



Identification of non-peptide CCR5 receptor agonists by structure-based virtual screening

Esther Kellenberger, Jean-Yves Springael, Marc Parmentier, Muriel Hachet-Haas, Jean-Luc Galzi, Didier Rognan

► To cite this version:

Esther Kellenberger, Jean-Yves Springael, Marc Parmentier, Muriel Hachet-Haas, Jean-Luc Galzi, et al.. Identification of non-peptide CCR5 receptor agonists by structure-based virtual screening. *Journal of Medicinal Chemistry*, 2007, 50 (6), pp.1294-1303. 10.1021/jm061389p . hal-02196836

HAL Id: hal-02196836

<https://hal.science/hal-02196836>

Submitted on 29 Jul 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Identification of non-peptide CCR5 receptor agonists by structure-based virtual screening

Esther Kellenberger,^{1} Jean-Yves Springael,² Marc Parmentier,² Muriel Hachet-Haas,³ Jean-Luc Galzi³ and Didier Rognan¹*

¹ Bioinformatics of the Drug, CNRS UMR 7175-LC1, F-67400 Illkirch

² Institut de Recherche Interdisciplinaire en Biologie Humaine et Moléculaire (IRIBHM),
Université Libre de Bruxelles, B-1070 Brussels, Belgium.

³ Dynamics of G protein-coupled receptors, CNRS UMR7175-LC1, F-67400 Illkirch

Short title: in silico discovery of CCR5 non-peptide agonists.

* To whom correspondence should be addressed: Phone: +33-3-90 24 42 21, Fax: +90 24 43 10, E-mail: esther.kellenberger@pharma.u-strasbg.fr

Abstract

A three-dimensional model of the chemokine receptor CCR5 has been built to fulfill structural peculiarities of its α -helix bundle and to distinguish known CCR5 antagonists from randomly-chosen drug-like decoys. In silico screening of a library of 1.6 million commercially-available compounds against the CCR5 model by sequential filters (drug-likeness, 2-D pharmacophore, 3-D docking, scaffold clustering) yielded a hit list of 59 compounds, out of which 10 exhibited a detectable binding affinity to the CCR5 receptor. Unexpectedly, most binders tested in a functional assay were shown to be agonists of the CCR5 receptor. A follow up database query based on similarity to the most potent binders identified three new CCR5 agonists. Despite a moderate affinity of all non-peptide ligands for the CCR5 receptor, one of the agonist was shown to promote efficient receptor internalization, which is a process therapeutically favorable for protection against HIV-1 infection.

Keywords: HIV, CCR5, GPCR, agonist, docking, virtual screening

Introduction

G protein-coupled receptors (GPCRs) are among the most investigated therapeutic target families. There are numerous reasons that make GPCRs attractive for the pharmaceutical industry: (i) GPCRs select their endogenous ligands from a uniquely broad chemical space¹ which serves the design of new receptor modulators;² (ii) GPCRs have an ubiquitous distribution and regulate a wide array of physiological and pathological processes;³ (iii) over the last 30 years, many GPCRs have proven to be druggable,⁴ and there is a reasonable hope that present and future drug discovery programs will result in the development of drugs acting on other receptors, especially those that are close to well-investigated targets.⁵ Notably, the high-resolution crystal structure of bovine rhodopsin⁶ has paved the way for structure-based design of GPCR ligands. Hence, virtual screening of chemical libraries, which has been successfully used over the last 10 years⁷ to design ligands for enzyme targets, may be applied to GPCR homology models based on the rhodopsin template. However, whether GPCR models are reliable starting points is still a matter of debate,⁸ mainly because most GPCRs exhibit low sequence similarity (typically less than 30% identity) with bovine rhodopsin. Four recent reports support the optimistic alternative: they describe the use of homology models of rhodopsin-like receptors for structure-based discovery of drug-like sub-micromolar antagonists.⁹⁻¹² In all cases, previous knowledge regarding the interaction between the GPCR and known ligands was necessary to fine tune the receptor model. Moreover, the choice of a relevant pharmacophore hypothesis was a key factor to downsize a hit list prior to molecular docking. Last, a visual inspection was necessary to ensure that key intermolecular interactions were established with selected hits.

Identifying GPCR agonists by structure-based *in silico* screening is more challenging, because GPCRs are expected to undergo significant conformational changes upon activation.¹³ Retrospective studies that involve the customization of a GPCR active state 3-D

models by either discrete rotation of key TM helices¹⁴ or experimentally-driven restrained molecular dynamics¹⁵ allowed the recovery of true agonists that were seeded with drug-like random decoys. However, the structure-based discovery of novel GPCR agonists by database screening has not been reported yet. Notably, designing non-peptide surrogates for endogenous GPCR peptide ligands is highly desirable.^{16,17} In the present paper, we describe the successful virtual screening of a ligand library against a presumed ground state GPCR homology model. By searching for antagonists of the chemokine CCR5 receptor as an attempt to derive anti-HIV agents,¹⁸ we report the unexpected identification of non-peptide CCR5 agonists.

Results and discussion

Modeling the antagonist-bound form of the CCR5 transmembrane (TM) domain

The 3-D structure of the CCR5 TM domain was built using restraints from the crystal structure of bovine rhodopsin bound to cis-retinal,⁶ combined with data from molecular dynamics simulations.^{19,20} The four amino acids of the first extracellular loop (ECL1) were included in the model since the small size of the peptide connecting TMs 2 and 3 greatly limits the arrangement of the two helices. Noteworthy is the TxP motif in TM2¹⁹; it induces a kink that significantly differs from the bending observed in the bovine rhodopsin template. As a consequence, the distance between TMs 3 and 5 is shortened.

An antagonist binding site was developed by the manual data-driven docking of five well characterized antagonists (Figure 1, compounds **1-5**) into the receptor cavity, then refined by energy-based optimization of the complexes. The cavity coated with one layer of water molecules bears a size of about 500 Å³ and a mouth of 70 Å² (as computed with Pocket²¹). Eighty percent of the side chains exposed at the cavity surface are hydrophobic or aromatic (Figure 2A). All exposed polar residues (Asn/Asp/Gln/Glu/Thr) are located near the pocket

mouth, except Asn6.52 and Glu7.39. The negatively charged residue Glu7.39, which is likely to recognize the protonable amine of the receptor antagonists, is found in the vicinity of an aromatic floor in the center of the cavity. This residue actually spans the junction between two hydrophobic subsites (Figure 2B). The first subsite is shallow and widely opened. It is delimited by TMs 1, 2, 3 and 7 (Ala1.31, Leu1.35, Tyr1.39, Trp2.60, Ala2.64, Trp94, Met3.24, Leu 3.28 and Tyr 3.32). The second subsite is more embedded and deeper. It involved TMs 3, 5 and 6 (Tyr3.32, Phe3.33, Phe3.36, Phe3.37, Leu5.47, Trp6.48, Tyr6.51, Val6.54 and Leu6.55).

The CCR5 TM cavity definition, as well as the predicted binding mode for compounds **1-5**, was principally assisted by the only mutagenesis study available when we build the model ²². The investigation of 106 positions allowed a coherent description of a transmembrane binding pocket that was consistent with studies published later. Four recent site directed mutagenesis studies agree that Glu7.39 in TM7, together with few specific aromatic and hydrophobic residues in TMs 1, 2, 3, 5 and 6, mediate the interaction between the CCR5 receptor and antagonists **1**,^{23,24} **2-3**²³⁻²⁵ and **4-5**²⁶ (Figure 2A). These positions are in general agreement with bovine rhodopsin amino acids whose side chains delimit the retinal binding site (Supplementary Figure S1). The only significant mismatch occurred in TM2, where the TxP motif is presumed to stress the helix kink in CCR5. In the refined model, the side chain of Trp2.60, whose mutation in alanine was shown to dramatically decrease the binding of all five antagonists, is directed toward the cavity whereas it would face the lipid bilayer in a homology model obtained by direct threading to the rhodopsin template. It is worth noting that experimental findings suggested that the mutated CCR5 receptors described in the literature remain, to some extent, functional, so that effects observed upon mutation may be due to the loss of key interaction(s) between the ligand and the receptor rather than to the incapability of the ligand to bind to a misfolded or unstable protein.

Customizing a CCR5 model for virtual screening

Preliminary virtual screening experiments demonstrated that the CCR5 model is reliable enough to discriminate known antagonists from randomly-chosen drug-like decoys. Docking of the "validation dataset" (containing 1,000 molecules, including seven known antagonists 1-7) performed by Gold and Surflex reached satisfactory balance of sensitivity (selection of true positives) and specificity (discard of true negatives). The selection of the 5% best scored molecules produced hit lists that contain about half of the receptor antagonists (Figure 3, Supplementary Table 1). In the case of Gold docking, the poses of antagonists found among the top scorers (typically, with fitness values higher than 47) were in general agreement with experimental data, whereas unrealistic poses got poorer scores. From a qualitative point of view, Gold better performed for antagonists that were not expected to be deeply buried in the second subsite. Although better statistics were obtained for Surflex docking, scores assigned to the antagonists did not reflect the docking quality. Typically, best ranked antagonists exhibited unrealistic poses. The most accurate poses, although not well scored, were obtained for compounds expected to be deeply buried in the second subsite. It was therefore difficult to understand the good performance of the program. We nevertheless decided to use Surflex for prospective screen because the properties and binding modes of molecules selected by this program would most probably differ from that found using Gold.

Virtual screening

The virtual screening flowchart is summarized in Figure 4. First, the ligands library setup (or "screening dataset") implied the merging of data available from various suppliers, followed by numerous filtering steps in order to discard non drug-like, toxic or reactive compounds as well as frequent hitters. A subset was extracted by picking molecules that meet

a simple 2-D pharmacophore (aromatic ring and protonable amine) that is shared by all known CCR5 antagonists. A total of 44,524 compounds were screened in parallel by high throughput docking using both Gold and Surflex. The post-processing procedure was set up in the light of the results obtained for the preliminary retrospective screening of the "validation dataset". For Gold prospective screen, the fitness threshold was thus set to 47. This cutoff yielded the selection of about 5% of the entire database. For Surflex prospective screen, we set the pKi threshold to 8 (together with a bump score limit of -40) so that about 5% of the docked database is selected too. In this way we could apply the same postprocessing rules to both docking runs. Less than hundred compounds were common to Surflex and Gold hitlists (and among them only 5 shared compatible poses), thereby advocating separate processing of Surflex and Gold runs rather than using consensus methods. Potential hits of the two lists were independently classified by graph-based maximum common substructures (MCS). A post-processing procedure retained only classes (MCS) enriched enough in well-scored molecules. Compound sampling followed by visual inspection of predicted binding mode returned a final list of 77 molecules.

Validated hits are mostly competitive agonists of the receptor

From the 77 virtual hits, 59 compounds could be purchased and tested experimentally for their ability to bind to the CCR5 receptor expressed in CHO_K1 cells and to modify the receptor functional response in an aequorin luminescence assay. The binding properties of the compounds to the CCR5 receptor were measured in a competition binding assay, using radiolabeled MIP1- β (CCL4) as a tracer, and the compounds at two concentrations (10 and 100 μ M). The functional properties of the binders were then tested (at 1, 10 and 100 μ M) by measuring intracellular calcium release, using a reporter assay based on light emission by the

apoequorin-coelenterazine complex in cell lines expressing the CCR5 receptor together with apoequorin.

Out of the 59 compounds tested, 10 exhibited detectable binding activity to the CCR5 receptor in the primary screening and 6 were picked based on their functional response. Five of the hits were common to both assays, and the 10 hits were confirmed in subsequent dose response curves. Unexpectedly, most binders were characterized as agonists in the aequorin-based functional assay. Although many of the compounds displayed weak affinities and/or potencies, four molecules, namely **8-11**, exhibited affinities in the high micromolar range (Figure 5, Table 1, and Supplementary Figure S2) and compounds **8**, **10** and **11** were able to stimulate CCR5-mediated calcium release at similar concentrations.

The three most potent ligands **8**, **9** and **11** share common pharmacophoric features: (i) a central nitrogen-rich scaffold (triazole, benzotetrahydropiperazinone, and benzimidazole), (ii) a protonable nitrogen atom at the extremity of an aliphatic side chain, (iii) one aromatic ring connected to the scaffold by linker groups of variable length. However, Tanimoto similarity coefficients estimated from FCFP_4 fingerprints²⁷ indicate that the investigated compounds (**1-11**) are rather diverse. Moreover, computed chemical fingerprints do not allow the distinction of CCR5 antagonists (compounds **1-7**, **9**) from agonists (compounds **8**, **10**, **11**).

The predicted binding mode of compounds **8**, **9** and **11** is described in Figure 6. Like antagonists, agonists are all anchored to the cavity by a salt bridge between the protonated amine of the ligand and Glu7.39 carboxylate. Aromatic rings of compounds **9** and **11** are embedded in the subsite 2 (TMs 3, 5 and 6). By contrast, aromatic interactions observed for compound **8** involve Trp and Tyr residues of subsite 1 (TMs 1, 2 and 3). In addition, the amide oxygen atom of **8** is H-bonded to the Trp93 side chain (EL1). It is however worth noting that compound **8** comes from the Gold screen whereas compounds **9** and **11** are hits from the Surflex screen, thereby the observed binding modes may be biased by docking

program hallmarks. Site-directed mutagenesis of residues predicted to interact with the described agonists is however mandatory to validate the hypothesis of overlapping binding sites.

Ground state GPCR homology models are suitable for retrieving full agonists.

CCR5 endogenous ligands are large peptides, namely chemokines like macrophage inflammatory protein -1 α (MIP-1 α , CCL3), MIP-1 β (CCL4), RANTES (CCL5) and MCP-2 (CCL7). The chemokine core domain contacts the receptor extracellular domains²⁸, while the N-terminus of several chemokines has been proposed to mediate receptor activation by interacting with the transmembrane helix bundle.²⁹ The CCR5 receptor TM cavity is known to accommodate small molecules that inhibit the receptor response to chemokines by an allosteric mechanism.^{30,31} The present work proves the utility of the modeled TM domain structure for the discovery of non-peptide modulators of the CCR5 receptor. Surprisingly, the putative ground state CCR5 receptor model resulted in the identification of agonists. Our finding hence hints that a single CCR5 structure might bind small molecules that can either favor the active or inactive state of the receptor. This assumption is supported by the recent crystal structure of the photoactivated deprotonated intermediate of bovine rhodopsin, which clearly demonstrated that the ground state and the described metaII activated state of the receptor are structurally similar³² and that the acknowledged large scale structural rearrangement of the receptor transmembrane domain upon activation is overestimated.

Agonist-induced CCR5 receptor internalization

Several reports have suggested that internalizing the CCR5 receptor may constitute a very efficient strategy to block HIV infection: (i) the capability to favor internalization of CCR5 has been demonstrated to be strongly related to the antiviral properties of RANTES

derivatives;³³ (ii) a lower susceptibility to HIV infection has been reported for individuals presenting multiple copies of the gene encoding the CCR5 agonist MIP-1 α ,³⁴ (iii) small chemical agonists promoting internalization of the receptor might prevent the selection of HIV mutants that would be able to interact with CCR5 bound to the allosteric inhibitors presently tested in clinical trials.

The potency of compounds **8**, **9** and **11** to induce the CCR5 internalization was investigated by monitoring the fluorescence of the receptor fused to the Enhanced Green Fluorescent Protein variant (EGFP).³⁵ All three compounds promote CCR5 receptor internalization, yet poor efficiency was observed for compounds **9** and **11** ($5 \pm 2\%$ and $15 \pm 2\%$ of receptor internalized after 15 min, respectively). On the other hand, the level of internalized receptors in presence of compound **8** at a concentration of 20 μ M reached 22% (Figure 7). By the way of comparison, the reference value measured in presence of chemokine RANTES (CCL5) at 200 nM was 37%. Molecule **8** thus appears to be a good hit compound for further development of CCR5-internalizing non-peptide agents (Figure 7).

Follow-up similarity search identifies other CCR5 receptor agonists

In an attempt to find more potent CCR5 receptor agonists, we have undertaken a ligand-based virtual screening to find novel agonists similar to three previously identified hits (compounds **8**, **10** and **11**). In order to optimize the chances to find novel hits,³⁶ different screening strategies using various templates and 2-D fingerprints (Table 2) were followed

Out of the starting 1,186,452 library compounds ("updated screening dataset"), 62,761 shared physicochemical properties as well as key pharmacophoric features exhibited by compounds **8**, **10**, and **11**. Four independent ligand-based screens of this dataset were performed using various 2-D fingerprint representations (circular FCFP-4 fingerprints, MPMFP, Unity 2D, see Table 2). Hundred compounds were manually selected from the four

hit lists. 17 close analogs of agonist **8** were chosen in order to derive SAR rules, and the 83 other compounds were picked to optimize chemical diversity according to ClassPharmer maximum common substructures. Experimental testing for CCR5 functional response revealed 7 additional true positives that were characterized by an EC₅₀ below 100 μ M. The three most potent compounds **12-14** were tested in a CCR5 binding assay, and shown to compete for MIP-1 β binding (Table 1, Figure S2 supplementary materials). Disappointingly, none of them displayed enhanced affinity for CCR5. The lack of activity of 17 close analogs of compound **8** could be either explained by a change in the distance between triazole and piperidine rings or by the addition of polar substituents to the terminal anilide moiety.

Non-peptide agonists of chemokine receptors

Very few non-peptide agonists are known for the members of the chemokine receptor family. Detailed characterization of a non-peptide low molecular-weight agonist is available for the CCR8 receptor subtype.³⁷ Non-peptidic modulators of the chemokine receptor are also described for the CCR3 receptor subtype³⁸. Interestingly, structure-activity relationships developed for the investigated series indicated that minor chemical modifications may switch compound activity from antagonist to agonist. Another striking example is reported for CCR3 receptor by De Lucca and co-workers³⁹, who observed that quaternary piperidinium salts were agonists while corresponding piperidine derivatives were antagonists of the receptor.

The present work describes several novel CCR5 receptor non-peptide antagonists and agonists (Figure 5), discovered by ligand- and structure-based virtual screening. All compounds share a protonable tertiary amine and several aromatic rings that are expected to interact with key residues in the TM domain of the receptor. Similar features were observed in the CCR5 agonist (YM-370749) that was very recently discovered by Saita and co-workers⁴⁰ by high-throughput screening. Like compound **8**, YM-370749 is able to induce the

intracellular accumulation of the CCR5 receptor. Interestingly, it did not stimulate undesirable chemotaxis. Provided that compound **8** does not trigger undesirable proinflammatory activities, it is an attractive candidate for the design of new class of anti-HIV agent, since its computed molecular properties are compatible with oral bioavailability.

Conclusions

A 3-D model of the CCR5 receptor was derived from the inactive dark state of bovine rhodopsin, then customized to discriminate true CCR5 antagonists from drug-like decoys and finally utilized for the *in silico* screening of commercially-available compound libraries. Out of 59 molecules selected by a series of logical filters of increasing complexity, 10 compounds were found to display micromolar affinities to the CCR5 receptor. Unexpectedly, six hits were found to be agonists of the target receptor, and no obvious chemical features could discriminate between agonist and antagonist compounds. Our findings show that the current CCR5 receptor model is suitable for retrieving both agonist and antagonist compounds, suggesting that the TM structure of the activated and inactivated states of the receptor are similar, like recently demonstrated for bovine rhodopsin.³² A follow-up ligand-based virtual screening identified three new CCR5 agonists exhibiting 2-D fingerprints which were similar to the first-generation hits. One of the non-peptide CCR5 agonist (compound **8**) was shown to promote efficient receptor internalization, which is a process therapeutically favorable for protection against HIV-1 infection.

Computational methods

Molecular modeling

The 3-D coordinates of a previously published 3-D model of the CCR5 α -helix bundle^{19,20} were kindly provided by C. Govaerts and L. Pardo. After addition of the first intracellular loop that constraints TMs 2 and 3 inner relationships, the starting model was first minimized in AMBER6.0⁴¹ in vacuum using the parm98 parameter set. Five cycles of 500 steps of steepest descent followed by 500 steps of conjugated gradient were applied as followed; first, all α carbons of the TM domain were constrained using a harmonic constant of 50 kcal/mol/Å. Next, the constant value was gradually reduced to zero, and finally, the last minimization round ended after 10,000 steps, resulting in RMS gradient of the potential energy in the 0.05-0.01 kcal/mol/Å range.

Further refinement aimed at the fit of non-peptide ligands into the TM cavity. An N-terminal binding pocket enclosing 33 residues (residues in red, Figure 2A) was defined by similarity with the 11-*cis*-retinal binding site found in the bovine rhodopsin crystal structure. It was optimized in order to accommodate a set of five known CCR5 antagonists: TAK-779, SCH-351125, SCH-350634, and representative structure of 1-amino-2-phenyl-4-(piperidin-1-yl)butane and 1,3,4-trisubstituted pyrrolidines from a MERCK series (Figure 1: compounds **1-5**). The ligand starting conformation consisted in the 3-D coordinates generated by Concord⁴² from IsisDraw⁴³ 2-D sketches. Their most likely ionization state at physiological pH was set manually. All antagonists were manually docked according to existing site-directed mutagenesis²² and detailed structure-activity relationships.⁴⁴⁻⁵¹ Both rotatable bonds of ligand and torsional χ angles of residues of the putative binding pocket were adjusted to avoid intermolecular steric clashes. The five reference antagonists were parameterized for Amber6.0 according to a previously-described procedure.⁵² Energy minimization of each complex was realized as describe above. A distance restraint ensured the interaction between the positively charged nitrogen atom of the ligand and any of the carboxylate oxygen atoms of Glu7.39. At last, the coordinates of the five receptor-bound minimized ligands were merged

into the TAK779-bound minimized CCR5 structure and the resulting supramolecule was optimized by energy minimization (see above parameters) using the AMBER belly approach that allowed the ligand atoms to remain at fixed positions during protein refinement.

Structure-based virtual screening

Databases setting

Starting from 1,620,316 commercially available compounds (sd format) collected from eight suppliers (Asinex, Bionet, Chembridge, Chemdiv, InterBioScreen, Maybridge, Specs, Tripos-Leadquest), 431,029 unique drug-like compounds were selected by properties and functional groups⁵³ using the Filter1.0 program,⁵⁴ then ionized at physiological pH using Filter again and finally converted into 3-D coordinates (mol2 format) by Concord.⁴² A subsequent filtering step by 2-D pharmacophore search afforded 44,524 compounds sharing an aromatic ring and a protonable nitrogen atom ("screening dataset").

Seven known CCR5 antagonists (Figure 1, Compounds 1-7) together with 993 molecules randomly chosen from the screening dataset were incorporated in a separate library to create a "validation dataset" of 1,000 molecules.

High throughput docking

All molecules of the databases were docked into the final CCR5 model using Gold2.0⁵⁵ and Surflex1.1.⁵⁶ Library screening settings of Gold were used to generate a maximum of 10 poses for each molecule. Poses were evaluated using Goldscore. The binding cavity was defined from the centre of mass of 33 residues previously selected to define the antagonist-binding pocket. Early termination of the genetic algorithm was allowed whether the top three ranked poses were within 1.5 Å rms deviations.

Surflex docking was performed using the "whole molecule" approach (-whole), a maximum of 100 simultaneously considered posed fragment in each stage of the incremental construction process (-maxconf 100) and default settings for all other parameters. The protocol based on protein residues that form the CCR5 binding pocket as defined above contains 451 points.

Post-processing and hit triage

Two independent hitlists were obtained by selecting the top ranked molecules proposed by Gold and Surflex (Figure 4). Defining a Goldscore threshold of 47 yielded the selection of a first list of 2,410 molecules from Gold poses. A second set of 2,724 molecules was retained from Surflex poses by combining a Surflex score threshold value of 8 and a penetration threshold value of -40. The two hitlists were separately clustered according to common graph-based scaffolds using the ClassPharmer 3.5 program.⁵⁷ The clustering procedure was performed using the medium setting for homogeneity and redundancy parameters, and allowed fuzzy ring closure. All molecules discarded by the docking score-based selection were imported into each classification. Classes were then prioritized according to distribution of docking scores and the proportion of imported compounds in each class. Compound sampling retrieved the best scored molecule for each selected class in order to produce a hit list. Last, relevance of the potential hits was assessed using visual check of the receptor-hit geometry. More especially, the visual analysis relied on 1) surface complementarity between the ligand and the protein, 2) the percentage of the ligand surface buried in the complex, 3) the amount and the properties of ligand surface that is exposed to solvent, 4) the number of occluded polar atoms of both ligand and protein in the complex, 5) the establishment of a salt bridge to the carboxylate of Glu7.39 and 6) the strain energy of the bound ligand. Altogether,

two lists of 38 and 43 compounds, respectively, were selected from Gold and Surflex docking runs and merged to afford a final list of 77 compounds to purchase.

Ligand-based virtual screening for hit optimization

1,186,452 unique drug-like and commercially available compounds were filtered as described above. The stepwise ligand-based screen of the database ("updated screening dataset") was performed as follow; first, a substructure search yielded 89,121 compounds that contain a positively charged nitrogen atom together with an aromatic bond and ring-membered sp² nitrogen atom. Subsequent filtering according to physicochemical properties of compounds **8**, **10** and **11** (molecular weight ranging from 300 to 480, less than 10 rotatable bonds and more than two aromatic rings) selected 62,761 compounds. Last, similarity searches to compounds **8**, **10** and **11** was carried out using a Tanimoto coefficient (Tc) as similarity metric and three different fingerprints, namely FCFP_4 circular fingerprints as implemented in PipelinePilot 5.1,²⁷ MPMFP⁵⁸ as implemented in MOE 2006.06 and UNITY 2D fingerprint as implemented in Sybyl7.1⁴². Virtual hits were merged and clustered using ClassPharmer 3.5.⁵⁷ 100 compounds were finally picked based on human expertise out of which 83 could be purchased.

Experimental procedures

Cell lines

A construct encoding wild-type CCR5 in the pEFIN3 bicistronic vector was transfected in a CHO-K1 cell line expressing an apoeaquorin variant targeted to mitochondria, and a stable cell line was established as described.⁵⁹ CHO-K1 cells were cultured in HAM's F12 medium supplemented with 10% fetal calf serum (Life Technologies), 100 units/ml penicillin and 100 µg/ml streptomycin (Life Technologies).

Binding assays

CHO-K1 cells expressing wild-type or mutant CCR5 were collected from plates with Ca^{2+} and Mg^{2+} -free PBS supplemented with 5 mM EDTA, gently pelleted by centrifuging for 2 min at $1000 \times g$ and resuspended in binding buffer (50 mM Hepes pH 7.4, 1 mM CaCl_2 , 5 mM MgCl_2 , 0.5% BSA). Competition binding assays were performed in Minisorb tubes (Nunc), with 40,000 cells in a final volume of 0.1 ml. The mixture contained 0.1 nM [^{125}I]-MIP-1 β as tracer, and variable concentrations of competitors (R&D Systems). Total binding was measured in the absence of competitor, and non-specific binding was measured with a 100 fold excess of unlabelled ligand. Samples were incubated for 90 minutes at 27° C, then bound tracer was separated by filtration through GF/B filters pre-soaked in 0.1% bovine serum albumine (Sigma). Filters were counted in a β -scintillation counter. Assays were performed in triplicate.

Aequorine assay

Functional response to ligands was analyzed by measuring the luminescence of aequorin as described.⁵⁹ Cells were collected from plates with Ca^{2+} and Mg^{2+} -free DMEM supplemented with 5 mM EDTA. They were then pelleted for 2 min at $1000 \times g$, resuspended in DMEM at a density of 5×10^6 cells/ml and incubated for 2 h in the dark in the presence of 5 μM coelenterazine H (Molecular Probes). Cells were diluted 5 fold before use. Agonists in a volume of 50 μl DMEM were added to 50 μl of cell suspension (50,000 cells) and luminescence was measured for 30 sec in a Berthold Luminometer. Assays were performed in duplicate.

Internalization of EGFP-CCR5 Receptors Measured Using Anti-GFP Antibodies

HEK 2193 cells expressing EGFP-CCR5 receptor were suspended at a concentration of 10^6 cells/ml in 24 ml of HEPES-BSA buffer. The cell suspension was distributed into 4 x 50 ml tubes equipped with magnetic stirring and incubated at 37°C. At time-point 0, control or test compounds were added to each sample. The final concentration of RANTES (CCL5) was 200 nM. The final concentration of test compounds was chosen according to measured IC₅₀ determined (20 μ M for compound **8** and 100 μ M for compounds **9** and **11**). At time points 0, 5, 15 and 30 and 45 min, 1ml of each cell suspension was collected, pelleted by centrifuging for 3 min at $2000 \times g$ at 4 °C and resuspended in 1 ml ice-cold PBS-1% BSA. Immunolabeling of EGFP-CCR5 receptors expressed at the surface was performed by incubation for 30 min on ice with a monoclonal mouse anti-GFP (Roche Molecular Biochemicals; 1/100 dilution). Cell were then washed 3 times with 1 ml of ice-cold PBS-1% BSA and secondary labeling was done using R-phycoerythrin-conjugated AffiniPure F(ab)₂ fragment goat antimouse IgG (Immunotech, Marseille; 1/100). Samples were washed twice in 1 ml of PBS, fixed for 5 min in ice-cold PBS-4% paraformaldehyde, resuspended in 1 ml of ice-cold PBS, and stored overnight. Phycoerythrin staining was quantified by flow cytometric analysis (FACScalibur, Becton-Dickinson) using 5,000 cells per sample. The mean intensity of phycoerythrin fluorescence was calculated using CELLQuest (Becton-Dickinson) software, after subtraction of non-specific staining by the two antibodies measured on non-transfected cells. The amount of CCR5 receptors present at the cell surface was deduced from the quantification of phycoerythrin intensities.

Acknowledgements

This work was supported by the Belgian programme on Interuniversity Poles of Attraction initiated by the Belgian State, Prime Minister's Office, Science Policy Programming, the LifeSciHealth (grant LSHB-CT-2003-503337 and LSHB-CT-2005-518167) programme of

the European Community, the Fonds de la Recherche Scientifique Médicale of Belgium and the French Agence Nationale de Recherche sur le SIDA (ANRS). We sincerely thank the Centre Informatique National de l'Enseignement Supérieur (CINES, Montpellier, France) for providing us with computing time on the SGI3800 supercomputer. The scientific responsibility is assumed by the authors.

Supplementary information

Supporting information comprises Figure S1 (Sequence alignment of human CCR5 and bovine rhodopsin) and S2 (Characterization of the hits resulting from screening in binding and functional assays), Table S1 (Comparison of Gold and Surflex performance in the screening of the "validation dataset" against CCR5 model), and Chart A (Filtering rules in OpenEye's Filter program). This information is available free of charge via the Internet at <http://pubs.acs.org>

References

- (1) Klabunde, T.; Hessler, G. Drug design strategies for targeting G-protein-coupled receptors. *Chembiochem* **2002**, *3*, 928-944.
- (2) Clardy, J.; Walsh, C. Lessons from natural molecules. *Nature* **2004**, *432*, 829-837.
- (3) Kristiansen, K. Molecular mechanisms of ligand binding, signaling, and regulation within the superfamily of G-protein-coupled receptors: molecular modeling and mutagenesis approaches to receptor structure and function. *Pharmacol. Ther.* **2004**, *103*, 21-80.
- (4) Wise, A.; Jupe, S. C.; Rees, S. The Identification of Ligands at Orphan G-Protein Coupled Receptors. *Annu. Rev. Pharmacol. Toxicol.* **2004**, *44*, 43-66.
- (5) Crossley, R. The design of screening libraries targeted at G-protein coupled receptors. *Curr. Top. Med. Chem.* **2004**, *4*, 581-588.
- (6) Palczewski, K.; Kumasaka, T.; Hori, T.; Behnke, C. A.; Motoshima, H.; Fox, B. A.; Trong, I. L.; 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.
- (7) Shoichet, B. K. Virtual screening of chemical libraries. *Nature* **2004**, *432*, 862-865.
- (8) Oliveira, L.; Hulsen, T.; Lutje Hulsik, D.; Paiva, A. C.; Vriend, G. Heavier-than-air flying machines are impossible. *FEBS Lett.* **2004**, *564*, 269-273.
- (9) Becker, O. M.; Marantz, Y.; Shacham, S.; Inbal, B.; Heifetz, A.; Kalid, O.; Bar-Haim, S.; Warshaviak, D.; Fichman, M.; Noiman et, a. G protein-coupled receptors: in silico drug discovery in 3D. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 11304-11309.
- (10) Evers, A.; Klabunde, T. Structure-based drug discovery using GPCR homology modelling: Successful virtual screening for antagonists of the alpha1A adrenergic receptor. *J. Med. Chem.* **2005**, *48*, 1088-1097.

- (11) Evers, A.; Klebe, G. Ligand-supported homology modeling of g-protein-coupled receptor sites: models sufficient for successful virtual screening. *Angew Chem. Int. Ed. Engl.* **2004**, *43*, 248-251.
- (12) Varady, J.; Wu, X.; Fang, X.; Min, J.; Hu, Z.; Levant, B.; Wang, S. Molecular modeling of the three-dimensional structure of dopamine 3 (D3) subtype receptor: discovery of novel and potent D3 ligands through a hybrid pharmacophore- and structure-based database searching approach. *J. Med. Chem.* **2003**, *46*, 4377-4392.
- (13) Kobilka, B. Agonist binding: a multistep process. *Mol. Pharmacol.* **2004**, *65*, 1060-1062.
- (14) Bissantz, C.; Bernard, P.; Hibert, M.; Rognan, D. Protein-based virtual screening of chemical databases. II. Are homology models of G-Protein Coupled Receptors suitable targets? *Proteins* **2003**, *50*, 5-25.
- (15) Gouldson, P. R.; Kidley, N. J.; Bywater, R. P.; Psaroudakis, G.; Brooks, H. D.; Diaz, C.; Shire, D.; Reynolds, C. A. Toward the active conformations of rhodopsin and the beta2-adrenergic receptor. *Proteins* **2004**, *56*, 67-84.
- (16) Tyndall, J. D.; Pfeiffer, B.; Abbenante, G.; Fairlie, D. P. Over one hundred peptide-activated G protein-coupled receptors recognize ligands with turn structure. *Chem. Rev.* **2005**, *105*, 793-826.
- (17) Jones, R. M.; Boatman, P. D.; Semple, G.; Shin, Y. J.; Tamura, S. Y. Clinically validated peptides as templates for de novo peptidomimetic drug design at G-protein-coupled receptors. *Curr. Opin. Pharmacol.* **2003**, *3*, 530-543.
- (18) Maeda, K.; Nakata, H.; Ogata, H.; Koh, Y.; Miyakawa, T.; Mitsuya, H. The current status of, and challenges in, the development of CCR5 inhibitors as therapeutics for HIV-1 infection. *Curr. Opin. Pharmacol.* **2004**, *4*, 447-452.

- (19) 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.
- (20) 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.
- (21) Edelsbrunner, H.; Koehl, P. The weighted-volume derivative of a space-filling diagram. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 2203-2208.
- (22) 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.
- (23) Seibert, C.; Ying, W.; Gavrilov, S.; Tsamis, F.; Kuhmann, S. E.; Palani, A.; Tagat, J. R.; Clader, J. W.; McCombie, S. W.; Baroudy, B. M.; Smith, S. O.; Dragic, T.; Moore, J. P.; Sakmar, T. P. Interaction of small molecule inhibitors of HIV-1 entry with CCR5. *Virology* **2006**, *349*, 41-54.
- (24) Maeda, K.; Das, D.; Ogata-Aoki, H.; Nakata, H.; Miyakawa, T.; Tojo, Y.; Norman, R.; Takaoka, Y.; Ding, J.; Arnold, G. F.; Arnold, E.; Mitsuya, H. Structural and molecular interactions of CCR5 inhibitors with CCR5. *J. Biol. Chem.* **2006**, *281*, 12688-12698.
- (25) Tsamis, F.; Gavrilov, S.; Kajumo, F.; Seibert, C.; Kuhmann, S.; Ketas, T.; Trkola, A.; Palani, A.; Clader, J. W.; Tagat, J. R.; McCombie, S.; Baroudy, B.; Moore, J. P.; Sakmar, T. P.; Dragic, T. Analysis of the mechanism by which the small-molecule

- CCR5 antagonists SCH-351125 and SCH-350581 inhibit human immunodeficiency virus type 1 entry. *J. Virol.* **2003**, *77*, 5201-5208.
- (26) Castonguay, L. A.; Weng, Y.; Adolfsen, W.; Di Salvo, J.; Kilburn, R.; Caldwell, C. G.; Daugherty, B. L.; Finke, P. E.; Hale, J. J.; Lynch, C. L.; Mills, S. G.; MacCoss, M.; Springer, M. S.; DeMartino, J. A. Binding of 2-aryl-4-(piperidin-1-yl)butanamines and 1,3,4-trisubstituted pyrrolidines to human CCR5: a molecular modeling-guided mutagenesis study of the binding pocket. *Biochemistry* **2003**, *42*, 1544-1550.
- (27) SciTegic Inc., San Diego, CA 92123-1365, USA.
- (28) 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.
- (29) Blanpain, C.; Doranz, B. J.; Bondue, A.; Govaerts, C.; De Leener, A.; Vassart, G.; Doms, R. W.; Proudfoot, A.; Parmentier, M. The core domain of chemokines binds CCR5 extracellular domains while their amino terminus interacts with the transmembrane helix bundle. *J. Biol. Chem.* **2003**, *278*, 5179-5187.
- (30) 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.
- (31) Oppermann, M. Chemokine receptor CCR5: insights into structure, function, and regulation. *Cell. Signal.* **2004**, *16*, 1201-1210.
- (32) Salom, D.; Lodowski, D. T.; Stenkamp, R. E.; Le Trong, I.; Golczak, M.; Jastrzebska, B.; Harris, T.; Ballesteros, J. A.; Palczewski, K. Crystal structure of a photoactivated

- deprotonated intermediate of rhodopsin. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 16123-16128.
- (33) Hartley, O.; Gaertner, H.; Wilken, J.; Thompson, D.; Fish, R.; Ramos, A.; Pastore, C.; Dufour, B.; Cerini, F.; Melotti, A.; Heveker, N.; Picard, L.; Alizon, M.; Mosier, D.; Kent, S.; Offord, R. Medicinal chemistry applied to a synthetic protein: development of highly potent HIV entry inhibitors. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 16460-16465.
- (34) Gonzalez, E.; Kulkarni, H.; Bolivar, H.; Mangano, A.; Sanchez, R.; Catano, G.; Nibbs, R. J.; Freedman, B. I.; Quinones, M. P.; Bamshad, M. J.; Murthy, K. K.; Rovin, B. H.; Bradley, W.; Clark, R. A.; Anderson, S. A.; O'Connell R, J.; Agan, B. K.; Ahuja, S. S.; Bologna, R.; Sen, L.; Dolan, M. J.; Ahuja, S. K. The influence of CCL3L1 gene-containing segmental duplications on HIV-1/AIDS susceptibility. *Science* **2005**, *307*, 1434-1440.
- (35) Valenzuela-Fernandez, A.; Palanche, T.; Amara, A.; Magerus, A.; Altmeyer, R.; Delaunay, T.; Virelizier, J. L.; Baleux, F.; Galzi, J. L.; Arenzana-Seisdedos, F. Optimal inhibition of X4 HIV isolates by the CXC chemokine stromal cell-derived factor 1 alpha requires interaction with cell surface heparan sulfate proteoglycans. *J. Biol. Chem.* **2001**, *276*, 26550-26558.
- (36) Clark, D. E.; Higgs, C.; Wren, S. P.; Dyke, H. J.; Wong, M.; Norman, D.; Lockey, P. M.; Roach, A. G. A virtual screening approach to finding novel and potent antagonists at the melanin-concentrating hormone 1 receptor. *J. Med. Chem.* **2004**, *47*, 3962-3971.
- (37) Haskell, C. A.; Horuk, R.; Liang, M.; Rosser, M.; Dunning, L.; Islam, I.; Kremer, L.; Gutierrez, J.; Marquez, G.; Martinez, A. C.; Biscone, M. J.; Doms, R. W.; Ribeiro, S. Identification and characterization of a potent, selective nonpeptide agonist of the CC chemokine receptor CCR8. *Mol. Pharmacol.* **2006**, *69*, 309-316.

- (38) Anderskewitz, R.; Bauer, R.; Bodenbach, G.; Gester, D.; Gramlich, B.; Morschhauser, G.; Birke, F. W. Pyrrolidinohydroquinazolines--a novel class of CCR3 modulators. *Bioorg. Med. Chem. Lett.* **2005**, *15*, 669-673.
- (39) De Lucca, G. V.; Kim, U. T.; Johnson, C.; Vargo, B. J.; Welch, P. K.; Covington, M.; Davies, P.; Solomon, K. A.; Newton, R. C.; Trainor, G. L.; Decicco, C. P.; Ko, S. S. Discovery and structure-activity relationship of N-(ureidoalkyl)-benzyl-piperidines as potent small molecule CC chemokine receptor-3 (CCR3) antagonists. *J. Med. Chem.* **2002**, *45*, 3794-3804.
- (40) Saita, Y.; Kodama, E.; Orita, M.; Kondo, M.; Miyazaki, T.; Sudo, K.; Kajiwarra, K.; Matsuoka, M.; Shimizu, Y. Structural basis for the interaction of CCR5 with a small molecule, functionally selective CCR5 agonist. *J. Immunol.* **2006**, *177*, 3116-3122.
- (41) Case, D. A.; Pearlman, D. A.; Caldwell, J. W.; Cheatham III, D. E.; Ross, W. S.; Simmerling, C. L.; Darden, T. A.; Merz, K. M.; Stanton, R. V.; Cheng, A. L.; Vincent, J. J.; Crowley, M.; Tsui, V.; Radmer, R. J.; Duan, Y.; Pitera, J.; Massova, I.; Seibel, G. L.; Singh, U. C.; Weiner, P. K.; Kollman, P. A. AMBER 6, University of California, San Francisco. **1999**.
- (42) TRIPOS, Assoc., Inc., St-Louis, MO 63144-2319, USA
- (43) Elsevier MDL, San Leandro, CA 94577, USA.
- (44) Finke, P. E.; Meurer, L. C.; Oates, B.; Shah, S. K.; Loebach, J. L.; Mills, S. G.; MacCoss, M.; Castonguay, L.; Malkowitz, L.; Springer, M. S.; Gould, S. L.; DeMartino, J. A. Antagonists of the human CCR5 receptor as anti-HIV-1 agents. Part 3: a proposed pharmacophore model for 1-[N-(methyl)-N-(phenylsulfonyl)amino]-2-(phenyl)-4-[4-(substituted)piperidin-1-yl]butanes. *Bioorg. Med. Chem. Lett.* **2001**, *11*, 2469-2473.

- (45) Hale, J. J.; Budhu, R. J.; Holson, E. B.; Finke, P. E.; Oates, B.; Mills, S. G.; MacCoss, M.; Gould, S. L.; DeMartino, J. A.; Springer, M. S.; Siciliano, S.; Malkowitz, L.; Schleif, W. A.; Hazuda, D.; Miller, M.; Kessler, J.; Danzeisen, R.; Holmes, K.; Lineberger, J.; Carella, A.; Carver, G.; Emini, E. 1,3,4-Trisubstituted pyrrolidine CCR5 receptor antagonists. Part 2: lead optimization affording selective, orally bioavailable compounds with potent anti-HIV activity. *Bioorg. Med. Chem. Lett.* **2001**, *11*, 2741-2745.
- (46) Finke, P. E.; Oates, B.; Mills, S. G.; MacCoss, M.; Malkowitz, L.; Springer, M. S.; Gould, S. L.; DeMartino, J. A.; Carella, A.; Carver, G.; Holmes, K.; Danzeisen, R.; Hazuda, D.; Kessler, J.; Lineberger, J.; Miller, M.; Schleif, W. A.; Emini, E. A. Antagonists of the human CCR5 receptor as anti-HIV-1 agents. Part 4: synthesis and structure-activity relationships for 1-[N-(methyl)-N-(phenylsulfonyl)amino]-2-(phenyl)-4-(4-(N-(alkyl)-N-(benzyl)oxycarbonyl)amino)piperidin-1-yl)butanes. *Bioorg. Med. Chem. Lett.* **2001**, *11*, 2475-2479.
- (47) Dorn, C. P.; Finke, P. E.; Oates, B.; Budhu, R. J.; Mills, S. G.; MacCoss, M.; Malkowitz, L.; Springer, M. S.; Daugherty, B. L.; Gould, S. L.; DeMartino, J. A.; Siciliano, S. J.; Carella, A.; Carver, G.; Holmes, K.; Danzeisen, R.; Hazuda, D.; Kessler, J.; Lineberger, J.; Miller, M.; Schleif, W. A.; Emini, E. A. Antagonists of the human CCR5 receptor as anti-HIV-1 agents. part 1: discovery and initial structure-activity relationships for 1 -amino-2-phenyl-4-(piperidin-1-yl)butanes. *Bioorg. Med. Chem. Lett.* **2001**, *11*, 259-264.
- (48) Finke, P. E.; Meurer, L. C.; Oates, B.; Mills, S. G.; MacCoss, M.; Malkowitz, L.; Springer, M. S.; Daugherty, B. L.; Gould, S. L.; DeMartino, J. A.; Siciliano, S. J.; Carella, A.; Carver, G.; Holmes, K.; Danzeisen, R.; Hazuda, D.; Kessler, J.; Lineberger, J.; Miller, M.; Schleif, W. A.; Emini, E. A. Antagonists of the human

- CCR5 receptor as anti-HIV-1 agents. Part 2: structure-activity relationships for substituted 2-Aryl-1-[N-(methyl)-N-(phenylsulfonyl)amino]-4-(piperidin-1-yl)butanes. *Bioorg. Med. Chem. Lett.* **2001**, *11*, 265-270.
- (49) Hale, J. J.; Budhu, R. J.; Mills, S. G.; MacCoss, M.; Malkowitz, L.; Siciliano, S.; Gould, S. L.; DeMartino, J. A.; Springer, M. S. 1,3,4-Trisubstituted pyrrolidine CCR5 receptor antagonists. Part 1: discovery of the pyrrolidine scaffold and determination of its stereochemical requirements. *Bioorg. Med. Chem. Lett.* **2001**, *11*, 1437-1440.
- (50) Palani, A.; Shapiro, S.; Clader, J. W.; Greenlee, W. J.; Cox, K.; Strizki, J.; Endres, M.; Baroudy, B. M. Discovery of 4-[(Z)-(4-bromophenyl)- (ethoxyimino)methyl]-1'-[(2,4-dimethyl-3- pyridinyl)carbonyl]-4'-methyl-1,4'- bipiperidine N-oxide (SCH 351125): an orally bioavailable human CCR5 antagonist for the treatment of HIV infection. *J. Med. Chem.* **2001**, *44*, 3339-3342.
- (51) Tagat, J. R.; Steensma, R. W.; McCombie, S. W.; Nazareno, D. V.; Lin, S. I.; Neustadt, B. R.; Cox, K.; Xu, S.; Wojcik, L.; Murray, M. G.; Vantuno, N.; Baroudy, B. M.; Strizki, J. M. Piperazine-based CCR5 antagonists as HIV-1 inhibitors. II. Discovery of 1-[(2,4-dimethyl-3-pyridinyl)carbonyl]-4- methyl-4-[3(S)-methyl-4-[1(S)-[4-(trifluoromethyl)phenyl]ethyl]-1-piperazi nyl]- piperidine N1-oxide (Sch-350634), an orally bioavailable, potent CCR5 antagonist. *J. Med. Chem.* **2001**, *44*, 3343-3346.
- (52) Rognan, D.; Mukhija, S.; Folkers, G.; Zerbe, O. NMR-restrained docking of a peptidic inhibitor to the N-terminal domain of the phosphoenolpyruvate:sugar phosphotransferase enzyme I. *J. Comput. Aided. Mol. Des.* **2001**, *15*, 103-115.
- (53) Charifson, P. S.; Walters, W. P. Filtering databases and chemical libraries. *J. Comput. Aided. Mol. Des.* **2002**, *16*, 311-323.
- (54) OpenEye Scientific Software, Santa Fe, NM 87507, USA

- (55) Verdonk, M. L.; Cole, J. C.; Hartshorn, M. J.; Murray, C. W.; Taylor, R. D. Improved protein-ligand docking using GOLD. *Proteins* **2003**, *52*, 609-623.
- (56) Jain, A. N. Surflex: fully automatic flexible molecular docking using a molecular similarity-based search engine. *J. Med. Chem.* **2003**, *46*, 499-511.
- (57) Bioreason Inc., Santa Fe, NM 87507, USA
- (58) Godden, J. W.; Stahura, F. L.; Bajorath, J. Anatomy of fingerprint search calculations on structurally diverse sets of active compounds. *J. Chem. Inf. Model.* **2005**, *45*, 1812-1819.
- (59) 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.
- (60) Chemical Computing Group, H3A 2R7 Montreal, Canada.
- (61) Ballesteros, J. A.; Weinstein, H. Integrated methods for the construction of three dimensional models and computational probing of structure-function relations in G-protein coupled receptors. *Methods Neurosci.* **1994**, *25*, 366-428.

Table 1: Experimental activities of compounds 8-14. EC₅₀ and IC₅₀ values were calculated by non-linear regression (Graphpad Prism software), respectively from dose-response curves in the aequorin functional assay, and competition binding assays using [¹²⁵I]-labeled MIP-1β as tracer. Part of the functional and binding data is provided as supplementary data. When full curves could not be obtained due to the low affinity/efficacy of the compounds, the values are reported as > 100.

Compound	Type	Origin	IC ₅₀ , μM	EC ₅₀ , μM
8	Agonist	SBVS ^a	17.0	3.0
9	Antagonist	SBVS	49.0	n.d. ^b
10	Agonist	SBVS	79.4	>100
11	Agonist	SBVS	22.4	1.9
12	Agonist	LBVS ^c	269.2	13.5
13	Agonist	LBVS	14.1	2.0
14	Agonist	LBVS	5.8	1.1

^a SBVS: structure-based virtual screening

^b n.d. = not determined

^c LBVS: ligand-based virtual screening

Table 2: Ligand-based virtual screening.

Hit list	Fingerprint	Reference	Similarity	number of hits		Validated hits ^b
		Ligands	threshold ^a	virtual	Tested	
1	FCFP_4 ^c	8	0.47	100	32	13
2	FCFP_4	8, 10, 11	0.35	328	41	none
3	MPMFP ^d	8	0.97	72	21	12, 14
4	UNITY2D ^e	8	0.65	146	46	12, 13

^a Similarity is expressed by the Tanimoto coefficient obtained by comparing the virtual hit with the reference. For hit list 2, compounds were selected by similarity to at least two out of the three reference ligands.

^b detectable binding and EC₅₀ lower than 20 μ M.

^c Pipeline Pilot 4.1²⁷ circular fingerprint using a functional atomic description (e.g. H-bond donor) and a layer of 4 connectivity neighbors.

^d Fingerprint from Godden et al.⁵⁸ implemented in MOE 2006.06⁶⁰

^e Fingerprint implemented in Unity⁴²

Figure legends

Figure 1: Structure of CCR5 antagonists used in the receptor refinement protocol (molecules **1-5**) and in the test database (molecules **1-7**). **1:** Takeda TAK-779, **2:** Schering-Plough SCH-351125, **3:** Schering-Plough SCH-350634, **4:** representative structure of 1-amino-2-phenyl-4-(piperidin-1-yl)butane MERCK series, **5:** representative structure of 1,3,4-trisubstituted pyrrolidines MERCK series, **6:** selected structure of benzanilides from GlaxoSmithKline CCR5 patent (WO 0040239), **7:** representative structure of 1,3-propanediamine-based Pfizer series (EP 1013276)

Figure 2: CCR5 TM binding cavity. **A.** Amino acid sequence. Residues with side chain pointing toward the cavity are written in red and pointed using Ballesteros numbering⁶¹, excepted in the extracellular domain 1 (EL1). Numbers indicated underneath the sequence summarize experimental mapping of receptor interaction site for non-peptide antagonists **1-5**; *1*: residues important for the efficiency of TAK-779,^{22,23,24} *2*: residues important for the efficiency of Schering-Plough compounds AD101 and SCH-351125^{23,24,25} and *3*: residues important for the binding Merck 2-aryl-4-(piperidin-1-yl)butanamines and 1,3,4-trisubstituted pyrrolidines.²⁶ **B.** The Connolly surface of the CCR5 receptor cavity (colored according to the lipophilic potential) is displayed together with the ribbon diagram of the seven TM helices. Side chains of key residues highlighted in the sequence are depicted using line representation. The bottom view is rotated about a vertical axis by 180° relative to top view.

Figure 3: Enrichment in CCR5 antagonists. The cumulative percentage of known actives recovered by virtual screening is indicated as a function of the top-scoring fraction of the "validation dataset" selected for generating a hit list. Gold (green line) and Surflex (red line) results are compared to random picking (black solid line) and optimum (black dotted line).

Figure 4: Virtual screening flowchart

Figure 5: Structure of CCR5 new ligands. Molecules **8** and **9-11** are Gold and Surflex hits from the structure-based screening, respectively. Molecules, **12-14** are hits from the ligand-based screening.

Figure 6: View of complexes between the CCR5 receptor cavity and agonists **8** (magenta) and **11** (cyan), and antagonists **1** (TAK-779, orange) and **9** (green). Ligands and side chains of protein residues that are less than 4 Å apart are shown inside the transparent Connolly surface of the protein channel using ball and stick and stick representation, respectively. Oxygen atoms are displayed in red, nitrogen atoms in dark blue and sulfur atoms in yellow. For the sake of clarity, protein carbon atoms are colored in white and ligand carbon atoms followed the above mentioned color code.

Figure 7: Characterization of compound **8** in binding and functional assays. **A.** Aequorin-based functional assay. The functional response of the CCR5 receptor to the ligand was tested on a CHO-K1 cell line co-expressing CCR5 and apoaequorin. The data were analyzed by non linear regression using the Prism software (Graphpad Software, San Diego, CA 92130 USA) and a sigmoidal dose-response model. The displayed data represent the mean \pm S.E.M. for duplicate data points **B.** Competition binding assays performed on a CHO-K1 cell line expressing CCR5 using 125 I-MIP-1 β as tracer and the ligands as competitor. The data were normalized for total binding (100%) in the absence of competitor and non specific binding (0%), in the presence of 300 nM MIP-1 β . The data were analyzed by non linear regression using the Prism software and a single-site competition model. The displayed data represent the mean \pm S.E.M. for triplicate data points. **C.** Time-dependent endocytosis of EGFP-CCR5 in HEK 2193 cells induced by 200nM CCL5 (closed triangle) and 20 μ M molecule **8** (closed ring) at 37°C. The data were normalized for the amount of receptors at the surface of

untreated cells (100%). The displayed data represent the mean \pm S.E.M. for triplicate data points.

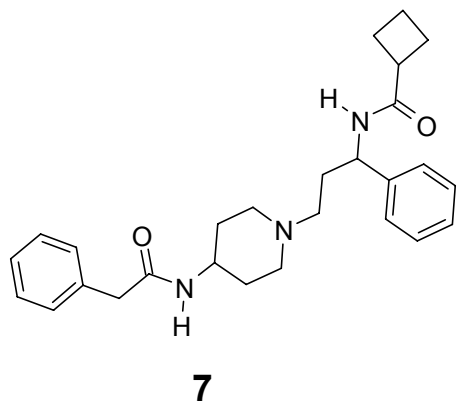
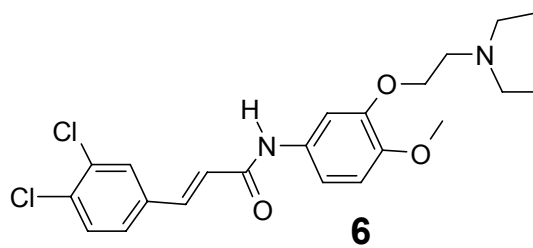
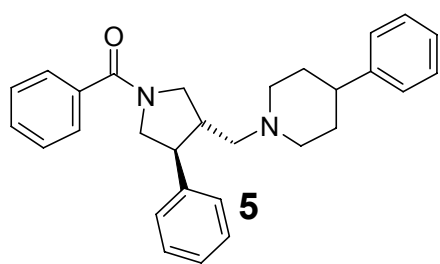
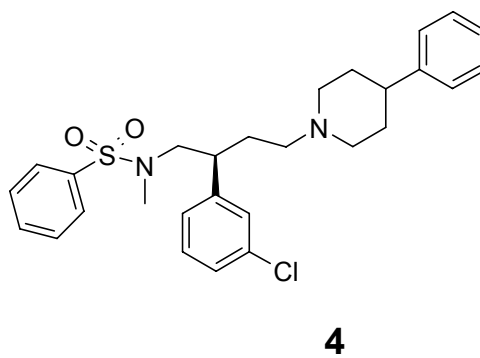
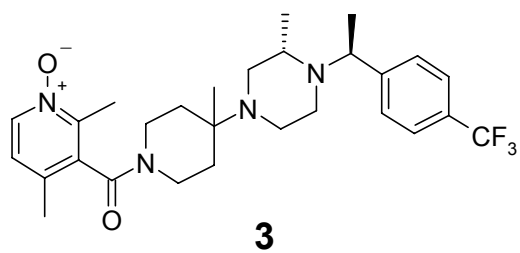
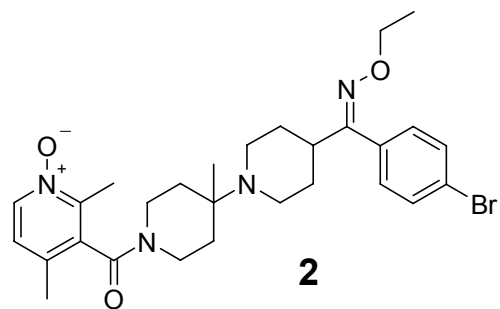
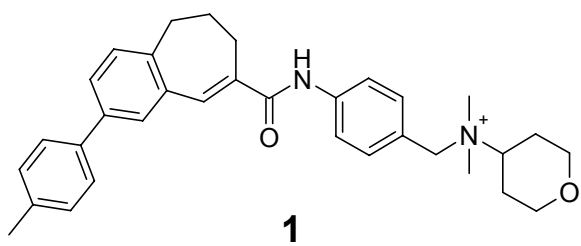


Figure 1

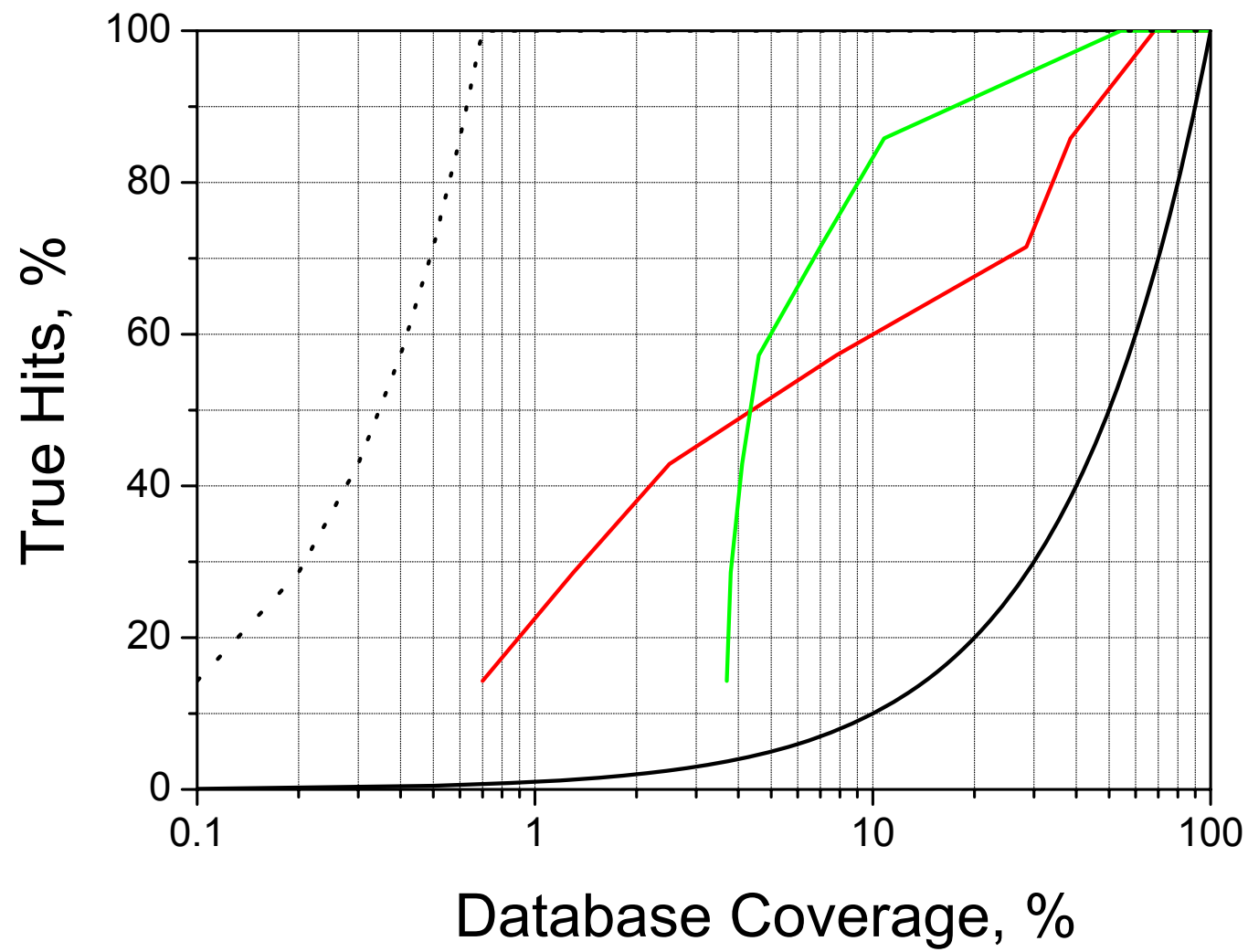


Figure 3

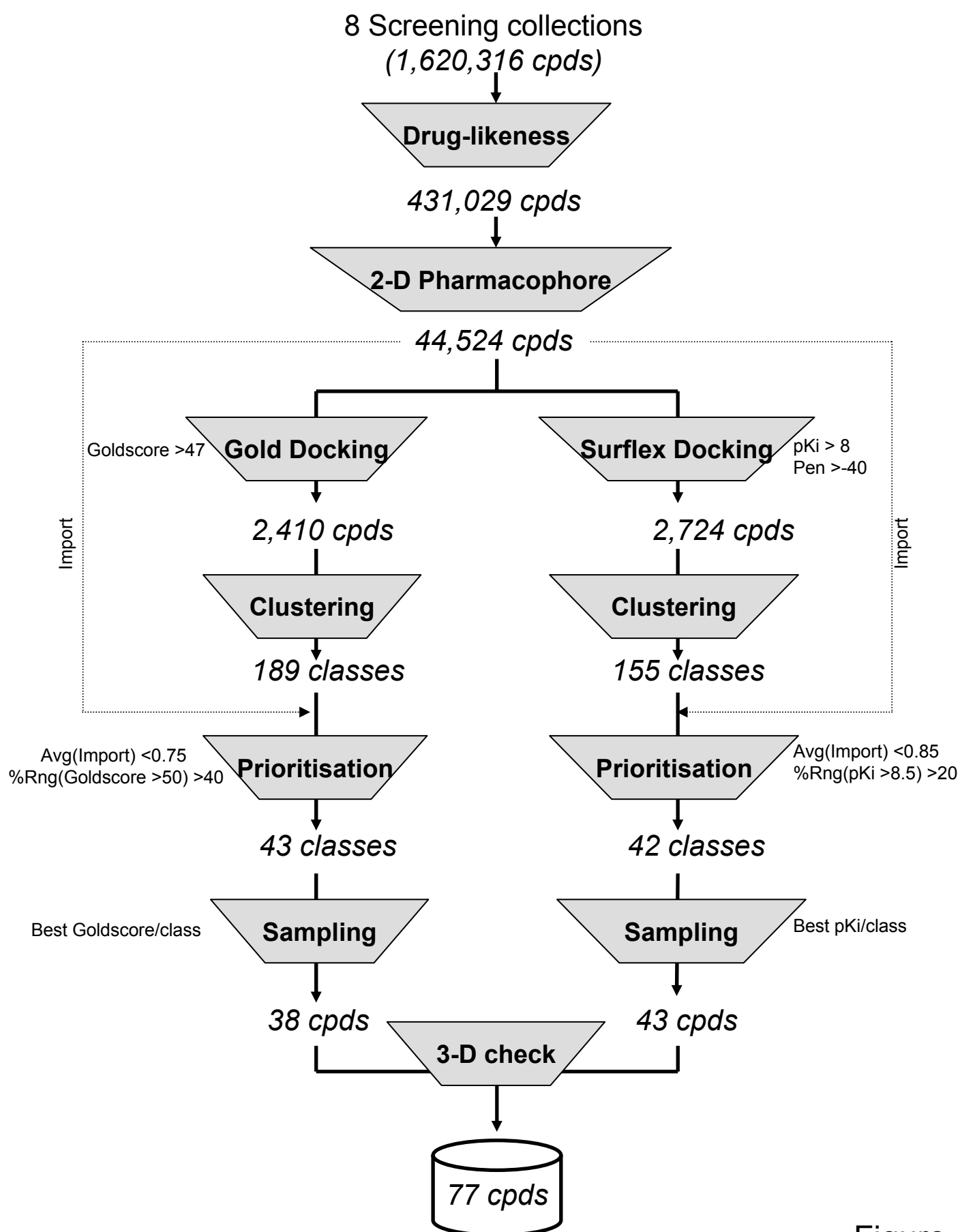
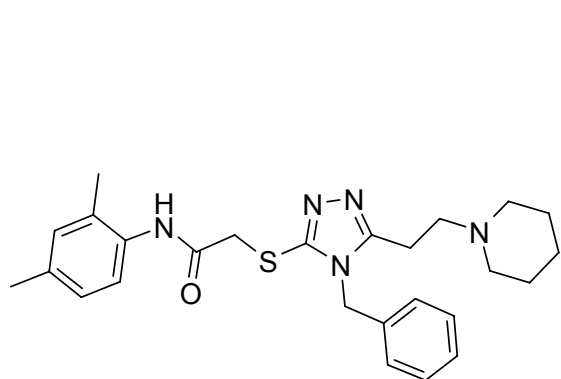
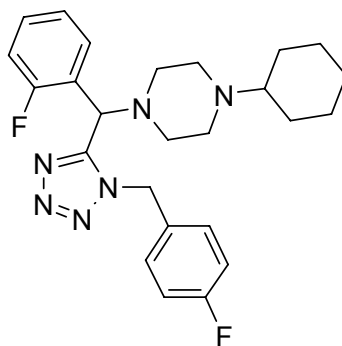


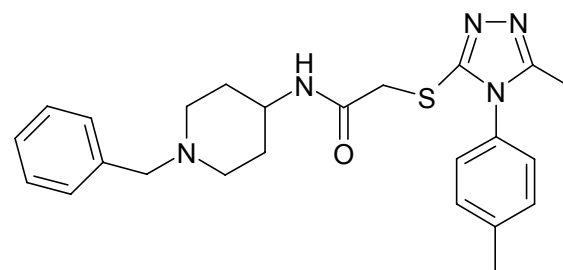
Figure 4



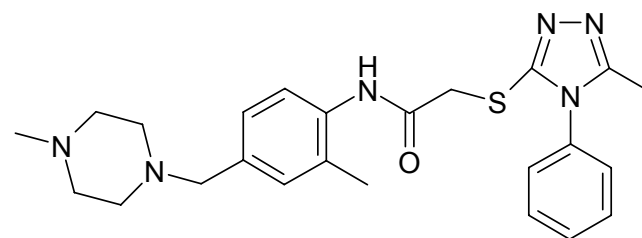
8



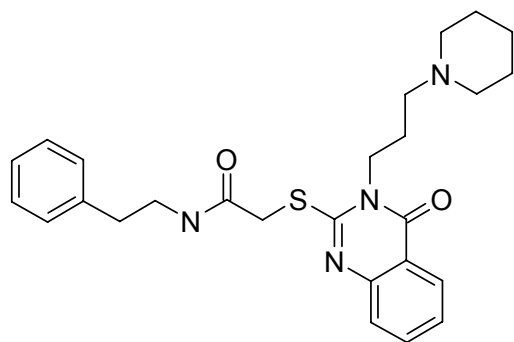
10



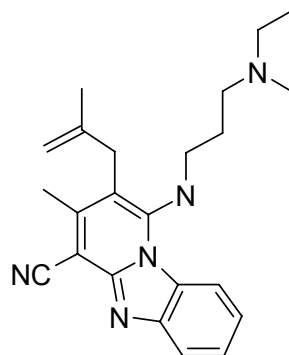
12



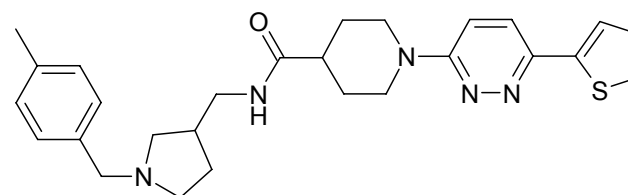
13



9

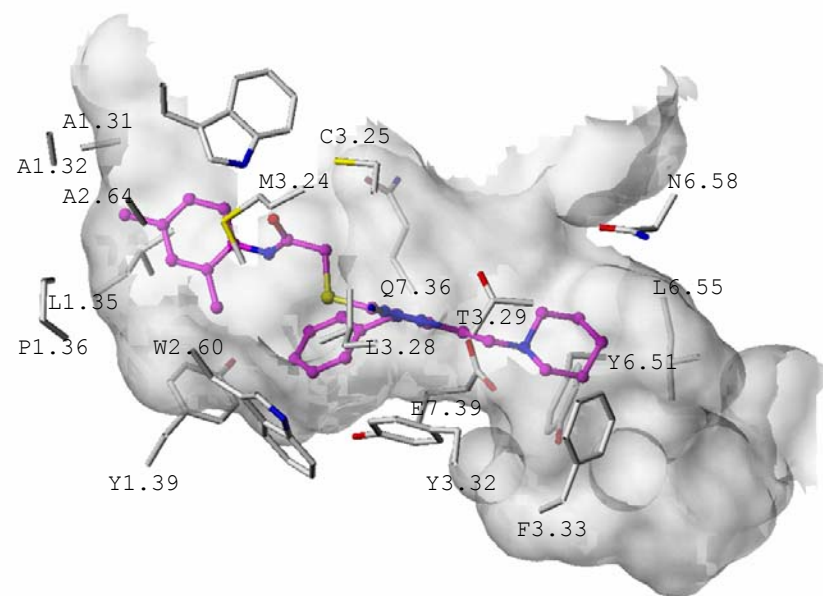


11

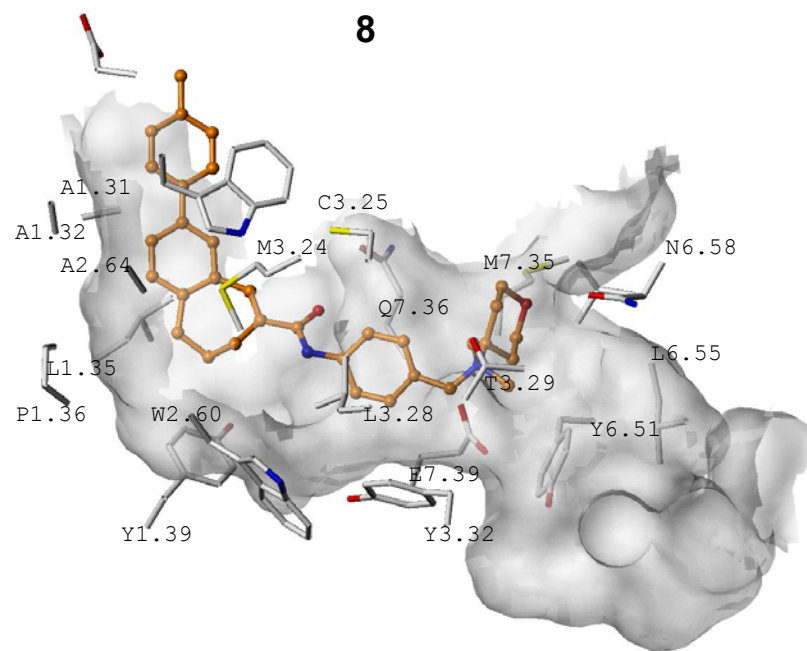


14

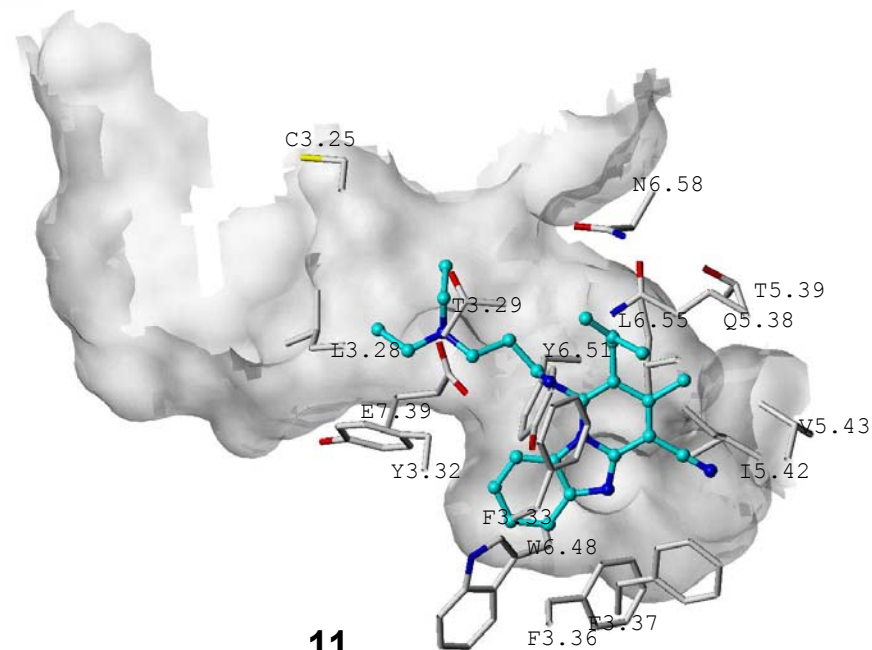
Figure 5



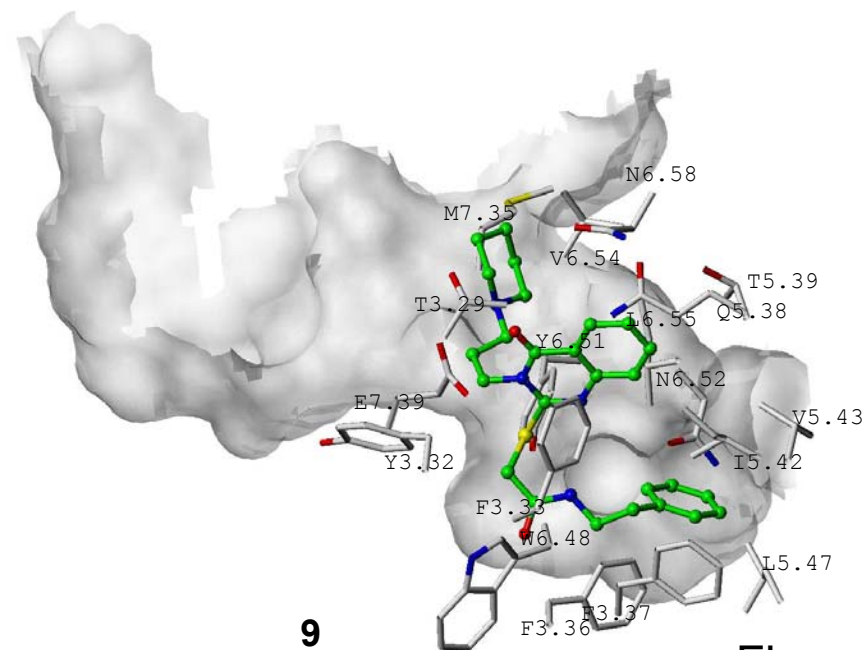
8



1



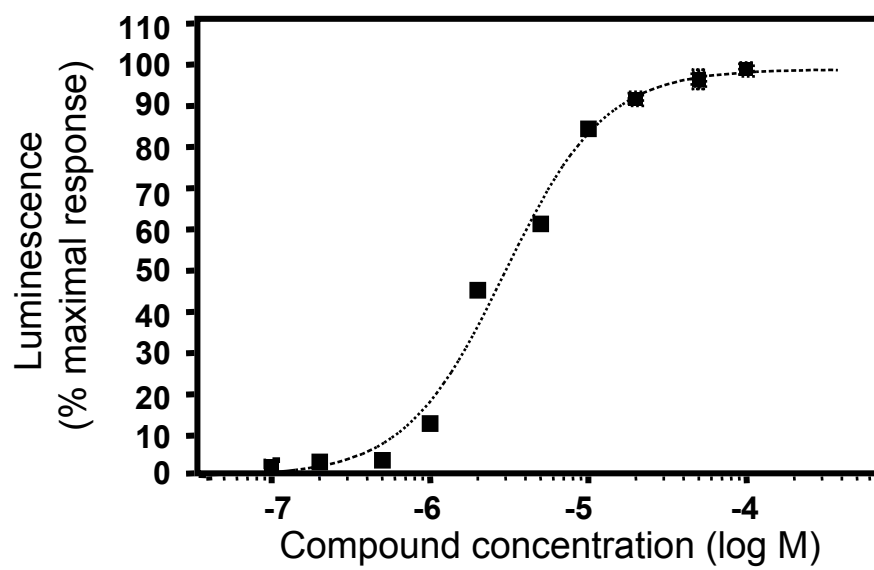
11



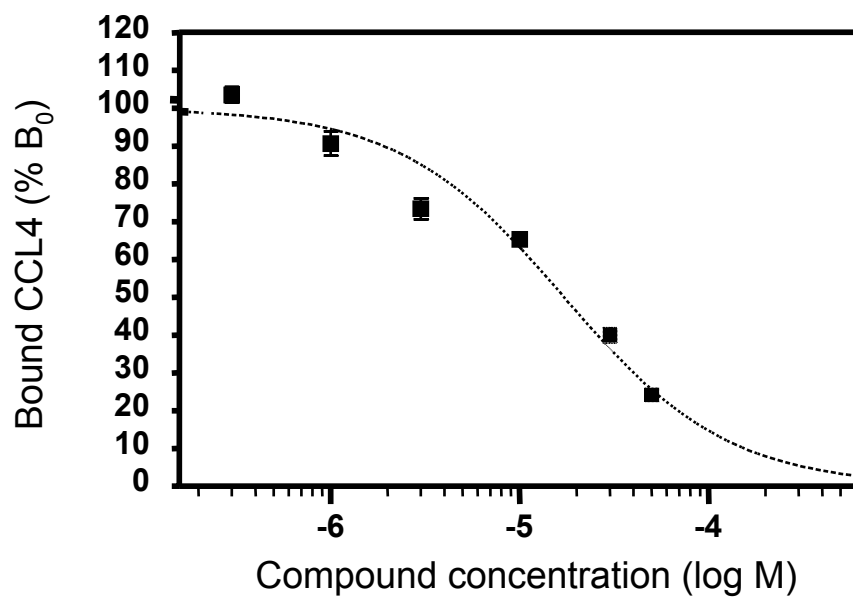
9

Figure 6

A



B



C

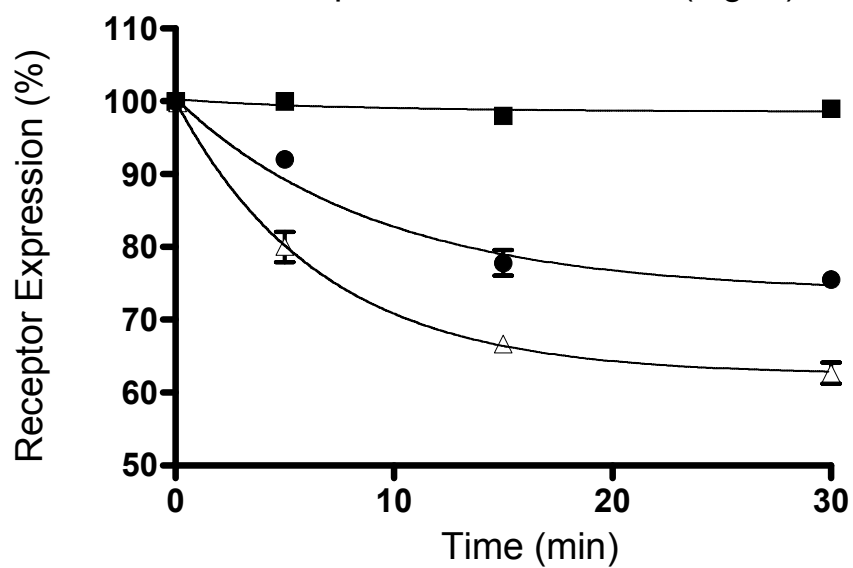
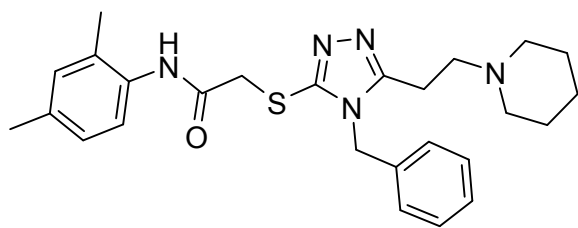


Figure 7

Table of Content Graphics



8 (IC₅₀ = 20 μ M)