Modeling the Structural Basis of Human CCR5 Chemokine Receptor Function: From Homology Model Building and Molecular Dynamics Validation to Agonist and Antagonist Docking

A. Fano,[†] D. W. Ritchie,[‡] and A. Carrieri*,[†]

Dipartimento Farmaco-Chimico, Università degli Studi di Bari, Via Orabona 4, 70125, Bari, Italy, and Department of Computing Science, University of Aberdeen, Aberdeen AB24 3UE, Scotland, United Kingdom

Received November 10, 2005

This article describes the construction and validation of a three-dimensional model of the human CCR5 receptor using a homology-based approach starting from the X-ray structure of the bovine rhodopsin receptor. The reliability of the model is assessed through molecular dynamics and docking simulations using both natural agonists and a synthetic antagonist. Some important structural and functional features of the receptor cavity and the extracellular loops are identified, in agreement with data available from site-directed mutagenesis. The results of this study help to explain the structural basis for the recognition, activation, and inhibition processes of CCR5 and may provide fresh insights for the design of HIV-1 entry blockers.

INTRODUCTION

Acquired immune deficiency syndrome (AIDS) has become a deadly global disease. Current therapies, based on the combined use of viral protease and reverse transcriptase inhibitors, can reduce mortality rates in infected people but are still not able to eliminate the virus from the organism and hence prevent the progression and associated debilitating aspects of the disease. Therefore, there is an ongoing need for novel therapeutics which can prevent the entry of human immunodeficiency virus (HIV-1) into its target cells.

HIV viral entry is mediated by specific interactions between the gp120 viral envelope glycoprotein and plasmatic membrane receptors of the target cell.[1] These cause conformational changes in both the glycoprotein and the membrane receptors, which in turn leads to virus−cell fusion. Several previous studies have elucidated the role of CD4, whose interaction with gp120 is essential but not sufficient for the disease to become established.[2,3] Besides CD4, some chemokine receptors (CCRs), members of the G-protein-coupled-receptor superfamily (GPCRs), have recently been identified as cotargets necessary for viral entry into the cell.[4] CCRs and their natural chemokine ligands, RANTES, MIP-1α, MIP-1β, and MCP2, regulate signaling among immune cells and hence represent potential target systems for preventing virus−cell fusion. Different CCRs have previously been identified and characterized.[5] The involvement of one of them depends on the viral strain involved, but the majority of strains (the R5 isolates) are sexually transmitted and act on CCR5 during the entire course of the disease.[6] Conversely, the variant strains (the X4 and R5X4 isolates), which act on CXCR4, or on both CXCR4 and CCR5, cause a loss of T-cells during the progression of the disease but are less involved initially.[7] Besides this, other evidence[8] shows that the transition from the R5 to X4 strains is also suppressed in vivo. The importance of CCR5 in the transmission of HIV-1 can also be appreciated from further experimental evidence,[9] whereby individuals who are homozygous for the Δ32 CCR5 allele, which causes an inoperative CCR5, are resistant to HIV infection but are otherwise generally healthy. This suggests that functional inhibition of CCR5 could help protect against infection without provoking damage to patients and, therefore, that blocking viral entry using small-molecule antagonists selective for this receptor might provide a new and more effective type of anti-HIV drug.

Unfortunately, no experimentally determined three-dimensional (3D) CCR structure is currently available. Therefore, the main ways to investigate the properties of CCR5 and its interaction with chemokines are currently based on site-directed mutagenesis (SDM) or molecular modeling techniques, as reported in previous studies on CCR/chemokine interactions[10−12] and other GPCR systems.[13] Although homology modeling of CCR5 has been described previously,[14−17] we decided to consolidate the information available from this earlier work into a comprehensive and up-to-date model by incorporating extensive molecular dynamics simulations (MD), flexible docking of a synthetic antagonist TAK779, and soft protein−protein docking with the large (~70 KD) natural agonists RANTES and MIP-1β using a novel docking protocol. We validated our homology-built CCR5 structural model through exhaustive conformational sampling, and we assessed its quality and reliability by analyzing four low-energy conformations with respect to the available SDM data and by considering the side-chain interactions that occur at important regions of the receptor throughout the MD trajectory. Docking calculations using these conformations give predicted binding modes which are consistent with the experimental SDM data in each case. The results of this study help to explain the structural basis for the recognition, activation, and inhibition processes of CCR5 and may provide fresh insights for the design of HIV-1 entry blockers.

MATERIALS AND METHODS

CCR5 Model Building. Our model of human CCR5 was built using the first crystal structure of bovine rhodopsin to

* Corresponding author tel.: +39 080 544 2638; fax: +39 080 544 2230; e-mail: carrieri@farmchim.uniba.it.
[†] Università degli Studi di Bari.
[‡] University of Aberdeen.

10.1021/ci050490k CCC: $33.50
Published on Web 03/18/2006

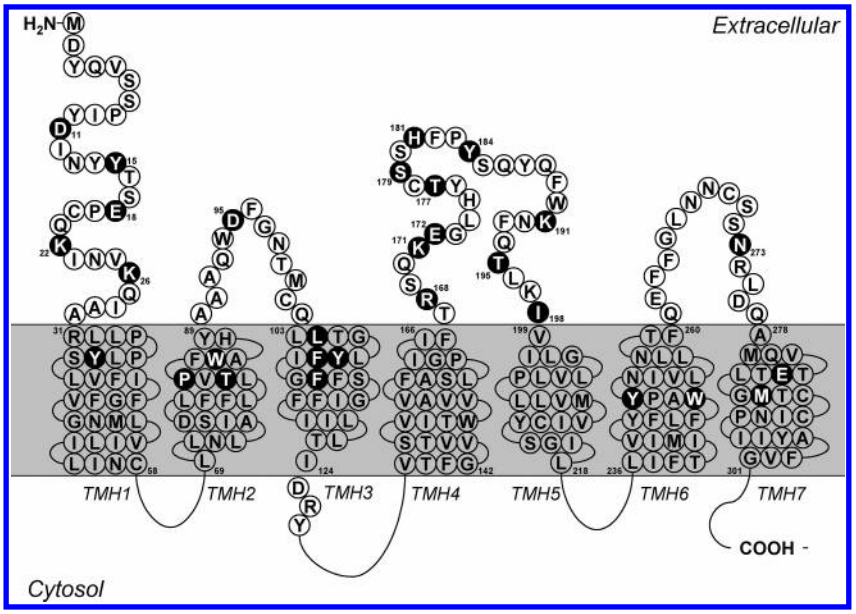

Figure 1. Illustration of the human CCR5 sequence (N-terminal, extracellular; C-terminal, intracellular) showing the disposition of secondary structure features with respect to the cell membrane (gray). Black-filled circles correspond to key residues mentioned in the text.

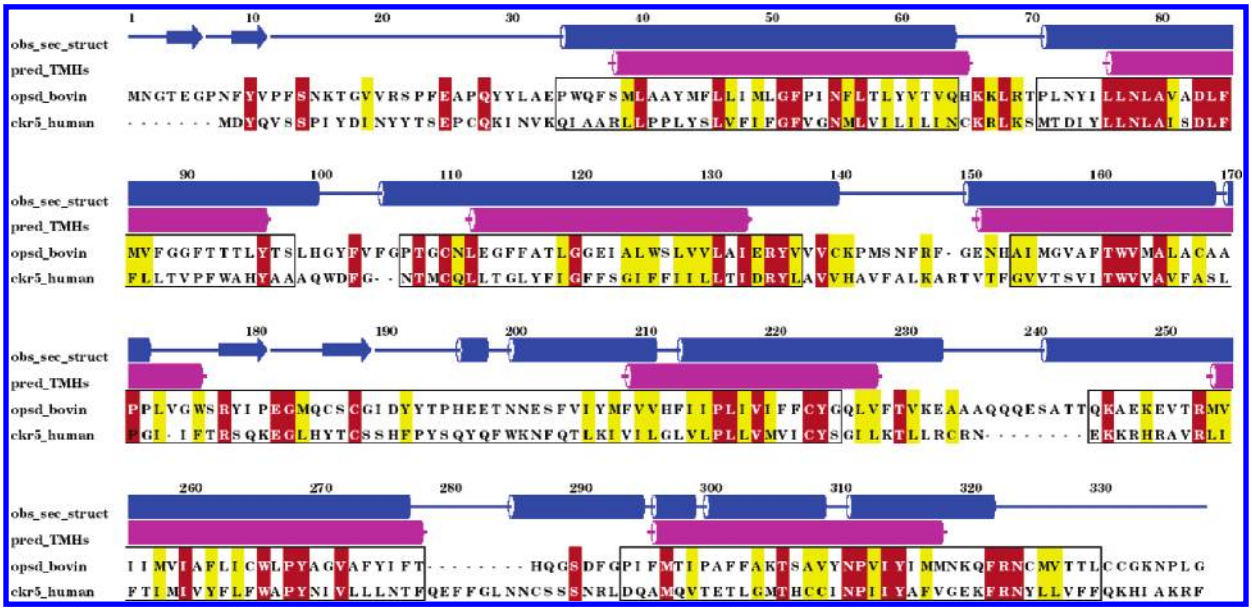

Figure 2. Pairwise alignment of the bovine rhodopsin and human CCR5 sequences, calculated using the PAM350 matrix in CLUSTALW with a gap-open penalty of 10.0 and a gap-extension penalty of 0.1; secondary structure gap penalties were also applied.

become available. This 2.8 Å resolution structure, solved by Palczewski et al.,[18] consists of a transmembrane (TM) domain of seven α helices connected by three extracellular loops (ECLs) and three intracellular loops (ICLs). Despite the low sequence identity between the two proteins (<20%), bovine rhodopsin is currently the best available template for homology modeling CCR5. However, the sequence identity increases to ∼30% when considering only the transmembrane helices (TMHs), and several of the amino acid residues essential for maintaining CCR5's architecture and receptor function are highly conserved (for a full review, see ref 19). Hence, the structure of Palczewski et al. provides a feasible and indeed appropriate template for modeling GPCRs.

Here, residues are numbered according to their positions in the human CCR5 sequence retrieved from the SwissProt database (CKR5_HUMAN).[20] The *transmem* annotation of this sequence file was used when defining the residues of the seven TMHs. Applying the PHD-htm[21] secondary structure prediction algorithm to the sequence gave the same secondary structure elements as those of the bovine rhodopsin X-ray structure in all seven α helices, and also in ECL2, which contains a small antiparallel β sheet. In this extended loop, the sequence homology between CCR5 and the template is somewhat higher (12 residues out of 31) than that in the remaining nonconserved structural regions. Figure 1 shows the disposition of these secondary structure elements with respect to the cell membrane, and Figure 2 shows the CLUSTALW[22] sequence alignment of CCR5 and the rhodopsin template.

The CCR5 sequence contains four cysteines which form two disulfide links, Cys20−Cys269 and Cys101−Cys178, the latter being highly conserved in all GPCRs and in the rhodopsin template as well. Because the Cys101−Cys178 disulfide bridge also involves the β sheet of ECL2, we

decided to build that loop by homology from the template structure. MODELLER[23] 6.2 was used to assemble the initial CCR5 framework for residues 1−320. The final 32 C-terminal residues were not modeled because these were absent in the template. The Cα Cartesian coordinates of the seven TMHs and ECL2 were copied from the corresponding template atoms (PDB code 1F88, chain A) according to the sequence alignment shown in Figure 2, while the N-terminal domain and the remaining loops were built de novo using MODELLER's *loop* subroutine. Side-chain atom coordinates were added from MODELLER's standard residues library.

To construct a preliminary input structure, a set of 500 3D conformations were generated using MODELLER's *refine_2* fast simulated annealing procedure. To ensure full coverage of possible CCR5 conformations in this step, all side-chain rotamers were fully randomized. However, the loop search included a spatial restraint in order to maintain the second disulfide bridge between Cys20 (N-terminal) and Cys269 (ECL3), which is known to be necessary for chemokine binding.[24] All generated conformations were then checked using PROCHECK.[25] The best conformation according to the fraction of residues in the core region of the Ramachandran plot was selected and remodeled, this time using the *loop* subroutine for the N-terminal fragment and the three ECLs. This gave 100 new conformations, of which the best PROCHECK conformation was selected as the input structure for MD refinement.

MD Loop Refinement. To find reliable low-energy CCR5 conformations, MD was performed on the above starting structure using a previously validated protocol.[26] All MD calculations used the all-atom force field of Weiner et al.,[27] as implemented in AMBER 6.0, using a distance cutoff for all nonbonded interactions of 10 Å and a distance-dependent dielectric factor of 4*r*. Our MD protocol repeatedly applies alternate cycles of low (300 K) and high (600 K) temperature trajectories until no new low-energy conformations are found. Here, the initial structure was stepwise minimized using 5000 cycles of steepest descent followed by 5000 cycles of conjugate gradient using a force constant of 100 kcal/mol Å^2 on all Cα atoms. This was followed by four minimization steps each of 2500 cycles of conjugate gradient in which the force constant was set to 50, 25, 10, and 5 kcal/mol Å^2. The N- and C-terminal domains and all loops were then relaxed using a final round of conjugate gradient minimization until the root-mean-square (RMS) slope fell below 0.001 kcal/mol Å. During this last step, all TMH core residues were restrained by a 5 kcal/mol Å^2 force constant, except in the first and last turns of each helix, in which the restraint was set to 3 kcal/mol Å^2. The resulting structure showed the same helix kinks around proline and glycine residues as observed in the template structure. Although the above procedure might seem unusual, it has been shown that the use of a simple membrane mimic during GPCR modeling is not sufficient to explore the receptor movements during the activation process.[28,29] The specific constraints used here take into account more explicitly the TM location of our receptor. Moreover, it has been proved that the in vacuo approach gives comparable results with similar computations using the more expensive IMM1 implicit membrane/water model.[30]

To explore more fully the conformational space accessible to the minimized structure obtained above, we then applied the SHAKE algorithm[31] to all bonds in a 2000 ps MD simulation at 300 K, updating the nonbonded interaction list every 25 fs. Snapshots of the MD trajectory were taken and minimized every 2 fs to give an initial set of 1000 sample conformations. These were then filtered by rejecting any conformation more than 25 kcal/mol above the global minimum and clustered by first superimposing all TMH Cα atoms and then clustering the ECL Cα coordinates using NMRCLUST.[32] No predefined RMS threshold was used during this procedure. The lowest-energy conformation from each of nine structurally different clusters was then selected and submitted to a 100 ps MD simulation at 600 K in which conformations were sampled and minimized every 1 fs. This produced a second set of 9 × 100 new conformations, and the above clustering procedure was repeated on the combined set. No new low-energy minima were found, presumably because of the presence of the two disulfide bridges, which limit the flexibility of the extracellular domain.

To enhance the quality of the 3D models achieved thus far and to explore the feasibility of favorable side-chain interactions occurring between residues identified from SDM as essential for receptor activity, the minimum energy conformation was submitted to a 3000 ps MD run at 300 K. This time, the core of each TMH (residues 35−54 for TMH1, 73−85 for TMH2, 107−120 for TMH3, 146−162 for TMH4, 203−214 for TMH5, 240−256 for TMH6, and 282−297 for TMH7) was gently constrained by applying a force constant of 0.1 kcal/mol Å^2 while the rest of the protein was allowed to move. Conformations were sampled and minimized every 10 fs after a 400 ps equilibration period. The resulting 260 conformations were then clustered into 11 groups. Considering an energy threshold of 5 kcal/mol above the global minima, four low-energy conformations were identified. Three of these (models 1, 2, and 4) were part of a 22-member cluster including the lowest-energy structure. The fourth (model 3) was part of a different cluster of 32 members. All other conformations were discarded because of their high energies (>20 kcal/mol) above the global minimum. Because the energy differences between the selected four conformations are small ($\Delta E_{max} = 3.39$ kcal/mol), we consider them to be equally plausible and energetically accessible conformations. As can be seen in Figure 3, the main secondary structure features as assessed by STRIDE[33] are essentially conserved in the four models, not only in the TMHs, as might have been expected, but also in the loop regions. Ramachandran plots of these models show that very few residues (<1.0%) have disallowed backbone conformations and almost 80% of residues fall within the allowed dihedral angle ranges.

Molecular Docking. Once a theoretical model has been built, one of the most appropriate methods to explore its predicted structural features is through docking simulations with selected ligands. Unfortunately, this task is not always easy, especially when the receptor has more than one binding site or when the molecules to be docked are large. Previously, we succeeded in docking small endogenous ligands to the binding sites of other GPCRs.[26,34] However, the current case was somewhat more challenging, particularly for docking the large agonists and, therefore, two different approaches were used. Although previous docking experiments with the small synthetic antagonist TAK779 have been described,[14] we decided to perform a similar flexible ligand binding

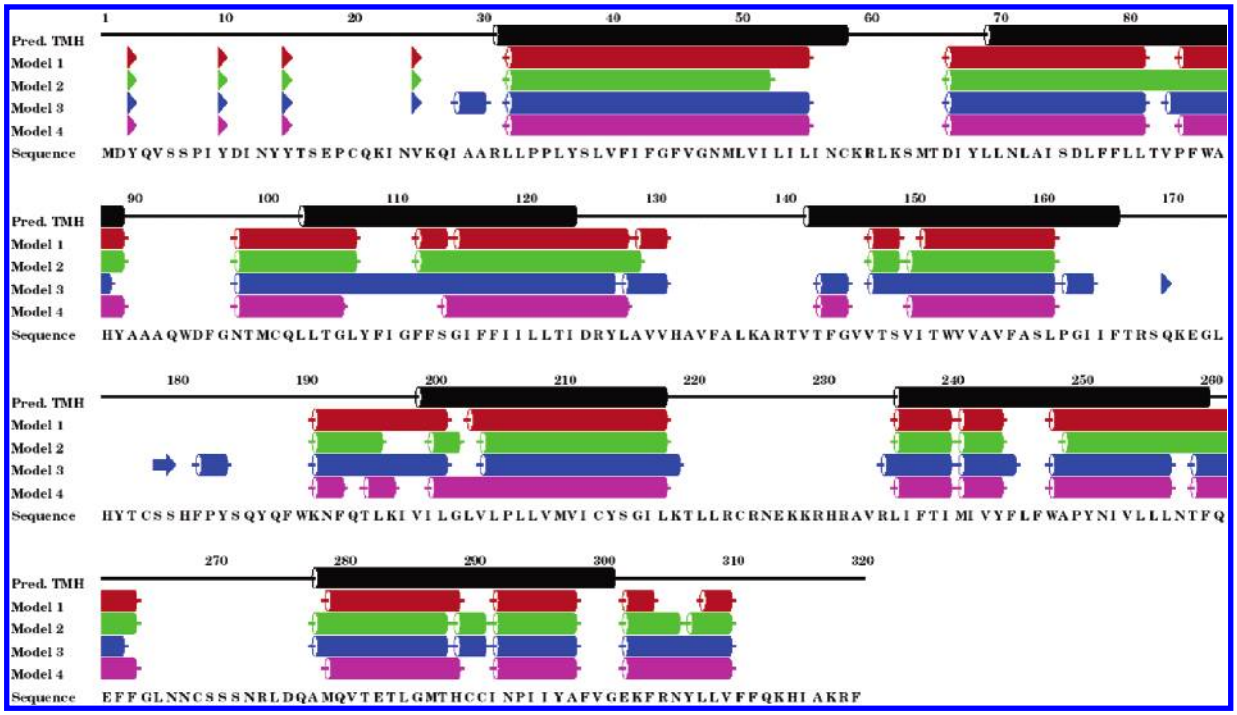

Figure 3. Secondary structure motifs as calculated by STRIDE for the four lowest-energy CCR5 conformations.

calculation here using our more exhaustively sampled CCR5 MD structures.

The 3D structure of TAK779 was built using QXP[35] using standard bond lengths and valences. A 3000-step Monte Carlo simulation followed by a Polak−Ribiere conjugate gradient minimization using the mixed AMBER/MM2 force field was carried out in order to select a low-energy starting conformation for the docking run. The ligand was initially placed in the receptor binding site as determined by alanine-scanning mutagenesis,[36] located near the CCR5 extracellular surface in the TM cavity formed from residues Tyr37 (TMH1), Thr82, Trp86 (THM2), Tyr108 (TMH3), and Glu283 (TMH7). TAK779 was then docked to the binding site using QXP's DYNDOCK module using 500 steps of flexible dynamic docking with the CCR5 Cα trace held fixed and using a single distance constraint between the phenoxy hydrogen of Tyr37 and pyrane oxygen of TAK779.

For chemokine docking, we used a modified version of Hex[37] called GridHex[38] to soft dock the natural protein agonists. The GridHex scoring function aims to identify favorable chemical interactions between pairs of protein surfaces by correlating molecular interaction fields (MIFs) from GRID,[39] without requiring that the proteins fit together perfectly. This approach was used recently in rounds 4, 5, and 7 of the CAPRI blind docking experiment.[40,41] However, because there is no direct way to determine which of the several types of MIF provide the driving force for binding in any particular complex, we use a principal component analysis (PCA) technique to select the most significant MIFs to use for each specific complex, as described below. In the current study, we started from the NMR average structures taken from the Protein Data Bank (PDB)[42]—RANTES, PDB code 1RTO, chain A;[43] MIP-1β, PDB code 1HUM, chain A.[44] These structures were first minimized in AMBER using 5000 steps of steepest descent, followed by 5000 steps of conjugate gradient with the Cα trace kept fixed with a 1000 kcal/mol force constant. This was followed by full minimization until the RMS slope fell below 0.001 kcal/mol Å.

Table 1. GRID Probe Types Used in the PCA Analysis

GRID probes	atom
C3	methyl CH_3 group
C1=	sp^2 CH aromatic or vinyl
N3+	sp^3 amine NH_3 cation
N1	neutral flat NH
N:=	sp^2 N with lone pair
O::	sp^2 carboxy oxygen atom
O1	alkyl hydroxy OH group
O	sp^2 carbonyl oxygen
S1	neutral SH group
DRY	hydrophobic probe

Prior to docking, GRID MIFs were calculated on a 0.8 Å grid for each protein using each of the 10 probe moieties listed in Table 1. These probes represent the main types of physicochemical (i.e., steric, electrostatic, hydrophobic, and hydrogen bond) interaction that occur in protein complexes. The MIF arrays were then imported into GOLPE[45] and pretreated by truncating any positive potentials to zero in order to ignore repulsive interactions. To focus the analysis on the chemokine binding site, all CCR5 data beyond a radius of 18 Å from Lys171 (taken as the center of mass of the ECLs) was discarded, as was any data having only one value (i.e., presenting a skewed distribution). GOLPE's PCA facility was then used to decompose the MIF data into two small matrices of "loadings" and "scores." The loadings give the weights of the original variables in the analysis, and the scores give the contribution each probe makes to the new uncorrelated variables, or principal components (PCs). Plotting the probe scores against the PCs allows those probes that explain most of the variance to be identified.

Figure 4 shows the score plots for the first two principle components for each protein. These plots show that the probes may be grouped into essentially three main types:

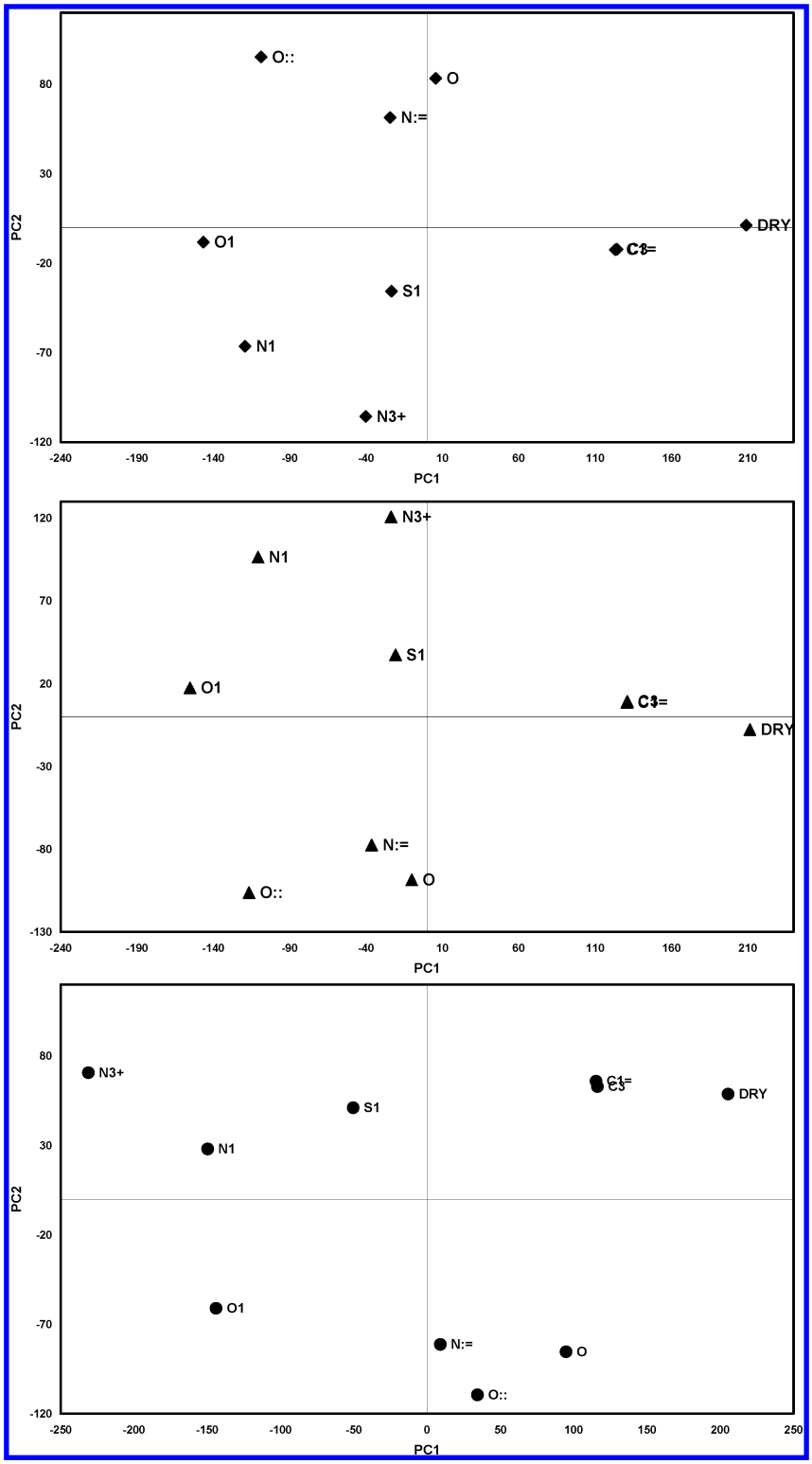

Figure 4. GRID probe PCA plots using the probe types listed in Table 1 for CCR5 (top, diamonds), RANTES (middle, triangles), and MIP-1β (bottom, circles).

the positively charged hydrogen-bond-donating group, the negatively charged hydrogen-bond-accepting group, and the hydrophobic group. It is well-known in PCA that the greater the difference in the scores, the better is the ability to discriminate the properties of the objects under study. In the current case, the PCA shows that the N3+, O::, and DRY probes are the most discriminating. Hence, the MIFs corresponding to these three probes were selected and used in the GridHex docking correlation. The chemical contribution to the correlation energy is calculated as

$$E_{\mathrm{MIF}} = \frac{1}{2}\int [\tau_{\mathrm{A}}(\phi_{\mathrm{B}}^{\mathrm{N3+}} + \phi_{\mathrm{B}}^{\mathrm{O::}} + \phi_{\mathrm{B}}^{\mathrm{DRY}}) + \tau_{\mathrm{B}}(\phi_{\mathrm{A}}^{\mathrm{N3+}} + \phi_{\mathrm{A}}^{\mathrm{O::}} + \phi_{\mathrm{A}}^{\mathrm{DRY}})]\, \mathrm{d}V$$

where the functions φ_{A} and φ_{B} represent the above MIFs

Table 2. Summary of SDM Data Relating to Human CCR5[a]

residue	location	effect of mutation
Asp11	N-terminal	strong reduction in functional response; no binding of RANTES and MIP-1α
Tyr15		reduced binding of MIP-1α
Glu18		strong reduction in chemokine bindingand of functional response
Lys22		reduced binding of MIP-1α
Lys26		reduced binding of MIP-1α
Tyr37	TMH1	strong reduction in binding of receptor antagonists
Pro84	TMH2	reduced functional response; moderate reduction in antagonist binding
Trp86		strong lowering of the binding of receptor antagonists
Asp95	ECL1	regulation of the binding of MIP-1α
Leu104	TMH3	essential for receptor activation
Tyr108		strong reduction in binding of receptor antagonists
Phe109		essential for receptor activation
Phe112		essential for receptor activation
Arg168		reduction in binding of MIP-1α but not RANTES
Lys171		no binding to mAB 2D7
Glu172		reduction in binding of MIP-1α but not RANTES; no binding to mAB 2D7
Ser179	ECL2	altered coreceptor function; no virus-cell fusion; no infection
His181		no binding of RANTES and MIP-1α
Tyr184		no binding of RANTES and MIP-1α; no coreceptor function
Lys191		strong reduction in binding of RANTES and MIP-1α
Thr195		no coreceptor function
Ile198		moderate reduction in binding of receptor antagonists
Tyr251	TMH6	moderate reduction in binding of receptor antagonists
Glu283	TMH7	moderate reduction in binding of receptor antagonists
Met287		essential for the binding of MIP-1β

[a] See main text for references.

and τ_A and τ_B represent the steric density functions of the receptor and ligand, respectively.

In the docking step, RANTES and MIPS-1β were docked onto each of the four low-energy CCR5 conformations using a six-dimensional rotational−translational search with one angular constraint (receptor $\beta_1 \leq 45°$), which allows the ligand to spin freely while loosely constraining it to remain near the receptor binding site. The calculated MIF energy and basic Hex shape scores were combined in the ratio of 25:75 to give an overall pseudo-energy for each pose.

All energy minimizations, MD simulations, and GridHex dockings were run on a Pentium IV 2.56 GHz computer, whereas the QXP dockings were run on a SGI Origin300 R14000 server.

RESULTS AND DISCUSSION

Following the exhaustive conformational sampling of our homology-built CCR5 model structure, we assessed its quality and reliability by analyzing the generated 3D structures with respect to the available SDM data listed in Table 2 and by examining side-chain interactions during the MD trajectory and in the ligand-binding sites.

Analysis of Low-Energy MD Conformations. Some important common structural features were observed in the low-energy CCR5 conformations. For example, all four structures exhibit similar hydrogen-bonding patterns in the TM cavity close to Tyr108, Glu283, and Tyr251. As previously reported,[36] these residues are essential to confer CCR5 its biological role of HIV coreceptor. In our model 1, the phenoxy group of Tyr108 is hydrogen-bonded to the Glu283 carboxy group and also interacts in a face-to-face π-stacking arrangement with the aromatic ring of Tyr251. Simultaneously, the Glu283 carboxy group hydrogen-bonds the side chains of Thr177 and Ser179 in ECL2. It has been shown previously[46] that this loop contains some of the residues required for efficient binding to the natural agonists. It has also been suggested that Ser179 might face a negatively charged residue,[47] as observed here.

Further important structural aspects of the TM cavity can be inferred by considering the combined mechanism by which hydrogen bonds are made and broken on going between our low-energy conformations. For example, the hydrogen bond between Glu283 and Tyr108 is easily broken in model 1 and, after an appropriate side-chain flip (see models 2 and 3 in Figure 5), only the interactions between Glu283 and the ECL2 residues Thr177 and Ser179 are retained. We hypothesize that this conformation of ECL2 might correspond to an early activated state during ligand binding. Similar transitions leading to signal transduction across the intracellular domain have been described in other GPCR models.[26,48] From these observations, the importance of Glu283 in stabilizing the receptor into a low-energy conformation by tightly packing TMH3, TMH6, TMH7, and ECL2 seems overwhelming. On the other hand, Glu283 has also been identified as a hot-spot residue in the binding of positively charged antagonists.[36] To explore this further, we mapped our model of this binding site using a charged sp^3 nitrogen GRID probe. This produced a favorable energy profile around Glu283, which supports the charge stabilization hypothesis (data not shown).

Near the above positively charged TM cavity, an inner cavity can be identified comprising the polar and hydrophobic residues Tyr37 (TMH1), Thr82, Val83, Pro84, Trp86, Tyr89 (TMH2), Leu104 (TMH3), and Thr284 and Met287 (TMH7). Among these residues, Thr82, Val83, and Pro84 resemble the typical TXP motif, which is highly conserved within the CCR family. The role of this motif seems to be to induce a strong deformation of the TMH2 helical axis, which is kinked toward TMH1. It is known that the P84A mutation causes a decreased binding affinity for chemokines

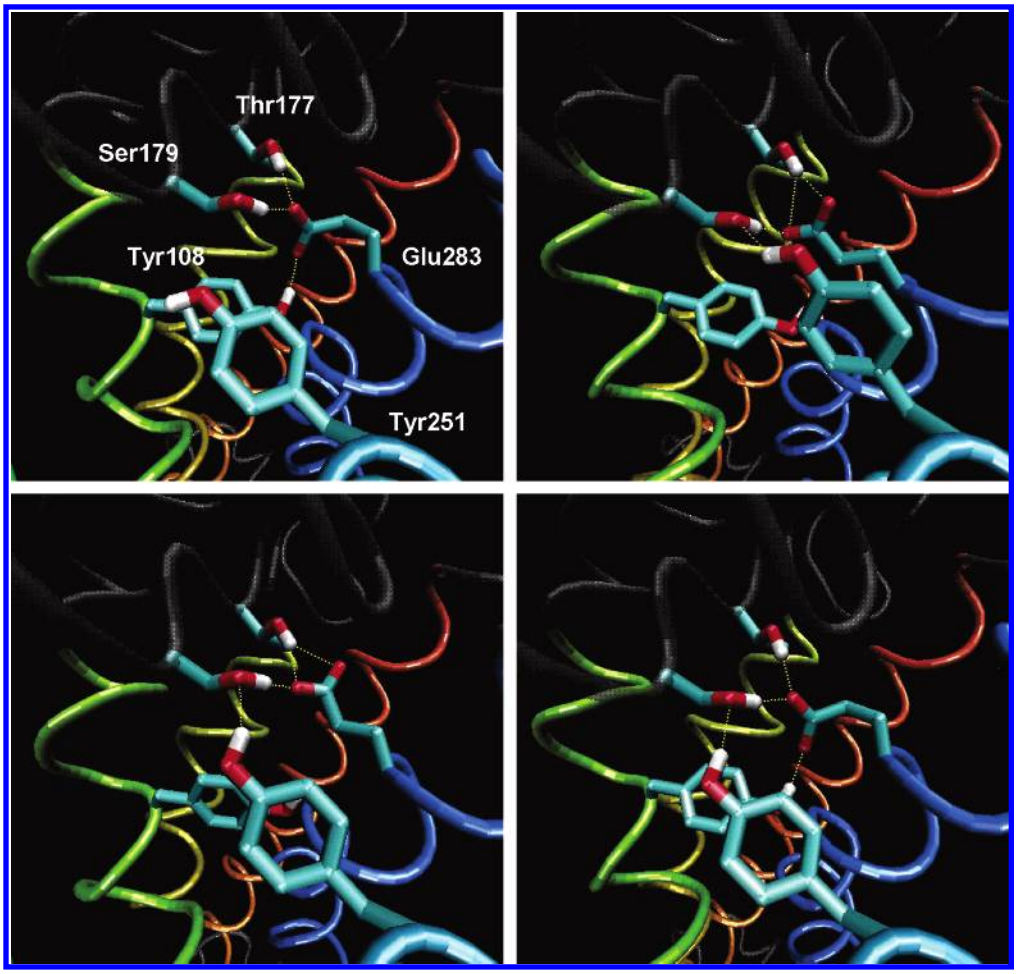

Figure 5. Side-chain conformations in the TM binding site of the four calculated low-energy CCR5 conformations (model 1, top left; model 2, top right; model 3, bottom left; model 4, bottom right).

and nearly abolishes the functional response of the receptor.[49] In our model, this cavity is lined with the polar/hydrophobic residues Try37, Thr82, Val83, and Pro84. Hence, any mutation of Pro84 would dramatically alter the shape of this cavity and thus explain the loss of affinity mentioned above.

Just below the inner cavity is an aromatic pocket formed from the highly conserved residues Phe109, Phe112, and Phe113 (TMH3), as well as Trp248 (TMH6). These residues have been previously suggested[50] to be part of an aromatic cluster governing the mechanism of action in CCRs. Our model suggests that this cluster might be activated by small structural modifications to the hydrogen-bonding pattern described above or by the binding of a small lipophilic molecule, for example.

Analysis of Key SDM Residues. One challenging aspect of the current work was building the extracellular regions of CCR5 sufficiently well to model its interaction with the natural chemokine ligands. For example, the role of the N-terminal domain of CCR5 as a primary, perhaps nonspecific, binding site has been demonstrated previously for several agonists, and mutating to alanine any of the key residues Asp11, Tyr15, Glu18, Lys22, or Lys26 is known to hamper or eliminate chemokine binding (especially RANTES and MIP-1α).[51−56] Significantly, in all of our low-energy MD conformations, the charged or polar functional groups of these residues are exposed to the solvent and are involved in a hydrogen-bond network or π-cation interactions (see Figure 6). It can be postulated that the above hydrogen-bonding network is essential for maintaining this part of the receptor in a stable structural motif and for substrate recognition.

Other notable interactions are also observed in the modeled ECLs, for example, the Asp95−Lys171 salt bridge and the hydrogen bonds between His181, Lys191, and Gln186, as well as those between Glu172, Tyr184, and Asn273. Because the ECLs, especially ECL2, are important and specific chemokine recognition sites, these residues are crucial for characterizing the shape and potential of the extracellular surface. In our models, the ECL residues seem to be arranged in such a way to account for the available SDM evidence. From the solvent accessibility data shown in Table 3, it can be seen that some polar residues of ECL2, primarily Thr177 and Ser179, are completely buried in the entry channel of the TM cavity. As discussed above, these residues form part of an important hydrogen-bonding network. It therefore seems reasonable to suppose that the binding of CCR5 antagonists within the TM domain could disrupt key ECL2 interactions and hence explain how some antagonists can allosterically prevent chomokine binding, as in the case of TAK779.

MD Trajectory Population Analysis. During the 3000 ps MD run, side-chain movements were monitored along the trajectory and the time evolution of the positions of structurally essential residues was analyzed. First, we considered the making and breaking of hydrogen bonds between those residues buried within the α-helical bundle that define the

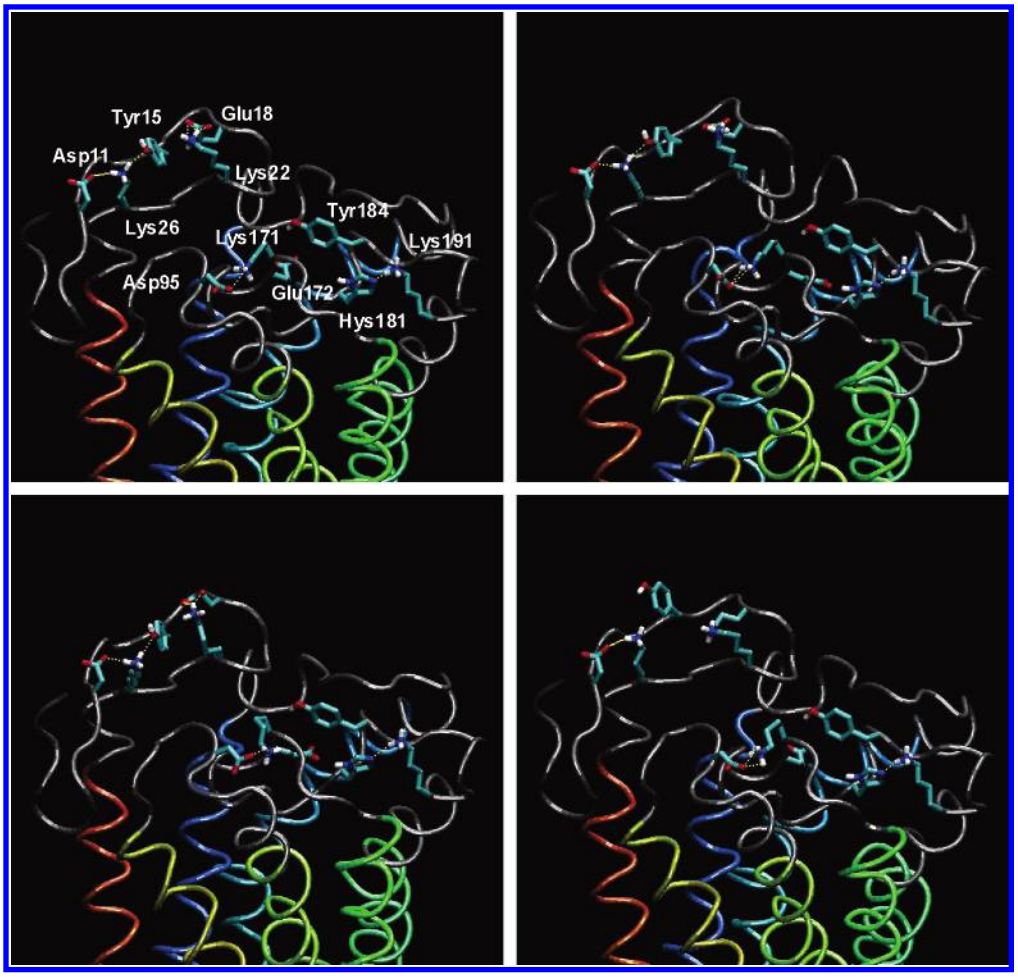

Figure 6. ECL side-chain conformations in the four calculated low CCR5 energy conformations (model 1, top left; model 2, top right; model 3, bottom left; model 4, bottom right).

positively charged TM binding site. From the hydrogen-bond population occupancies listed in Table 4, it can be seen that Glu283 is almost permanently trapped in polar interactions with nearby residues. In our low-energy models, the two Glu283 carboxy oxygens interact mainly with the TMH channel hydroxyl groups of Thr177, Ser179, and Tyr251 but also to a lesser extent with that of Tyr108. Tyr108 also binds to the Glu283 backbone carbonyl and interacts with Tyr251 largely through edge-to-face or face-to-face π-stacking arrangements (not shown). In agreement with the work of Dragic et al.,[36] this supports the evidence to suggest that Tyr108, which is located just between the positively charged binding site and the aromatic cluster identified above, is a key residue for CCR5 receptor function.

A significant interaction is also seen between the indole ring of Trp86 and the Asp95 backbone atoms. In the 3000 ps MD run, the triptophan NH group is hydrogen-bonded to the aspartate carbonyl group for around half of the simulation time. It might be postulated that this polar interaction is needed to maintain the structural motif of the ECL2 and the overall 3D shape of the antagonist binding site. Experimental support for this hypothesis is given by the W86A mutant, which is know to permit HIV-1 entry but is almost insensitive to TAK779 inhibition.[36] Hydrogen bonds also occur at the CCR5 extracellular surface between the side chains of Asp11 and Lys26 and those of Glu18 and Lys22. These pairs of residues are exposed to the solvent and are separated by the Tyr15 phenoxy group, which also hydrogen-bonds Lys26 (see Figure 6). As discussed further below, these residues might represent a first anchoring point for the endogenous agonists.

Other putative chemokine binding sites can also be identified around some charged or aromatic residues of the extracellular loops. For example, the 3000 ps MD trajectory shows that Asp95 and Lys171 form a permanent salt bridge, which might eventually be disrupted once the mostly basic chemokine epitope residues bind to the receptor. Similarly, hydrogen-bonding interactions can be seen between His181 and Lys191 and between Glu172 and Tyr184. This polar network might be essential for stabilizing the ECLs in their bioactive conformation. Indeed, mutations here generally disrupt chemokine binding,[57] probably because of the consequent misfolding of this part of the receptor surface.

Antagonist Docking. TAK779, synthesized by Takeda (see Figure 7),[58] is the first known antagonist with a specific and high affinity for CCR5 (K_i = 10 nm). It has been shown that this molecule inhibits HIV-1 replication by blocking the binding of viral gp120 to CCR5, thus preventing virus−cell fusion. TAK779 also inhibits natural ligand binding. Binding studies of TAK779 to CCR5 mutants have shown that it binds within the CCR5 TM helices.[36] Using QXP, we successfully docked TAK779 into this TM binding site, as shown in Figure 8. Significantly, TAK779 adopts an L-like shape in the global minimum energy conformation and binds to CCR5 largely through polar and hydrophobic interactions. The lipohilic methylphenylbenzocycloheptenyl part of the

Table 3. Average Solvent Accessibilities[a] (%) of the N-Terminal and ECL Residues

Nter	abs.[b]	rel.[c]	ECL1	abs.[b]	rel.[c]	ECL2	abs.[b]	rel.[c]	ECL3	abs.[b]	rel.[c]
Met1	140.13	72.20	Ala90	0.00	0.00	Thr167	0.00	0.00	Gln261	15.44	8.65
Asp2	73.37	52.25	Ala91	32.90	30.50	Arg168	83.41	34.95	Glu262	2.70	1.58
Tyr3	5.06	2.40	Ala92	18.66	17.30	Ser169	12.25	10.50	Phe263	114.75	57.55
Gln4	133.68	74.93	Gln93	0.00	0.00	Gln170	73.48	41.15	Phe264	138.4	69.40
Val5	145.88	96.33	Trp94	95.77	38.40	Lys171	8.52	4.28	Gly265	14.42	18.00
Ser6	13.22	11.35	Asp95	32.44	23.13	Glu172	0.00	0.00	Leu266	68.15	38.15
Ser7	7.17	6.13	Phe96	38.28	19.20	Gly173	0.00	0.00	Asn267	89.54	62.23
Pro8	68.29	50.18	Gly97	30.78	38.43	Leu174	0.00	0.00	Asn268	52.92	36.75
Ile9	112.02	63.98	Asn98	65.93	45.80	Hid175	0.00	0.00	Cys269	6.32	4.73
Tyr10	25.63	12.05	Thr99	132.23	94.95	Tyr176	0.00	0.00	Ser270	24.31	20.85
Asp11	35.49	25.25	Met100	75.88	39.08	Thr177	0.00	0.00	Ser271	64.51	55.4
Ile12	59.79	34.13	Cys101	0.00	0.00	Cys178	0.00	0.00	Ser272	16.98	14.58
Asn13	81.28	56.45	Gln102	19.93	11.15	Ser179	0.24	0.20	Asn273	0.00	0.00
Tyr14	85.27	40.08				Ser180	0.29	0.25	Arg274	1.78	0.75
Tyr15	72.37	34.00				Hid181	4.23	2.33	Leu275	73.68	41.25
Thr16	99.19	71.20				Phe182	0.80	0.40	Asp276	94.39	67.23
Ser17	100.45	86.25				Pro183	0.00	0.00	Gln277	60.27	33.78
Glu18	41.18	23.93				Tyr184	1.47	0.70			
Pro19	89.46	65.70				Ser185	0.00	0.00			
Cys20	13.45	10.00				Gln186	45.53	25.50			
Gln21	72.80	40.78				Tyr187	114.03	53.63			
Lys22	70.92	35.30				Gln188	125.47	70.30			
Ile23	7.13	4.05				Phe189	144.36	72.38			
Asn24	11.03	7.68				Trp190	136.41	54.73			
Val25	11.52	7.60				Lys191	11.81	5.88			
Lys26	26.30	13.13				Asn192	6.68	4.65			
Gln27	32.10	17.98				Phe193	81.21	40.70			
Ile28	123.79	70.70				Gln194	35.50	19.88			
Ala29	45.46	42.10				Thr195	0.00	0.00			
Ala30	0.00	0.00				Leu196	47.84	26.78			
Arg31	1.59	0.65				Lys197	26.43	13.15			
						Ile198	2.01	1.15			

[a] Calculated by NACCESS[61] using a probe radius of 1.4 Å. [b] Absolute solvent accessibility in Å^2. [c] Relative solvent accessibility, calculated relative to the accessibility of the corresponding residue when placed in an extended ALA−X−ALA tripeptide.

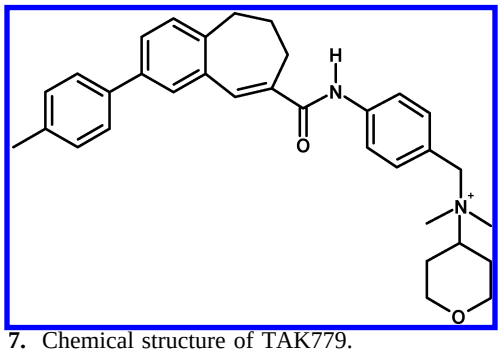

Figure 7. Chemical structure of TAK779.

ligand is deeply buried in the TM barrel and makes extensive van der Waals contacts with Phe109, Phe112, Phe113 (TMH3), Ile198 (TMH5), and Trp248 (TMH6). The polar ammonium and pyran ring groups of the ligand stabilize binding through a salt bridge with the Glu283 carboxy group and a charge-reinforced hydrogen bond with the Tyr37 hydroxyl group, respectively. As a consequence of this binding mode, the TAK779 anilido benzene ring makes a face-to-face aromatic stacking contact with the Tyr108 phenoxy group. These interactions all contribute favorably to our calculated binding energy of −28.71 kcal/mol. A similar binding mode was also achieved using the AUTOGROUP module of the GRID software (data not shown).

Agonist Docking. If identifying a consistent binding mode is difficult for small ligands, a stronger effort has to be made when docking macromolecules because of the complexity of the protein−protein interaction process and the generally greater approximations used in current algorithms. To help avoid making false-positive predictions, we once again used experimental evidence to guide the calculations. For example, previous SDM binding studies using monoclonal antibodies

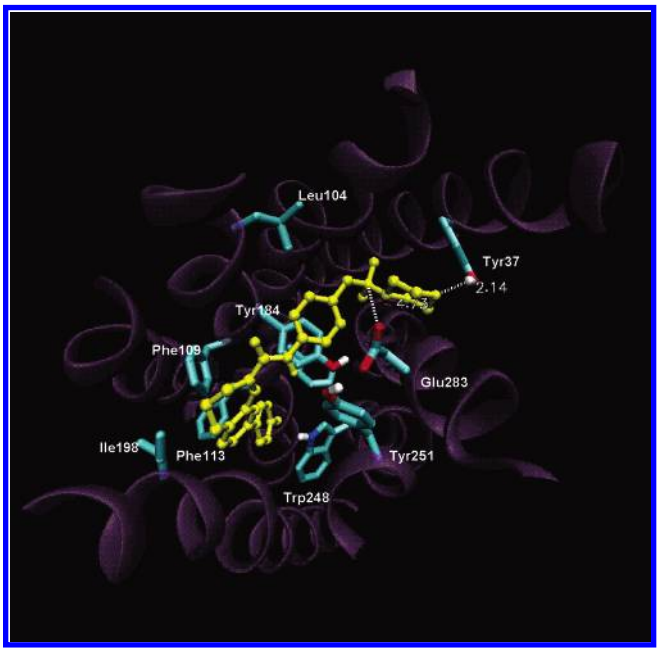

Figure 8. TAK779 (yellow) docked into the CCR5 TM binding site (mauve ribbon). Key CCR5 residues (cyan) are annotated.

suggest that CCR5 interacts with chemokines through an exposed extracellular structural motif comprising part of the N-terminal domain and ECL2.[59] In particular, it has been shown for human CCR5 that the negatively charged hydrophobic residues Asp11, Glu18, and Asp95 are critical in chemokine binding.[52−53,60] For example, alanine scanning of the amino-terminal domain shows that binding of RANTES is completely blocked if Asp11 and Glu18 are mutated.[51] Other charged residues such as Asp95 (ECL1), Arg168,

Table 4. Hydrogen-Bond Population Occupancy (%) in the 3000 ps MD Simulation

	Lys22			Lys26			Tyr37	Trp86	Tyr108	Lys171		
	Hζ1	Hζ2	Hζ3	Hζ1	Hζ2	Hζ3	Hη	Hε1	Hη	Hζ1	Hζ2	Hζ3
Asp11 Oδ1				20.67	18.33	32.53						
Asp11 Oδ2				5.33	17.13	3.47						
Tyr15 Oη	0.60	0.53	0.80	17.20	24.07	14.73						
Glu18 Oδ1	21.73	19.53	23.87									
Glu18 Oδ2	22.33	19.27	23.60									
Tyr37 OH												
Asp95 Oδ1								5.07		14.20	11.20	9.80
Asp95 Oδ2										24.07	25.27	23.13
Tyr108 Oη												
Glu172 Oε1												
Glu172 Oε1												
Thr177 Oγ1									0.13			
Ser179 Oγ									0.80			
His181 Nε2												
Tyr184 Oη												
Tyr251 OH									4.47			
Ser270 Oγ												
Gln273 Oδ1												
Gln283 Oε1							0.07		8.53			
Gln283 Oε2									11.87			
Thr284 Oγ1							30.60					

	Thr177	Ser179	Tyr184	Lys191			Tyr251	Ser270	Gln273		Thr284
	Hγ1	Hγ	Hη	Hζ1	Hζ2	Hζ3	Hη	Hγ	Hδ1	Hδ2	Hγ1
Asp11 Oδ1											
Asp11 Oδ2											
Tyr15 Oη											
Glu18 Oδ1											
Glu18 Oδ2											
Tyr37 OH											29.40
Asp95 Oδ1											
Asp95 Oδ2											
Tyr108 Oη	0.27	0.60					6.60				
Glu172 Oε1			20.00					15.47		32.67	
Glu172 Oε1			5.87					19.67		19.53	
Thr177 Oγ1		8.47									
Ser179 Oγ	19.60						14.87				
His181 Nε2				23.80	29.33	20.67					
Tyr184 Oη									2.80		
Tyr251 OH		7.80									
Ser270 Oγ											
Gln273 Oδ1			6.73								
Gln283 Oε1	59.80	46.67					32.23				
Gln283 Oε2	60.40	46.60					28.33				
Thr284 Oγ1											

Lys171, and Lys191 (ECL2) are also implicated in chemokine binding. Indeed, Arg168 is important for binding all agonists except RANTES. Further SDM results show that chemokines also interact with CCR5 through the basic and hydrophobic residues Phe12, Arg17, Arg44, and Lys45 (and also Arg47 in the case of RANTES). Hence, we expected to find that mapping the CCR5 and chemokine surfaces with appropriate GRID probes would show complementary arrangements of MIFs in the binding orientation. Our PCA score plots (Figure 4) tend to confirm that both electrostatic *and* hydrophobic interactions are discriminating in these interactions.

Although docking calculations cannot reproduce the precise physical interaction between a receptor and its ligand, we still expected to find macromolecular complexes which exhibit at least some of the favorable interactions mentioned above. Furthermore, because chemokines block HIV-1 entry by antagonizing the binding of viral gp120 to CCR5, we also expected to find binding modes in which the agonists block a considerable region of the CCR5 extracellular surface. Considering the many millions of orientations enumerated and evaluated by GridHex, and bearing in mind the inevitable inaccuracies in our models, it was thus pleasing to find that docking solutions satisfying the above criteria were found at the relatively low cluster rank (cluster 11) for docking RANTES to receptor model 2 and at the somewhat higher rank (cluster 73) for docking MIP-1β to receptor model 4. This suggests that docking model-built protein structures is indeed feasible if SDM information is available to help eliminate false-positive dockings.

In these SDM-compatible binding modes, the buried solvent-accessible interface areas were calculated to be 1850.7 and 1721.2 Å^2 for RANTES and MIP-1β, respectively, indicating that the two proteins make extensive contacts compared to similar known complexes ($\sim$1100 Å^2). Given the high sequence similarity (70%) between RANTES and MIP-1β, we expected to find similar calculated binding modes. This was indeed the case, with both RANTES and MIP-1β binding to CCR5 with their antiparallel β sheets facing ECL1 and ECL2, and with their C-terminal α helices fully exposed to the solvent, as shown in Figure 9. This figure also illustrates how CCR5/RANTES binding is stabilized by the salt bridges between Asp11 and Lys45, Asp95 and Arg47, and Glu18 and Arg17. The interaction between the charged nitrogen of Lys191 and the aromatic ring of Phe12 provides additional π-cation stabilization. Similar interactions are seen

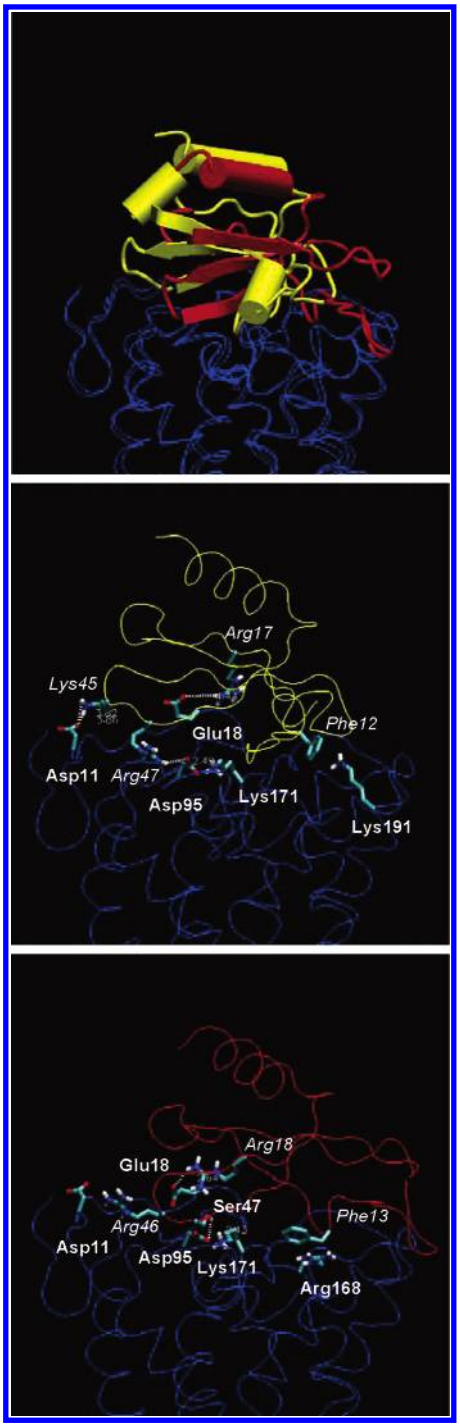

Figure 9. Docked CCR5/chemokine complexes. Top: CCR5 in blue, RANTES in yellow, MIP-1β in red. Middle: CCR5/RANTES complex with key chemokine annotated in italics. Bottom: CCR5/MIP-1β.

in the CCR5/MIP-1β complex. In this case, Glu18 still hydrogen-bonds Arg18 (Arg17 in RANTES), but the corresponding Asp11 hydrogen bond seems to take place over a larger distance. This seems plausible, considering that mutating Asp11 affects RANTES binding but not the binding of other chemokines.[51] Additionally, RANTES Phe13 now contributes to a charge-transfer complex with the CCR5 guanidinium group of Arg168, which is known to be important, especially for binding RANTES,[57] and binding is further stabilized by the hydrogen bond between Ser47 and Asp95. In both docking models, Tyr184 does not interact with the chemokines, which can probably be explained by its low exposure to the solvent. The previously reported importance of this residue might then be more related to structural rather than functional aspects. For example, it could still be important in stabilizing the bioactive conformation of ECL2 necessary for agonist binding.

CONCLUSIONS

A 3D model of the human CCR5 was homology-built from the bovine rhodopsin X-ray structure. The resulting model, consisting of the TMHs and all ECLs and ICLs, was validated by MD conformational analysis, which showed it to be consistent with the currently available SDM data. To fully exploit this model, and to give further insights into the molecular basis of the initiation and development of HIV-1 infection, we docked one antagonist and two agonists to four low-energy CCR5 conformations using a novel docking protocol. These docked complexes are also compatible with the available SDM evidence and hence provide a consistent structural model for agonist and antagonist binding. The results of our combined modeling, dynamics, and docking study provide new structural insights into CCR5/chemokine interactions, which may be useful in the rational design of HIV-1 entry blockers.

REFERENCES AND NOTES

(1) Cormier, E. G.; Persuh, M.; Thompson, D. A.; Lin, S. W.; Sakmar, T. P.; Olson, W. C.; Dragic, T. Specific Interaction of CCR5 Amino-Terminal Domain Peptides Containing Sulfotyrosines with HIV-1 Envelope Glycoprotein gp120. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 5762−5777.
(2) Farzan, M.; Vasilieva, N.; Schnitzler, C. E.; Chung, S.; Robinson, J.; Gerard, N. P.; Gerard, C.; Choe, H.; Sodroski, J. A Tyrosine-Sulfated Peptide Based on the N Terminus of CCR5 Interacts with a CD4-enhanced Epitope of the HIV-1 gp120 Envelope Glycoprotein and Inhibits HIV-1 Entry. *J. Biol. Chem.* **2000**, *275*, 33516−334521.
(3) Cormier, E. G.; Tran, D. N.; Yukhayeva, L.; Olson, W. C.; Dragic T. Mapping the Determinants of the CCR5 Amino-Terminal Sulfopeptide Interaction with Soluble Human Immunodeficiency Virus Type 1 gp120-CD4 Complexes. *J. Virol.* **2001**, *75*, 5541−5549.
(4) Cocchi, F.; DeVico, A. L.; Garzino-Demo, A.; Arya, S. K.; Gallo, R. C.; Lusso, P. Identification of RANTES, MIP-1 Alpha, and MIP-1 Beta as the Major HIV-Suppressive Factors Produced by CD8+ T Cells. *Science* **1995**, *270*, 1811−1815.
(5) Horuk, R.; Ng, H. P. Chemokine Receptor Antagonists. *Med. Res. Rev.* **2000**, *20*, 155−168.
(6) Starr-Spires, L. D., Collman, R. G. HIV-1 Entry and Entry Inhibitors as Therapeutic Agents. *Clin. Lab. Med.* **2002**, *22*, 681−701.
(7) Yi, Y.; Shaheen, F.; Collman, R. G. Preferential Use of CXCR4 by R5X4 Human Immunodeficiency Virus Type 1 Isolates for Infection of Primary Lymphocytes. *J. Virol.* **2005**, *79*, 1480−1486.
(8) Meister, S.; Otto, C.; Papkalla, A.; Krumbiegel, M.; Pohlmann, S.; Kirchhoff, F. Basic Amino Acid Residues in the V3 Loop of Simian Immunodeficiency Virus Envelope Alter Viral Coreceptor Tropism and Infectivity but Do Not Allow Efficient Utilization of CXCR4 as Entry Cofactor. *Virology* **2001**, *284*, 287−296.
(9) Ditzel, H. J.; Rosenkilde, M. M.; Garred, P.; Wang, M.; Koefoed, K.; Pedersen, C.; Burton, D. R.; Schwartz, T. W. The CCR5 Receptor Acts as an Alloantigen in CCR5Δ32 Homozygous Individuals: Identification of Chemokines and HIV-1-Blocking Human Antibodies. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 5241−5245.
(10) Rojo, D.; Suetomi, K.; Navarro, J. Structural Biology of Chemokine Receptors. *Biol. Res.* **1999**, *32*, 263−272.
(11) Rusconi, S.; Scozzafava, A.; Mastrolorenzo, A.; Supuran, C. T. New Advances in HIV Entry Inhibitors Development. *Curr. Drug Targets: Infect. Disord.* **2004**, *4*, 339−355.
(12) Horuk, R. Development and Evaluation of Pharmacological Agents Targeting Chemokine Receptors. *Methods* **2003**, *29*, 369−75.
(13) Klabunde, T.; Hessler, G. Drug Design Strategies for Targeting G-Protein-Coupled Receptors. *ChemBioChem* **2002**, *3*, 928−944.
(14) Paterlini, G. Structure Modeling of the Chemokine Receptor CCR5: Implications for Ligand Binding and Selectivity. *Biophys. J.* **2002**, *83*, 3012−3031.

(15) Huang, X. Q.; Jiang, H. L.; Luo, X. M.; Chen, K. X.; Ji, R. Y.; Cao, Y.; Pei, G. Building Three-Dimensional Structures of HIV-1 Coreceptor CCR5 and Its Interaction with Antagonist TAK779 by Comparative Molecular Modeling. *Acta Pharmacol. Sin.* **2000**, *21*, 521−528.

(16) Xu, Y.; Liu, H.; Niu, C.; Luo, C.; Luo, X.; Shen, J.; Chen, K.; Jiang, H. Molecular Docking and 3D QSAR Studies on 1-Amino-2-phenyl-4-(piperidin-1-yl)-butanes Based on the Structural Modeling of Human CCR5 Receptor. *Bioorg. Med. Chem.* **2004**, *12*, 6193−6208.

(17) Kazmierski, W.; Bifulco, N.; Yang, H.; Boone, L.; DeAnda, F.; Watson, C.; Kenakin, T. Recent Progress in Discovery of Small-Molecule CCR5 Chemokine Receptor Ligands as HIV-1 Inhibitors. *Bioorg. Med. Chem.* **2003**, *11*, 2663−2676.

(18) Palczewski, K.; Kumasaka, T.; Hori, T.; Behnke, C. A.; Motoshima, H.; Fox, B. A.; Le, T. I.; Teller, D. C.; Okada, T.; Stenkamp, R. E.; Yamamoto, M.; Miyano, M. Crystal Structure of Rhodopsin a G Protein-Coupled Receptor. *Science* **2000**, *289*, 739−745.

(19) Fanelli, F.; De Benedetti, P. G. Computational Modeling Approaches to Structure−Function Analysis of G-Protein-Coupled Receptors. *Chem. Rev.* **2005**, *105*, 3297−3351.

(20) Bairoch, A.; Apweiler, R. The SWISS−PROT Protein Sequence Data Bank and Its Supplement TrEMBL in 1998. *Nucleic Acids Res.* **1998**, *26*, 38−42.

(21) Burkhard, R.; Casadio, R.; Fariselli, P.; Chris, S. Prediction of Helical Transmembrane Segments at 95% Accuracy. *Protein Sci.* **1995**, *4*, 521−533.

(22) Thompson, J. D.; Higgins, D. G.; Gibson, T. J. CLUSTAL W: Improving the Sensitivity of Progressive Multiple Sequence Alignment through Sequence. *Nucleic Acids Res.* **1994**, *22*, 4673−4680.

(23) Sali, A.; Blundell, T. L. Comparative Protein Modelling by Satisfaction of Spatial Restraints. *J. Mol. Biol.* **1993**, *234*, 779−815.

(24) Blanpain, C.; Lee, B.; Vakili, J.; Doranz, B. J.; Govaerts, C.; Migeotte, I.; Sharron, M.; Dupriez, V.; Vassart, G.; Doms, R. W.; Parmentier, M. Extracellular Cysteines of CCR5 Are Required for Chemokine Binding.; but Dispensable for HIV-1 Coreceptor Activity. *J. Biol. Chem.* **1999**, *274*, 18902−18908.

(25) Laskowski, R. A.; MacArthur, M. W.; Moss, D. S.; Thornton, J. M. PROCHECK: A Program To Check the Stereochemical Quality of Protein Structures. *J. Appl. Crystallogr.* **1993**, *26*, 283−291.

(26) Carrieri, A.; Centeno, N. B.; Rodrigo, J.; Sanz, F.; Carotti, A. Theoretical Evidence of a Salt Bridge Disruption as the Initiating Process for the Alpha1d-adrenergic Receptor Activation: A Molecular Dynamics and Docking Study. *Proteins: Struct., Funct., Genet.* **2001**, *43*, 382−394.

(27) Weiner, S. J.; Kollman, P. A.; Case, D. A.; Singh, U. C.; Ghio, C.; Alagona, G.; Profeta, S., Jr.; Weiner, P. A New Force Field for Molecular Mechanical Simulation of Nucleic Acids and Proteins. *J. Am. Chem. Soc.* **1984**, *106*, 765−784.

(28) Rohrig, U. F.; Guidoni, L.; Rothlisberger, U. Early Steps of the Intramolecular Signal Transduction in Rhodopsin Explored by Molecular Dynamics Simulations. *Biochemistry* **2002**, *41*, 10799−809.

(29) Rohrig, U. F.; Guidoni, L.; Rothlisberger, U. Solvent and Protein Effects on the Structure and Dynamics of the Rhodopsin Chromophore. *ChemPhysChem* **2005**, *6*, 1836−1847.

(30) Lazaridis, T. Implicit Solvent Simulations of Peptide Interactions with Anionic Lipidmembranes. *Proteins* **2005**, *58*, 518−27.

(31) Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. Numerical Integration of the Cartesian Equation of Motion of a System Constraints: Molecular Dynamics of N-Alkanes. *J. Comput. Phys.* **1977**, *23*, 327−341.

(32) Kelley, L.; Gardner, S.; Sutcliffe, M. An Automated Approach for Clustering an Ensemble of NMR-Derived Protein Structures into Conformationally-Related Subfamilies. *Protein Eng.* **1996**, *9*, 1063−1065.

(33) Frishman, D.; Argos P. Knowledge-Based Protein Secondary Structure Assignment. *Proteins: Struct., Funct., Genet.* **1995**, *23*, 556−579.

(34) Gentili, F.; Ghelfi, F.; Giannella, M.; Piergentili, A.; Pigini, M.; Quaglia, W.; Vesprini, C.; Crassous, P. A.; Paris, H.; Carrieri, A. α2-Adrenoreceptors Profile Modulation. 2. Biphenyline Analogues as Tools for Selective Activation of the α2C-Subtype. *J. Med. Chem.* **2004**, *47*, 6160−73.

(35) McMartin, C.; Bohacek, R. S. QXP: Powerful Rapid Computer Algorithms for Structure-Based Drug Design. *J. Comput.-Aided Mol. Des.* **1997**, *11*, 333−344.

(36) Dragic, T.; Trkola, A.; Thompson, D. A.; Cormier, E. G.; Kajumo, F. A.; Maxwell, E.; Lin, S. W.; Ying, W.; Smith, S. O.; Sakmar, T. P.; Moore, J. P. A Binding Pocket for a Small Molecule Inhibitor of HIV-1 Entry within the Transmembrane Helices of CCR5. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 5639−5644.

(37) Ritchie, D. W.; Kemp, G. J. L. Protein Docking Using Spherical Polar Fourier Correlations. *Proteins: Struct., Funct., Genet.* **2000**, *39*, 178−194.

(38) Fano, A. Using Molecular Interaction Fields to Improve the Scoring of Protein−Protein Interactions: Application to CCR5-Chemokines System and HIV-Blockers, from Pharmacophore, to Theoretical Model, to Docking. Ph.D. Thesis, Faculty of Pharmacy, University of Bari, Bari, Italy, 2004.

(39) Goodford, P. J. A Computational Procedure for Determining Energetically Favourable Binding Sites on Biologically Important Macromolecules. *J. Med. Chem.* **1985**, *28*, 849−857.

(40) Janin, J. Assessing Predictions of Protein−Protein Interaction: The CAPRI Experiment. *Protein Sci.* **2005**, *14*, 278−283.

(41) Mendez, R.; Leplae, R.; Lensink, M. F.; Wodak, S. J. Assessment of CAPRI Predictions in Rounds 3−5 Shows Progress in Docking Procedures. *Proteins: Struct., Funct., Bioinf.* **2005**, *60*, 150−169.

(42) Sussman, J. L.; Lin, D.; Jiang, J.; Manning, N. O.; Prilusky, J.; Ritter, O.; Abola, E. E. Protein Data Bank (PDB): Database of Three-Dimensional Structural Information of Biological Macromolecules. *Acta Cryst.* **1998**, *54*, 1078−1084.

(43) Skelton, N. J.; Aspiras, F.; Ogez, J.; Schall, T. J. Proton NMR Assignments and Solution Conformation of RANTES, a Chemokine of the C-C Type. *Biochemistry* **1995**, *34*, 5329−5342.

(44) Lodi, P. J.; Garrett, D. S.; Kuszewski, J.; Tsang, M. L.; Weatherbee, J. A.; Leonard, W. J.; Gronenborn, A. M.; Clore, G. M. High-Resolution Solution Structure of the Beta Chemokine hMIP-1 Beta by Multidimensional NMR. *Science* **1994**, *263*, 1762−1767.

(45) *GOLPE*, version 4.5; Multivariate Informatic Analysis Srl: Perugia, Italy, 1999.

(46) Nardese, V.; Longhi, R.; Polo, S.; Sironi, F.; Arcelloni, C.; Paroni, R.; DeSantis, C.; Sarmientos, P.; Rizzi, M.; Bolognesi, M.; Pavone, V.; Lusso, P. Structural Determinants of CCR5 Recognition and HIV-1 Blockade in RANTES. *Nat. Struct. Biol.* **2001**, *8*, 611−615.

(47) Chabot, D. J.; Broder, C. C. Substitutions in a Homologous Region of Extracellular Loop 2 of CXCR4 and CCR5 Alter Coreceptor Activities for HIV-1 Membrane Fusion and Virus Entry. *J. Biol. Chem.* **2000**, *275*, 23774−23782.

(48) Scheer, A.; Fanelli, F.; Costa, T.; De Benedetti, P. G.; Cotecchia, S. Constitutively Active Mutants of the α_{1B}-Adrenergic Receptor: Role of Highly Conserved Polar Amino Acids in Receptor Activation. *EMBO J.* **1996**, *15*, 3566−3578.

(49) Govaerts, C.; Blanpain, C.; Deupi, X.; Ballet, S.; Ballesteros, J. A.; Wodak, S. J.; Vassart, G.; Pardo, L.; Parmentier, M. The TXP Motif in the Second Transmembrane Helix of CCR5. A Structural Determinant of Chemokine-Induced Activation. *J. Biol. Chem.* **2001**, *276*, 13217−13225.

(50) Govaerts, C.; Bondue, A.; Springael, J. Y.; Olivella, M.; Deupi, X.; Le Poul, E.; Wodak, S. J.; Parmentier, M.; Pardo, L.; Blanpain, C. Activation of CCR5 by Chemokines Involves an Aromatic Cluster between Transmembrane Helices 2 and 3. *J. Biol. Chem.* **2003**, *278*, 1892−1903.

(51) Blanpain, C.; Doranz, B. J.; Vakili, J.; Rucker, J.; Govaerts, C.; Baik, S. S.; Lorthioir, O.; Migeotte, I.; Libert, F.; Baleux, F.; Vassart, G.; Doms, R. W.; Parmentier, M. Multiple Charged and Aromatic Residues in CCR5 Amino-Terminal Domain Are Involved in High Affinity Binding of Both Chemokines and HIV-1 Env Protein. *J. Biol. Chem.* **1999**, *274*, 34719−34727.

(52) Dragic, T.; Trkola, A.; Lin, S. W.; Nagashima, K. A.; Kajumo, F.; Zhao, L.; Olson, W. C.; Wu, L.; Mackay, C. R.; Allaway, G. P.; Sakmar, T. P.; Moore, J. P.; Maddon, P. J. Amino-Terminal Substitutions in the CCR5 Coreceptor Impair gp120 Binding and Human Immunodeficiency Virus Type 1 Entry. *J. Virol.* **1998**, *72*, 279−285.

(53) Farzan, M.; Choe, H.; Vaca, L.; Martin, K.; Sun, Y.; Desjardins, E.; Ruffing, N.; Wu, L.; Wyatt, R.; Gerard, N.; Gerard, C.; Sodroski, J. A Tyrosine-Rich Region in the N Terminus of CCR5 Is Important for Human Immunodeficiency Virus Type 1 Entry and Mediates an Association Between gp120 and CCR5. *J. Virol.* **1998**, *72*, 1160−1164.

(54) Genoud, S.; Kajumo, F.; Guo, Y.; Thompson, D.; Dragic, T. CCR5-Mediated Human Immunodeficiency Virus Entry Depends on an Amino-Terminal gp120-Binding Site and on the Conformational Integrity of All Four Extracellular Domains. *J. Virol.* **1999**, *73*, 1645−1648.

(55) Rabut, G. E.; Konner, J. A.; Kajumo, F.; Moore, J. P.; Dragic, T. Alanine Substitutions of Polar and Nonpolar Residues in the Amino-Terminal Domain of CCR5 Differently Impair Entry of Macrophage- and Dualtropic Isolates of Human Immunodeficiency Virus Type 1. *J. Virol.* **1998**, *72*, 3464−3468.

(56) Navenot, J. M.; Wang, Z. X.; Trent, J. O.; Murray, J. L.; Hu, Q. X.; DeLeeuw, L.; Moore, P. S.; Chang, Y.; Peiper, S. C. Molecular Anatomy of CCR5 Engagement by Physiologic and Viral Chemokines and HIV-1 Envelope Glycoproteins: Differences in Primary Structural Requirements for RANTES.; MIP-1 Alpha.; and vMIP−II Binding. *J. Mol. Biol.* **2001**, *313*, 1181−1193.

(57) Blanpain, C.; Doranz, B. J.; Bondue, A.; Govaerts, C.; De Leener, A.; Vassart, G.; Doms, R. W.; Proudfoot, A.; Parmentier, M. The CoreDomain of Chemokines Binds CCR5 Extracellular Domains while Their Amino Terminus Interacts with the Transmembrane Helix Bundle. *J. Biol. Chem.* **2003**, *278*, 5179−5187.

(58) Baba, M.; Nishimura, O.; Kanzaki, N.; Okamoto, M.; Sawada, H.; Iizawa, Y.; Shiraishi, M.; Aramaki, Y.; Okonogi, K.; Ogawa, Y.; Meguro, K.; Fujino, M. A Small-Molecule, Nonpeptide CCR5 Antagonist with Highly Potent and Selective anti-HIV-1 Activity. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 5698−5703.

(59) Olson, W. C.; Rabut, G. E.; Nagashima, K. A.; Tran, D. N.; Anselma, D. J.; Monard, S. P.; Segal, J. P.; Thompson, D. A.; Kajumo, F.; Guo, Y.; Moore, J. P.; Maddon, P. J.; Dragic, T. Differential Inhibition of Human Immunodeficiency Virus Type 1 Fusion, gp120 Binding, and CC-Chemokine Activity by Monoclonal Antibodies to CCR5. *J. Virol.* **1999**, *73*, 4145−4155.

(60) Ross, T. M.; Bieniasz, P. D.; Cullen, B. R. Multiple Residues Contribute to the Inability of Murine CCR-5 To Function as a Coreceptor for Macrophage-Tropic Human Immunodeficiency Virus Type 1 Isolates. *J. Virol.* **1998**, *72*, 1918−1924.

(61) Hubbard, S. J.; Thornton, J. M. *"NACESS", Computer Program*; Department of Biochemistry and Molecular Biology: University College of London, 1993.

CI050490K