Genetically Encoded Photo-cross-linkers Map the Binding Site of an Allosteric Drug on a G Protein-Coupled Receptor

Amy Grunbeck,† Thomas Huber,† Ravinder Abrol,‡ Bartosz Trzaskowski,‡ William A. Goddard, III,‡ and Thomas P. Sakmar*†

‡Laboratory of Molecular Biology & Biochemistry, The Rockefeller University, 1230 York Ave., New York, New York 10065, United States

§Materials and Process Simulation Center (MC 139-74), California Institute of Technology, 1200 E. California Blvd., Pasadena, California 91125, United States

Supporting Information

ABSTRACT: G protein-coupled receptors (GPCRs) are dynamic membrane proteins that bind extracellular molecules to transduce signals. Although GPCRs represent the largest class of therapeutic targets, only a small percentage of their ligand-binding sites are precisely defined. Here we describe the novel application of targeted photo-cross-linking using unnatural amino acids to obtain structural information about the allosteric binding site of a small molecule drug, the CCR5-targeted HIV-1 co-receptor blocker maraviroc.

Many G protein-coupled receptor (GPCR)-modulated signaling pathways are involved in human disease, and GPCRs are a major target class for small molecule therapeutics and biologicals.1 Some GPCR-targeted drugs bind to orthosteric sites and inhibit the binding interactions with endogenous agonist ligands that are necessary to form productive signaling complexes. However, GPCR-mediated signaling can also be affected by allosteric modulators.2,3 Targeting potential allosteric sites on GPCRs opens up new avenues for structure-based drug design.1 To facilitate future drug development and to understand the mechanism of action of existing drugs, it is important to identify the binding sites of both orthosteric and allosteric GPCR modulators.4,5 However, it is not straightforward to identify the precise binding sites and mechanisms of action of GPCR ligands. For example, it is known that GPCRs bound to antagonists or agonists display different packing within the 7-helix bundles, and it is likely that receptor activation proceeds through a series of agonist-GPCR conformations. In addition, computational predictions suggest that binding of various antagonists to CCR5 stabilizes different packings within the 7-helix bundle, which might lead to changes in function (e.g., coupling to β-arrestin or G protein). Here, we demonstrate the application of a chemically precise technique in live cells to define the binding site of a small molecule drug to an important GPCR drug target, and we reconcile the results with computational predictions.

Maraviroc, a small molecule HIV-1 entry inhibitor, is a GPCR allosteric modulator with inverse agonist activity (Figure 1a).6 CC chemokine receptor 5 (CCR5), the molecular target for maraviroc,7 is the primary co-receptor required for HIV-1 cellular entry.8,9 Maraviroc is the first GPCR-specific HIV-1 entry inhibitor to be approved by the FDA for therapeutic use, but its precise receptor–drug binding site interactions have not been defined, and the mechanism for how the drug blocks HIV-

Received: February 8, 2012
Accepted: March 28, 2012
The targeted photo-cross-linking technology we used to investigate the maraviroc binding site is an adaptation of a method we previously developed to identify residues in CXC chemokine receptor 4 (CXCR4) that were within about 3 Å of its binding partner in order to form a covalent complex after photoactivation. This distance constraint is also consistent with earlier studies using benzophenone-based cross-linkers. The nature of the photoreactive species formed from phenylazide is controversial, but we make the assumption that the carbonyl in the benzophenone moiety of BzF needs to be within 3 Å from its binding partner in order to form a covalent bond.

Several earlier studies proposed models of the maraviroc binding site in CCR5 based on scanning mutagenesis, molecular dynamics simulations, or structure–activity relationship studies. Some of this work was based on studies of other structurally related CCR5 pharmacophores such as TAK-779. From these earlier data, we chose 8 positions in CCR5 for additional study (Figure 1c). We chose these positions with the expectation, based on earlier mutagenesis data, that when mutated to either BzF or azF, a loss-of-function phenotype would not be observed. Our aim was to create a set of UAA mutant receptors that would bind maraviroc normally and in addition contain reactive side chains that might cross-link upon UV excitation.

We expressed each of the selected CCR5 mutants in HEK293T cells under conditions designed to incorporate either BzF or azF at the position of interest. Cells expressing the CCR5 UAA mutants were incubated with [3H]-maraviroc and then exposed to UV light. Cells were lysed in detergent buffer solution, and receptors were immunopurified, run on an SDS-PAGE gel, transferred to a PVDF membrane, and immunoblotted to determine receptor expression. The sample lanes were then cut from the PVDF membrane, and each sample lane was cut into three different molecular weight segments. These membrane segments were then put into scintillation vials with scintillation fluid and then counted on a beta-scintillation counter to quantify the amount of radioactivity in each segment.

The targeted photo-cross-linking technology we used to investigate the maraviroc binding site is an adaptation of a method we previously developed to identify residues in CXC chemokine receptor 4 (CXCR4) that were within about 3 Å of a fluorescently labeled analogue of its peptide ligand, T140. The results from this experiment were benchmarked using the crystal structure of CXCR4 bound to a peptide ligand, CVX15, homologous to T140. The obvious limitation of this earlier work was that in essence we mapped the binding site of maraviroc on CCR5 to a minimally altered receptor.

The targeted photo-cross-linking technology that probes the binding of [3H]-labeled maraviroc to a minimally altered receptor.

The targeted photo-cross-linking technology we used to investigate the maraviroc binding site is an adaptation of a method we previously developed to identify residues in CXC chemokine receptor 4 (CXCR4) that were within about 3 Å of a fluorescently labeled analogue of its peptide ligand, T140. The results from this experiment were benchmarked using the crystal structure of CXCR4 bound to a peptide ligand, CVX15, homologous to T140. The obvious limitation of this earlier work was that in essence we mapped the binding site of maraviroc on CCR5 to a minimally altered receptor.

The targeted photo-cross-linking technology that probes the binding of [3H]-labeled maraviroc to a minimally altered receptor.

The targeted photo-cross-linking technology we used to investigate the maraviroc binding site is an adaptation of a method we previously developed to identify residues in CXC chemokine receptor 4 (CXCR4) that were within about 3 Å of a fluorescently labeled analogue of its peptide ligand, T140. The results from this experiment were benchmarked using the crystal structure of CXCR4 bound to a peptide ligand, CVX15, homologous to T140. The obvious limitation of this earlier work was that in essence we mapped the binding site of maraviroc on CCR5 to a minimally altered receptor.

The targeted photo-cross-linking technology we used to investigate the maraviroc binding site is an adaptation of a method we previously developed to identify residues in CXC chemokine receptor 4 (CXCR4) that were within about 3 Å of a fluorescently labeled analogue of its peptide ligand, T140. The results from this experiment were benchmarked using the crystal structure of CXCR4 bound to a peptide ligand, CVX15, homologous to T140. The obvious limitation of this earlier work was that in essence we mapped the binding site of maraviroc on CCR5 to a minimally altered receptor.

The targeted photo-cross-linking technology we used to investigate the maraviroc binding site is an adaptation of a method we previously developed to identify residues in CXC chemokine receptor 4 (CXCR4) that were within about 3 Å of a fluorescently labeled analogue of its peptide ligand, T140. The results from this experiment were benchmarked using the crystal structure of CXCR4 bound to a peptide ligand, CVX15, homologous to T140. The obvious limitation of this earlier work was that in essence we mapped the binding site of maraviroc on CCR5 to a minimally altered receptor.

The targeted photo-cross-linking technology we used to investigate the maraviroc binding site is an adaptation of a method we previously developed to identify residues in CXC chemokine receptor 4 (CXCR4) that were within about 3 Å of a fluorescently labeled analogue of its peptide ligand, T140. The results from this experiment were benchmarked using the crystal structure of CXCR4 bound to a peptide ligand, CVX15, homologous to T140. The obvious limitation of this earlier work was that in essence we mapped the binding site of maraviroc on CCR5 to a minimally altered receptor.
higher tritium levels in the 25−50 kDa segment. (b) The same positions in CCR5 were also tested for cross-linking using BzF. This data set is displayed as in panel a. Positions 28 and 109 were found in this cross-linking scan to have significantly higher tritium levels in the 25−50 kDa segment.

W86 was replaced by azF. However, when BzF was incorporated at position I28, only a slight increase above background was detected. A significant increase in tritium levels was detected, however, when BzF was incorporated at position F109, a site that did not cross-link when azF was present.

For these 4 samples (I28azF-CCR5-maraviroc; W86azF-CCR5-maraviroc; I28BzF-CCR5-maraviroc; F109BzF-CCR5-maraviroc) the elevated levels of tritium were detected only in the 25−50 kDa gel fraction, which includes CCR5 with an apparent molecular mass of approximately 37 kDa. Since the molecular mass of maraviroc is under 500 Da, tritium only appears in the 25−50 kDa fraction if the [3H]-maraviroc is covalently bound to CCR5. These cross-linking results were also found to be UV-treatment-specific (Supplementary Figures 1 and 2). From this information, we conclude that the elevated levels of tritium in the 25−50 kDa gel fractions are a result of a direct UV-dependent cross-link between the UAA incorporated into CCR5 and [3H]-maraviroc. The cross-linking patterns observed with BzF and azF may differ as a result of either the structural orientation or reactivity of their unique side chains. Alternatively, the structural difference between the azido group and benzoyl group might have had different effects on ligand affinity for the mutant CCR5.

We performed ligand-binding experiments with [3H]-maraviroc to determine if introducing a UAA at any of the positions we tested altered maraviroc binding. Only the W86BzF-CCR5 mutant showed significantly decreased levels of binding to maraviroc (Supplementary Figure 3). This result suggests that W86BzF-CCR5 may have failed to cross-link to maraviroc because of decreased ligand binding affinity. On the other hand, we also detected different cross-linking results at position 109 for the two cross-linkers. The binding data suggest that maraviroc is able to bind to both the F109BzF-CCR5 and F109azF-CCR5 mutants. Therefore, we can conclude that the difference in cross-linking between the two photoactivable groups at position 109 was a result of a difference in reactivity of the photoactivatable cross-linking groups. Other relevant factors that might affect the reactivity of these cross-linking groups include the chemical bonds of maraviroc near to the reactive moiety and the specific spatial orientation of the side chain.

We further evaluated our cross-linking results using a newly predicted structure of the CCR5−maraviroc complex, which was obtained by using the GenSeMBLE structure prediction methodology.24 This methodology (summarized in the Supporting Information) is a Monte Carlo based method aimed at sampling all reasonable packings of the 7-helix bundle (we examined 10 trillion) and selecting an ensemble of 100 low-energy packings likely to play a role in binding of various ligands and in the activation process. The lowest 20 of these structures were used to predict the most stable CCR5−maraviroc complex, which was used to understand the experimental observations. We selected this predicted structure to calculate 100 different conformations of the BzF and azF side chains at each of the eight positions in CCR5 that were tested.

Figure 2. continued
in the cross-linking studies. Distance measurements were then made from either the carbonyl in BzF or the nitrogen adjacent to the phenyl ring in azF to the nearest possible non-hydrogen atom in maraviroc. We then calculated a probability distribution for the side chain at each of the indicated positions to come within 3 Å of contacting a non-hydrogen atom in maraviroc (Figure 3). According to the predicted structures, the only positions in CCR5 with a reasonable probability of being within 3 Å of the bound maraviroc were W86, Y108, and F109.

The experimental cross-linking data are consistent with the predictions based on the W86 and F109 positions in the CCR5–maraviroc complex. On the other hand, our predicted model suggested that Y108 would be within cross-linking distance of bound maraviroc, but no experimental cross-link was observed under the conditions tested. This discrepancy between predictions based on the complex model and the experimental data could mean that the complex model is not entirely correct (although the other sites were accurately predicted). One possible explanation is that CCR5 might adopt a different low-energy conformation for Y108BzF and Y108azF mutants. This is plausible because we found that the Y108A mutant does adopt a different transmembrane bundle packing when bound to maraviroc (data not shown). If the CCR5–maraviroc complex for the UAA-containing protein does not change, a second possibility for the discrepancy could be that the photo-cross-linkers at position 108 were not able to react due to either a suboptimal orientation of the side chain or the lack of an appropriate target bond in maraviroc. Both of these possible explanations are plausible because Y108 side chain lies in a tight region near maraviroc, whereas both W86 and F109 reside in a more open region. Thus, to accommodate the BzF and azF UAAAs at the Y108 position, maraviroc may have to move to a different (but may be proximal to original) binding site not accessible to the photo-cross-linkers. Our computational methods can be used to test these suggestions, but we have not yet done so. In addition, position 28, which formed cross-links in both the 128azF-CCR5 and 128BzF-CCR5 mutants, was not identified in the model to be within cross-linking distance of maraviroc. This result again could indicate limitations of the model. However, given the location of I28 in the N-terminal tail of CCR5, we favor the possibility that the experimental cross-link could suggest an additional maraviroc-binding site, which has also been predicted before by others and us.6,25 The location of I28, is consistent with a maraviroc “docking site” on the extracellular surface of CCR5 and might be relevant for understanding its mechanism of action with respect to its function as an HIV-1 entry blocker. To further evaluate the exact orientation of maraviroc in the CCR5 ligand-binding pocket, we are in the process of identifying the locations of the cross-links on the maraviroc molecule. In addition, we will examine the effect of UAAAs on the structure of CCR5 conformations.

We report here, to our knowledge, the first demonstration of a direct chemical cross-link between a GPCR and its native small-molecule ligand. We show that the targeted photo-cross-linking technology using UAA mutagenesis can be applied to identify the binding site of a small molecule ligand and that tritium is a sensitive detection tag for small-scale cell-based assays. We used the cross-linking results to evaluate a model of the CCR5–maraviroc complex and confirmed that maraviroc binds within the transmembrane helix bundle. This method should prove to be valuable for obtaining structural information about the binding site of GPCR allosteric modulators. In principle, the method should be applicable to any receptor–ligand complex where the receptor can be heterologously expressed in mammalian cells in culture and isotopically labeled ligands are available.

METHODS

Materials. [3H]-Maraviroc was a generous gift from Bill Goddard and PharmSelex. 1D4 monoclonal antibody was obtained from the National Cell Culture Center. The goat anti-mouse antibody conjugated to horseradish peroxidase was purchased from VWR (catalogue no. 95059-094).

Site-Directed Mutagenesis and Plasmid Construction. The suppressor tRNA and BzF and azF amino-acyl tRNA synthetase plasmids were constructed as previously described.3,15 The human CCR5 gene was in a pcDNA3.1(+) plasmid and contained a C-terminal 1D4-epitope, TETSQVAPA. The amber stop codons were introduced into CCR5 using the Quikchange Lightning Site-Directed Mutagenesis kit (Stratagene).

Expression of CCR5 Unnatural Amino Acids Mutants in Mammalian Cells. CCR5 UAA containing mutants were expressed in HEK293T cells as previously described.12,35 In brief, HEK293T cells were transfected with three cDNA simultaneously using the lipofectamine plus reagent (Invitrogen). These cDNA contained the genes for the suppressor tRNA, amino-acyl tRNA synthetase, and human CCR5 amber mutant. The ratio of DNA in micrograms was 1:0.1:1 CCR5 amber mutant/tRNA synthetase/suppressor tRNA. For transfection of...
one 10 cm plate, 3.5 μg of the CCR5 amber mutant DNA was used. WT CCR5 was transfected using 0.14 μg per 10 cm² plate in order to obtain comparable expression levels to the CCR5 UAA mutants. Cells were then grown in media containing 10% FBS with 0.5 mM of UAA in Dulbecco’s Modified Eagle’s Medium (DMEM; 4.5 g/L glucose, 2 mM glutamine; Gibco). Cells were used for photo-cross-linking experiments 48 h post-transfection.

**Photo-cross-linking of Cells Expressing CCR5 UAA Mutants to [3H]-Maraviroc.** Two days post-transfection HEK293T cells were suspended in Dulbecco’s phosphate buffered saline (DPBS; Gibco). Cells were then spun down and resuspended in Hank’s Buffered Salt Solution (HBSS; pH 7.5) containing 20 mM HEPES with 0.2% BSA and 100 nM [3H]-Maraviroc. Cell suspensions were then incubated for 2 h at 37 °C on a nutator. After the incubation cell suspensions were transferred to a 24 well plate and exposed to a Maxima ML-3500S UV-A light (Spectronics Corporation) in a 4 °C cold room on ice for 15 min. After UV light exposure cells were transferred to eppendorf tubes and pelleted, and supernatant was removed. Cell pellets were stored at −20 °C until further analysis.

**Western Blot Analysis.** Cell pellets were solubilized in 1% (w/v) n-dodecyl β-β-maltoside in 20 mM Tris pH 7.5 containing protease inhibitors for 1 h at 4 °C on a nutator. After solubilization the lysate was then spun down at 20,000 × g for 5 min. The supernatant was then bound to sepharose beads conjugated to the 1D4 antibody overnight at 4 °C. The next day the beads were washed three times with the 1% n-dodecyl β-β-maltoside buffer. Samples were then eluted from the beads by shaking the beads for 1 h at 37 °C in 1% SDS. Fifteen microliters of the eluted sample was then mixed with 5 μL of 4X NuPAGE LDS sample buffer containing DTT and then run on an SDS-PAGE gel (NuPAGE Novex 4–12% Bis-Tris Gel). The remaining eluted sample was stored at −20 °C for further characterization. The SDS-PAGE gel was then transferred to a polyvinylidene difluoride (PVDF) membrane for immunoblotting. The PVDF membrane was blocked in 5% milk in 1X TBST for 1 h at RT on a shaker followed by incubation with the primary monoclonal 1D4 antibody overnight at 4 °C. After the incubation the beads were washed three times with the 1% n-dodecyl β-β-maltoside buffer. Samples were then eluted from the beads by shaking the beads for 1 h at 37 °C in 1% SDS. Fifteen microliters of the eluted sample was then mixed with 5 μL of 4X NuPAGE LDS sample buffer containing DTT and then run on an SDS-PAGE gel (NuPAGE Novex 4–12% Bis-Tris Gel). The remaining eluted sample was stored at −20 °C for further characterization. The SDS-PAGE gel was then transferred to a polyvinylidene difluoride (PVDF) membrane for immunoblotting. The PVDF membrane was blocked in 5% milk in 1X TBST for 1 h at RT on a shaker followed by incubation with the primary monoclonal 1D4 antibody overnight at 4 °C.

**Quantification of [3H]-Maraviroc Bound and Cross-Linked to CCR5 Mutants.** After the PVDF membrane was exposed to film, the membrane was then cut into segments to quantify the amount of tritium present. The membrane was cut between each lane to separate each sample; in addition each sample was cut into three different molecular weight segments as specified. Each of these membrane segments were then transferred to individual vials with scintillation fluid and counted on a LKB Wallac 1209 Rackbeta Liquid Scintillation Counter (Perkin-Elmer). To quantify the amount of [3H]-maraviroc that was bound to each of the CCR5 UAA mutants, 15 μL of the original eluted sample from the 1D4 mAb sepharose beads was transferred directly into a vial with scintillation fluid. All of the scintillation vials were mixed thoroughly before being counted on the beta-scintillation counter. Each vial was counted over 10 min. After counting, the counts per minute (cpm) per cm² of PVDF membrane was then calculated for each membrane segment by dividing the cpm for a particular sample by the area of the membrane segment in cm². The amount of [3H]-maraviroc each CCR5 UAA was able to bind was normalized to the other mutants by setting the WT sample as 100% and then calculating for the UAA mutants what percentage they were of WT.

**Molecular Modeling of CCR5 UAA Mutants.** The CCR5–maraviroc complex model obtained as described above was used to predict 100 conformations of the azF and BzF side chains for each site examined in CCR5 using modeler. These structures were analyzed for their potential to form a cross-link with maraviroc by determining the distance from the carbonyl in BzF or the nitrogen adjacent to the phenyl ring in azF to the nearest non-hydrogen atom in maraviroc. The contacts were calculated from the cumulative distribution function at a distance of 3 Å. The distance measurements and molecular graphics were made using VMD.27

**REFERENCES**


