dynamic interaction process of CCR5 with gp120.

Interaction analysis of CCR5 with gp120–CD4

For all these three complex models, the hydrogen bonding, electrostatic interaction, van der Waals contact, and interaction energy between extracellular domain of CCR5 and gp120 were investigated. The results are as follows. (i) Hydrogen bond formation between CCR5 and gp120 was observed in all of these models, but the regions and residues involved in hydrogen bond formation are different. For example, hydrogen bonds between Nt of CCR5 and the bridging sheet of gp120 could be found in all models, but those between ECL2 and V3 loop base could only be detected in model II and III. In model I, the residues involved in hydrogen bond formation are Asp2, Val5, Ser6 and Tyr10 in Nt of CCR5 and Thr202, Thr198, Lys432 and Met434 in bridging sheet of gp120. In model II, the residues involved in hydrogen bond formation are Asp2, Val5, Ser6, Ser7, Asp11, Gln21 in Nt, Tyr176 in ECL2 of CCR5, and Ser199, Val200, Thr202, Gln203, Cys205, Lys432 in bridging sheet and Ala299 in V3 loop of gp120. In model III, the hydrogen bond residues are Asp2, Tyr3, Asp11, Tyr14 in Nt of CCR5, Cys119, Thr202, Gln203, Lys207, Pro438 in the bridging sheet of gp120. (ii) All the complex models show that the basic gp120 surface (Fig. 1c) complements the acidic CCR5 surface (Fig. 1a and b), but only a few salt bridges are observed. In model I, no salt bridge was observed between CCR5 and gp120; in model II, there are two salt bridges: Asp2–Lys117and Glu18–Lys421, and in model III, only one salt bridge exists between Asp11–Lys207. (iii) Van der Waals contacts were measured with pairwise interresidual distances less than 2.5 Å. The Nt region has a few close contacts with the bridging sheet in all three models. ECL2 has no contacts with the V3 loop base in model I, few contacts with it in model III, and a few contacts in model II. Contacts were also observed between ECL3 and bridging sheet in model II. (iv) The interaction energy between each extracellular segment on CCR5 with gp120 was calculated for each of the three models, which revealed the relative contributions of different regions on CCR5 to gp120 binding. For model I, II and III, The interaction energies between gp120 and Nt are −80.1, −127.9 and −104.2 kcal mol^{-1}, between gp120 and ECL2 are −0.88, −63.0 and –46.1 kcal mol^{-1}, and ECL3 are −0.22, −18.0 and −0.28, respectively. The

Table 1 Potential interaction sites between CCR5 and gp120 as predicted from structural analysis of the models of CCR5-gp120-CD4 complex

Nt of CCR5	Asp2, Tyr3,Val5,Ser6, Tyr10, Asp11, Tyr14, Tyr15, Glu18, Gln21
ECL1 of CCR5	Not found
ECL2 of CCR5	His175, Tyr176, Ser180, Pro183
ECL3 of CCR5	Leu266, Ser270
Bridging sheet of gp120	Lys117,Lys121, Leu125, Ser199, Val200, Gln203, Thr202, Lys207, Lys432
V3 loop base of gp120	Pro417, Arg419, Ile420, Lys421, Gln422

ECL1 of CCR5 contributes almost no interaction energy to gp120 binding in all three models. These results reveal that the Nt and ECL2 are important regions involved in gp120 binding, which is in agreement with available experimental data [17, 18, 19, 20, 21, 22, 23, 24, 25, 26]. In addition, the loss in solvent-accessible surface (SAS) area was calculated for each residue in the three models. In model I, residues with loss more than 40% in SAS area upon the formation of the gp120–CCR5 complex are Asp2, tyr3, Val5 and Pro8 in Nt of CCR5 and Lys121 and Ile201 in bridging sheet of gp120; In model II, residues with substantial SAS area loss are Asp2, Tyr3, Val5, Ser7, Tyr10, Asp11, Tyr14 and Glu18 in Nt, His175 and Tyr176 in ECL2, Lys117, Lys121, Leu125, Ser199, Thr202, Lys207 in bridging sheet, and Pro417, Arg419, Lys421 and Gln422 near the base of V3 loop. In model III, such residues are Asp2, Tyr3, Val5, Asp11and Tyr14 in Nt, His175 in ECL2, Lys117, Thr202, Cys205 in bridging sheet and Arg419 near the base of the V3 loop. These residues were selected as the potential binding sites. Together with the interaction information of hydrogen bond, van der Waals contact and electrostatic interaction between CCR5 and gp120, the residues that are most likely to be involved in interaction between CCR5 and gp120 in the complex were finally obtained (Table 1).

Discussion

In our latest study [Liu S.Q. and Shi X.F. et al., submitted], two models of CCR5 were constructed by means of molecular modeling and molecular dynamics. The molecular dynamics simulation for the extracellular domain of CCR5 revealed that when two disulfide bonds Cys20–269 and Cys101–178 were retained, there were primarily two conformational states in Nt region. In one state, the Nt region projects from and stays on the top of the extracellular domain (activated state). In the other state the Nt adsorbs on the surface of the ECLs (ground state). Integrating the feature of Nt conformational motion with available experimental data, a two-step gp120–CCR5 binding mechanism was proposed: first, the Nt region of CCR5 projects from the extracellular domain and adopts an appropriate orientation being ready to recognize envelope glycoprotein. Second, the binding of gp120 to Nt induces the conformational change of Nt, which makes gp120 interact further with ECLs of CCR5, resulting in virus entry. Here we further test this mechanism through the approaches of SA docking and molecular dynamics simulation. The gp120–CD4 structure used for docking was extracted from the X-ray crystal structure of an HIV-1 gp120 core in ternary complex with CD4 and the Fab fragment of the CD4 induced antibody 17b. The gp120 core has deletions of 52 and 19 residues from the N and C termini, respectively, Gly–Ala–Gly tripeptide substitutions for 67 V1/V2 loop residues and 32 V3 loop residues, and the removal of all sugar groups beyond the linkages between the two core *N*-acetylglucosamine residues [13, 14]. Despite these modifications, the capacity of core gp120 to interact with CD4 and antibodies against CD4 induced epitopes is preserved at or near wild-type levels [30], so the core gp120 retains structural integrity [14], and can be used for generating structural model in complex with CCR5. Three complex models of gp120–CD4–CCR5 in different interaction intermediates were built to explore the dynamic nature of gp120 binding to CCR5. In the absence of an experimentally determined structure, these models provide alternative templates for understanding the dynamic process of virus entry and structural details of the coreceptor–virus glycoprotein interaction. As discussed later, the models of CCR5 complex in gp120–CD4 proposed here are also useful both for explaining available data and for matching the interaction sites in CCR5 with those in gp120.

Electrostatic analysis reveals the dipole characteristics of CCR5 and core gp120. For CCR5, the negative pole is located in the extracellular domain, but a strong positive pole is found in the cytoplasmic domain (Fig. 1a and b). For gp120, the strong negative region flanks the left side of the molecule and covers the interface for CD4 binding (Fig. 1c), while the basic region is located at the right side and the bottom of the molecule, and the bottom part faces away from the virus envelope and toward the target cell membrane (Fig. 1c). Electrostatic forces are long-range effective forces; we therefore guess that the extracellular acidic region of CCR5 may play a role in the initial attraction and interaction with the basic region of gp120. On the basis of electrostatic potential maps, the probable binding interface between CCR5 and gp120 for docking was established. Actually, our docking results revealed that the acidic Nt region of CCR5 is crucial in gp120 binding because it makes contact with the bridging sheet region in all the three complex models, which is consistent with experimental data [22, 31].

It has been shown in studies using CCR5 chimeras that multiple extracellular regions of CCR5 are involved in HIV-1 binding and entry [18, 32]. However, among a large number of point mutations in these regions of CCR5, only a few were found to impair gp120 binding significantly [22, 23, 24]. For examples, previous exper-

iments using chimeras constructed between human CCR5 and either murine CCR5 or CCR2b indicated the important functional role of the Nt of CCR5 in HIV-1 binding [18, 33]. Recently, studies using site-directed mutagenesis [22], truncation and deletion [31] of residues within the Nt region revealed that the charged aspartic acid residues at positions 2 and 11 (D2A and D11A) and a glutamic acid residue at position 18 (E18A), as well as aromatic residues Tyr3, Tyr10, Tyr14 and Tyr15 are involved in CCR5 coreceptor function. These results are in general consistent with our structural models, which also indicate the importance of these acidic residues and tyrosine-rich site (Table 1). On the other hand, in gp120, the site of interaction with CCR5 induced by CD4 binding involves the highly conserved residues of gp120, mutational analysis showed that the basic or polar gp120 residues near or within the bridging sheet and the V3 loop are important for CCR5 binding [15]. Specifically, in our three complex models of gp120–CD4–CCR5, the Nt of CCR5 is close to or lies a cross the base of the four-stranded bridging sheet and the Nt of each models has interaction with some basic or polar residues within the four strands of gp120 (Table 1). However, the interaction intensities between Nt and gp120 are different among the three models. In complex model I (Fig. 2a), only a few direct interatomic contacts are observed between four residues of CCR5 Nt region and five residues of the gp120 bridging sheet, including 12 van der Waals contacts and four hydrogen bonds. In complex model II (Fig. 2c), six residues within the Nt are involved in direct interatomic contacts with nine residues of bridging sheet. The number of van der Waals contacts and hydrogen bonds is 14 and eight, respectively. In complex model III (Fig. 2b), the number of residues involved in direct contacts between CCR5 and gp120 is five in Nt and six in bridging sheet. These results indicate that models I and II have the weakest and the strongest interaction intensity between CCR5 Nt and gp120 bridging sheet, respectively, while model III has moderate interaction intensity.

Using monoclonal antibody (mAbs) competition binding experiments, Lee B et al. [26] found ECL2-specific mAbs were more efficient in blocking chemokine binding than those mAbs directed to other extracellular segments, and Nt mAbs blocked gp120–CCR5 binding more effectively than ECL2 mAbs, but surprisingly, ECL2 mAbs were more potent inhibitors of viral infection than Nt mAbs. These results imply that the binding sites of chemokines and gp120 on CCR5 are distinct but overlapping, and suggest that Nt is more important for gp120 binding while the ECL2 is more important for inducing conformational changes in envelope glycoprotein that lead to membrane fusion and virus entry. Interestingly, in our modeling studies on gp120–CD4–CCR5, only ECL2 and ECL3 can make contacts with gp120 among the three extracellular loops of CCR5 ECL1-3 in model II, and ECL2 contributes more coreceptor–gp120 interaction energy than ECL3, while ECL1 contributes almost no interaction energy and none of the ECL1 residues is involved in direct contact with gp120 in all three models. These results suggested that, in addition to Nt, the ECL2 also participates in interaction with gp120. Like the interaction between Nt and gp120, the interaction intensities between ECL2 and gp120 are also different among the three complex models. In model I (Fig. 1a), although the Nt contacts with bridging sheet, ECL2 makes no direct contact with gp120 and a large gap exists between ECL2 and gp120. In model II (Fig. 2c), interatomic contacts are observed between two CCR5 residues in ECL2 and three gp120 residues near or within the V3 loop base. In model III (Fig. 2b), ECL2 is close to the region around the V3 loop base but few van der Waals contacts and hydrogen bonds are observed between them. These results indicate that the initial recognition of CCR5 with gp120 involved no interactions between ECL2 and gp120. Gradually, the interaction between them occurred and was enhanced by conformational changes of Nt and ECL2. Furthermore, the solvent-accessible area excluded upon interaction is 334, 516, 421 $Å^2$ from CCR5 and 344, 530, 428 $Å^2$ from gp120 for model I, II and III, respectively, which come mainly from the Nt (334, 425, and 375 $Å^2$ for model I, II and III, respectively) and ECL2 (0, 79, 46 for model I, II and III, respectively) in CCR5, and the regions near or within bridging sheet (344, 441, 382 $Å^2$) and V3 loop base in gp120 (0, 89, 46 $Å^2$), indicating the formation of large interface are mainly between Nt and the bridging sheet, as well as between ECL2 and the V3 loop base. Models II displayed the most significant loss in SAS area that appeared to be consistent with its most favored coreceptor–gp120 interaction energy. The loss in SAS area for these three models is also consistent with their interaction intensities between CCR5 and gp120: model I<model III<model II.

To summarize, we have built structural models of CCR5 in complex with gp120 and CD4 through the approaches of protein structure modeling, docking and molecular dynamics simulation. Three models provide snapshots that represent different intermediates of CCR5–gp120 interaction: the first intermediate represented by model I, in which the Nt of CCR5 recognizes and binds to the bridging sheet of gp120, is the initial state of recognition of CCR5 with gp120. The second intermediate represented by model III, in which the ECL2 of CCR5 begins to interact with the V3 loop base of gp120, may be a snapshot of the transition state of the binding process. The third intermediate represented by model II, in which CCR5 interacts fully with gp120, including the full interactions between Nt and the bridging sheet, ECL2 and the V3 loop base, and a few contacts between ECL3 and bridging sheet, probably represents a “fusion-active” state prior to virus entry. The three snapshots of the gp120–CD4–CCR5 ternary complex described here reveal many molecular aspects of HIV-1 entry, including the atomic resolution structure of CCR5, the sites on CCR5 and gp120 involved in interaction with each other, and the dynamic interaction process of CCR5 with gp120. On the basis of these models, the infection process of HIV-1 can be proposed as the following: the binding of CD4 to gp120 induces a conformational change in gp120 and

makes a full formation or exposure of the binding site for CCR5. The extended acidic Nt region in CCR5 first interacts with the basic bridging sheet in gp120, this interaction induces conformational changes of Nt and leads the V3 loop region of gp120 close to ECL2. The further interactions between ECL2 and the V3 loop base make conformational changes in gp120 that alter the relation between gp120 and gp41, resulting in the exposure of the gp41 ectodomain and interaction with the target cell membrane, which leads to membrane fusion and ultimate virus entry. It is worth pointing out that the electrostatic interactions between CCR5 and gp120 are important for the initial recognition of Nt with gp120 and the inducement of conformational changes in CCR5 and gp120. In the absence of experimental determined 3D structural complex of gp120–CD4–CCR5, our complex models explained the structural basis of available biological data about CCR5 and gp120, and support the two-step gp120–CCR5 binding mechanism. The separate interaction of the two sites (Nt and ECL2) in CCR5 with gp120 may assist the development of selective inhibitors to intervene in virus–coreceptor interactions.

PDB access IDs

1OPN is for the complex structure of gp120–CD4–CCR5 model I, 1OPW for model II, and 1OPT for model III.

Acknowledgements This work was supported in part by the National Natural Science Grant (China) (No. 19947006) and partially by the National Key Foundation Research Grant (No. 2002AA231031) in China (863)

References

1. Dalgleish AG, Beverley PC, Clapham PR, Crawford DH, Greaves MF, Weiss RA (1984) Nature 312:763–766
2. Klatzmann D, Champagne E, Chamaret S, Gruest J, Guetard D, Hercend T, Gluckman JC, Montagnier L (1984) Nature 312:767–768
3. Feng Y, Broder CC, Kennedy PE, Berger EA (1996) Science 272:872–877
4. Dragic T, Litwin V, Allaway GP, Martin SR, Huang Y, Nagashima KA, Cayanan C, Maddon PJ, Koup RA, Moore JP, Paxton WA (1996) Nature 381:667–673
5. Deng H, Liu R, Ellmeier W, Choe S, Unutmaz D, Burkhart M, Marzio PD, Marmon S, Sutton RE, Hill CM, Davis CB, Pepier SC, Schall TJ, Littman DR, Landau NR (1996) Nature 381:661–666
6. Alkhatib G, Combadiere C, Broder CC, Feng Y, Kennedy PE, Murphy PM (1996) Science 272:1955–1958
7. Strader CD, Fong TM, Tota MR, Underwood D (1994) Annu Rev Biochem 63:101–132
8. Dean M, Carrington M, Winkler C, Huttley GA, Smith MW, Allikmets R, Goedert JJ, Buchbinder SP, Vittinghoff E, Gomperts E, Donfield S, Vlahov D, Kaslow R, Saah A, Rinaldo C, Detels R, O'Brien SJ (1996) Science 273:1856–1862
9. Huang Y, Paxton WA, Wolinsky SM, Neumann AU, Zhang L, He T, Kang S, Ceradini D, Jin Z, Yazdanbakhsh K, Kunstman K, Erickson D, Dragon E, Landau NR, Phair J, Ho DD, Koup RA (1996) Nat Med 2:1240–1243
10. Samson M, Libert F, Doranz BJ, Rucker J, Liesnard C, Farber CM, Saragosti S, Lapoumeroulie C, Cognaux J, Forceille C, Muyldermans G, Verhofstede C, Burtonboy G, Georges M, Imai T, Rana S, Yi Y, Smyth RJ, Collman RG, Doms RW, Vassart G, Parmentier M (1996) Nature 382:722–725
11. Wu L, Gerard NP, Wyatt R, Choe H, Parolin C, Ruffing A, Borsetti A, Cardoso AA, Desjardin E, Newman W, Gerard C, Sodorski J (1996) Nature 384:179–183
12. Trkola A, Dragic T, Arthos J, Binley JM, Olson WC, Allaway GP, Cheng-Mayer C, Robinson J, Maddon PJ, Moore JP (1996) Nature 384:184–186
13. Kwong PD, Wyatt R, Robinson J, Sweet RW, Sodroski J, Hendrickson WA (1998) Nature 393:648–659
14. Wyatt R, Kwong PD, Desjardins E, Sweet RW, Robinson J, Hendrickson WA, Sodroski JG (1998) Nature 393:705–711
15. Rizzuto CD, Wyatt R, Hernandez-Ramos N, Sun Y, Kwong PD, Hendrickson WA, Sodroski JA (1998) Science 280:1949–1953
16. Walker JE, Saraste M (1996) Curr Opin Struct Biol 6:457–459
17. Gosling J, Monteclaro FS, Atchison RE, Arai H, Tsou CL, Goldsmith MA, Charo IF (1997) Proc Natl Acad Sci USA 94:5061–5066
18. Atchison RE, Gosling J, Monteclaro FS, Franci C, Digilio L, Charo IF, Goldsmith MA (1996) Science 274:1924–1926
19. Picard L, Simmons G, Power CA, Meyer A, Weiss RA, Clapham PR (1997) J Virol 71:5003–5011
20. Alkhatib G, Ahuja SS, Light D, Mummidi S, Berger EA, Ahuja SK (1997) J Biol Chem 272:19771–19776
21. Samson M, LaRosa G, Libert F, Paindavoine P, Detheux M, Vassart G (1997) J Biol Chem 272:24934–24941
22. Dragic T, Trkola A, Lin SW, Nagashima KA, Kajumo F, Zhao L, Olson WC, Wu L, Mackay CR, Allaway GP, Sakmar TP, Moore JP, Maddon PJ (1998) J Virol 72:279–285
23. Rabut GE, Konner JA, Kajumo F, Moore JP, Dragic T (1998) J Virol 72:3464–3468
24. Farzan M, Choe H, Vaca L, Martin K, Sun Y, Desjardins E, Ruffing N, Wu L, Wyatt R, Gerard N, Gerard C, Sodroski J (1998) J Virol 72:1160–1164
25. Navenot JM, Wang ZX, Trent JO, Murray JL, Hu QX, DeLeeuw L, Moore PS, Chang Y, Peiper SC (2001) J Mol Biol 313:1181–1193
26. Lee B, Sharron M, Blanpain C, Doranz BJ, Vakili J, Setoh P, Berg E, Liu G, Guy HR, Durell SR, Parmentier M, Chang CN, Price K, Tsang M, Doms RW (1999) J Biol Chem 274:9617–9626
27. Genoud S, Kajumo F, Guo Y, Thompson D, Dragic T (1999) J Virol 73:1645–1648
28. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE (2000) Nucleic Acids Research 28:235–242
29. Guex N, Peitsch MC (1997) Electrophoresis 18:2714–2723
30. Binley JM, Wyatt R, Desjardins E, Kwong PD, Hendrickson W, Moore JP, Sodroski J (1998) AIDS Res Hum Retrovir 14:191–198
31. Blanpain B, Doranz BJ, Vakili J, Rucker J, Govaerts C, Baikc SSW, Lorthioir O, Migeotte I, Libert F, Baleux F, Vassart G, Doms RW, Parmentier M (1999) J Biol Chem 274:34719–34727
32. Doranz BJ, Lu ZH, Rucker J, Zhang TY, Sharron M, Cen YH, Wang ZX, Guo HH, Du JG, Accavitti MA, Doms RW, Peiper SC (1997) J Virol 71:6305–6314
33. Rucker J, Samson M, Doranz BJ, Libert F, Berson JF, Yi Y, Smyth RJ, Collman RG, Broder CC, Vassart G, Doms RW, Parmentier M (1996) Cell 87:437–446
34. Kraulis PJ (1991) J Appl Crystallogr 24:946–950