JOURNAL OF CHEMICAL INFORMATION AND MODELING

LETTER

pubs.acs.org/jcim

Ligand and Decoy Sets for Docking to G Protein-Coupled Receptors

Edgar A. Gatica and Claudio N. Cavasotto*,†

School of Biomedical Informatics, University of Texas Health Science Center at Houston, 7000 Fannin St., Houston, Texas 77030, United States

Supporting Information

ABSTRACT: We compiled a G protein-coupled receptor (GPCR) ligand library (GLL) for 147 targets, selecting for each ligand 39 decoy molecules, collected in the GPCR Decoy Database (GDD). Decoys were chosen ensuring a ligand−decoy similarity of six physical properties, while enforcing ligand−decoy chemical dissimilarity. The performance in docking of the GDD was evaluated on 19 GPCRs, showing a marked decrease in enrichment compared to bias-uncorrected decoy sets. Both the GLL and GDD are freely available for the scientific community.

INTRODUCTION

With ∼280 family members and more than 39 identified major drug targets,[1] class A G protein-coupled receptors (GPCRs) are targeted by ∼25% of the marketed drugs,[2] and finding new drug lead candidates remains an active area of research in drug discovery. The recent progress in GPCR crystallization, with more than 35 structures of seven different proteins solved in the past four years, opens the opportunity to rationalize site mutagenesis data of both agonists and antagonists and greatly expands the horizon of GPCR structure-based drug discovery.[3,4] Moreover, by providing a diverse set of structural templates, this array of new structures enhances the possibility of developing more accurate homology models of GPCR agonist- and antagonist-bound binding sites,[5−9] which can then be used for high-throughput docking (HTD). The quality of homology models is ultimately judged by the presence of known (for example, from mutagenesis data) or conserved ligand−protein interactions and their performance in docking.[5,10−12]

The development and benchmarking of HTD protocols for GPCRs using either experimental structures or homology models requires an adequate set of ligands and decoys. The necessity of the latter is dictated by the need to avoid artificial enrichment[13−15] by generating sets of decoys with similar physical properties to those of the ligands but that are chemically diverse. The first condition avoids the artificially high enrichment caused by the separation of simple physical properties, whereas the second reduces the likelihood of the decoy sets containing actual binders (which would artificially lower the enrichment). The Directory of Useful Decoys (DUD) for 40 diverse targets was developed following those principles, exhibiting in docking consistently lower enrichments than bias-uncorrected decoy sets.[13] Recently, using the Virtual Decoy Set (VDS) of *in silico* generated decoy molecules, similar results were obtained.[16] In the VDS, the condition of decoy molecules' synthetic feasibility is removed. However, to date, there is no tailored decoy set for docking onto GPCR targets, which would be extremely welcomed in view of the current progress in their crystallization, the need to evaluate and use homology models, and the need to benchmark docking programs.

In this work, we compiled a GPCR ligand library (GLL) of 25 145 ligands for 147 GPCR targets, generating for each of them an associated decoy set with a 39 molecules per ligand ratio, collected in the GPCR Decoy Database (GDD) containing ∼1 million molecules. Decoys were generated ensuring a ligand−decoy similarity of six physical properties, while enforcing ligand−decoy chemical dissimilarity. The performance in docking of the GDD was evaluated on 19 GPCR targets and compared to other decoy sets, finding a significant decrease in enrichment when bias-corrected decoy sets from GDD are used. The use of different charge-related physical properties in decoy set generation was also investigated. Both the GLL and GDD are freely available for the scientific coummunity at http://cavasotto-lab.net/Databases/GDD/.

RESULTS AND DISCUSSION

GPCR Ligand Library (GLL). For each human class A GPCR target listed on the GLIDA database[17] and a given biological activity filter (agonist or antagonist), a list of IDs was extracted and cross-referenced with PubChem to obtain structure files in mol format. Ligands not present in PubChem were not retrieved. Ligands were then prepared with LigPrep[18] and their MACCS fingerprints and seven physical properties calculated: molecular weight (MW), formal charge, number of rotatable bonds (RBs), number of hydrogen bond acceptors (HBAs) and donors (HBDs), octanol−water partition coefficient (log *P*), and topological polar surface area (tPSA; see Methods). Considering that many GPCR ligands are borderline regarding the Lipinski rules,[19] but with the desire of keeping the drug-likeness of the ligand set within a reasonable range, an upper limit of 15 RBs, 12 HBAs, or 7 HBDs was imposed. Native ligands from crystallized GPCRs not present in the set extracted from PubChem were manually added to their corresponding subset. An additional set of 22 β_2AR antagonists taken from ref 20 was also added. All ligand subsets (separated by target/ligand activity) thus generated with five or more ligands are compiled in the GLL, with a total of 25 145 ligands for 147 human class A GPCR targets (cf. Table S1 in the Supporting Information for a

† Address after January 2012: Instituto de Biomedicina de Buenos Aires − Max Planck Society Partner (IBioBA-MPSP), Godoy Cruz 2201, C1425FQA, Buenos Aires, Argentina

Published: December 14, 2011

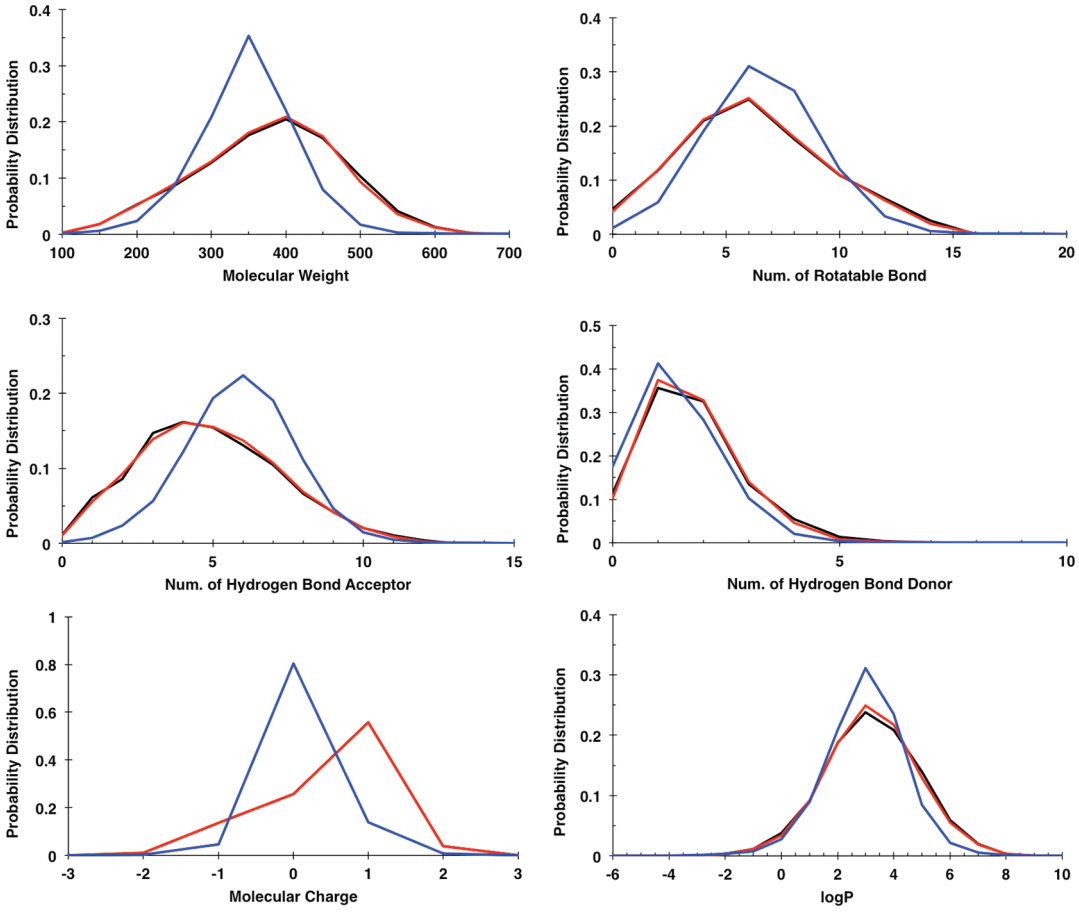

Figure 1. Physical property distributions for the GPCR ligand library (GLL), the GPCR decoy database (GDD), and the random decoy set extracted from ZINC (RZ). Color code: GLL, black; GDD, red; ZINC, blue. The distributions of formal charge for GLL and GDD perfectly overlap; thus only the red line is seen.

complete list of individual GPCRs with the number of agonists and antagonists included in GLL). The GLL is publicly available at http://cavasotto-lab.net/Databases/GDD/.

GPCR Decoys Database (GDD). For each target/ligand-activity subset in GLL, an associated decoy subset containing 39 decoy molecules per ligand was generated. Initially, all molecules (~20 million) from the ZINC database[21] (http://zinc.docking.org, accessed Jan. 2011) were downloaded in their pH 7.0 form and their fingerprints and physical properties calculated as with the GPCR ligands. For each ligand, decoy molecules from ZINC were selected satisfying these criteria: (i) a molecular weight of ±50, (ii) the same number of rotatable bonds and hydrogen bond donors and acceptors, (iii) a log *P* of ±1.0, and (iv) an identical formal charge. The first three are similar to the VDS,[16] while the last constraint stems from the fact that for several GPCR targets, the vast majority of the ligands have nonzero charge. The importance of accounting for formal charge in decoy set generation has been highlighted by Irwin[22] and quantitatively assessed by Mysinger and Shoichet using the DUD set.[23] Duplicate decoy candidates (Tanimoto coefficient (Tc) of 1.0) were discarded. In order to ensure structural diversity between ligands and their associated decoys, decoy candidates with a Tc ≥ 0.75 with respect to *any* ligand in the corresponding GLL subset were removed. From all candidates satisfying criteria ii−iv, the 39 molecules were selected on the basis of the closest match for MW. If 39 molecules meeting those criteria could not be found, the search was repeated relaxing the restrictions on RBs, HBAs, and HBDs by ±1 and tolerating a log *P* of ±1.5. If the total was still less than 39, up to two additional searches were attempted, each one increasing by 1 the margin on rotatable bonds. The decoy molecules eventually found during the additional searches were added to the original subset. Since decoys are selected on a "first come, first served" basis, and to ensure that a diverse set of ligands could select decoys first, within each subset, ligands with Tc < 0.98 to each other were selected and placed at the top of the decoy selection queue.

In cases where 39 decoys could not be found for a given ligand, both the ligand and the incomplete set of associated decoys are not present in the corresponding subsets of GLL and GDD, respectively. This could stem from two causes: (i) The chemical space of the source (ZINC) is not diverse enough; in this case, the inclusion of too few decoys in GDD could lead to an unrealistic higher enrichment. (ii) The decoys have been picked up by another ligand.

The comparison of the six physical properties used in the generation of GLL and GDD is displayed in Figure 1, showing a clear similarity of GLL and GDD in terms of physical properties.

The GDD grouped by target/ligand-activity is publicly available at http://cavasotto-lab.net/Databases/GDD/.

Docking Benchmark. To evaluate the impact of the type of decoy sets in GPCR docking, HTD was conducted on 19 GPCR crystal structures with noncovalently bound ligands: β_2 adrenergic

Table 1. High-Throughput Docking onto 19 Experimentally Solved GPCR Structures: Impact of the Type of Decoy in the Enrichment

GPCR target					EF_{max}		EF_2		EF_{10}		
protein	binding site conformation	PDB code	resolution (Å)	ref	RZ	GDD	RZ	GDD	RZ	GDD	docking RMSD (Å)
β_2AR	antagonist	2RH1	2.4	26	33.3	7.4	13.0	5.9	4.8	3.5	0.5
		3NY8	2.8	27	36.9	13.3	17.2	7.1	5.2	3.9	1.0
		3NY9	2.8	27	36.7	17.1	14.0	5.4	4.9	3.6	0.7
		3NYA	3.2	27	33.3	26.7	12.5	6.9	4.1	3.5	0.9
	agonist	3P0G	3.5	28	38.9	38.2	32.9	21.0	8.6	7.9	1.4
β_1AR	antagonist	2VT4	2.7	29	36.0	5.0	13.6	3.1	5.2	2.6	0.8
		2YCW	3.0	30	30.0	3.8	10.7	2.9	5.0	3.0	0.7
		2YCZ	3.6	30	35.0	6.5	13.6	5.2	4.8	3.1	0.9
	agonist	2Y01	2.5	31	35.0	12.7	19.6	10.8	8.2	5.2	0.6
		2Y02	2.6	31	38.9	34.3	27.1	14.6	8.4	5.7	1.8
		2Y03	2.6	31	38.3	20.0	24.4	8.9	8.5	5.9	2.0
		2Y04	2.8	31	26.7	4.8	17.0	4.1	7.3	3.9	0.5
D_3R	antagonist	3PBL	2.9	32	4.7	1.7	4.6	0.3	2.8	1.5	1.1
A_{2A}AR	antagonist	3EML[a]	2.6	33	17.1	12.9	9.7	7.2	4.4	3.3	0.4
		3EML[b]	2.6	33	37.1	32.8	12.1	11.7	3.4	3.1	0.4
	agonist	3QAK	2.7	34	10.5	2.3	6.8	1.2	3.9	2.1	N/A[c]
		2YDO	3.0	35	38.9	31.1	25.2	20.9	6.5	6.1	0.9
		2YDV	2.6	35	38.5	36.0	24.0	18.5	6.5	6.0	0.7
H_1R	antagonist	3RZE	3.1	36	38.1	32.0	25.2	16.8	7.6	7.1	0.5

[a] Excluding crystallographic waters. [b] Including eight crystallographic waters. [c] The corresponding native ligand was excluded from GLL since it exceeded the limit on number of rotatable bonds.

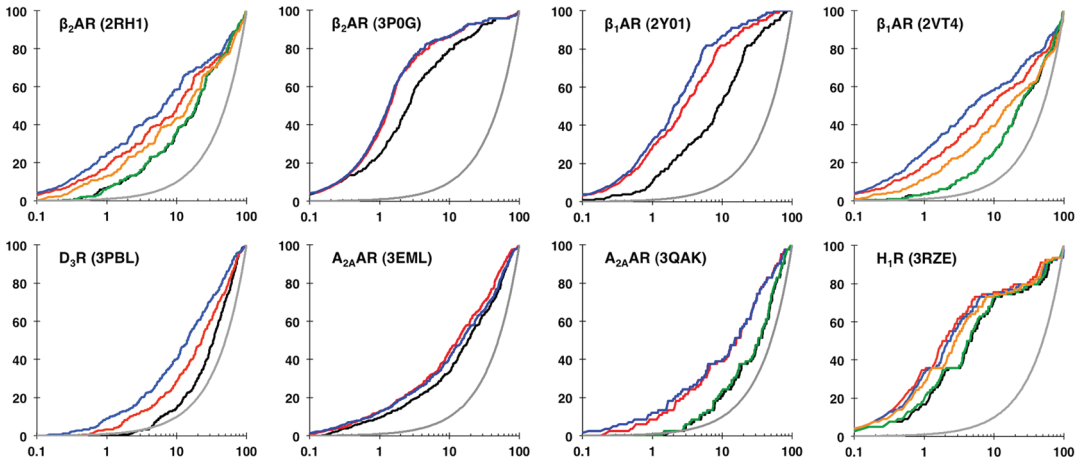

Figure 2. Eight enrichment curves corresponding to docking onto 19 GPCR targets, showing the percent of the screened database (*x* axis) versus the percent of ligands recovered. Ligands were taken from GLL. Decoys were taken from GDD and random sets from ZINC (RZ) and ChemBridge (RCB), mantaining the 39 decoys per ligand ratio. In selected structures, decoys from GDD* and a random set of molecules with a formal charge of +1 from ZINC (RZ1) were also docked. The docking to A_{2A}AR (3EML) displayed was performed without crystallographic waters. The remaining 11 enrichment curves are shown in Figure S1 (Supporting Information). Color code: GDD, black; RZ, red; RCB, blue; GDD*, green; RZ1, orange; random selection, gray.

receptor (β_2AR), β_1 adrenergic receptor (β_1AR) and A_{2A} adenosine receptor (A_{2A}AR) in the agonist and antagonist bound form, and histamine H_1 receptor (H_1R) and dopamine D_3 receptor (D_3R) only in the antagonist bound form (Table 1). Docking was performed using Glide[24] with standard precision (SP) scoring.[25] For each target, docking sets were generated taking the corresponding ligands from GLL and using decoys from GDD and a random collection from ZINC (RZ; 39 decoys were selected per ligand to keep the same ligand/decoy ratio). For the sake of comparison, decoys randomly selected from ChemBridge (RCB) were also docked to representative structures. Enrichment factors (maximum, at 2 and 10%) are summarized in Table 1, together with the RMSD of the top-scoring native ligands. It can be seen that all ligands docked with a RMSD better than 2 Å, and better than 1 Å for >70% of them. The enrichment curves are shown in Figure 2 and Figure S1 (Supporting

Table 2. Impact of Imposing Formal Charge Contraints on Decoy Generation

GPCR target			EF_{max}				EF_2				EF_{10}			
protein	binding site conformation	PDB code	RZ	RZ1	GDD*	GDD	RZ	RZ1	GDD*	GDD	RZ	RZ1	GDD*	GDD
β_2AR	antagonist	2RH1	33.3	19.0	7.0	7.4	13.0	11.3	6.4	5.9	4.8	4.4	3.6	3.5
β_1AR	antagonist	2VT4	36.0	17.5	5.7	5.0	13.6	9.5	3.1	3.1	5.2	3.9	2.6	2.6
A_{2A}AR	agonist	3QAK	10.5	N/A[a]	2.4	2.3	6.8	N/A[a]	1.2	1.2	3.9	N/A[a]	2.3	2.1
H_1R	antagonist	3RZE	38.1	35.0	32.0	32.0	25.2	18.7	17.4	16.8	7.6	7.4	7.3	7.1

[a] A_{2A}AR ligands have an average formal charge of ~0; thus RZ1 was not docked to this structure

Table 3. Impact of Imposing tPSA Contraints on Decoy Generation

GPCR target			EF_{max}				EF_2				EF_{10}			
protein	binding site conformation	PDB code	GDD	GDD*	GDD(p)	GDD*(p)	GDD	GDD*	GDD(p)	GDD*(p)	GDD	GDD*	GDD(p)	GDD*(p)
β_2AR	antagonist	2RH1	7.4	7.0	7.9	7.5	5.9	6.4	6.4	6.4	3.5	3.6	3.6	3.7
β_1AR	antagonist	2VT4	5.0	5.7	5.0	8.0	3.1	3.1	3.1	3.1	2.6	2.6	2.6	2.7
A_{2A}AR	agonist	3QAK	2.3	2.4	2.5	2.5	1.2	1.2	1.2	1.2	2.1	2.3	2.2	2.3
H_1R	antagonist	3RZE	32.0	32.0	32.0	32.0	16.8	17.4	16.1	14.2	7.1	7.3	7.1	7.2

Information), and the associated ROC curves are shown in Figure S2 (Supporting Information). A comparison of the six physical descriptors for the ligand and decoy sets used in the docking experiments, per target/ligand activity, is shown in Figure S3 (Supporting Information).

Enrichments and ROC curves (cf. Table 1 and Figures 2 and S1 and S2, Supporting Information) show a marked decrease in the number of actives recovered when decoy molecules match the physicochemical properties of the ligand set. Enrichments for RCB are slightly higher than RZ. This shows that biased enrichments are obtained using a physically dissimilar decoy set, in agreement with what has been obtained in other cases.[13,16] It should be mentioned that these enrichment values cannot be plainly compared with those of earlier works, mainly due to differences in the decoys sets and/or docking programs. For the antagonist bound β_2AR (PDB code 2RH1), the enrichments using random selections are comparable, though lower, than previously reported.[8,9,20] It should be noted, however, that in those cases, ligands with poor docking scores had been removed from the docking set (cf. ref 20). For the A_{2A}AR antagonist bound (PDB code 3EML), enrichments with RZ and RCB are in line with previous works.[8,37] Given that most of the GPCR crystal structures were solved recently, for most of the receptors, this is the first time that extensive HTD results are reported.

In order to assess the impact of imposing constraints on the formal charge, two alternative decoy sets were generated in the following way (keeping the 39 decoys per ligand ratio): (i) randomly selecting molecules with a formal charge of +1 from ZINC (RZ1), as recently done in a similar fashion;[5] (ii) using the same algorithm used to generate GDD, but with the restriction on the formal charge completely removed (GDD*). These sets were evaluated on four structures: β_2AR (2RH1), β_1AR (2VT4), and H_1R (3RZE) antagonist bound and A_{2A}AR agonist bound (3QAK). Since A_{2A}AR ligands have an average formal charge of ~0, RZ1 was not docked to this structure. The distribution of physical properties of GDD and GDD* for those four receptors is also compared in Figure S3 (Supporting Information). The enrichment values, together with those of RZ and GDD, are displayed in Table 2 and the enrichment curves in Figure 2. Comparing the results obtained using RZ and RZ1, it is seen that the enforcement of a single constraint on the formal charge has a visible impact. The higher enrichments of GDD* compared to GDD are moderate, thus suggesting that constraints on HBAs, HBDs, and log *P* would guarantee a reasonable similarity in terms of charge-related physical properties, though optimal results are obtained when a constraint on the formal charge is also added.

To further explore the use of charge distribution-related properties, we generated decoy sets adding a tPSA constraint of ± 30 Å^2 to the decoy generation algorithm of GDD and GDD*. The two new libraries, GDD(p) and GDD*(p), respectively, were generated and docked to the four structures of Table 2. The enrichment values and curves are shown in Table 3 and Figure S4 (Supporting Information), respectively. The comparison of the distributions for tPSA is displayed in Figure S5 (Supporting Information).

The enrichment curves show a very similar performance using any of the four decoys sets, for each of the four receptors, which can be understood from Figure S5 (Supporting Information): even without imposing the tPSA constraint, GDD and GDD* have already a similar tPSA distribution to the corresponding ligand set, confirming what has been said above regarding the use of HBA, HBD, and log *P* as a sufficient requirement to ensure ligand−decoy set physical similarity. From Table 3 and Figure S4 (Supporting Information), it is seen that adding the tPSA constraint to GDD* slightly improves the performance (lower enrichments) in one case, and this could be considered an alternative to impose the formal charge constraint, since GDD*(p) and GDD enrichments are very similar. Imposing both formal charge and tPSA constaints [GDD(p)], on average, slightly deteriorates the performance. In this work, we preferred to impose formal charge rather than tPSA constraints in our final working decoy sets (GDD) for purely practical reasons.

dx.doi.org/10.1021/ci200412p |*J. Chem. Inf. Model.* 2012, 52, 1–6

Docking onto A_{2A}AR antagonist bound (3EML) was performed without and with eight crystallographic waters. From the results in Table 1, it is seen that the presence of water molecules clearly improved enrichments for all decoy sets considered, in agreement with an earlier work.[37] The impact of using a tailored decoy set is consistent in both cases.

Two distinct possible biases were highlighted in the decoy generation process:[22] the incomplete representation of the ligand chemical space and the incomplete representation of the repository chemical space (from which decoy molecules are drawn). Regarding the ligand chemical space, ~55% of the targets listed on Table S1 (Supporting Information) have more than 50 ligands. Although this does not guarantee per se adequate chemical diversity, it may ensure that in the general case, enough ligand chemical diversity is achieved. Regarding the molecule repository (ZINC), and being aware of its limitations,[22] it has to be acknowledged that we were able to generate decoy sets closely matching ligands' properties even in cases where a large number of ligands have nonzero charge, as happens with several GPCRs. Encouraging results were also obtained by generating *in silico* the decoy candidates,[16] which has the advantage of allowing a wider coverage of the chemical space. The approach presented here, however, does not need molecule generator software, and a new set of decoys could be easily regenerated any time the molecule repository is updated.

It was recently shown that it is possible to discover new ligands which are structurally different but physically similar to existing ligands.[38,39] This highlights the importance of periodically updating the ligand library to reduce the chances of actual unknown ligands being selected as decoys.

METHODS

Small-Molecule Libraries Preparation. Ligands and molecules from the ChemBridge MW-Set (ChemBridge, San Diego, CA) were prepared using LigPrep,[18] generating the appropriate protonation state at pH 7 and the most probable tautomer. Chiral ligands were prepared in their stereochemical form if available, either from 3D or 2D mol files. For decoy molecules, both chiralities (R and S) were generated at each chiral center. For ligands and decoy molecules, MACCS fingerprints and seven physical properties, MW, formal charge, RBS, HBAs, HBDs, log *P*, and tPSA were calculated using OpenBabel[40] (http://openbabel.com, version 2.3.0., accessed Feb. 2011).

Structure Selection and Docking Setup. GPCR crystal structures of noncovalently bound small molecules were used in the docking benchmark (cf. Table 1). Structures with mutations were not considered. Of the three available structures of β_2AR bound to carazolol (PDB codes 2RH1, 2R4R, and 2R4S), and of the three structures of β_1AR bound to cyanopindolol (PDB codes 2VT4, 2YCX, and 2YCY), the best resolution structure in each case was selected (2RH1 and 2VT4, respectively). Although 2YCX exhibited structural differences in helices 1 and 6, those were far from the binding site. For human β_1AR, the turkey structures were used, which differ just in one residue within the binding site (L101 (turkey numbering) vs I118 (human)).[31] The CXCR4 chemokine receptor was not included in the set, since fewer than five ligands could be retrieved from GLIDA/PubChem; thus no CXCR4 associated subset in GLL is present. For H_1R, both doxepin isomers present in the crystal structure (PDB code 3RZE) were included in the corresponding GLL subset. The cocrystallized ligand in the A_{2A}R agonist bound structure (PDB code 3QAK) has 20 rotatable bonds and thus was excluded from GLL; no associated decoys were generated for this ligand. Water molecules were removed from all structures except in 3EML, in which docking was performed without and with eight crystallographic water molecules in the neighborhood of the ligand. Although water molecules were present in the vicinity of the ligand in 2YDO and 2YDV (A_{2A}AR agonist bound), to ensure consistency with 3QAK, no waters were included in the dockings. Whenever present in the crystal structure, residues from the antibodies used for crystallization were removed. The GPCR structures were prepared using the Protein Preparation workflow in Maestro.[41] For the structures with missing loops or side chains, Prime[42] was used to fill them in. The correct protonation and tautomeric state of the cocrystallized ligands were selected manually. The hydrogen bonding network was optimized automatically, and the entire structure was energy minimized.

Docking Protocol. Docking was perfomed using Glide and the SP scoring function. The docking grid was generated using default settings, so the location and size of the docking "box" were based on the centroid and size of the bound ligand. Enrichments at x% were calculated as $EF_x = (Hits_x/N_x)/(Hits_{total}/N_{total})$, where $Hits_x$ is the number of ligands in the top x% of the screened database, N_x is the size of x% of the total database, and $Hits_{total}$ and N_{total} are the corresponding values referring to the complete database.

ASSOCIATED CONTENT

Supporting Information. Complete list of human GPCR targets and the number of their associated agonist and antagonists in the GPCR Ligand Library (GLL), enrichment and ROC curves for docking to 19 GPCR targets, and physical property distributions for ligand and decoy sets. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author
*E-mail: cnc@Cavasotto-lab.net.

ACKNOWLEDGMENT

The authors are grateful to Drs. John Irwin and Brian Shoichet for allowing parts of the ZINC library to be posted online.

REFERENCES

(1) Lagerström, M. C.; Schiöth, H. B. Structural diversity of G protein-coupled receptors and significance for drug discovery. *Nat. Rev. Drug Discovery* **2008**, *7*, 339–357.

(2) Hopkins, A. L.; Groom, C. R. The druggable genome. *Nat. Rev. Drug Discovery* **2002**, *1*, 727–30.

(3) Congreve, M.; Langmead, C. J.; Mason, J. S.; Marshall, F. H. Progress in structure based drug design for G protein-coupled receptors. *J. Med. Chem.* **2011**, *54*, 4283–4311.

(4) Tautermann, C. S.; Pautsch, A. The implication of the first agonist bound activated GPCR X-ray structure on GPCR in silico modeling. *ACS Med. Chem. Lett.* **2011**, *2*, 414–418.

(5) Cavasotto, C. N.; Orry, A. J.; Murgolo, N. J.; Czarniecki, M. F.; Kocsi, S. A.; Hawes, B. E.; O'Neill, K. A.; Hine, H.; Burton, M. S.; Voigt, J. H.; Abagyan, R. A.; Bayne, M. L.; Monsma, F. J., Jr. Discovery of novel chemotypes to a G-protein-coupled receptor through ligand-steered

homology modeling and structure-based virtual screening. *J. Med. Chem.* **2008**, *51*, 581–588.
(6) Michino, M.; Abola, E.; Brooks, C. L., 3rd; Dixon, J. S.; Moult, J.; Stevens, R. C. Community-wide assessment of GPCR structure modelling and ligand docking: GPCR Dock 2008. *Nat. Rev. Drug Discovery* **2009**, *8*, 455–463.
(7) Moro, S.; Deflorian, F.; Bacilieri, M.; Spalluto, G. Ligand-based homology modeling as attractive tool to inspect GPCR structural plasticity. *Curr. Pharm. Des.* **2006**, *12*, 2175–2185.
(8) Phatak, S. S.; Gatica, E. A.; Cavasotto, C. N. Ligand-steered modeling and docking: A benchmarking study in Class A G-Protein-Coupled Receptors. *J. Chem. Inf. Model.* **2010**, *50*, 2119–2128.
(9) Vilar, S.; Ferino, G.; Phatak, S. S.; Berk, B.; Cavasotto, C. N.; Costanzi, S. Docking-based virtual screening for GPCRs ligands: not only crystal structures but also in silico models. *J. Mol. Graphics Modell.* **2011**, *29*, 614–623.
(10) Cavasotto, C. N. Homology models in docking and high-throughput docking. *Curr. Top. Med. Chem.* **2011**, *11*, 1528–1534.
(11) Cavasotto, C. N.; Phatak, S. S. Homology modeling in drug discovery: current trends and applications. *Drug Discovery Today* **2009**, *14*, 676–683.
(12) Villoutreix, B. O.; Eudes, R.; Miteva, M. A. Structure-based virtual ligand screening: recent success stories. *Comb. Chem. High Throughput Screening* **2009**, *12*, 1000–1016.
(13) Huang, N.; Shoichet, B. K.; Irwin, J. J. Benchmarking sets for molecular docking. *J. Med. Chem.* **2006**, *49*, 6789–801.
(14) Verdonk, M. L.; Berdini, V.; Hartshorn, M. J.; Mooij, W. T.; Murray, C. W.; Taylor, R. D.; Watson, P. Virtual screening using protein-ligand docking: avoiding artificial enrichment. *J. Chem. Inf. Comput. Sci.* **2004**, *44*, 793–806.
(15) Bissantz, C.; Folkers, G.; Rognan, D. Protein-based virtual screening of chemical databases. 1. Evaluation of different docking/scoring combinations. *J. Med. Chem.* **2000**, *43*, 4759–4767.
(16) Wallach, I.; Lilien, R. Virtual decoy sets for molecular docking benchmarks. *J. Chem. Inf. Model.* **2011**, *51*, 196–202.
(17) Okuno, Y.; Tamon, A.; Yabuuchi, H.; Niijima, S.; Minowa, Y.; Tonomura, K.; Kunimoto, R.; Feng, C. GLIDA: GPCR--ligand database for chemical genomics drug discovery--database and tools update. *Nucleic Acids Res.* **2008**, *36*, D907–D912.
(18) *LigPrep*, version 2.3; Schrödinger, LLC: New York, 2009.
(19) Lipinski, C. A.; Lombardo, F.; Dominy, B. W.; Feeney, P. J. Experimental and Computational Approaches to Estimate Solubility and Permeability in Drug Discovery and Development Settings. *Adv. Drug Delivery Rev.* **2001**, *46*, 3–26.
(20) Vilar, S.; Karpiak, J.; Costanzi, S. Ligand and structure-based models for the prediction of ligand-receptor affinities and virtual screenings: Development and application to the beta(2)-adrenergic receptor *J. Comput. Chem.* **2010**, *31*, 707–720.
(21) Irwin, J. J.; Shoichet, B. K. ZINC--a free database of commercially available compounds for virtual screening. *J. Chem. Inf. Model.* **2005**, *45*, 177–182.
(22) Irwin, J. J. Community benchmarks for virtual screening. *J. Comput.-Aided Mol. Des.* **2008**, *22*, 193–199.
(23) Mysinger, M. M.; Shoichet, B. K. Rapid context-dependent ligand desolvation in molecular docking. *J. Chem. Inf. Model.* **2010**, *50*, 1561–1573.
(24) Friesner, R. A.; Banks, J. L.; Murphy, R. B.; Halgren, T. A.; Klicic, J. J.; Mainz, D. T.; Repasky, M. P.; Knoll, E. H.; Shelley, M.; Perry, J. K.; Shaw, D. E.; Francis, P.; Shenkin, P. S. Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. *J. Med. Chem.* **2004**, *47*, 1739–1749.
(25) *Glide*, version 5.5; Schrödinger, LLC: New York, 2009.
(26) Cherezov, V.; Rosenbaum, D. M.; Hanson, M. A.; Rasmussen, S. G.; Thian, F. S.; Kobilka, T. S.; Choi, H. J.; Kuhn, P.; Weis, W. I.; Kobilka, B. K.; Stevens, R. C. High-resolution crystal structure of an engineered human beta2-adrenergic G protein-coupled receptor. *Science* **2007**, *318*, 1258–1265.
(27) Wacker, D.; Fenalti, G.; Brown, M. A.; Katritch, V.; Abagyan, R.; Cherezov, V.; Stevens, R. C. Conserved binding mode of human beta2 adrenergic receptor inverse agonists and antagonist revealed by X-ray crystallography. *J. Am. Chem. Soc.* **2010**, *132*, 11443–11445.
(28) Rasmussen, S. G.; Choi, H. J.; Fung, J. J.; Pardon, E.; Casarosa, P.; Chae, P. S.; Devree, B. T.; Rosenbaum, D. M.; Thian, F. S.; Kobilka, T. S.; Schnapp, A.; Konetzki, I.; Sunahara, R. K.; Gellman, S. H.; Pautsch, A.; Steyaert, J.; Weis, W. I.; Kobilka, B. K. Structure of a nanobody-stabilized active state of the beta(2) adrenoceptor. *Nature* **2011**, *469*, 175–180.
(29) Warne, T.; Serrano-Vega, M. J.; Baker, J. G.; Moukhametzianov, R.; Edwards, P. C.; Henderson, R.; Leslie, A. G.; Tate, C. G.; Schertler, G. F. Structure of a beta1-adrenergic G-protein-coupled receptor. *Nature* **2008**, *454*, 486–491.
(30) Moukhametzianov, R.; Warne, T.; Edwards, P. C.; Serrano-Vega, M. J.; Leslie, A. G.; Tate, C. G.; Schertler, G. F. Two distinct conformations of helix 6 observed in antagonist-bound structures of a beta1-adrenergic receptor. *Proc. Natl. Acad. Sci. U.S.A.* **2011**, *108*, 8228–8232.
(31) Warne, T.; Moukhametzianov, R.; Baker, J. G.; Nehme, R.; Edwards, P. C.; Leslie, A. G.; Schertler, G. F.; Tate, C. G. The structural basis for agonist and partial agonist action on a beta(1)-adrenergic receptor. *Nature* **2011**, *469*, 241–244.
(32) Chien, E. Y.; Liu, W.; Zhao, Q.; Katritch, V.; Han, G. W.; Hanson, M. A.; Shi, L.; Newman, A. H.; Javitch, J. A.; Cherezov, V.; Stevens, R. C. Structure of the human dopamine D3 receptor in complex with a D2/D3 selective antagonist. *Science* **2011**, *330*, 1091–1095.
(33) Jaakola, V. P.; Griffith, M. T.; Hanson, M. A.; Cherezov, V.; Chien, E. Y.; Lane, J. R.; Ijzerman, A. P.; Stevens, R. C. The 2.6 angstrom crystal structure of a human A2A adenosine receptor bound to an antagonist. *Science* **2008**, *322*, 1211–1217.
(34) Xu, F.; Wu, H.; Katritch, V.; Han, G. W.; Jacobson, K. A.; Gao, Z. G.; Cherezov, V.; Stevens, R. C. Structure of an agonist-bound human A2A adenosine receptor. *Science* **2011**, *332*, 322–7.
(35) Lebon, G.; Warne, T.; Edwards, P. C.; Bennett, K.; Langmead, C. J.; Leslie, A. G.; Tate, C. G. Agonist-bound adenosine A2A receptor structures reveal common features of GPCR activation. *Nature* **2011**, *474*, 521–525.
(36) Shimamura, T.; Shiroishi, M.; Weyand, S.; Tsujimoto, H.; Winter, G.; Katritch, V.; Abagyan, R.; Cherezov, V.; Liu, W.; Han, G. W.; Kobayashi, T.; Stevens, R. C.; Iwata, S. Structure of the human histamine H1 receptor complex with doxepin. *Nature* **2011**, *475*, 65–70.
(37) McRobb, F. M.; Capuano, B.; Crosby, I. T.; Chalmers, D. K.; Yuriev, E. Homology modeling and docking evaluation of aminergic G protein-coupled receptors. *J. Chem. Inf. Model.* **2010**, *50*, 626–637.
(38) Carlsson, J.; Coleman, R. G.; Setola, V.; Irwin, J. J.; Fan, H.; Schlessinger, A.; Sali, A.; Roth, B. L.; Shoichet, B. K. Ligand discovery from a dopamine D(3) receptor homology model and crystal structure. *Nat. Chem. Biol.* **2011**, *7*, 769–778.
(39) Kolb, P.; Rosenbaum, D. M.; Irwin, J. J.; Fung, J. J.; Kobilka, B. K.; Shoichet, B. K. Structure-based discovery of beta2-adrenergic receptor ligands. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106*, 6843–6848.
(40) O'Boyle, N.; Banck, M.; James, C.; Morley, C.; Vandermeersch, T.; Hutchison, G. Open Babel: An open chemical toolbox. *J. Cheminf.* **2011**, *3*, 33.
(41) *Maestro*, version 9.0; Schrödinger, LLC: New York, 2009.
(42) *Prime*, version 2.2; Schrödinger, LLC: New York, 2009.