Published in final edited form as:

Chembiochem. 2014 July 21; 15(11): 1614–1620. doi:10.1002/cbic.201402056.

Anti-HIV Small-Molecule Binding in the Peptide Subpocket of the CXCR4:CVX15 Crystal Structure

Dr. Bryan D. Cox, Anthony R. Prosser, Brooke M. Katzman, Dr. Ana A. Alcaraz, Prof. Dennis C. Liotta, Dr. Lawrence J. Wilson, and Dr. James P. Snyder^[a]

^[a]Department of Chemistry, Emory University, 1521 Dickey Drive, Atlanta, GA 30322 (USA)

Abstract

The CXC chemokine receptor 4 (CXCR4) is involved in chemotaxis and serves as a coreceptor for T-tropic HIV-1 viral entry, thus making this receptor an attractive drug target. Recently, crystal structures of CXCR4 were reported as complexes with the small molecule IT1t and the CVX15 peptide. Follow-up efforts to model different antagonists into the small molecule CXCR4:IT1t crystal structure did not generate poses consistent with either the X-ray crystal structure or site-directed mutagenesis (SDM). Here, we compare the binding pockets of the two CXCR4 crystal structures, revealing differences in helices IV, V, VI, and VII, with major differences for the His203 residue buried in the binding pocket. The small molecule antagonist AMD11070 was docked into both CXCR4 crystal structures. An AMD11070 pose identified from the CXCR4:CVX15 model presented interactions with Asp171, Glu288, Trp94, and Asp97, consistent with published SDM data, thus suggesting it is the bioactive pose. A CXCR4 receptor model was optimized around this pose of AMD11070, and the resulting model correlated HIV-1 inhibition with MM-GBSA docking scores for a congeneric AMD11070-like series. Subsequent NAFIS NMR results successfully linked the proposed binding pose to an independent experimental structure. These results strongly suggest that not all small molecules will bind to CXCR4 in a similar manner as IT1t. Instead, the CXCR4:CVX15 crystal structure may provide a binding locus for small organic molecules that is more suitable than the secondary IT1t site. This work is expected to provide modeling insights useful for future CXCR4 antagonist and X4-tropic HIV-1 based drug design efforts.

Keywords

chemokine receptors; computational chemistry; conformational analysis; molecular modeling; NMR spectroscopy

Introduction

The GPCR CXC-motif chemokine receptor 4 (CXCR4) plays crucial roles in HIV-1 infection and cell motility, making it a target of intense medicinal interest. CXCR4 interacts with the CXCL12/SDF-1 (stromal cell-derived factor 1) chemokine to regulate cell

Correspondence to: James P. Snyder.

Supporting information for this article is available on the WWW under <http://dx.doi.org/10.1002/cbic.201402056>.

migration. Accordingly, dysfunction of the CXCR4/SDF-1 axis is associated with cancer metastasis, angiogenesis, and WHIM syndrome.^[1] This dimeric protein is also utilized by T-tropic HIV-1 variants as a coreceptor for viral entry. Several peptides and small-molecules have been developed that target CXCR4 to exploit these biological roles. The SDF-1-based peptide homologue CTCE-9908 (Chemokine Therapeutics Corp.) is in phase I/II clinical trials for breast cancer.^[2] One small molecule, IT1t,^[3] was recently cocrystallized with CXCR4^[4] and is complemented by two other small-molecule agents that have progressed to the clinic: Plerixafor (AMD3100, Mozobil) is an FDA-approved CXCR4 antagonist used in stem cell mobilization,^[5] while AMD11070 was clinically examined for T-tropic HIV-1 viral entry inhibition (Scheme 1).^[6] Due to its medicinal importance, CXCR4 has undergone structural investigation to aid in drug discovery. Prior to the publication of the crystal structure, homology models were used to structurally investigate CXCR4, frequently using bovine rhodopsin and β -adrenergic receptors as templates.^[7–9] Structural analysis of the CXCR4 homology models led to ideas involving multiple sites within the binding pocket, a deduction complemented by docking drug-like molecules into both homology models and the CXCR4 crystal structure.^[10, 11]

The generally poor performance of the homology models in previous pose prediction was likely due to the lack of a sequentially similar template crystal structure.^[12] In 2010, Wu et al. reported the first crystal structures of the CXCR4-T4L fusion protein cocrystallized with the small molecule IT1t and the peptide CVX15, providing direct confirmation of differential binding sites.^[4] Virtual screening of the CXCR4:IT1t binding pocket increased hit rates relative to homology models.^[13] However, docking IT1t back into the crystal resulted in large RMSD values compared to the solved structures,^[12, 14] and docking small molecule antagonists into the CXCR4:IT1t structure resulted in poses lacking critical interactions identified by site-directed mutagenesis.^[15, 16] Inspection of the CXCR4 crystal structures revealed clear differences in residues important for small molecule binding, and these observations implied a need to examine these differences in more detail in order to generate insights that go beyond the results of constrained docking.^[17]

We recently engaged in the development of GPCR binding models for our CXCR4 antagonist program and focused on forms of the receptor provided in the two crystal structures of CXCR4 with IT1t and CVX15.^[18] Although initial considerations for small molecule modeling might find the IT1t binding site attractive, other studies demonstrated difficulty regenerating the bound conformation of IT1t.^[14] Alternatively, we hypothesized that some small molecules interact more favorably in the CVX15-bound form of the receptor and a priori serve as peptide mimetics. To test this hypothesis, we compared conformational variations in the binding sites of the CXCR4:IT1t and CXCR4:CVX15 crystal structures for use in docking studies. The small molecule AMD11070 is an attractive model compound for these studies, as site-directed mutagenesis (SDM) data supports molecular interactions with the CXCR4 receptor (in the absence of an AMD11070-bound crystal structure). We identified a pose derived from docking AMD11070 into the CXCR4:CVX15 peptide crystal structure that agrees with independent SDM, as opposed to our results from docking AMD11070 into the CXCR4:IT1t crystal structure. By optimizing the receptor model around the CXCR4:CVX15-derived pose, we observed a correlation with HIV-1 inhibition potency for an AMD11070-like congeneric series. The NMR-guided^[19] conformational

profile of AMD11070 in solution reveals an ensemble of conformations, one of which is similar to the docked AMD11070, as determined by shape-based analysis. These results indicate that AMD11070 binds within the main binding pocket of the CXCR4:CVX15 crystal structure, which is consistent with the bioactive pose and inconsistent with its presence in the secondary small molecule IT1t pocket. Accordingly, the peptide crystal structure should be considered a potential receptor model for any novel small molecule CXCR4 antagonist.

Results and Discussion

Binding pocket analysis and molecular modeling

The CXCR4 crystal structures possess seven transmembrane helices numbered consecutively from the N terminus and a dimerization interface between helices V and VI. CXCR4 was crystallized separately with two ligands in the extracellular binding pocket—the small molecule IT1t and the CVX15 peptide. Structural examination of the binding pockets for CXCR4 shows key differences between the IT1t- and CVX15-bound crystal structures. Analysis of protein–ligand contacts revealed 35 significant residues in the binding cleft in contact with IT1t and/or CVX15, and the side-chain conformations of these residues were compared by heavy-atom RMSD (Table S1 in the Supporting Information). Residues at the extracellular ends of the helices display large RMSD values between the two crystal structures (Figure 1). For example, residues Asp262 (VI) and His281 (VII) are key residues for cyclam antagonist binding,^[20] and these residues exhibit heavy-atom RMSD values of 1.3 and 3.9 Å between the two ligand protein complexes, respectively. The interior residues for helices I–III and VII have low RMSD values, thus suggesting that this subpocket is ordered and rigid, in agreement with previous molecular dynamics simulations.^[21] Interior residues for helices IV, V, and VI deliver moderate RMSD values. For example, residues Val196 and Phe199 are core residues in helix V, with RMSD values of 0.8 and 1.6 Å, respectively. Two residues on the interior of the binding pocket on helix V (Gln200 and His203) furnish large RMSD values; this indicates a large structural fluctuation between the two ligand–receptor complexes (Figure 2). In the CXCR4:IT1t crystal structure, these residues do not interact with the ligand but are involved in hydrogen bonding interactions with helices III and VI. To accommodate the naphthylalanine 3 (Nal3) residue in the CVX15 peptide, His203 rotates away from helix III towards helix VI. Gln200 also adapts to the Nal3 residue by rotating away from Tyr255/Ile259 towards Asp262 on helix VI. Thus, the analysis reveals that the binding pocket of CXCR4 reorganizes at the extracellular ends of the helices and at the core of helices IV–VI in response to the differential binding by CVX15 and IT1t.

The next level of analysis compared the poses of small molecules docked into the ligand pockets presented by the CXCR4:IT1t and CXCR4:CVX15 crystal structures. We chose AMD11070 because it lacks a large, flexible cyclam ring (a common moiety present in AMD3100-like compounds, Scheme 1), and SDM data established key residues for binding to CXCR4. AMD11070 is a molecule composed of three groups attached to a central tertiary nitrogen: 1) a nitrogen-containing tetrahydroquinoline ring, 2) a butyl amine chain, and 3) a benzimidazole heterocycle. Previous SDM identified four key residues implicated in the

binding of AMD11070 to CXCR4: Glu288, Asp171, Asp97, and Trp94.^[10] Notably, the Asp171Asn (D171N) mutation results in a more than 1000-fold loss in potency, thus indicating that AMD11070 engages in a strong interaction with this residue.

To obtain an assessment of this and other contacts, AMD11070 was docked into the CXCR4:IT1t and CXCR4:CVX15 receptor models using Glide, and the per-residue energies of interaction were calculated for the four key residues. The corresponding energy values for the top three poses for each wild-type crystal structure are listed in Table 1, and the poses are provided in Figure S2. The docking scores were similar for each ligand–protein complex. Only the top scoring pose from the CXCR4:CVX15 model displayed a strong interaction with Asp171, and this pose accommodated all interactions identified by SDM (Figure 3A). In this pose, the protonated butyl amine engages in a strong electrostatic interaction with Asp171 and also forms hydrogen bond interactions with the hydroxy group on Thr117 and the backbone carbonyl of His113 (Figure 3 B). In concert, the benzimidazole donates a hydrogen bond to Glu288 (Figure 3 C). The interactions of the butyl amine and benzimidazole moieties are similar to those in one of the three poses from the homology model docking by Wong et al.^[10] Unlike the previous Wong et al. poses, the tetrahydroquinoline ring in our work forms two weak interactions. A carbon on the aromatic ring is 3.6 Å from a carboxyl oxygen on Asp97, while the dihedral angle between the aromatic C–H on AMD11070 with the C–O on Asp97 is 140°, parameters that describe a weak C–H...O hydrogen bond.^[22] The aromatic ring also forms a hydrophobic interaction with Trp94, with offset ring centroids that appear to narrowly avoid a π – π stacking interaction (Figure 3 D). In other words, this pose presents a similar pharmacophore to that from previous homology model docking but in a significantly different conformation.

In our pose, the butyl amine side chain has one *gauche* interaction for C2–C3, at the relatively low energy cost of ~0.8 kcal mol⁻¹.^[23] The strongest interaction from this pose was the electrostatic interaction between the butyl amine chain of AMD11070 and Asp171, which is consistent with SDM. It is clear that the approximately 1000-fold loss in potency for the Asp171Asn mutation for AMD11070 likely arises from the loss of the electrostatic interaction between Asp171 and the protonated butyl amine side chain of AMD11070.

Indeed, comparing the energy of interaction with the pK_i^{MUT} (change in SDF-1 K_i upon mutation) for each residue provides a linear correlation with $r^2=0.98$ (Figure 4). A single-point quantum chemical calculation at the B3LYP/6-31G* level of theory in an implicit water solvent model revealed that the AMD11070 pose from the CXCR4:CVX15 crystal structure is 8.1 kcal mol⁻¹ lower in energy than the best AMD11070 pose from the CXCR4:IT1t crystal structure (data not shown). These observations suggest that docking AMD11070 into the CXCR4:CVX15 structure yields a satisfying representation of the bioactive pose that matches SDM results and captures a low-energy ligand conformation. Interestingly, juxtaposition of this pose with the ordered water molecules from the CXCR4:CVX15 crystal structure reveal two water molecules (1601 and 1604) occupying the same space as the benzimidazole of AMD11070 (Figure S5). It is likely that either AMD11070 replaces these water molecules upon binding, or the waters remain to stabilize interactions around the ligand's benzimidazole ring (Figure 3C).

Study of the putative bioactive pose of AMD11070 in the CXCR4:IT1t crystal structure is complicated, because Asp171 is sterically inaccessible in this structure. His203 forms a favorable electrostatic interaction with Arg188 in the CXCR4:IT1t crystal structure. A van der Waals violation occurs between the AMD11070 butyl amine and Arg188 (distance of 2.8 Å) when attempting to engage Asp171. In contrast, His203 rotates away from Arg188 in the CXCR4:CVX15 structure, thereby exposing Asp171 to the ligand binding pocket and an association with the butyl amine side chain of AMD11070. The His203 residue is located on helix V at the dimerization interface (Figure 2), and this residue can, in principle, play a role in communication across the interface. The orientations of Met205 and Phe201 on helix V differ between the IT1t- and CVX15-bound CXCR4 crystal structures and might be influenced by rotation of the flanking His203 residue. These observations suggest that the interaction between the AMD11070 butyl amine side chain with Asp171 leads to a cascade of structural changes through His203 that may affect CXCR4 dimerization and signaling.^[24]

To support the bioactive pose of AMD11070 obtained from the CXCR4:CVX15 peptide structure, we tested the receptor models in their ability to predict experimental HIV-1 activity. The congeneric series reported by Skerlj et al., in which AMD11070-like compounds vary in the alkyl amine chain from two to six carbons,^[25] was chosen as a provisional test of the binding hypothesis because the compounds contain the same pharmacophore features as AMD11070, and the activity of these compounds span a range of three log units. Post-Glide docking energies were rescored using Prime MM-GBSA, as it often increases the quality of a correlation.^[26] Docking and rescoring this series of structures into the CXCR4:CVX15 crystal structure failed to parallel the HIV-1 potency (Figure 5 A). Thus, an AMD11070-optimized receptor model was created by relaxing the CXCR4:CVX15 receptor model around the putative bioactive conformation using Prime MM-GBSA. Docking the series into this model (CXCR4:AMD11070) yielded a satisfying correlation between HIV-1 viral entry inhibition and the MMGBSA G binding energy (Figure 5 B), with the four-carbon chain scoring better than those of other chain lengths. The single-run correlation typically yielded an r^2 value of ~0.5; however, an improved correlation ($r^2=0.87$) was obtained by averaging over several sets of docking attempts (Table S3), which agrees with the principle of consensus docking where repeated measurements approach the actual value.^[27, 28] The high degree of correlation suggests that the optimized CXCR4:CVX15 receptor describes the binding of this AMD11070-like congeneric series. The same optimization was repeated by including the three water molecules in the peptide binding pocket of the CXCR4:CVX15 crystal structure. Docking the AMD11070-like series into this receptor locus correlated HIV-1 entry inhibition potency with Prime MM-GBSA G values with an $r^2=0.70$ (Figure S4), but the range of Prime MM-GBSA G energy was much lower. This is consistent with previous reports demonstrating that inclusion of crystallographic waters does not improve the enrichment of CXCR4 receptor models.^[29] Following this same methodology, a receptor model was created that optimized the AMD11070 pose from the CXCR4:IT1t crystal structure. Docking and scoring the AMD11070-like congeneric series into this model did not yield a correlation between HIV-1 potency and Prime MMGBSA G values (Table S3). These results suggest that only the CXCR4:CVX15 receptor model was optimized around the bio-active pose of AMD11070.

Comparison of the proposed bioactive pose of AMD11070 with experimental conformations

Because ROCS (rapid overlay of chemical structures) provides excellent 3D similarity comparisons for small molecule conformations, this method was employed to match the bioactive pose of AMD11070 with experimentally determined conformations. The crystal structure of the AMD11070 free-base solved by Crawford et al.^[30] overlays reasonably well with the bioactive pose of AMD11070, with a Tanimoto shape score of 0.64 (Figure 6 A). The general orientations of the three substituents are similar between the two poses, but there are significant differences. The benzimidazole ring is slightly rotated between the bioactive and small molecule crystal conformers. The tetra-hydroquinoline ring rotates nearly 180° in the small molecule crystal structure relative to the bioactive pose. The crystal form presents an extended butyl amine chain in contrast to the bio-active pose, which sustains a single *gauche* twist.

Solution conformers of AMD11070 were obtained by the NMR NAMFIS method,^[19] in which the averaged intramolecular H–H distances derived from NOESY NMR crosspeaks are de-convoluted into a set of conformations associated with individual mole fractions (or populations). Input for analysis consists of a “complete set” of in silico generated conformers and the distance-scaled NOE crosspeaks. Thus, an exhaustive in silico search for AMD11070 yielded a total of 10 755 conformations (4397 in AMBER*, 2861 in MMFFs, and 3497 in OPLS-2005). 2D NOESY analysis of the CXCR4 antagonist furnished 31 NOE crosspeaks, which were converted to H–H distances by using an *ortho*-phenyl NOE crosspeak and the known distance as an internal ruler (Table S5). NAMFIS was used to intersect the H–H distances and the combined conformational database of 10 755 structures to provide 13 conformers of AMD11070 with populations ranging from 0.6 % to 24.6 %. A satisfying quality of fit was indicated by the sum of square distances (SSD =56). Of the 13 conformers, two solution conformers of AMD11070 had high Tanimoto shape scores when overlaid with the free-base crystal structure (NAMFIS-4 and NAMFIS-9, both with Tanimoto shape scores of 0.71, Figure S3). Notably, these conformers have the same orientation of the tetrahydroquinoline ring as the small molecule crystal structure. This result supports the proposition that the solution ensemble of AMD11070 conformers contains the small molecule X-ray form. All 13 solution conformers of AMD11070 were aligned with the bioactive pose, and the Tanimoto scores with respect to shape and color were calculated (Table S6). The bioactive pose was almost identical to that of the NAMFIS-3 conformer (14.3 %), exhibiting a Tanimoto shape score of 0.79 (Figure 6 B). The tetrahydroquinoline and benzimidazole rings of the NAMFIS-3 conformer align faithfully with those of the bioactive pose. The main difference between the AMD11070 bioactive pose and this solution conformer is the orientation of the butyl amine side chain: the bioactive pose exhibits one *gauche* twist, and the NAMFIS-3 conformer contains two *gauche* twists. This outcome demonstrates that the empirically-derived solution-based AMD11070 conformational pool contains conformers similar to both the small molecule X-ray structure and the most favorable docked ligand in the CXCR4:CVX15 binding site. The result is consistent with previous analyses along these lines^[31] and supports the docked AMD11070 structure as a viable low-energy binding conformation.

Conclusions

The extracellular ligand binding site of CXCR4, a class A GPCR, is composed of surface-exposed tethers at the N terminus utilized by the CXCL12/SDF-1 chemokine and two subpockets within the seven-transmembrane (7-TM) helical bundle—a major pocket composed of helices III–VII and a minor pocket composed of helices I–III and VII.^[32] The crystal structures of CXCR4 are unique in that they were solved with ligands occupying the minor pocket (IT1t) and with a ligand spanning both subpockets (CVX15). Not surprisingly, previous modeling of small molecule antagonists of CXCR4 focused exclusively on the minor pocket site in the small molecule CXCR4:IT1t crystal structure. As SDM indicates that CXCR4 antagonists AMD3100, AMD11070, and AMD3465 interact with residues in both the major and minor pockets, we explored both the IT1t and CVX15 crystal structures as possible models for antagonist binding. Analysis reveals that the minor binding pocket of CXCR4 presents similar residue conformations in the two crystal structures; however, significant conformational differences are observed in the major binding pocket. Modeling with both crystal structures yielded a pose of AMD11070 in good agreement with SDM—our proposed bioactive pose—obtained surprisingly from docking into the CXCR4:CVX15 structure. The results presented here suggest that both CXCR4 crystal structures (IT1t and CVX15) should be considered for structurally uncharacterized compounds that interact with this receptor. The approach presented here may aid in the rational design of future CXCR4 antagonists for HIV-1 viral entry inhibition. Further work in this area using this approach for analysis of other CXCR4 antagonists^[18] is in progress.

Experimental Section

X-ray structures

Both the CXCR4:CVX15 (PDB ID: 3OE0) and the CXCR4:IT1t (PDB ID: 3ODU) crystal structures were examined as targets for predictive docking. The 3ODU CXCR4:IT1t crystal structure was chosen for its superior resolution (2.5 Å), compared to the three other available CXCR4:IT1t crystal structures (3OE6, 3OE8, and 3OE9, each with resolution >3.0 Å).

Binding pocket analysis

Ligand–protein contacts (LPC) were derived with LPC software, and contacts of structural units (CSU) were derived with CSU software.^[33] The CSU server analyzes protein–protein interactions, while the LPC server treats small molecule–protein interactions. LPC was used to analyze contacts in the CXCR4:IT1t crystal structure as well as the CVX15 peptide residues Nal3, Dpr8, Cir12, and Dpr16 in contact with the receptor. CSU was also employed for analysis of contacts between the common amino acid residues in the CVX15 peptide (Arg1, Arg2, Cys4, Tyr5, Gln6, Lys7, Pro9, Tyr10, Arg11, Cys13, Arg14, and Gly15) and CXCR4 in 3OE0. To calculate the RMSD for CXCR4 residues in the ligand binding pockets, the 7-TM region of chain A from the CXCR4:IT1t structure was aligned with the 7-TM region of the CXCR4:CVX15 structure using visual molecular dynamics (VMD).^[34] The heavy-atom RMSD was calculated for each residue using the RMSD Calculator tool in VMD.

Docking methods

Unless otherwise stated, the default parameters in the Maestro docking modules located in the Schrodinger Suite (v. 9.2) were used.^[35] As no conserved structured waters or fatty acids are found in the crystal structures, nonligand molecules were removed from each structure. The receptors were prepared using the Protein Preparation Wizard in the Schrodinger Suite that assigns bond orders, adds hydrogen atoms, and creates disulfide bonds. The hydrogen-bonding network was optimized at neutral pH. Receptor grids were constructed with Glide using a $10 \times 10 \times 10$ Å boundary box spanning the entire ligand binding pocket and centered on the centroid of either IT1t or the Arg2–Nal3 residues of CVX15. The Arg2–Nal3 residues in the CXCR4:CVX15 structure were chosen because these two residues are mutationally diagnostic^[36] and they share the greatest contact surface with CXCR4 (Table S2 and Figure S1). Errors arising outside the binding pocket, as defined by the grids, were ignored. The CXCR4:AMD11070 receptor grid was generated by optimizing the receptor around the bioactive pose reported below. CXCR4 residues within 15.0 Å from AMD11070 were minimized using Prime MM-GBSA,^[37] and a grid was constructed in Glide from this structure with a $10 \times 10 \times 10$ Å boundary box surrounding the centroid of the ligand. The AMD11070-like ligands were prepared using the 2D Sketcher, Lig-Prep, and the MMFFs force field in Maestro, while maintaining a positive charge on the butyl amine, which was calculated to be the dominant protonation species (~60–75 %) at physiologic pH by ACD software^[38] (Figure S6). Low-energy conformers of AMD11070 monocation were generated using Macromolecule with MMFFs molecular mechanics. The global minimum was found 49 times from the 1100 structures using 5000 steps. Conformers with an energy greater than 21.0 kJ mol^{-1} (5 kcal mol^{-1}) were rejected, resulting in 17 low-energy conformations of AMD11070 in the singly protonated state. Ligands were docked flexibly into the CXCR4 receptor grids using Glide Standard Precision (Glide SP),^[39] and the Epic state-penalties were added to the Glide score. To treat the experimental data obtained from a racemic mixture for the AMD11070-like congeneric series (Figure 5), both enantiomers were docked, and the best scoring enantiomer was refined using Prime MM-GBSA^[37] with ligand strain and a distance of 0 Å.

NAMFIS analysis

AMD11070 was synthesized as described by Crawford et al. as an off-white solid (0.25 g, 55 %).^[30] The calculated $[M+H]^+$ mass of AMD11070 ($C_{21}H_{28}N_5$) was m/z 350.23392, and high-resolution MS (ESI) showed a peak at m/z 350.23399 that we assigned as AMD11070. HPLC revealed greater than 95 % purity with 256 nm UV/Vis detection (Agilent Zorbax 3.5 µM XDB C18 4.6×50 mm). The NMR sample was prepared by dissolving 11.0 mg of AMD11070 in 0.8 mL of $CDCl_3$ (Cambridge Isotope Laboratories) and subsequently degassing and sealing the tube. All NMR spectra were recorded on either a Varian Inova 600 spectrometer equipped with a 5 mm ID-ZPFG ($^1H/X$) probe or on a Varian Unity Plus 600 spectrometer equipped with a 5 mm TRES-ZPFG ($^1H/^{13}C/^{15}N$) probe. Both spectrometers use VNMR 6.1C software.^[40] All 2D experiments were conducted with temperature regulation set at 298 K. The chemical shifts were assigned to AMD11070 by standard 2D 1H , 1H and 1H , ^{13}C experiments. Spectra were processed and analyzed using the MestReNova v. 8.1.2 software package.^[41] 1H chemical shifts (Table S4) are referenced to the residual solvent peak. A NOESY spectrum was recorded at 300 ms mixing time, which is expected

to be in the linear range for small molecules. Interproton distances were calculated from crosspeak volumes using an internal calibration distance of 2.48 Å between H2 and H3. To ensure complete relaxation between scans, the relaxation delay (d1) was set to 6 s (approximately three times the longest T1 of ~2.5 s). The 31 NOE-derived distances are presented in Table S5. MCMM conformational searches (50000 steps each) of AMD11070 were performed with three individual force fields (AMBER*, MMFFs, and OPLS-2005) using the MacroModel module in Maestro.^[42] The published crystal structure stereochemistry of AMD11070^[30] was corrected to match the biologically active form (*R* to *S*) and was used as the starting structure. Searches were performed on three forms of the structure: the two stereoisomers of the protonated tertiary nitrogen and the corresponding neutral form. The GBSA/H₂O solvation model was used, along with a relaxed 30 kJ mol⁻¹ energy cutoff, resulting in a total of 10 755 conformations (4397 in AMBER*, 2861 in MMFFs, and 3497 in OPLS-2005). The global minimum was found between 22 and 214 times for the different searches, assuring complete coverage of conformational space.^[43]

ROCS analysis of AMD11070 conformations

ROCS, v. 3.2,^[44] was used to align the NAMFIS output conformations to the proposed bioactive pose or the X-ray structure with and without color atoms.^[17, 45,46] A simple ROCS run was performed for each query with the NAMFIS structures treated as a single-conformer molecule database. Color optimization was used for the two queries employing color atom scoring; whereas only shape optimization was used for the other two.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We are grateful to OpenEye Scientific Software (Santa Fe, New Mexico) for kindly providing us with access to the ROCS software. We would also like to thank Dr. Shaoxiong Wu, director of the Emory NMR facilities, for his assistance.

References

1. Wong D, Korz W. Clin Cancer Res. 2008; 14:7975. [PubMed: 19088012]
2. Huang EH, Singh B, Cristofanilli M, Gelovani J, Wei C, Vincent L, Cook KR, Lucci A. J Surg Res. 2009; 155:231. [PubMed: 19482312]
3. Thoma G, Streiff MB, Kovarik J, Glickman F, Wagner T, Beerli C, Zerwes HG. J Med Chem. 2008; 51:7915. [PubMed: 19053768]
4. Wu B, Chien EY, Mol CD, Fenalti G, Liu W, Katritch V, Abagyan R, Brooun A, Wells P, Bi FC, Hamel DJ, Kuhn P, Handel TM, Cherezov V, Stevens RC. Science. 2010; 330:1066. [PubMed: 20929726]
5. De Clercq E. Biochem Pharmacol. 2009; 77:1655. [PubMed: 19161986]
6. Moyle G, DeJesus E, Boffito M, Wong RS, Gibney C, Badel K, Mac-Farland R, Calandra G, Bridger G, Becker S. XACTS Team. Clin Infect Dis. 2009; 48:798. [PubMed: 19193109]
7. Neves MA, Simões S, Sá e Melo ML. J Comput-Aided Mol Des. 2010; 24:1023. [PubMed: 20960031]
8. Pettersson S, Pérez-Nueno VI, Mena MP, Clotet B, Esté JA, Borrell JI, Teixidó J. ChemMedChem. 2010; 5:1272. [PubMed: 20533501]

9. Bhattacharya S, Lam AR, Li H, Balaraman G, Niesen MJ, Vaidehi N. *Proteins Struct Funct Bioinf.* 2013; 81:729.
10. Wong RS, Bodart V, Metz M, Labrecque J, Bridger G, Fricker SP. *Mol Pharm.* 2008; 74:1485.
11. Mooring SR, Liu J, Liang Z, Ahn J, Hong S, Yoon Y, Snyder JP, Shim H. *ChemMedChem.* 2013; 8:622. [PubMed: 23468189]
12. Kufareva I, Rueda M, Katritch V, Stevens RC, Abagyan R. *GPCR Dock 2010 participants. Structure.* 2011; 19:1108. [PubMed: 21827947]
13. Mysinger MM, Weiss DR, Ziarek JJ, Gravel S, Doak AK, Karpiak J, Heveker N, Shoichet BK, Volkman BF. *Proc Natl Acad Sci USA.* 2012; 109:5517. [PubMed: 22431600]
14. Beuming T, Sherman W. *J Chem Inf Model.* 2012; 52:3263. [PubMed: 23121495]
15. Kawatkar SP, Yan M, Gevariya H, Lim MY, Eisold S, Zhu X, Huang Z, An J. *Exp Biol Med.* 2011; 236:844.
16. Zhang C, Hou T, Feng Z, Li Y. *Curr Comput-Aided Drug Des.* 2013; 9:60. [PubMed: 22734712]
17. Pérez-Nueno VI, Ritchie DW, Rabal O, Pascual R, Borrell JI, Teixidó J. *J Chem Inf Model.* 2008; 48:509. [PubMed: 18298095]
18. Truax VM, Zhao H, Katzman BM, Prosser AR, Alcaraz AA, Saindane MT, Howard RB, Culver D, Arrendale RF, Gruddanti PR, Evers TJ, Natchus MG, Snyder JP, Liotta DC, Wilson LJ. *ACS Med Chem Lett.* 2013; 4:1025. [PubMed: 24936240]
19. Cicero D, Barbato G, Bazzo R. *J Am Chem Soc.* 1995; 117:1027.
20. Gerlach LO, Skerlj RT, Bridger GJ, Schwartz TW. *J Biol Chem.* 2001; 276:14153. [PubMed: 11154697]
21. Rodríguez D, Gutiérrez-de-Terán H. *Proteins Struct Funct Bioinf.* 2012; 80:1919.
22. Desiraju GR. *Acc Chem Res.* 1996; 29:441. [PubMed: 23618410]
23. Szasz GJ, Sheppard N, Rank DH. *J Chem Phys.* 1948; 16:704.
24. Vila-Coro AJ, Rodríguez-Frade JM, De Ana AM, Moreno-Ortiz MC, Martínez-A C, Mellado M. *FASEB J.* 1999; 13:1699. [PubMed: 10506573]
25. Skerlj RT, Bridger GJ, Kaller A, McEachern EJ, Crawford JB, Zhou Y, Atsma B, Langille J, Nan S, Veale D. *J Med Chem.* 2010; 53:3376. [PubMed: 20297846]
26. Lyne PD, Lamb ML, Saeh JC. *J Med Chem.* 2006; 49:4805. [PubMed: 16884290]
27. Klon AE, Glick M, Thoma M, Acklin P, Davies JW. *J Med Chem.* 2004; 47:2743. [PubMed: 15139752]
28. Charifson PS, Corkery JJ, Murcko MA, Walters WP. *J Med Chem.* 1999; 42:5100. [PubMed: 10602695]
29. Planesas JM, Pérez-Nueno VI, Borrell JI, Teixidó J. *J Mol Graphics Modell.* 2012; 38:123.
30. Crawford JB, Chen G, Gauthier D, Wilson T, Carpenter B, Baird IR, McEachern E, Kaller A, Harwig C, Atsma B. *Org Process Res Dev.* 2008; 12:823.
31. Thepchattri P, Eliseo T, Cicero DO, Myles D, Snyder JP. *J Am Chem Soc.* 2007; 129:3127. [PubMed: 17323946]
32. Roumen L, Scholten DJ, de Kruijf P, de Esch IJP, Leurs R, de Graaf C. *Drug Discovery Today Technol.* 2012; 9:e281.
33. Sobolev V, Sorokine A, Prilusky J, Abola EE, Edelman M. *Bioinformatics.* 1999; 15:327. [PubMed: 10320401]
34. Humphrey W, Dalke A, Schulten K. *J Mol Graphics.* 1996; 14:33.
35. L. Schrodinger 9.2 ed. New York, NY: 2011.
36. Tamamura H, Omagari A, Oishi S, Kanamoto T, Yamamoto N, Peiper SC, Nakashima H, Otaka A, Fujii N. *Bioorg Med Chem Lett.* 2000; 10:2633. [PubMed: 11128640]
37. Maestro, Schrodinger LLC 2.1 ed. New York, NY: 2009.
38. ACD/pKa, Version 12.5. Advanced Chemistry Development, Inc; Toronto, ON, Canada: 2012. <http://www.acdlabs.com/>
39. Friesner RA, Banks JL, Murphy RB, Halgren TA, Klicic JJ, Mainz DT, Repasky MP, Knoll EH, Shelley M, Perry JK. *J Med Chem.* 2004; 47:1739. [PubMed: 15027865]
40. Agilent (Varian), 6.1C ed. Palo Alto, CA: 2001.

41. MestReNova, M. Research, version 8.1.2, 2013.
42. Prime, Schrodinger LLC 9.9 ed. New York, NY: 2011.
43. Hand CE, Auzanneau FI, Honek JF. Carbohydr Res. 2006; 341:1164. [PubMed: 16630596]
44. Hawkins PC, Skillman AG, Nicholls A. J Med Chem. 2007; 50:74. [PubMed: 17201411]
45. Nicholls A, MacCuish NE, MacCuish JD. J Comput-Aided Mol Des. 2004; 18:451. [PubMed: 15729846]
46. Grant JA, Gallardo M, Pickup BT. J Comput Chem. 1996; 17:1653.

Figure 1. Heavy-atom RMSD values of CXCR4 binding pocket residues between the CXCR4:IT1t and CXCR4:CVX15 crystal structures. Residues in blue exhibit small RMSD values (0–1 Å), residues in light red show moderate RMSD values (1–2 Å), and residues in dark red display large RMSD values (>2 Å).

Figure 2.
Conformational change of Gln200 and His203 on helix V observed between the small molecule CXCR4:IT1t (light gray) and the CXCR4:CVX15 (black) crystal structures.

Figure 3.

A) The proposed bioactive pose of AMD11070 (gray/blue ball and stick) from modeling into the CXCR4:CVX15 crystal structure (gray ribbon). Interactions between AMD11070 and key residues are highlighted: B) Ala171 (helix IV), His113, and Thr117 (helix III), C) Glu288 (helix VII), and D) Asp97 and Tyr94 (helix II).

Figure 4.

Correlation of per-residue energy of interaction from AMD11070 with pK_i^{MUT} (decrease in SDF-1 α K_i upon indicated mutation taken from Wong et al. 2008) from A) CXCR4:IT1t, pose 1 or B) CXCR4:CVX15, pose 1.

1a ($n=2$)
1b ($n=3$)
1c ($n=4$)
1d ($n=5$)
1e ($n=6$)

Figure 5.

Correlation of MM-GBSA ΔG_{Bind} for an AMD11070-like congeneric series.

Chemical structures of the series of interest docked into A) CXCR4:CVX15 receptor and B) CXCR4:AMD11070 receptor.

Figure 6. Comparison of the bioactive pose of AMD11070 from the CXCR4:CVX15 crystal structure (lighter) with A) the AMD11070 small molecule crystal structure (darker; Tanimoto shape score of 0.639); B) NAMFIS-3 (darker; 14.3 %, Tanimoto shape score of 0.792).

Scheme 1.
Chemical structures of selected CXCR4 antagonists.

Table 1

Analysis of the top three poses of AMD11070 docked into CXCR4.

Crystal structure	Glide score	Per-residue interaction energy [kcal mol ⁻¹]			
		Glu288	Asp171	Asp97	Trp94
CXCR4:IT1t	-7.7	-29.0	-15.6	-50.1	-7.2
CXCR4:IT1t	-7.2	-33.2	-14.9	-53.4	-6.9
CXCR4:IT1t	-7.2	-30.8	-14.5	-25.1	-6.5
CXCR4:CVX15	-6.4	-32.9	-49.2	-17.7	-2.6
CXCR4:CVX15	-6.0	-35.6	-16.3	-34.4	-2.7
CXCR4:CVX15	-5.8	-20.4	-15.2	-50.7	-4.1