Molecular Docking Screening Using Agonist-Bound GPCR Structures: Probing the A$_{2A}$ Adenosine Receptor

David Rodríguez[†,‡,§], **Zhang-Guo Gao**[‖], **Steven M. Moss**[‖], **Kenneth A. Jacobson**[‖,*], and **Jens Carlsson**[†,‡,§,*]

[†]Science for Life Laboratory, Stockholm University, Box 1031, SE-171 21 Solna, Sweden

[‡]Swedish e-Science Research Center (SeRC), SE-100 44 Stockholm, Sweden

[§]Department of Biochemistry and Biophysics and Center for Biomembrane Research, Stockholm University, SE-106 91 Stockholm, Sweden

[‖]Molecular Recognition Section, Laboratory of Bioorganic Chemistry, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, Maryland 20892, United States

Abstract
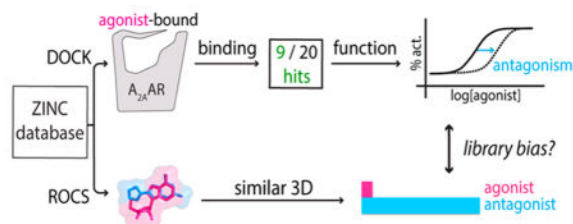
Crystal structures of G protein-coupled receptors (GPCRs) have recently revealed the molecular basis of ligand binding and activation, which has provided exciting opportunities for structure-based drug design. The A$_{2A}$ adenosine receptor (A$_{2A}$AR) is a promising therapeutic target for cardiovascular diseases, but progress in this area is limited by the lack of novel agonist scaffolds. We carried out docking screens of 6.7 million commercially available molecules against active-like conformations of the A$_{2A}$AR to investigate whether these structures could guide the discovery of agonists. Nine out of the 20 predicted agonists were confirmed to be A$_{2A}$AR ligands, but none of these activated the ARs. The difficulties in discovering AR agonists using structure-based methods originated from limited atomic-level understanding of the activation mechanism and a chemical bias toward antagonists in the screened library. In particular, the composition of the screened library was found to strongly reduce the likelihood of identifying AR agonists, which

[*]Corresponding Authors: jens.carlsson@dbb.su.se. kennethj@helix.nih.gov.

Notes

ASSOCIATED CONTENT

Supporting Information

Tables S1–S7. This material is available free of charge via the Internet at http://pubs.acs.org.

reflected the high ligand complexity required for receptor activation. Extension of this analysis to other pharmaceutically relevant GPCRs suggested that library screening may not be suitable for targets requiring a complex receptor–ligand interaction network. Our results provide specific directions for the future development of novel A_{2A}AR agonists and general strategies for structure-based drug discovery.

---

1. INTRODUCTION

G protein-coupled receptors (GPCRs) constitute the largest group of eukaryotic cell-surface receptors and are responsible for signal transduction across the membrane.[1] These receptors share a topology characterized by seven transmembrane (TM) helices and exist in multiple conformational states that regulate intracellular signaling pathways via interactions with G proteins and other effectors. The conformational equilibrium of a GPCR can be modulated by extracellular ligands that bind to the orthosteric binding site.[2] Ligands are called agonists if they stimulate receptor activation of intracellular signaling pathways, while antagonists block the binding of other molecules without significantly altering the receptor basal activity. The development of GPCR ligands has received considerable attention from the pharmaceutical industry, and close to 30% of all marketed drugs target these receptors.[3]

Adenosine receptors (ARs) have been intensively studied for their therapeutic relevance.[4,5] The AR family consists of four subtypes (A_1, A_{2A}, A_{2B}, and A_3) that signal via different G proteins (A_1 and A_3 through G_i and the A_2 subtypes through G_s). A large number of A_{2A}AR antagonists, primarily based on adenine-like (e.g. *N*-[9-chloro-2-(2-furanyl) [1,2,4]triazolo[1,5-*c*]quinazolin-5-amine, CGS15943, **9**) or xanthine scaffolds (e.g., 1,3-dipropyl-8-arylxanthine, XAC, **10**) (Figure 1), have been developed, and several of these are currently being tested in clinical trials against Parkinson's disease.[5] The therapeutic potential of AR agonists has also been recognized by the pharmaceutical industry. The short-acting agonists adenosine (**1**) and its close analogue regadenoson are widely used to evaluate coronary artery disease during myocardial imaging.[6] The only class of non-nucleoside AR agonists was reported in 2003,[7] and at least one of these, the 2-amino-3,5-dicyanopyridine-containing compound capadenoson (**3**), is undergoing clinical trials.[8] AR agonists are relevant for the treatment of cerebral and cardiac ischemia, epilepsy, thrombosis, and arterial hypertension,[5,9] but drug development in this area is limited by the lack of novel scaffolds.

A breakthrough in the AR field was the determination of a crystal structure for the A_{2A} subtype, which revealed the binding mode of the antagonist 4-[2-[7-amino-2-(2-furyl)-1,2,4-triazolo-[1,5-*a*][1,3,5]triazin-5-ylamino]ethyl]phenol (ZM241385, **7**).[10] This enabled the use of structure-based approaches in ligand discovery and optimization.[11–14] Encouragingly, molecular docking screens of chemical libraries against A_{2A}AR crystal structures have been remarkably successful, providing high hit rates and novel ligands.[11–13] All large-scale docking screens against the A_{2A}AR have been carried out against structures crystallized in an inactive state, and accordingly, the discovered ligands have been found to be antagonists in functional assays. This suggested that a bias toward discovering ligands with specific functional properties was encoded in GPCR crystal structures. Access to structures crystallized in complex with agonists should then provide an opportunity to discover ligands

with the ability to activate GPCRs. This hypothesis was recently supported by the discovery of agonists from docking screens against active-like structures of adrenergic and serotonin receptors.[15,16]

Recently, high-resolution structures of the $A_{2A}AR$ in complex with nucleoside agonists were determined.[17,18] In this work, we evaluated the ability of the active-like receptor conformations to guide the discovery of novel $A_{2A}AR$ agonists. On the basis of prospective molecular docking screens, we predicted 20 potential non-nucleoside agonists and evaluated these experimentally in binding and functional assays. The results revealed challenges for structure-based discovery of GPCR agonists associated with the chemical space covered by commercial libraries and understanding of receptor activation.

2. MATERIALS AND METHODS

2.1. Residue Numbering

The Ballesteros–Weinstein residue numbering scheme for GPCRs[19] is indicated in superscript and used throughout the text. According to this notation, residues are numbered following the *X.YY* pattern, where *X* denotes the TM helix and *YY* is a sequence-based number centered at the value 50, which is assigned to the most conserved residue of each TM helix. Residues belonging to an extracellular loop (EL) region are also indicated in superscript.

2.2. Molecular Docking Screens

Crystal structures of the human $A_{2A}AR$ in complex with three agonists (PDB codes 3QAK, 2YDO, and 2YDV)[17,18] and an antagonist (PDB code 3EML)[10] were prepared for docking with DOCK3.6.[20] All non-protein atoms and, in the case of the 3EML and 3QAK structures, the T4–lysozyme fusion protein were removed. Unless stated otherwise, the ionizable residues in the binding site were set to their most probable protonation state at pH 7. The tautomeric states of binding-site histidines were set on the basis of the hydrogen-bonding network. His250[6.52], His278[7.43], and His264[7.29(EL3)] were protonated at Nε, Nδ, and both side-chain nitrogens, respectively. Residue Glu13[1.39] was set to a half-protonated state and Asp52[2.50] was neutralized on the basis of their interactions with the surrounding environment. The ligand conformational sampling implemented in DOCK3.6 is based on superimposition of the docked ligand onto predefined matching spheres in the binding pocket.[21] Two sets of such spheres, derived from either the ribose or adenine groups of the cocrystallized ligands, were used in two separate docking calculations to achieve maximal sampling of the two receptor subpockets. The degree of conformational exploration is determined by the bin size, bin overlap, and distance tolerance parameters, which were set to 0.3, 0.1, and 1.4 Å, respectively. For each ligand solution that passed a steric filter, the binding energy was calculated as the sum of the electrostatic and van der Waals interaction energies,[20] corrected for ligand desolvation.[22] The desolvation penalty was estimated from a precalculated grid based on the transfer free energy of each docked molecule from aqueous solution to a low-dielectric medium.[22,23] Delphi[24] was used to map the electrostatic potential in the binding site using partial charges from a united-atom AMBER force field,[25] except for Asn253[6.55], His278[7.43], and Ser277[7.42]. For the agonist-bound crystal structures,

the side-chain dipole moments of these three residues were increased to favor hydrogen bonding (Table S1 in the Supporting Information), as previously described.[11,15,16] In the case of the inactive-state structure, only the side-chain partial charges of Asn253[6.55] were modified, in agreement with a previous screen with DOCK3.6.[11] To increase sampling in the ribose-binding pocket, two hydroxyl rotamers were explored for Ser277[7.42] in the docking screen against the 2YDO structure. Finally, 100 steps of rigid-body energy minimization were carried out for the top-scoring conformation of each molecule.

Molecules annotated as $A_{2A}AR$ agonists (63) and antagonists (375) were extracted from the WOMBAT database.[11,26] To enable retrospective analysis of docking screens for the different $A_{2A}AR$ crystal structures, property-matched decoys were generated with the DUD-E resource (http://dude.docking.org).[27] For each ligand, this method selects 50 decoy molecules that have similar physicochemical properties (molecular weight, predicted LogP, number of rotatable bonds, hydrogen-bond donors and acceptors, and net charge) and are topologically dissimilar to the ligands. The resulting decoy set had 25 200 molecules, which were used to test the ability of the crystal structures to enrich agonists or antagonists. Partial charges and transfer free energies for the docked molecules were calculated with AMSOL.[28,29] The van der Waals parameters were derived from an all-atom AMBER force field.[30] In total, up to 600 conformations for each docked molecule were pregenerated with the software OMEGA.[31,32]

The lead-like set of 6.7 million commercially available compounds from the ZINC database[33] (accessed in January 2013) were docked to the four $A_{2A}AR$ crystal structures. For each of the three agonist-bound structures, the top-ranked 20 000 molecules from the docking screens were filtered on the basis of interactions with residues involved in ligand binding and activation of the $A_{2A}AR$. Compounds were required to form polar contacts with Asn253[6.55] (3.5 Å cutoff) and either Ser277[7.42] or His278[7.43] (4 Å cutoff). For compounds that were top-ranked in several docking screens, the binding modes obtained for the 2YDO crystal structure were prioritized, followed by 2YDV and 3QAK. The compounds that passed the interaction filter were inspected visually to select compounds for experimental evaluation. In this step, energy terms not explicitly included in the DOCK3.6 scoring function, e.g., ligand internal energy and receptor desolvation, were taken into account as described previously.[11,13,16,34] Twenty compounds were selected on the basis of their complementarity to the receptor binding site, rank difference between agonist- and antagonist-bound $A_{2A}AR$ structures, chemical novelty, and commercial availability.

2.3. Radioligand Binding Assays

Tested compounds were purchased from the vendors Chembridge, Enamine, and VitasM. Compound purity was confirmed as 95% by liquid chromatography–mass spectrometry or NMR spectroscopy. Radioligand binding assays were performed as previously described[11] using membrane preparations from CHO or HEK293 cells stably expressing the human A_1AR, $A_{2A}AR$, or A_3AR. The agonist radioligands [^{3}H](R)-N^6-(phenylisopropyl)-adenosine (R-PIA, 1 nM), [^{3}H]2-[p-(2-carboxyethyl)-phenylethylamino]-5′-N-ethylcarboxamidoadenosine (CGS21680, 10 nM), and [^{125}I]N^6-(4-amino-3-iodobenzyl)-adenosine-5′-N-methyluronamide (I-AB-MECA, 0.5 nM) were used for the A_1AR, $A_{2A}AR$,

and A_3AR, respectively. Binding parameters were calculated using Prism 6 software (GraphPAD, San Diego, CA, USA). IC_{50} values obtained from competition curves were converted to K_i values using the Cheng–Prusoff equation. Data are expressed as mean ± standard error.

2.4. Functional Assays

CHO cells stably expressing the human ARs were cultured in DMEM supplemented with 10% fetal bovine serum, 100 units/mL penicillin, 100 μg/mL streptomycin, and 2 μmol/mL glutamine. Cells were plated in 96-well plates in 100 μL of medium. After 24 h, the medium was removed, and the cells were washed three times with 100 μL of DMEM containing 50 mM HEPES (pH 7.4). For testing of the $A_{2A}AR$ agonist activity, cells were treated for 20 min with agonists in the presence of rolipram (10 μM) and adenosine deaminase (3 units/ mL). For testing of the A_1AR and A_3AR agonist activity, cells were first incubated with agonists for 20 min, and forskolin (10 μM) was then added, after which the mixture was incubated for an additional 15 min. To test antagonist activity, compounds were added 20 min before the addition of agonists. The reactions were terminated by removal of the supernatant, and cells were lysed upon the addition of 100 μL of lysis buffer (0.3% Tween-20). For determination of cAMP production, an ALPHAScreen cAMP kit was used according to manufacturer's instructions.

2.5. 2D Molecular Similarity Calculations

Ligands with binding constants (K_i, EC_{50}, IC_{50}) < 10 μM for any of the AR subtypes were extracted from the ChEMBL15 database.[35] The software ScreenMD from ChemAxon[36] was used to calculate the similarity Tanimoto coefficients (T_c) of the tested compounds to the ARs ligands using ECFP4 2D fingerprints.

2.6. 3D Molecular Similarity Calculations with ROCS

The ligand-based virtual screening tool Rapid Overlay of Chemical Structures (ROCS)[37,38] was used to calculate the number of commercially available compounds that were similar to representative AR ligands and to agonists of the adrenergic, serotonin, opioid, and P2Y receptors. ROCS relies on the detection of molecules with 3D properties similar to those of a reference compound (molecular query). Briefly, conformers of each compound were superimposed and scored on the basis of their overlap with the query in terms of shape (ShapeTanimoto) and functional groups (ColorTanimoto). All of the 3D similarity calculations were ranked according to the TanimotoCombo metric, which is the sum of the ShapeTanimoto and Color-Tanimoto scores and thus ranges from 0 to 2.

In a first step, conformers of the fragment- and lead-like sets from the ZINC database,[33] 7.4 million molecules in total, were generated with OMEGA.[32] A maximum of 100 conformations were retained for each compound on the basi s of their mmf94s_NoEstat force field energies. Representative $A_{2A}AR$ ligands with different functional properties were selected as molecular queries to quantify differences in library bias toward antagonists and agonists of this receptor. We selected the antagonists **8**, CGS15943, **11**, and DPCPX; the partial agonists LUF5833, LUF5834, and LUF5835; and the full agonists adenosine and 5′-N-ethylcarboxamidoadenosine (NECA) (Figure 1). The conformations of ZM241585, XAC,

adenosine, and NECA were extracted from crystal structures (PDB codes 4EIY, 3REY, 2YDO, and 2YDV, respectively),[18,39,40] and these were also used as templates to generate the conformations for the other AR ligands. Compounds **8** and **11** were derived from relevant substructures of ZM241585 and XAC, respectively. The ligand structures for the remaining AR ligand queries were generated with OMEGA using the same parameters as for the ZINC library, and then ROCS was used to overlay them on the most relevant cocrystallized ligands: CGS15943 on ZM241385; DPCPX on XAC; and LUF5833, LUF5834, and LUF5835 on adenosine. In each case, the top-scored conformer from the superimposition was used as the molecular query. To assess the library bias toward ligands of GPCR families other than ARs, we also carried out ROCS calculations using agonists of the adrenergic receptor (ADR), serotonin receptor, opioid receptor (OR), and P2Y receptor. The 3D coordinates of adrenaline were extracted from a crystal structure of the β_2-ADR (PDB code 4LDO).[41] The conformers of serotonin and adenosine diphosphate (ADP) were extracted from the substructures of the cocrystallized agonist ergotamine and the partial agonist 2-methylthioadenosine-5′-triphosphate, bound to the 5-HT$_{1B}$ and P2Y$_{12}$ receptors (PDB codes 4IAR and 4PY0), respectively.[42,43] The conformers of the synthetic OR agonist meperidine (Table S2 in the Supporting Information) were generated as described above for the AR ligands, and a structure similar to that observed for the phenylpiperidine moiety of the cocrystallized ligand bound to the κ-OR (PDB code 4DJH)[44] was selected as a query for the ROCS screens. Default parameters were used in the ligand-based virtual screens of the ZINC database with ROCS, which included the optimization of the shape-based superimpositions with the ImplicitMillsDean force field ("-optchem true" flag).

3. RESULTS AND DISCUSSION

3.1. Retrospective Molecular Docking Screens against Agonist- and Antagonist-Bound A$_{2A}$AR Crystal Structures

The three available agonist-bound crystal structures of the A$_{2A}$AR were first analyzed to identify receptor–ligand interactions responsible for receptor activation. Slightly different conformations were observed for the orthosteric binding site depending on the cocrystallized nucleoside agonist (adenosine, NECA, and 6-(2,2-diphenylethylamino)-9-((2R,3R,4S,5S)-5-(ethylcarbamoyl)-3,4-dihydroxytetrahydrofuran-2-yl)-N-(2-(3-(1-(pyridin-2-yl)piperidin-4-yl)ureido)ethyl)-9H-purine-2-carboxamide (UK-432097), corresponding to PDB codes 2YDO, 2YDV, and 3QAK, respectively). Compared with inactive-like structures, the binding sites of agonist-bound receptors were more enclosed, in particular in the region recognizing the ribosyl moiety of the ligands. This was achieved through an upward shift of TM3, affecting the position of Thr88$^{3.36}$, and side-chain conformational changes for His250$^{6.52}$, Ser277$^{7.42}$, and His278$^{7.43}$ (Figure 2). Hydrogen bonds between the ligand and Asn253$^{6.55}$, His278$^{7.43}$, Ser277$^{7.42}$, and Glu169$^{5.30(EL2)}$ were conserved in all of the A$_{2A}$AR crystal structures with agonists bound. Among these, Glu169$^{5.30(EL2)}$ was not considered to be key for activation because this residue is not conserved across all AR subtypes, which are likely activated through a similar molecular mechanism. Ligand interactions with Asn253$^{6.55}$ were conserved for both agonist- and antagonist-bound structures, suggesting that interactions with this residue are key for ligand binding but cannot alone activate the receptor. The ribosyl moiety, which is crucial for agonist activity,

formed hydrogen bonds to Ser277$^{7.42}$ and His278$^{7.43}$ in each of the three structures. These observations agreed well with mutagenesis data demonstrating the importance of Asn253$^{6.55}$ for the binding of both A$_{2A}$AR antagonists and nucleoside agonists, whereas Ser277$^{7.42}$ and His278$^{7.43}$ typically affect only agonist binding.[45] Thr88$^{3.36}$ and His250$^{6.52}$ were not considered to be crucial for activation because adenosine does not interact directly with either of these residues (Figure 2). This was further supported by the fact that nucleoside agonists without a 5′-hydroxyl group, which is located in the close vicinity of His250$^{6.52}$ and Thr88$^{3.36}$ in the A$_{2A}$AR–adenosine crystal structure, act as full agonists.[46] On the basis of these observations, interactions with Asn253$^{6.55}$ and the formation of hydrogen bonds with Ser277$^{7.42}$ and His278$^{7.43}$ were considered to be the most important for A$_{2A}$AR activation.

The three agonist-bound and one antagonist-bound A$_{2A}$AR crystal structures (Figure 2) were prepared for molecular docking in DOCK3.6.[22] As in the previous DOCK screens against the A$_{2A}$AR, an increased dipole moment for the Asn253$^{6.55}$ side-chain led to more accurate predictions of ligand binding modes.[11,13] Docking of agonists was also improved for the active-like structures when the side-chain dipole moments for residues Ser277$^{7.42}$ and His278$^{7.43}$ were increased. Docking scores were significantly enhanced with the modified electrostatic potential, and these grids were used throughout this work (Table S1 in the Supporting Information). Sets of A$_{2A}$AR agonists and antagonists together with property-matched decoys[27] were docked to the four A$_{2A}$AR structures. Ligand conformations were sampled in the rigid receptor structures, and a physics-based scoring function was used to predict the binding energies for the screened molecules.[22] Encouragingly, all three active-like crystal structures displayed strong enrichment of agonists over decoys on the basis of their receiver operating characteristic (ROC) curves and enrichment factors at 1% of the ranked database (EF1%) (Figure 3). These structures also had significantly better enrichment of agonists compared with the antagonist-bound receptor conformation. For example, the 2YDO structure enriched agonists 63-fold better than random, whereas the corresponding value for the antagonists was only 6-fold. Conversely, the inactive-state crystal structure (PDB code 3EML) selected antagonists better than agonists. Additionally, the enrichment of A$_{2A}$AR agonists by the active-like receptor structures was on average 3.5-fold higher than for antagonists. Despite the relatively small structural changes in the binding site for the agonist-bound crystal structures, these receptor conformations appeared to be more suitable for identifying agonists on the basis of a retrospective assessment.

3.2. Prospective Screens for Novel A$_{2A}$AR Agonists

Molecular docking screens of commercially available libraries were performed with the goal of discovering new non-nucleoside ligands that could activate the A$_{2A}$AR. A set of 6.7 million lead-like molecules from the ZINC database[33] were docked to the orthosteric sites of the agonist- and antagonist-bound structures.[10,17,18] To assess whether it was reasonable to expect that AR ligands would be discovered from the screen, we quantified the number of known A$_{2A}$AR agonists and antagonists that would have been ranked among the top 0.3% of the molecules in the database. Encouragingly, we found that several known agonists, representing both the nucleoside and non-nucleoside scaffolds, were top-ranked in the

screens based on the active-like conformations of the $A_{2A}AR$. For these structures, a higher percentage of the agonists compared with antagonists was also recovered (Table S3 in the Supporting Information). As AR antagonists were also found to be among the top-ranked compounds in screens against the active-like receptor conformation, we used an additional filter that identified top-ranked compounds that interacted with residues important for $A_{2A}AR$ activation. This filter was based on polar contacts with Asn253[6.55] together with either Ser277[7.42], His278[7.43], or both of these residues and resulted in 600–1200 candidate molecules for each of the three agonist-bound crystal structures. Compounds were primarily selected from the screen against the 2YDO structure, as it had the best enrichment of agonists (Figure 3). We also prioritized 2YDV over 3QAK because of the alternative side-chain conformation observed for Glu169[5.30(EL2)] in the latter structure, which was stabilized by bulky substituents of the druglike ligand UK-432097 (Figure 2C). The binding poses for molecules identified in prospective screens were visually inspected to select candidates for experimental testing. Compounds satisfying the contact criteria were deeply buried in the orthosteric site of the receptor and typically had an aromatic group that overlaid with the adenine moiety of the cocrystallized agonists. Similar to the nucleoside agonists, the compounds also occupied the ribose-binding pocket. A set of 20 commercially available compounds was selected for experimental testing on the basis of their complementarity to the $A_{2A}AR$ binding site (Table 1 and Table S4 in the Supporting Information). Twelve compounds were identified on the basis of their docking poses in the 2YDO crystal structure. The remaining eight molecules were selected from the docking screens against the structures 2YDV (two compounds) and 3QAK (six compounds). The predicted agonists formed hydrogen bonds to Asn253[6.55] and extended into the ribose-binding pocket with polar moieties. Thus, the interactions formed by the predicted ligands resembled those of known $A_{2A}AR$ agonists. Additionally, the vast majority of the selected compounds had a significantly better docking rank for the agonist-bound crystal structure compared with the antagonist-bound structure, suggesting that they would stabilize an active conformation of the $A_{2A}AR$.

3.3. Binding and Functional Assays for Predicted $A_{2A}AR$ Agonists

Radioligand binding assays were performed for the 20 selected compounds at the human A_{2A}-, A_1- and A_3AR subtypes. Nine compounds displayed binding affinities better than 10 μM at the $A_{2A}AR$, which corresponds to a hit rate of 45%. Full concentration–response curves and K_i values were determined for all compounds that showed more than 50% inhibition at any of the three assayed AR subtypes (Table 1). Seven ligands were found to display submicromolar $A_{2A}AR$ affinities. Compound **17** was the most potent compound, with a K_i value of 16 nM. Eight out of the nine discovered ligands were retrieved from the docking screen against the 2YDO structure, whereas the remaining compound originated from the 2YDV screen. None of the six molecules that were selected on the basis of the 3QAK crystal structure were found to bind to the $A_{2A}AR$. Experimental data for all of the tested compounds are shown in Table S4 in the Supporting Information.

Assays measuring levels of intracellular cAMP production were carried out to determine the efficacy of the compounds that displayed significant radioligand inhibition for at least one of the ARs. The potential agonist activities of the compounds at the A_1-, A_{2A}- and A_3ARs were

assessed over a full range of concentrations (10^{-10} to 10^{-5} M). Only compound **23** induced significant activation of the $A_{2A}AR$ at 10 μM. Further characterization showed that the changes in cAMP levels were maintained in CHO cells that did not overexpress the $A_{2A}AR$, suggesting that the observed effect was due to interactions with a different, as yet unidentified target or receptor (Table S5 in the Supporting Information). The ability of the most potent ligands to antagonize agonist-induced cAMP activity at each receptor subtype was also characterized directly (Figure 4). All nine identified ligands were thus functional antagonists of the $A_{2A}AR$.

3.4. Structure–Activity Relationships for Discovered $A_{2A}AR$ Ligands

To further understand the functional properties of the discovered ligands, their predicted binding modes were analyzed on the basis of $A_{2A}AR$ crystal structures, structure–activity relationships, and molecular docking of non-nucleoside agonists.

Seven of the discovered ligands shared a 2-amino-3-cyanopyridine substructure (compounds **13** and **15**–**20**; Table 1 and Figure 5A–G). Compounds **15** and **17**, both of which have a tricyclic core, were the most potent ligands, with K_i values of 37 and 16 nM, respectively. These 2-amino-3-cyanopyridines were also ligands of the A_1- and A_3ARs, but compounds **15** and **17** displayed some selectivity for the $A_{2A}AR$ over the A_3 subtype (10- and 32-fold, respectively). The same overall binding mode was predicted for all seven compounds. The 2-amino-3-cyanopyr-idine group typically established two hydrogen bonds with Asn253[6.55] though the exocyclic amine (donor) and cyano (acceptor) moieties (Figure 5). Similar to the cocrystallized nucleoside agonists, these compounds were also predicted to form hydrogen bonds with residues in the ribose-binding pocket of the $A_{2A}AR$. In four cases (**13**, **15**, **16**, and **18**), a 2-furylmethanol group was predicted to form hydrogen bonds with the side-chain hydroxyl group of Ser277[7.42] in the ribose-binding pocket. The binding affinities for these four compounds at the $A_{2A}AR$ ranged from 37 to 1300 nM, among which compound **15** achieved the highest affinity. Compounds **17** and **19** instead have a hydroxyphenyl substituent with the –OH group at the meta and para position, respectively, which was predicted to affect their interactions in the ribose-binding site. Compound **17** forms a hydrogen bond to Ser277[7.42], whereas compound **19** interacts with His278[7.43]. The remaining ligand in the 2-amino-3-cyanopyridine series, compound **20**, was predicted to extend into the ribose-binding pocket with an imidazole moiety. The last two ligands, imidazo[1,5-*b*]pyridazin-7-amine **23** and pyrimidin-4-amine **31**, belong to two different scaffolds and are also the most novel of the discovered compounds. Compound **23** was a submicromolar ligand of the $A_{2A}AR$ (K_i = 260 nM) and was even more potent at the A_3AR subtype (K_i = 63 nM). In the predicted binding mode, the bicyclic core forms hydrogen bonds to Asn253[6.55] and Glu169[5.30(EL2)], whereas the hydroxyphenyl group was predicted to be within hydrogen-bonding distance of both Ser277[7.42] and His278[7.43]. The known $A_{2A}AR$ ligand most similar to compound **23** was identified by calculating the Tanimoto similarity coefficient (T_c) for all of the known AR ligands in the ChEMBL15 database[35] (Table S6 in the Supporting Information). The AR ligand closest to compound **23** was chemically dissimilar to this hit (T_c = 0.32), suggesting that it represents a novel scaffold. Compound **31** displayed a K_i value of 3.1 μM and had a low similarity coefficient (0.30), also indicating chemical novelty. In this case, a flexible ethanol moiety attached to the 4-

amino group of the pyrimidine ring was predicted to extend into the ribose-binding pocket and form a hydrogen bond to His278[7.43].

Some of the 2-amino-3-cyanopyridine ligands were similar to non-nucleoside agonists of the ARs.[47,48] For example, the known AR ligand closest to antagonist **17** was a nucleoside agonist (Table S6 in the Supporting Information). As crystal structures of the $A_{2A}AR$ in complex with a non-nucleoside agonist are not currently available, we docked two representative compounds, i.e., 3,5-dicyanopyridines LUF5834 and LUF5835, to the active-like structures of the receptor (Figure 6). Analogously to compounds **15** and **17**, the hydroxyphenyl groups of LUF5834 and LUF5835 were predicted to form hydrogen bonds to Ser277[7.42] and His278[7.43], respectively. As the flexible imidazole substituent of the non-nucleoside agonists was solvent-exposed, the main difference between the two sets of molecules appeared to be the additional nitrile group that extended into the ribose-binding pocket. Although this nitrile group was not predicted to be involved in direct interactions with the receptor, it may play a role in the activation mechanism by coordinating the bound water molecules that have been observed in the ribose-binding pocket of the $A_{2A}AR$ crystal structures, as suggested by Lane et al.[49] However, it should be noted that A_1- and $A_{2A}AR$ agonists with a 2-amino-3-cyanopyrimidine scaffold lacking this nitrile group have been developed.[50] We also identified an additional plausible binding mode for LUF5834 from our docking calculations (Figure 6B). In this case, no direct hydrogen bonds were formed with Ser277[7.42] or His278[7.43]. Instead, the imidazole moiety was deeply buried in the orthosteric site and hydrogen-bonded to His250[6.52] and Thr88[3.36].

One explanation of why the discovered compounds did not activate the ARs may be related to the fact that they were unable to form all the polar interactions required to shift the conformational equilibrium of the receptor toward an active state. On the basis of crystal structures, we found that polar interactions with Asn253[6.55] were sufficient to antagonize the $A_{2A}AR$. In addition, several antagonists were found to fill only a fraction of the binding-site volume. In contrast, the agonists occupied nearly the entire binding site, and judged by the active-like crystal structures and mutagenesis data, interactions with Asn253[6.55], Ser277[7.42], and His278[7.43] were key for receptor activation. Our discovered ligands typically formed hydrogen bonds to only one of the latter two residues. To quantify the number of commercially available compounds predicted to have the same interaction pattern as cocrystallized $A_{2A}AR$ agonists, we re-examined the top-ranked 50 000 molecules from each of the three docking screens. Compounds were required to establish hydrogen bonds (3.5 Å cutoff) with different donor/acceptor atoms to Ser277[7.42], His278[7.43], and Asn253[6.55], including a double hydrogen bond with the side-chain amide of the latter residue. Of the resulting 150 000 complexes, only 60 (0.04%) achieved this pattern of interactions (Table S7 in the Supporting Information). Of these 54 unique molecules, seven were nucleosides and 40 had azole cores resembling compounds **14**, **26**, **28**, and **29**, which were found to be inactive in binding assays (Table S4 in the Supporting Information). For comparison, more than 50% of the complexes had polar contacts with Asn253[6.55], which on the basis of crystal structures could be sufficient for antagonism. In addition to the interactions with Asn253[6.55], Ser277[7.42], and His278[7.43], cocrystallized nucleoside agonists were found to form additional hydrogen bonds with Thr88[3.36] and His250[6.52], which may also play a role in activation. In

light of these observations, the reason why the discovered ligands were not AR agonists could be related to the fact that activation requires a more complex receptor–ligand interaction network compared with antagonist binding.

3.5. Bias in Chemical Libraries toward Ligands of the A$_{2A}$AR and Other Class-A GPCRs

The fact that only one class of non-nucleoside agonists of the A$_{2A}$AR has been identified emphasizes the challenge of identifying agonists for this receptor. Analysis of crystal structures suggested that activation of the A$_{2A}$AR requires a more complex ligand structure than antagonism. We hypothesized that such compounds may be less abundant than antagonists in commercial libraries. This potential bias in chemical libraries toward ligands with different functional behaviors was analyzed for the A$_{2A}$AR and other pharmaceutically relevant GPCR families.

To quantify the chemical bias toward agonist- and antagonist-like chemotypes of the A$_{2A}$AR in the screened library, we calculated the numbers of compounds similar to known antagonists and agonists in terms of shape and polarity. The ZINC fragment-like and lead-like libraries were screened using ROCS, which superimposes each compound of the database onto a reference ligand and evaluates their similarity on the basis of shape and spatial arrangement of functional groups.[38] The AR ligands were carefully selected to ensure that differences in ligand complexity mainly originated from their efficacy profiles for the A$_{2A}$AR. For instance, the chemical complexity of ligands often increases as they are optimized for selectivity and ADMET properties, but this is not related to their functional profile. For these reasons, we focused the ROCS calculations on nonselective AR agonists and antagonists of similar and low molecular weight. Adenosine, NECA, LUF5833, LUF5834, and LUF5835 (compounds **1**, **2**, and **4**–**6**, respectively) were selected as agonists. As A$_{2A}$AR antagonists, we selected four compounds representing two distinct scaffolds from the ChEMBL data-base.[35] Compounds **8** (K_i = 160 nM) and CGS15943 (**9**) (K_i = 1.2 nM) were selected as adenine-like scaffolds, whereas 8-cyclopentyl-1,3-dipropylxanthine (DPCPX, **12**) (K_i = 129 nM) and **11** (K_i = 2.1 μM) were chosen from the group of xanthine derivatives. The similarities of each molecule in the ZINC database to the nine AR ligands were quantified using the TanimotoCombo metric. A TanimotoCombo value of 2 represents a perfect overlay of shape and polarity, whereas a value of 0 indicates no similarity. TanimotoCombo values   1.2 were considered to represent significant similarity, but the conclusions presented in this work are valid over a wide range of values.

The results from the ROCS calculations based on AR ligands are shown in Figure 7 and Table 2. The numbers of commercially available compounds from the ZINC database with significant similarity to the five A$_{2A}$AR agonists ranged from 210 to 1430. The ROCS search identified between 210 and 580 commercially available compounds that were similar to non-nucleoside agonists (LUF5833, LUF5834, and LUF5835). Compared with the agonists, significantly larger numbers of compounds in the ZINC library were found to share the same shape and polarity as the adenine- and xanthine-like antagonists. The difference was most pronounced for the search based on CGS15943, which returned 11-fold more compounds than adenosine. The searches based on xanthine derivatives (**11**, **12**) resulted in lower numbers of similar compounds, which were still 2-fold higher than for adenosine. In

this context, it should also be noted that the large number of additional known antagonist scaffolds of the $A_{2A}AR$[51] would likely increase the difference in library bias between agonists and antagonists even further.

In a second step, we investigated whether library bias could also affect docking screens for agonists of other GPCRs. The ROCS calculations were extended to include the ligands of the adrenergic, serotonin, P2Y, and opioid receptors (Figure 8 and Table S2 in the Supporting Information), which have received considerable interest as drug targets.[3] For the first three families, which had members crystallized in complex with agonists, the endogenous ligands were used in the ROCS calculations. A synthetic OR agonist was included, as compounds that activated the κ-OR receptor had been discovered in a recent docking screen.[52] Close to 7500 commercially available molecules were found to be similar to adrenaline, which activates the adrenergic receptors. A similar number of compounds were found to share the same shape and polarity as serotonin. Most of the retrieved compounds were fragment-like (MW < 250), in agreement with the relatively small sizes of these two neurotransmitters. For the OR agonist, which also had relatively low molecular weight, a total of 19 000 similar compounds were identified in the ZINC database. The numbers of compounds identified as similar to the adrenergic, serotonin, and opioid receptor agonists were 4- to 12-fold larger than for adenosine. ADP, an endogenous ligand of the $P2Y_{12}$ receptor, returned the lowest number of similar compounds of all queries (51 molecules), in agreement with the high molecular complexity of this ligand.

3.6. Challenges for Structure-Based Discovery of GPCR Agonists

The high hit rates obtained from structure-based screens against the $A_{2A}AR$[15,16] and other GPCRs (e.g., adrenergic and serotonin receptors) reflect the druggability of their orthosteric sites.[53] For this reason, it was not surprising that a large fraction of the tested compounds from our docking screen were $A_{2A}AR$ ligands. As the screens were carried out against agonist-bound structures, we also anticipated that several of them would be agonists. This was supported by the ability of molecular docking to enrich known agonists, including non-nucleoside scaffolds, both using the screened library and property-matched decoys as background (Figure 3 and Table S5 in the Supporting Information), which has also been observed in structure-based screens involving optimization of nucleoside agonists.[54] For these reasons, we also anticipated that several of the selected compounds from our docking screens would activate the $A_{2A}AR$. The finding that none of the discovered ligands activated the $A_{2A}AR$ suggests that screening success is determined not only by the druggability of the target but also by the composition of the chemical library and the understanding of the activation mechanism at the atomic level.

Although millions of molecules can be now tested in screening campaigns, these represent only a small fraction of the ~10^{60} possible druglike molecules. This limited coverage of chemical space will strongly influence success in ligand discovery efforts.[55] The composition of commercially available libraries has a general bias toward biogenic molecules and likely reflects the focus of drug discovery programs.[56] Previous studies have identified a bias toward GPCR ligands versus other drug target classes on the basis of 2D similarity methods.[57] Shoichet and co-workers estimated that there are 3- to 12-fold more

GPCR-like ligands in commercially available libraries compared with other targets such as kinases, ion channels, and proteases.[58] In particular, Carlsson et al.[11] estimated that there are close to 8-fold more ligands of ARs in chemical libraries compared with two enzyme targets (adenylyl cyclase and AmpC β-lactamase). The degree of library bias also appeared to correlate with hit rates from docking screens against each target; both were approximately 10-fold higher for the ARs compared with the enzymes. The unexpected result from our work, which was not assessed in previous studies, was that the library bias could depend on the sought functional profile of the ligand. The favorable bias toward $A_{2A}AR$ antagonists in chemical libraries was evident from the large number of molecules that had shapes and polarities similar to those of classical adenine- and xanthine-like AR ligands. Whereas we cannot exclude the possibility that the lack of agonists among the hits from our screen reflects limitations of docking or our understanding of receptor activation at the atomic level, our analyses suggest that there are significantly fewer AR agonists in commercially available libraries compared with antagonists. The observed library bias could likely, to a large extent, be explained by the higher molecular complexity of AR agonists compared with antagonists. We also extended our library bias calculations to other pharmaceutically relevant GPCRs. We were particularly interested in the β_2-ADR, 5-HT$_{1B}$ and P2Y$_{12}$ receptors, for which agonist-bound crystal structures have recently been determined. A large number of commercially available compounds appeared to have a composition of chemical groups necessary to activate the adrenergic and serotonin receptors. This is not only because there has been a significant medicinal chemistry focus on the aminergic receptor family but also because the endogenous ligands are fragment-sized molecules, for which commercially available libraries have better coverage of chemical space. In agreement with these observations, structure-based screens against the 5-HT$_{1B}$ receptor and β_2-ADR crystal structures resulted in the discovery of agonists.[15,16] A survey of the literature also revealed several ligands that could activate the two aminergic receptor families but shared little chemical similarity to the corresponding endogenous compounds (Figure 8). For the P2Y$_{12}$ receptor, which recognizes agonists that are even more complex than adenosine, a very low number of lead-like compounds were found to have similar shape and polarity as the endogenous ligand. Even if medicinal chemistry efforts have been mainly focused on P2Y$_{12}$ antagonists, this result is also supported by the fact that all known P2Y$_{12}$ agonists are nucleotides. In this context, it should be noted that an agonist of the κ-OR with a significantly less complex structure than the endogenous peptidic ligand was recently discovered in a structure-based screen.[52] This could be explained by the fact that only a fraction of the endogenous agonist may be buried in the orthosteric site of many peptide-binding receptors. A large number of fragment-like molecules have been found to activate ORs, and we also found a significant bias in the ZINC library toward such chemotypes (Table 2). Overall, we have observed a correlation between the molecular complexity of agonists and the number of chemotypes in commercially available libraries that can be expected to activate a specific receptor. This suggests that screens for novel agonists with significant molecular complexity will be strongly limited by the composition of the chemical library. One tempting solution would be to increase the number of complex molecules in the screening library. However, the probability of binding to a given pocket decreases dramatically with higher molecular complexity, which makes such compounds undesirable in screening libraries.[59] Instead, when agonists are required to fulfill a complex interaction

network, structure-guided modifications of antagonist scaffolds to transform them into agonists may be a more suitable approach. The use of fragment-based techniques to target the adenine- and ribose-binding subpockets separately, followed by linkage of fragments to create a new agonist, could be an interesting alternative strategy to library screening. The approach taken here to quantify library bias may also be useful to assess the suitability of a chemical library in screens for agonists of other GPCRs.

The revolution in structural biology for membrane proteins has contributed to deeper understanding of GPCR activation. Atomic-resolution structures of inactive- and active-like conformations of GPCRs have demonstrated that the family undergoes similar structural changes upon activation.[60,61] The ADRs have been most extensively characterized and include atomic-resolution structures in complex with inverse agonists, antagonists, and partial and full agonists.[1] Intriguingly, despite the large conformational rearrangements in the intracellular half of the receptor, the structural changes in the orthosteric site are remarkably subtle. The main difference between ADR agonists and antagonists appears to involve only additional hydrogen bonds between the agonists and serine side-chains in TM5 together with a small contraction of the binding site.[61] Subsequently, docking screens against a structure of the active β_2-ADR led to the identification of new agonists, all of which were predicted to interact with the serines in TM5.[15] In contrast to the β_2-ADR,[62] it should be noted that none of the available $A_{2A}AR$ crystal structures represent a fully active conformation with an intracellular partner. Additional conformational changes could be expected in an active $A_{2A}AR$ structure, which may have influenced the outcome of our docking screens. This suggests that a more detailed understanding of the receptor activation mechanism may be required in order to enable structure-based discovery of AR agonists. Water molecules have been found to be important for accurate modeling of ligand binding to the $A_{2A}AR$.[12,63–66] The $A_{2A}AR$–adenosine crystal structure also showed that several water molecules bridge ligand–receptor interactions.[18] For example, the water-mediated interaction between adenosine and His250[6.52] (Figure 2A) may be relevant for the receptor activation mechanism. Mutagenesis experiments have also revealed differences between nucleoside and non-nucleoside agonists. Intriguingly, Lane et al.[49] found that the binding affinity of the non-nucleoside agonist LUF5834 was essentially unchanged for several $A_{2A}AR$ mutants that significantly reduced the binding of the nucleoside agonist CGS21680. This could suggest that non-nucleoside agonists accomplish activation of the $A_{2A}AR$ through a different set (or a subset) of interactions compared with nucleoside ligands. Along those lines, our docking solutions for the agonist LUF5834 revealed two possible binding modes, in support of the conclusion that interactions either with Ser277[7.42] and His278[7.43] or with Thr88[3.36] and His250[6.52] could be responsible for activation of the $A_{2A}AR$. Crystal structures of the $A_{2A}AR$ bound to non-nucleoside agonists would be valuable in identifying the principal interactions responsible for activation and may shed light on the role of specific residues and water molecules in this process.

4. CONCLUSIONS

Identification of lead compounds with desired selectivity and functional profiles is a key step in drug development for GPCRs. Our docking screen for $A_{2A}AR$ agonists using active-like crystal structures identified novel and potent ligands. The absence of agonists among

the discovered ligands could be explained by the bias toward $A_{2A}AR$ antagonists versus agonists in chemical libraries, a finding that has important implications for drug discovery. Alternative approaches to library screening may be more efficient for the design of ligands required to fulfill a complex interaction network. The chemotypes from our work explore the ribose-binding pocket to a larger extent compared with ligands from previous in silico screens against the $A_{2A}AR$, and as this region is critical for activation, these leads may eventually serve as starting points for the development of novel agonists. Combined with an improved atomic-level understanding of receptor signaling, structure-based methods can be expected to make important contributions to the discovery of GPCR ligands with tailored pharmacological properties.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

ABBREVIATIONS

GPCR	G protein-coupled receptor
AR	adenosine receptor
TM	transmembrane
T_c	Tanimoto coefficient
cAMP	3′,5′-cyclic adenosine monophosphate
NECA	5′-N-ethylcarboxami-doadenosine
ADR	adrenergic receptor
ADP	adenosine diphosphate
OR	opioid receptor

References

1. Katritch V, Cherezov V, Stevens RC. Structure–Function of the G Protein-Coupled Receptor Superfamily. Annu Rev Pharmacol Toxicol. 2013; 53:531–556. [PubMed: 23140243]
2. Deupi X, Kobilka BK. Energy Landscapes as a Tool To Integrate GPCR Structure, Dynamics, and Function. Physiology. 2010; 25:293–303. [PubMed: 20940434]
3. Overington JP, Al-Lazikani B, Hopkins AL. How Many Drug Targets Are There? Nat Rev Drug Discovery. 2006; 5:993–996.
4. Chen JF, Eltzschig HK, Fredholm BB. Adenosine Receptors as Drug Targets—What Are the Challenges? Nat Rev Drug Discovery. 2013; 12:265–286.

5. Fredholm BB, IJzerman AP, Jacobson KA, Linden J, Muller CE. International Union of Basic and Clinical Pharmacology. LXXXI. Nomenclature and Classification of Adenosine Receptors—An Update. Pharmacol Rev. 2011; 63:1–34. [PubMed: 21303899]

6. Al Jaroudi W, Iskandrian AE. Regadenoson: A New Myocardial Stress Agent. J Am Coll Cardiol. 2009; 54:1123–1130. [PubMed: 19761931]

7. Diedrichs, N.; Henninger, K.; Hübsch, W.; Krämer, T.; Krahn, T.; Rosentreter, U.; Shimada, M.; Stasch, JP. Substituted 2-Thio-3,5-dicyano-4-phenyl-6-aminopyridines and Their Use as Adenosine Receptor-Selective Ligands. WO. 2003008384 A1. Jan 30. 2003

8. Tendera M, Gaszewska-Zurek E, Parma Z, Ponikowski P, Jankowska E, Kawecka-Jaszcz K, Czarnecka D, Krzeminska-Pakula M, Bednarkiewicz Z, Sosnowski M, Ochan Kilama M, Agrawal R. The New Oral Adenosine A_1 Receptor Agonist Capadenoson in Male Patients with Stable Angina. Clin Res Cardiol. 2012; 101:585–591. [PubMed: 22370739]

9. Yan L, Burbiel JC, Maass A, Muller CE. Adenosine Receptor Agonists: From Basic Medicinal Chemistry to Clinical Development. Expert Opin Emerging Drugs. 2003; 8:537–576.

10. Jaakola VP, Griffith MT, Hanson MA, Cherezov V, Chien EYT, Lane JR, IJzerman AP, Stevens RC. The 2. 6 Angstrom Crystal Structure of a Human A_{2A} Adenosine Receptor Bound to an Antagonist. Science. 2008; 322:1211–1217. [PubMed: 18832607]

11. Carlsson J, Yoo L, Gao ZG, Irwin JJ, Shoichet BK, Jacobson KA. Structure-Based Discovery of A_{2A} Adenosine Receptor Ligands. J Med Chem. 2010; 53:3748–3755. [PubMed: 20405927]

12. Katritch V, Jaakola VP, Lane JR, Lin J, IJzerman AP, Yeager M, Kufareva I, Stevens RC, Abagyan R. Structure-Based Discovery of Novel Chemotypes for Adenosine A_{2A} Receptor Antagonists. J Med Chem. 2010; 53:1799–1809. [PubMed: 20095623]

13. Chen D, Ranganathan A, IJzerman AP, Siegal G, Carlsson J. Complementarity between in Silico and Biophysical Screening Approaches in Fragment-Based Lead Discovery against the A_{2A} Adenosine Receptor. J Chem Inf Model. 2013; 53:2701–2714. [PubMed: 23971943]

14. Carlsson J, Tosh DK, Phan K, Gao ZG, Jacobson KA. Structure–Activity Relationships and Molecular Modeling of 1,2,4-Triazoles as Adenosine Receptor Antagonists. ACS Med Chem Lett. 2012; 3:715–720. [PubMed: 23342198]

15. Weiss DR, Ahn S, Sassano MF, Kleist A, Zhu X, Strachan R, Roth BL, Lefkowitz RJ, Shoichet BK. Conformation Guides Molecular Efficacy in Docking Screens of Activated β-2 Adrenergic G Protein Coupled Receptor. ACS Chem Biol. 2013; 8:1018–1026. [PubMed: 23485065]

16. Rodríguez D, Brea J, Loza MI, Carlsson J. Structure-Based Discovery of Selective Serotonin 5-HT_{1B} Ligands. Structure. 2014; 22:1140–1151. [PubMed: 25043551]

17. Xu F, Wu H, Katritch V, Han GW, Jacobson KA, Gao ZG, Cherezov V, Stevens RC. Structure of an Agonist-Bound Human A_{2A} Adenosine Receptor. Science. 2011; 332:322–327. [PubMed: 21393508]

18. Lebon G, Warne T, Edwards PC, Bennett K, Langmead CJ, Leslie AG, Tate CG. Agonist-Bound Adenosine A_{2A} Receptor Structures Reveal Common Features of GPCR Activation. Nature. 2011; 474:521–525. [PubMed: 21593763]

19. Ballesteros JA, 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. 1995; 25:366–428.

20. Lorber DM, Shoichet BK. Hierarchical Docking of Databases of Multiple Ligand Conformations. Curr Top Med Chem. 2005; 5:739–749. [PubMed: 16101414]

21. Shoichet BK, Kuntz ID. Matching Chemistry and Shape in Molecular Docking. Protein Eng. 1993; 6:723–732. [PubMed: 7504257]

22. Mysinger MM, Shoichet BK. Rapid Context-Dependent Ligand Desolvation in Molecular Docking. J Chem Inf Model. 2010; 50:1561–1573. [PubMed: 20735049]

23. Shoichet BK, Leach AR, Kuntz ID. Ligand Solvation in Molecular Docking. Proteins: Struct, Funct Bioinf. 1999; 34:4–16.

24. Nicholls A, Honig B. A Rapid Finite Difference Algorithm, Utilizing Successive Over-relaxation To Solve the Poisson–Boltzmann Equation. J Comput Chem. 1991; 12:435–445.

25. Weiner SJ, Kollman PA, Case DA, Singh UC, Ghio C, Alagona G, Profeta S, Weiner P. A New Force Field for Molecular Mechanical Simulation of Nucleic Acids and Proteins. J Am Chem Soc. 1984; 106:765–784.

26. Olah, M.; Rad, R.; Ostopovici, L.; Bora, A.; Hadaruga, N.; Hadaruga, D.; Moldovan, R.; Fulias, A.; Mractc, M.; Oprea, TI. WOMBAT and WOMBAT-PK: Bioactivity Databases for Lead and Drug Discovery. In: Schreiber, SL.; Kapoor, TM.; Wess, G., editors. Chemical Biology: From Small Molecules to Systems Biology and Drug Design. Vol. 2. Wiley-VCH; Weinheim, Germany: 2008. p. 760-786.

27. Mysinger MM, Carchia M, Irwin JJ, Shoichet BK. Directory of Useful Decoys, Enhanced (DUD-E): Better Ligands and Decoys for Better Benchmarking. J Med Chem. 2012; 55:6582–6594. [PubMed: 22716043]

28. Chambers CC, Hawkins GD, Cramer CJ, Truhlar DG. Model for Aqueous Solvation Based on Class IV Atomic Charges and First Solvation Shell Effects. J Phys Chem. 1996; 100:16385–16398.

29. Li J, Zhu T, Cramer CJ, Truhlar DG. New Class IV Charge Model for Extracting Accurate Partial Charges from Wave Functions. J Phys Chem A. 1998; 102:1820–1831.

30. Weiner SJ, Kollman PA, Nguyen DT, Case DA. An All Atom Force-Field for Simulations of Proteins and Nucleic-Acids. J Comput Chem. 1986; 7:230–252.

31. *OMEGA*, version 2.4.3. OpenEye Scientific Software; Santa Fe, NM: http://www.eyesopen.com

32. Hawkins PC, Skillman AG, Warren GL, Ellingson BA, Stahl MT. Conformer Generation with OMEGA: Algorithm and Validation Using High Quality Structures from the Protein Databank and Cambridge Structural Database. J Chem Inf Model. 2010; 50:572–584. [PubMed: 20235588]

33. Irwin JJ, Sterling T, Mysinger MM, Bolstad ES, Coleman RG. ZINC—A Free Tool To Discover Chemistry for Biology. J Chem Inf Model. 2012; 52:1757–1768. [PubMed: 22587354]

34. Carlsson J, Coleman RG, Setola V, Irwin JJ, Fan H, Schlessinger A, Sali A, Roth BL, Shoichet BK. Ligand Discovery from a Dopamine D_3 Receptor Homology Model and Crystal Structure. Nat Chem Biol. 2011; 7:769–778. [PubMed: 21926995]

35. Gaulton A, Bellis LJ, Bento AP, Chambers J, Davies M, Hersey A, Light Y, McGlinchey S, Michalovich D, Al-Lazikani B, Overington JP. ChEMBL: A Large-Scale Bioactivity Database for Drug Discovery. Nucleic Acids Res. 2012; 40:1100–1107.

36. *JChem*, version 5.11.4. ChemAxon; Budapest: http://www.chemaxon.com

37. *ROCS*, version 3.2.0.4. OpenEye Scientific Software; Santa Fe, NM: http://www.eyesopen.com

38. Hawkins PCD, Skillman AG, Nicholls A. Comparison of Shape-Matching and Docking as Virtual Screening Tools. J Med Chem. 2007; 50:74–82. [PubMed: 17201411]

39. Liu W, Chun E, Thompson AA, Chubukov P, Xu F, Katritch V, Han GW, Roth CB, Heitman LH, IJzerman AP, Cherezov V, Stevens RC. Structural Basis for Allosteric Regulation of GPCRs by Sodium Ions. Science. 2012; 337:232–236. [PubMed: 22798613]

40. Dore AS, Robertson N, Errey JC, Ng I, Hollenstein K, Tehan B, Hurrell E, Bennett K, Congreve M, Magnani F, Tate CG, Weir M, Marshall FH. Structure of the Adenosine A_{2A} Receptor in Complex with ZM241385 and the Xanthines XAC and Caffeine. Structure. 2011; 19:1283–1293. [PubMed: 21885291]

41. Ring AM, Manglik A, Kruse AC, Enos MD, Weis WI, Garcia KC, Kobilka BK. Adrenaline-Activated Structure of β_2-Adrenoceptor Stabilized by an Engineered Nanobody. Nature. 2013; 502:575–579. [PubMed: 24056936]

42. Zhang J, Zhang K, Gao ZG, Paoletta S, Zhang D, Han GW, Li T, Ma L, Zhang W, Muller CE, Yang H, Jiang H, Cherezov V, Katritch V, Jacobson KA, Stevens RC, Wu B, Zhao Q. Agonist-Bound Structure of the Human $P2Y_{12}$ Receptor. Nature. 2014; 509:119–122. [PubMed: 24784220]

43. Wang C, Jiang Y, Ma J, Wu H, Wacker D, Katritch V, Han GW, Liu W, Huang XP, Vardy E, McCorvy JD, Gao X, Zhou XE, Melcher K, Zhang C, Bai F, Yang H, Yang L, Jiang H, Roth BL, Cherezov V, Stevens RC, Xu HE. Structural Basis for Molecular Recognition at Serotonin Receptors. Science. 2013; 340:610–614. [PubMed: 23519210]

44. Wu H, Wacker D, Mileni M, Katritch V, Han GW, Vardy E, Liu W, Thompson AA, Huang XP, Carroll FI, Mascarella SW, Westkaemper RB, Mosier PD, Roth BL, Cherezov V, Stevens RC.

Structure of the Human -Opioid Receptor in Complex with JDTic. Nature. 2012; 485:327–332. [PubMed: 22437504]

45. Deflorian F, Kumar TS, Phan K, Gao ZG, Xu F, Wu HX, Katritch V, Stevens RC, Jacobson KA. Evaluation of Molecular Modeling of Agonist Binding in Light of the Crystallographic Structure of an Agonist-Bound A_{2A} Adenosine Receptor. J Med Chem. 2012; 55:538–552. [PubMed: 22104008]

46. Hou X, Kim HO, Alexander V, Kim K, Choi S, Park SG, Lee JH, Yoo LS, Gao ZG, Jacobson KA, Jeong LS. Discovery of New Human A_{2A} Adenosine Receptor Agonists: Design, Synthesis, and Binding Mode of Truncated 2-Hexynyl-4′-thioadenosine. ACS Med Chem Lett. 2010; 1:516–520. [PubMed: 21286238]

47. Beukers MW, Chang LC, von Frijtag Drabbe Kunzel JK, Mulder-Krieger T, Spanjersberg RF, Brussee J, IJzerman AP. New, Non-Adenosine, High-Potency Agonists for the Human Adenosine A_{2B} Receptor with an Improved Selectivity Profile Compared to the Reference Agonist N-Ethylcarboxamidoadenosine. J Med Chem. 2004; 47:3707–3709. [PubMed: 15239649]

48. Guo D, Mulder-Krieger T, IJzerman AP, Heitman LH. Functional Efficacy of Adenosine A_{2A} Receptor Agonists Is Positively Correlated to Their Receptor Residence Time. Br J Pharmacol. 2012; 166:1846–1859. [PubMed: 22324512]

49. Lane JR, Klein Herenbrink C, van Westen GJ, Spoorendonk JA, Hoffmann C, IJzerman AP. A Novel Nonribose Agonist, LUF5834, Engages Residues That Are Distinct from Those of Adenosine-like Ligands To Activate the Adenosine A_{2A} Receptor. Mol Pharmacol. 2012; 81:475–487. [PubMed: 22188926]

50. Louvel J, Guo D, Agliardi M, Mocking TA, Kars R, Pham TP, Xia L, de Vries H, Brussee J, Heitman LH, IJzerman AP. Agonists for the Adenosine A_1 Receptor with Tunable Residence Time. A Case for Nonribose 4-Amino-6-aryl-5-cyano-2-thiopyrimidines. J Med Chem. 2014; 57:3213–3222. [PubMed: 24669958]

51. Muller CE, Jacobson KA. Recent Developments in Adenosine Receptor Ligands and Their Potential as Novel Drugs. Biochim Biophys Acta. 2011; 1808:1290–1308. [PubMed: 21185259]

52. Negri A, Rives ML, Caspers MJ, Prisinzano TE, Javitch JA, Filizola M. Discovery of a Novel Selective κ-Opioid Receptor Agonist Using Crystal Structure-Based Virtual Screening. J Chem Inf Model. 2013; 53:512–526.

53. Shoichet BK, Kobilka BK. Structure-Based Drug Screening for G-Protein-Coupled Receptors. Trends Pharmacol Sci. 2012; 33:268–272. [PubMed: 22503476]

54. Tosh DK, Phan K, Gao ZG, Gakh AA, Xu F, Deflorian F, Abagyan R, Stevens RC, Jacobson KA, Katritch V. Optimization of Adenosine 5′-Carboxamide Derivatives as Adenosine Receptor Agonists Using Structure-Based Ligand Design and Fragment Screening. J Med Chem. 2012; 55:4297–4308. [PubMed: 22486652]

55. Virshup AM, Contreras-Garcia J, Wipf P, Yang W, Beratan DN. Stochastic Voyages into Uncharted Chemical Space Produce a Representative Library of All Possible Drug-like Compounds. J Am Chem Soc. 2013; 135:7296–7303. [PubMed: 23548177]

56. Hert J, Irwin JJ, Laggner C, Keiser MJ, Shoichet BK. Quantifying Biogenic Bias in Screening Libraries. Nat Chem Biol. 2009; 5:479–483. [PubMed: 19483698]

57. Keiser MJ, Setola V, Irwin JJ, Laggner C, Abbas AI, Hufeisen SJ, Jensen NH, Kuijer MB, Matos RC, Tran TB, Whaley R, Glennon RA, Hert J, Thomas KL, Edwards DD, Shoichet BK, Roth BL. Predicting New Molecular Targets for Known Drugs. Nature. 2009; 462:175–181. [PubMed: 19881490]

58. Kolb P, Rosenbaum DM, Irwin JJ, Fung JJ, Kobilka BK, Shoichet BK. Structure-Based Discovery of β_2-Adrenergic Receptor Ligands. Proc Natl Acad Sci US A. 2009; 106:6843–6848.

59. Hann MM, Leach AR, Harper G. Molecular Complexity and Its Impact on the Probability of Finding Leads for Drug Discovery. J Chem Inf Comput Sci. 2001; 41:856–864. [PubMed: 11410068]

60. Venkatakrishnan AJ, Deupi X, Lebon G, Tate CG, Schertler GF, Babu MM. Molecular Signatures of G-Protein-Coupled Receptors. Nature. 2013; 494:185–194. [PubMed: 23407534]

61. Lebon G, Warne T, Tate CG. Agonist-Bound Structures of G Protein-Coupled Receptors. Curr Opin Struct Biol. 2012; 22:482–490. [PubMed: 22480933]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

62. Rasmussen SG, DeVree BT, Zou Y, Kruse AC, Chung KY, Kobilka TS, Thian FS, Chae PS, Pardon E, Calinski D, Mathiesen JM, Shah ST, Lyons JA, Caffrey M, Gellman SH, Steyaert J, Skiniotis G, Weis WI, Sunahara RK, Kobilka BK. Crystal Structure of the β_2 Adrenergic Receptor–Gs Protein Complex. Nature. 2011; 477:549–555. [PubMed: 21772288]

63. Lenselink EB, Beuming T, Sherman W, van Vlijmen HW, IJzerman AP. Selecting an Optimal Number of Binding Site Waters To Improve Virtual Screening Enrichments against the Adenosine A_{2A} Receptor. J Chem Inf Model. 2014; 54:1737–1746. [PubMed: 24835542]

64. Sabbadin D, Ciancetta A, Moro S. Bridging Molecular Docking to Membrane Molecular Dynamics To Investigate GPCR–Ligand Recognition: The Human A_{2A} Adenosine Receptor as a Key Study. J Chem Inf Model. 2014; 54:169–183. [PubMed: 24359090]

65. Sabbadin D, Ciancetta A, Moro S. Perturbation of Fluid Dynamics Properties of Water Molecules During G Protein-Coupled Receptor–Ligand Recognition: The Human A_{2A} Adenosine Receptor as a Key Study. J Chem Inf Model. 2014; 54:2846–2855. [PubMed: 25245783]

66. Bortolato A, Tehan BG, Bodnarchuk MS, Essex JW, Mason JS. Water Network Perturbation in Ligand Binding: Adenosine A_{2A} Antagonists as a Case Study. J Chem Inf Model. 2013; 53:1700–1713. [PubMed: 23725291]
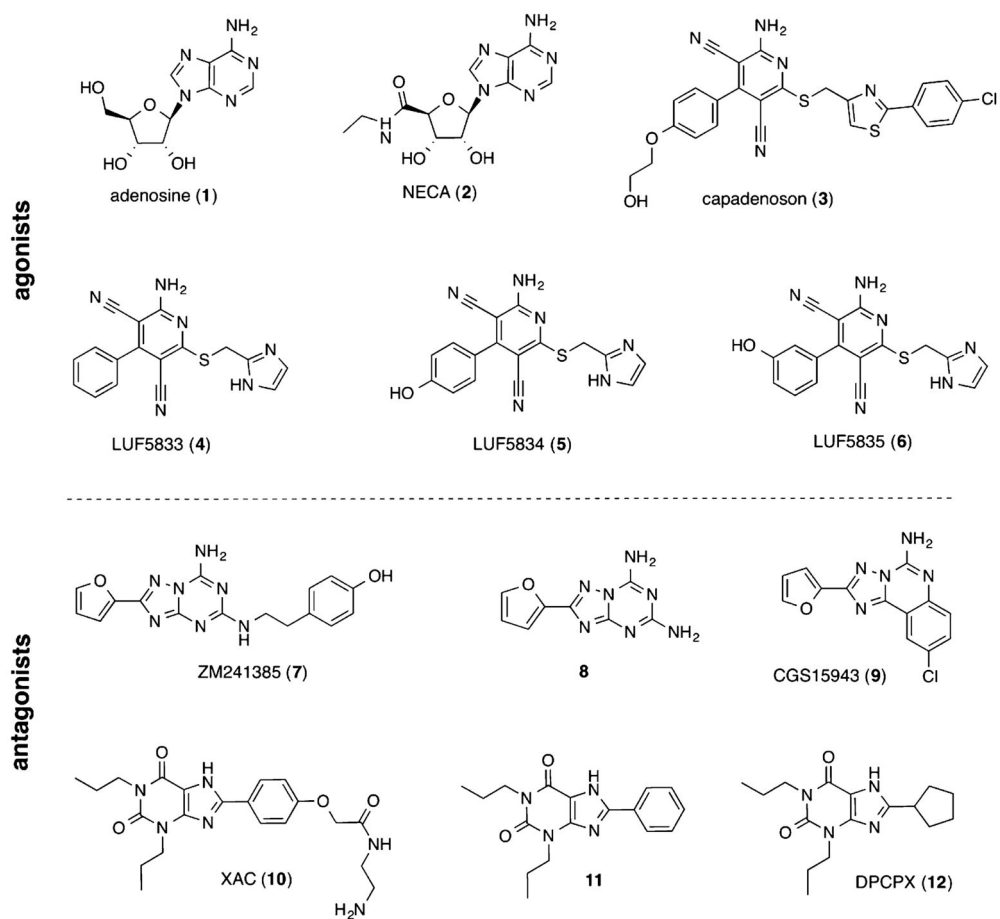
Figure 1.
Representative AR agonists and antagonists.
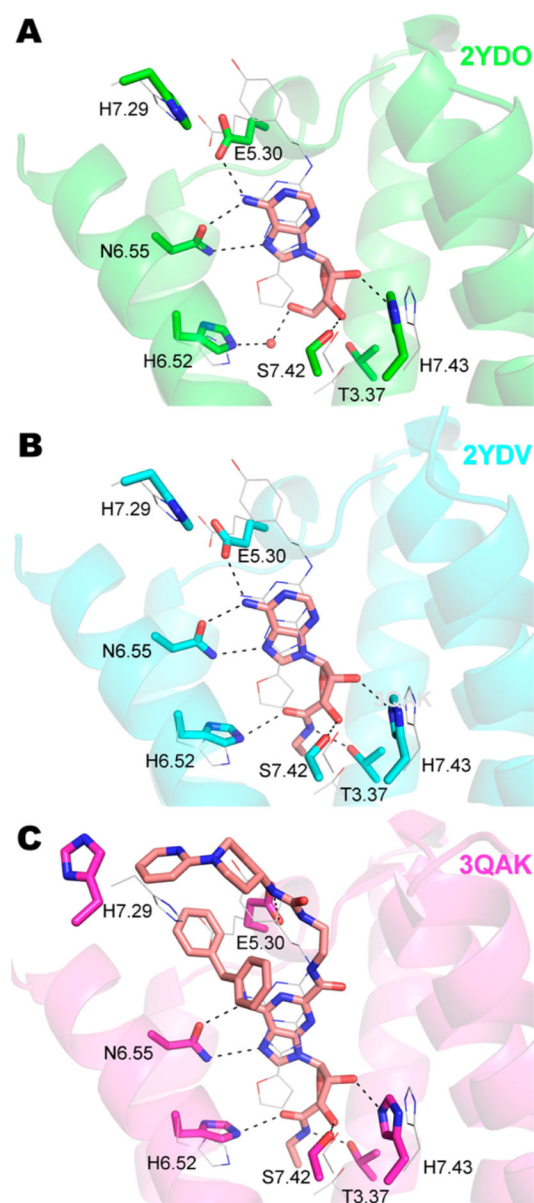
Figure 2.
Comparison of the binding sites for agonist- and antagonist-bound crystal structures of the
A_{2A}AR. The cocrystallized agonists are (A) adenosine, (B) NECA, and (C) UK-432097.
The agonists and selected residues are shown in sticks. In each panel, a crystal structure of
the A_{2A}AR in complex with the antagonist ZM241385 (PDB code 3EML) is shown in lines
with white carbons for comparison.
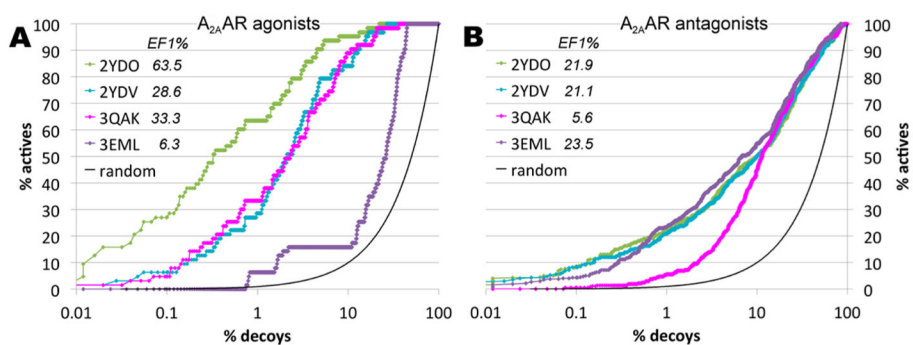
Figure 3.
ROC curves for enrichment of $A_{2A}AR$ agonists (A) and antagonists (B) using different crystal structures. In each panel, the solid black line represents the enrichment expected from random selection, and colored curves represent agonist-bound (2YDO, 2YDV, and 3QAK) and antagonist-bound (3EML) $A_{2A}AR$ structures. The enrichment factor value at 1% of the ranked database (EF1%) for each of the structures is also shown.
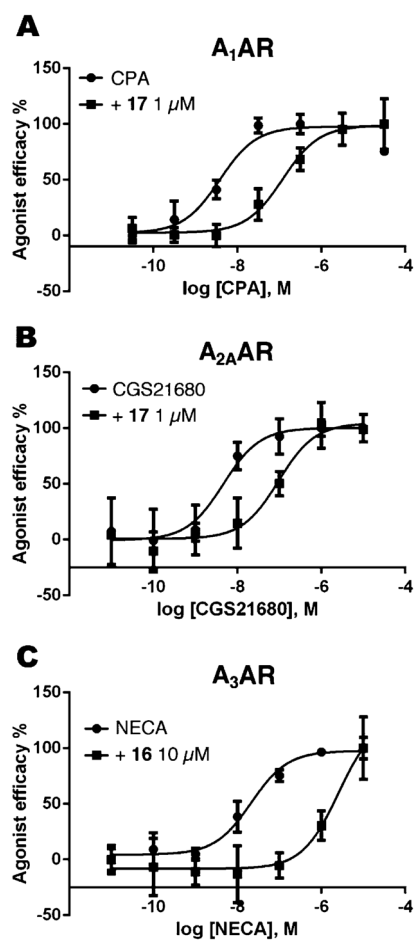
Figure 4.
Functional cAMP assays for representative ligands. Shifts of the dose–response curves induced by the discovered ligands for reference agonists are shown for (A) compound **17** at the A_1AR, (B) compound **17** at the $A_{2A}AR$, and (C) compound **16** at the A_3AR.
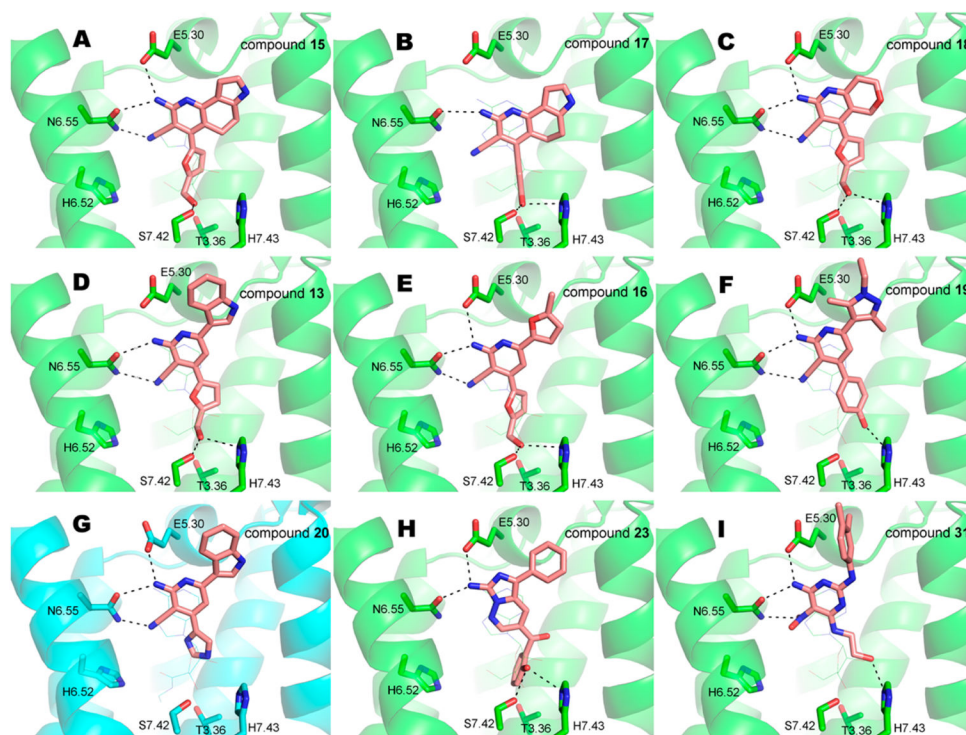
Figure 5.
Predicted A$_{2A}$AR binding modes for discovered ligands: (A) **15**, (B) **17**, (C) **18**, (D) **13**, (E) **16**, (F) **19**, (G) **20**, (H) **23**, (I) **31**. Receptor structures with PDB codes 2YDO and 2YDV are colored with green and cyan carbons, respectively. The cocrystallized agonists are shown in lines.
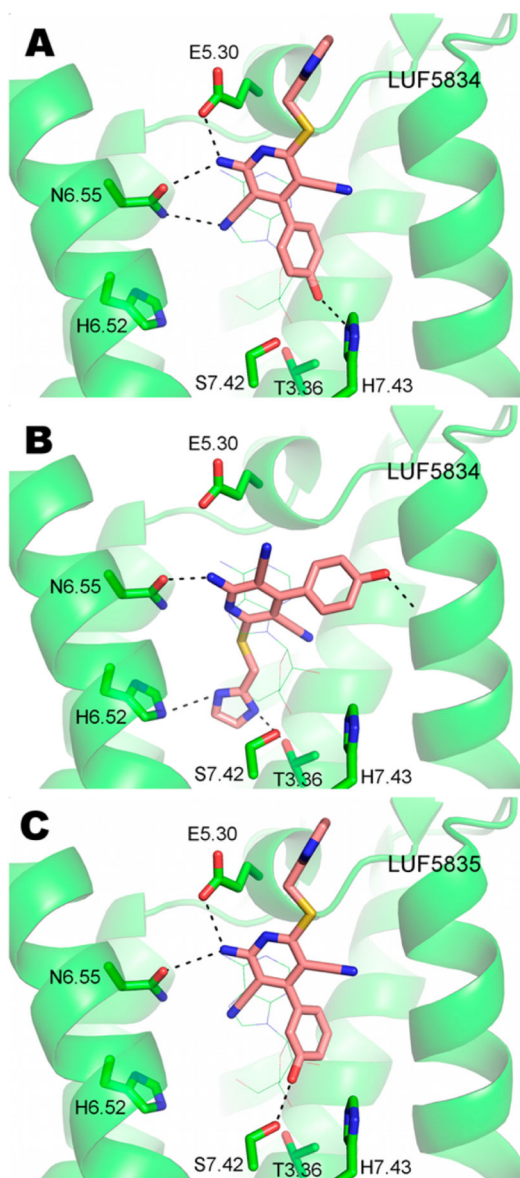
Figure 6.
Predicted binding modes for the non-nucleoside $A_{2A}AR$ agonists (A, B) LUF5834 and (C) LUF5835 bound to the $A_{2A}AR$ (PDB code 2YDO). Panels A and C represent docking solutions with the lowest predicted energies, whereas panel B depicts an alternative binding mode for LUF5834.
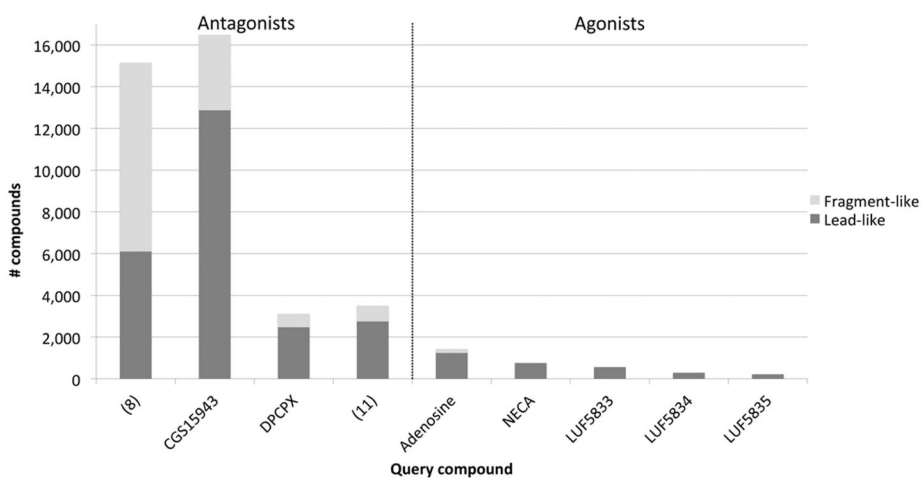
Figure 7.
Quantification of bias in commercial chemical libraries toward $A_{2A}AR$ ligands. Bars represent the numbers of commercially available compounds in the ZINC database that are similar to the molecular queries (representative AR agonists and antagonists) on the basis of ROCS TanimotoCombo scores.
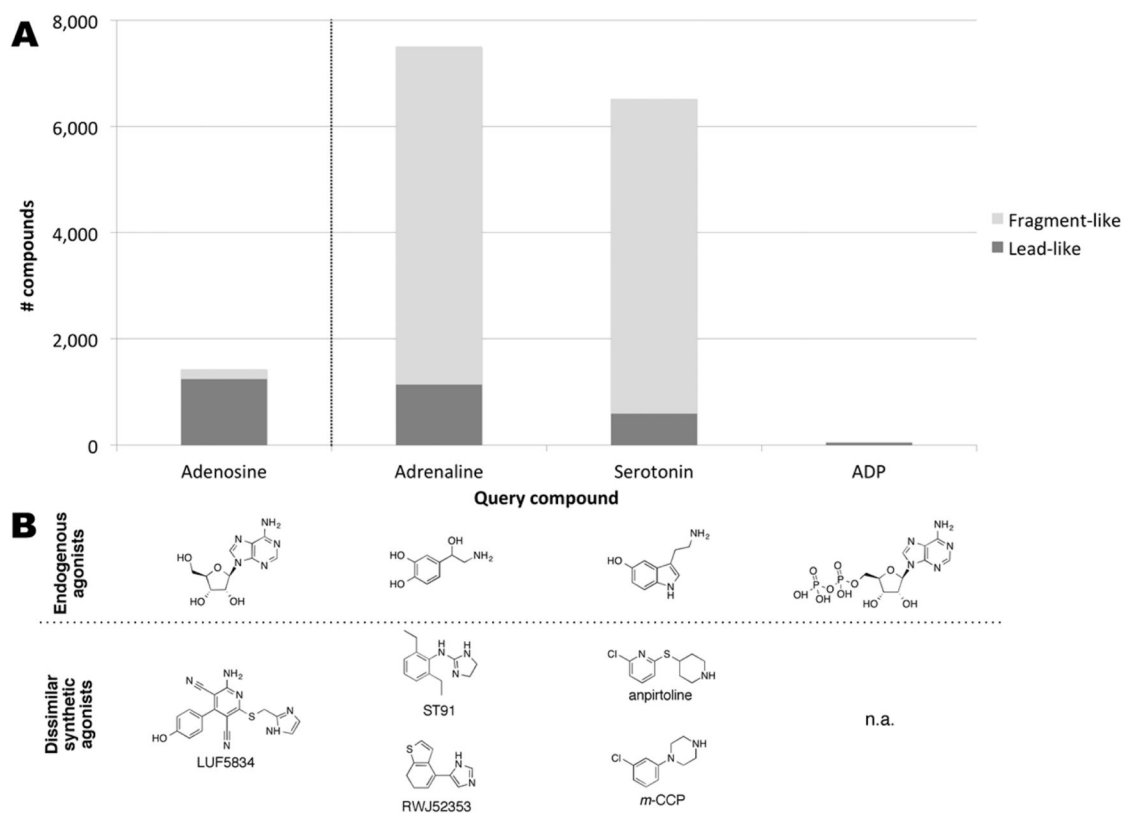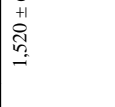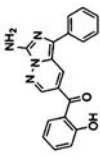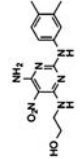
Figure 8.
Library bias in commercially available chemical libraries toward agonists of four GPCR families on the basis of 3D similarities to endogenous ligands. (A) Bars represent the numbers of commercially available compounds similar to the molecular queries on the basis of ROCS calculations. (B) 2D structures of the corresponding endogenous ligands, as well as examples of synthetic agonists that are chemically dissimilar to the natural agonists.

Table 1

Experimental Data and Chemical Novelty of the A$_{2A}$AR Ligands Discovered in the Docking Screens; Radioligand Binding Was Performed Using Membranes of Mammalian Cells Overexpressing One of Three Human AR Subtypes

ID	2D Structure	Active structure[a]	Docking rank[b]		%inh or K$_i$ (nM)[c]			T$_c$[d]
			Active	Inactive	A$_1$AR	A$_{2A}$AR	A$_3$AR	
13		2YDO	538	5,022	1,520 ± 60	1,290 ±280	128 ±17	0.45
15		2YDO	2,498	372,487	217 ± 63	37.2 ±7.8	378 ±129	0.38
16		2YDO	4,297	56,693	122 ± 27	175 ±30	22.1 ±8.0	0.68
17		2YDO	4,368	2,623,374	37.7 ± 3.8	15.9 ±4.7	513 ± 110	0.51
18		2YDO	10,936	582,757	783 ± 216	390 ± 50	3,790 ±1,330	0.40
19		2YDO	13,545	50,767	54 ± 2%	1,670 ±154	2,900 ±112	0.58
20		2YDV	1,325	22,428	1,310 ± 20	760 ± 50	412±61	0.45

ID	2D Structure	Active structure[a]	Docking rank[b]		%inh or K$_i$ (nM)[c]			T$_c$[d]
			Active	Inactive	A$_1$AR	A$_{2A}$AR	A$_3$AR	
23		2YDO	15,286	12,758	63.1 ± 31.9	262 ± 73	365 ± 58	0.32
31		2YDO	4,392	61,333	38 ± 1%	3,110 ± 260	363 ± 81	0.30

[a]PDB codes of the A2AAR active-like structures from which the ligands were selected.

[b]Docking ranks of the compounds.

[c]Data are expressed as mean ± standard error resulting from three independent experiments.

[d]Molecular similarities (T$_C$ values, ECFP4 fingerprints) to the closest AR ligands annotated in the ChEMBL15 database. The most similar known AR ligands are shown in Table S3 in the Supporting Information.

Table 2

Quantification of Library Bias in Commercially Available Chemical Libraries toward GPCR Ligands; the Numbers of Compounds in the Chemical Libraries Similar to Each Ligand Were Obtained from Calculations with ROCS

GPCR family	query compound[a]	efficacy	MW	similar molecules[b]		
				fragment-like	lead-like	total
ARs	8	antagonist	217.2	9044	6100	15144
	CGS15943		285.7	3616	12872	16488
	DPCPX		304.4	647	2467	3114
	11		312.4	759	2752	3511
	adenosine	agonist	267.2	190	1240	1430
	NECA		308.3	40	751	791
	LUF5833		332.4	27	552	579
	LUF5834		348.4	2	290	292
	LUF5835		348.4	3	208	211
ADRs	adrenaline	agonist	184.2	6366	1139	7505
5-HTRs	serotonin	agonist	176.2	5933	589	6522
P2YRs	ADP	agonist	427.2	1	50	51
ORs	meperidine	agonist	247.3	9952	9077	19029

[a] Molecular queries used in ROCS calculations.

[b] Numbers of molecules with TanimotoCombo scores 1.2 in ROCS calculations for 7.4 million commercially available compounds.