Published in final edited form as:

J Mol Graph Model. 2011 February ; 29(5): 614–623. doi:10.1016/j.jmglm.2010.11.005.

Docking-based virtual screening for ligands of G protein-coupled receptors: not only crystal structures but also *in silico* models

Santiago Vilar¹, Giulio Ferino¹, Sharangdhar S. Phatak², Barkin Berk¹, Claudio N. Cavasotto², and Stefano Costanzi^{1,*}

¹Laboratory of Biological Modeling, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, DHHS, Bethesda, MD 20892, USA.

²School of Health Information Sciences, The University of Texas Health Science Center at Houston, Houston, TX 77030, USA.

Abstract

G protein-coupled receptors (GPCRs) regulate a wide range of physiological functions and hold great pharmaceutical interest. Using the β_2 -Adrenergic receptor as a case study, this article explores the applicability of docking-based virtual screening to the discovery of GPCR ligands and defines methods intended to improve the screening performance. Our controlled computational experiments were performed on a compound dataset containing known agonists and blockers of the receptor as well as a large number of decoys. The screening based on the structure of the receptor crystallized in complex with its inverse agonist carazolol yielded excellent results, with a clearly delineated prioritization of ligands over decoys. Blockers generally were preferred over agonists; however agonists were also well distinguished from decoys. A method was devised to increase the screening yields by generating an ensemble of alternative conformations of the receptor that accounts for its flexibility. Moreover, a method was devised to improve the retrieval of agonists, based on the optimization of the receptor around a known agonist. Finally, the applicability of docking-based virtual screening also to homology models endowed with different levels of accuracy was proved. This last point is of uttermost importance, since crystal structures are available only for a limited number of GPCRs, and extends our conclusions to the entire superfamily. The outcome of this analysis definitely supports the application of computer-aided techniques to the discovery of novel GPCR ligands, especially in light of the fact that, in the near future, experimental structures are expected to be solved and become available for an ever increasing number of GPCRs.

1. Introduction

G protein-coupled receptors (GPCRs) constitute the largest superfamily of human membrane signaling proteins. They are implicated in a vast array of physiological functions

© 2010 Published by Elsevier Inc.

* address for correspondence: stefanoc@mail.nih.gov.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Supplementary data: Supplementary data associated with this article can be found in the online version.

and pathological conditions and, thus, are highly pursued as targets for pharmacological intervention [1,2].

Due difficulties inherent the crystallization of GPCRs, for years bovine rhodopsin has been the only member of the superfamily with an experimentally elucidated three-dimensional structure, and has been employed as a template for the construction of three-dimensional (3D) homology models. More recently, however, scientific breakthroughs yielded to the solution of the crystal structures of a few additional receptors, including the β_2 -Adrenergic receptor (β_2 -AR), while structures of additional receptors are expected to be solved in the near future [2,3].

In this work, using the β_2 -AR as a case study, we investigated the applicability of crystal structures and homology models endowed with different levels of accuracy to the identification of GPCR ligands, and defined methods intended to improve screening performance. We added a pool of known binders of the receptor to a large set of decoy compounds, and subsequently analyzed the ability of a series of controlled docking-based virtual screening experiments to prioritize ligands over decoys.

In particular, we: a) assessed the excellent results attainable with the crystal structure of the receptor in complex with the inverse agonist carazolol [4,5]; b) devised a method that managed to further improve the results through the generation of an ensemble of alternative conformations of the receptor that accounts for its flexibility – general importance: very rarely multiple crystal structures of a GPCR are available and can be used to account for its flexibility; c) defined a method to invert the tendency of virtual screening to prioritize blockers over agonists, by optimizing the structure of the receptor around bound agonists – general importance: to date, no GPCR has been crystallized in complex with an agonist; d) assessed the feasibility of the use of homology models in lieu of experimental structures, in the absence of the latter – general importance: for the vast majority of GPCRs crystal structures are not available. For this last point, we employed three rhodopsin-based β_2 -AR homology models, endowed with different levels of structural accuracy, that we have recently published [6]. A flowchart illustrating the different points of the study is shown in Figure 1.

2. Materials and methods

2.1. Compound dataset

The compound dataset used for the virtual screening experiments was composed of a set of 60 β_2 -AR ligands (29 agonists and 31 blockers) with pK_i values above 5, which we had collected from the literature in a previous work [7], and 55806 decoys, presumably inactive at the β_2 -AR, extracted from the ZINC Database [8,9]. After downloading the subset of lead-like compounds (1830871 molecules), a diverse pool of it was selected with the help of the QuaSAR-Cluster module of MOE [10]. In particular, the clustering was performed through the following operations: a) all compounds were sorted in alphabetical order, according to their names; b) 184 2D descriptors were calculated with the QuaSAR-Descriptor function; c) a Principal Components Analysis (PCA) was performed to reduce the number of variables to five components; d) finally, the database was clustered into 55806 groups with the QuaSAR-Cluster function and the first listed compound of each cluster was selected and added to the diverse subset of decoys.

Ligands and decoys were then subjected to an automatic preparation process, performed with the LigPrep tool of the Schrödinger package [11], generating all protonation and tautomeric states available within a pH range of 7.0 ± 2.0 .

2.2. Protein preparation

Crystal structure (2rh1) [4,5] and the *in silico* models of the β_2 -AR were subjected to the Protein Preparation Wizard workflow implemented in the Schrödinger package [11]. This added hydrogens, which were subsequently minimized using the OPLS_2005 force field and the Impact molecular mechanics engine, while heavy atoms were constrained. Furthermore, it optimized the protonation state of His residues and the orientation of hydroxyl groups, Asn residues, and Gln residues. More details on the protein preparation protocol can be found in a previously published work [6].

2.3. Molecular Docking

Molecular docking experiments were carried out by means of the Glide [12], as implemented in the Schrödinger package, considering the ligands as flexible but treating the receptor as a rigid structure. A cubing docking grid was centered on Val114, and was given a dimension sufficient to accommodate compounds with a length $\leq 15\text{\AA}$. The ligand-midpoint box was given a side of $10\text{\AA}$. No scaling factors were applied to the van der Waals (vdW) radii of the receptor atoms, while a scaling factor of 0.8 was applied to the non polar atoms of the ligands, defined as those with a partial charge lower than 0.15e. The HTVS (High Throughput Virtual Screening) and SP (Standard Precision) scoring functions of Glide were used, granting full flexibility to the ligands. A post-docking minimization, in which only the ligands were flexible, was performed on the output complexes in order to reduce the initially collected 25 poses per ligand to 5. A rescoring of the top ranking SP pose of each compound was then performed with the XP scoring function of Glide and with the London dG scoring function as implemented in the MOE [10]. Rescoring was always performed on the receptor-ligand complexes just as resulting from the SP calculation, without minimizations or relaxations.

2.4. Generation of alternative binding site conformations for receptor ensemble docking

The crystal structure of β_2 -AR (2rh1) was optimized in the presence of three different ligands: carazolol (the co-crystallized ligand), carvedilol and ritodrine, through a stochastic global energy minimization in the torsional coordinate space, in a similar manner to what has been already described in detail [13-15]. The system was represented through the ECEPP/3 force field, within the ICM platform [16]. The three positional and the three orientational coordinates of the ligand, together with the torsional coordinates of ligand and side-chains within $6\text{\AA}$ radius were considered free. The initial conformations of carvedilol and ritodrine chosen were those that showed maximum overlap with the co-crystallized carazolol. A quadratic restraint was imposed between the ligand charged amine and Asp113. No backbone relaxation was performed during the optimization process.

2.5. Induced-fit docking procedure. Protein conformation adapted to the agonist isoproterenol

A conformation of the receptor optimized in complex with the agonist isoproterenol was obtained with the Induced Fit docking procedure, as implemented in the Schrödinger package [11,12,17]. Specifically, the Induced Fit procedure involved three consecutive steps: a) a Glide-based docking of isoproterenol at the receptor, starting from the ligand pose extracted from the SP docking run (see section 2.3), with the SP scoring function and a vdW scaling factor of 0.5 applied to non polar atoms of receptor ligands, defined as those with partial charge lower than 0.25e for the receptor and lower than 0.15e for the ligands; b) a Prime-based optimization of the ligand-binding pocket for all the complexes obtained in step (a), granting flexibility to the ligand and all residues within an $8\text{\AA}$ radius from the ligand; c) a Glide-based docking of isoproterenol at each of the optimized ligand binding pockets obtained in step (b), with a vdW scaling factor of 0.8 for the non polar atoms of the ligand

only (only the complexes within 30 kcal/mol from the best were subjected to this final docking run. For steps (a) and (c) the docking grid was centered on Val114, and was given a size of 26Å.

3. Results

As explained in the Material and Methods section, a virtual compound dataset composed of 60 β_2 -AR ligands (29 agonists and 31 blockers) and 55,806 decoys extracted from the ZINC database, was docked to the receptor considering the flexibility of the ligands, but treating the protein as rigid [11,12]. The docking procedures and scoring functions that we applied are endowed with different computational time requirements, expected to be inversely correlated with their accuracy. The fastest and coarsest ones are intended to be used in the initial stages of virtual screening campaigns, in which datasets of millions of compounds are evaluated; in contrast, the slowest and most accurate ones are intended to be used in the subsequent stages of the campaigns, in which the datasets have been already considerably reduced in size. In our case, the average total CPU time required for the processing of 55866 compounds, was 13 hours for the HTVS docking, 94 hours for the SP docking, and 18 hours for the XP rescoring and 3 hours for the London dG rescoring – these performances were obtained using a single processor of a Quad Core Xeon X5482 system (3.20GHz, 2X 6MB L2Cache, 1600MHz). Notably, the actual time needed to complete the calculations can be considerably reduced using more than one processor in parallel.

The results were analyzed through two-dimensional Cartesian graphs in which the percentage of the screened dataset (% SD) and the yield – defined as the percentage of known ligands retrieved – were plotted on the axes X and Y, respectively (see Figures 2, 5, 7, and 8), and through the calculation of enrichment factors (see Figures 3 and 9), defined according to the following formula:

$$EF \quad (\%SD) = \frac{\frac{\text{number of active compounds up to a given \% SD}}{\text{number of compounds up to a given \% SD}}}{\frac{\text{total number of active compounds in screened database}}{\text{total number of compounds in screened database}}}$$

3.1. Virtual screening based on the crystal structure of the β_2 -AR

The results obtained when basing the screening on the β_2 -AR crystal structure were remarkable (Figures 2 and 3 and Table 1). Even the fast and coarse HTVS docking mode produced outstanding results, with a clear prioritization of ligands over decoys. As expected, the best results were obtained rescoring the SP results with the more accurate XP and London dG scoring functions, which both showed very high yields in ligand retrieval already in the top ranking portions of the dataset. For example, with the London dG scoring function, the yield reached 50% within the top 0.21% of the dataset, while, within the top 2.5% of the screened database, it reached 91.67% (see Table 1). Similar results were obtained with the XP scoring function. Notably, with the same two scoring functions, we achieved an EF above 300 within the top 10 compounds (i.e. the top 0.02 % of the dataset). The EF then gradually declines when considering increasingly large portions of the screened dataset (see Figure 3).

3.2. Improving virtual screening performance through receptor ensemble docking

Crystal structures offer only a static representation of the receptors, which, in fact, are intrinsically flexible entities. On the basis of our previous experience [18-20], we reasoned that generating alternative conformations of the ligand binding pocket, hence accounting for its flexibility, could improve virtual screening. Thus, through stochastic energy

minimizations of the side chains, alternative conformations of the β_2 -AR in complex with three representative ligands were generated, starting from the docking poses obtained as described in the previous section (Figure 4). To thoroughly explore different scenarios, Carazolol, which is the co-crystallized ligand, and two additional ligands were selected. Namely, these were ritodrine, which was well prioritized over the decoys when using the crystal structure, and carvedilol, which, instead, was not.

Following the receptor ensemble docking (RED) approach [18,20], the compound dataset was independently docked at each of the three alternative conformations. The results from these three docking runs were then combined with those obtained with the crystal structure and, for each ligand, the pose that obtained the maximum score was selected, thus giving to each docked ligand the possibility to choose its preferred receptor conformation.

The data clearly indicated that combining the results obtained with the three models and the crystal structures were always significantly improved over those obtained using the crystal structure alone (Figure 5). For all the scoring functions, but especially for HTVS, there was a noticeable improvement since the very early stages of the screening: within the top 0.5% of the screened database, the yield reached 15.0%, 26.67%, 55.0% and 73.33% for HTVS, SP, XP and London dG, when using crystal structure, while, with the same scoring functions, it reached 30.0%, 33.33%, 56.67% and 76.67%, when combining multiple structures. Similar conclusions remained still valid also when only one of the alternative conformations was used in addition to the crystal structure, as shown in Figure S1, S2, and S3 of the supplementary data for carazolol, carvedilol, and ritodrine, respectively. The structure that worked the best, both alone and in combination with the crystal structure, was the one optimized around the docked carazolol. On the other hand, the structure that, when taken alone, yielded the poorest results with all the scoring functions, with the exception of London dG, was the one optimized around the docked carvedilol, a ligand that, as we mentioned, received a poor score in the initial SP docking.

3.3. Prioritizing agonists or blockers

The β_2 -AR has been crystallized in complex with carazolol, an inverse agonist. Thus, it was not surprising to note that, although the virtual screening experiments based on the crystal structure effectively differentiate both agonists and blockers from non-binders, they generally prioritize blockers over antagonists. This tendency can be seen in Figure 2, where red and green dots represent blockers and agonists, respectively, and confirms what we reported in a recent publication [7].

Here, this tendency was successfully reversed, prioritizing agonists over blockers, by generating an alternative protein conformation optimized with the full agonist isoproterenol (Figure 6). Unlike for the generation of the alternative conformations described in section II, for this purpose a more drastic approach was employed, based on the Schrödinger's induced fit protocol, which, for the residues around the ligand, samples the degrees of freedom of the side chains in the dihedral space and grants flexibility to the backbone [21]. The analysis of the virtual screening conducted using this alternative structure, shown in Figure 7 and Table 2, clearly showed that with the HTVS, SP, and XP scoring functions agonists effectively migrated towards the top scoring positions of the database. As an example, using the XP scoring function, 11 agonists and 1 blocker were retrieved within the 50 top scoring compounds with the isoproterenol-adjusted structure, as opposed to the 4 agonists and 13 blockers retrieved with the plain crystal structure. Also the London dG scoring function tended to prioritize agonists over blockers with the isoproterenol-adjusted structure. However, the global retrieval yields fell significantly when using this structure, thus resulting in a lower absolute number of agonists found within the top scoring compounds, contrary to what detected with the other functions (see Table 2).

3.4. Virtual screening based on homology models of the β_2 -AR

Crystal structures exist only for very limited numbers of GPCRs [2,3]. However, these can be used as templates for the construction of homology models of the remaining members of the superfamily. Thus, it is of extreme importance to establish the applicability of such homology models to docking-based virtual screenings. For this purpose, here, the same docking-based virtual screening was also applied using three homology models recently published by Costanzi [6], dubbed model 1, model 2 and model 2-F290g+. From the structural point of view, the three models, when compared with the crystal structure showed different levels of accuracy: a) model 1 featured a second extracellular loop (EL2) built by homology to rhodopsin, buried into the opening of the interhelical cavity in a manner unnatural for the β_2 -AR, and showed six out of the twenty residues in the binding pocket in the wrong rotameric state; b) model 2 featured an EL2 built *de novo*, with a more native-like conformation, at least in the segment that lines the binding pocket, and only one residue, namely Phe290, in the wrong rotameric state; c) model 2-F290g+ was a modified version of model 2, featuring, as the only difference with the latter, residue Phe290 in the gauche+ conformation, to match the crystal structure.

As evident from Figure 8, the virtual screening experiments conducted at the three models resulted in a significant prioritization of ligands versus decoys, especially when using more sophisticated scoring functions. Not surprisingly, the homology models performed accordingly to their level of accuracy. With the London dG rescoring of the SP docking, 50% of the true ligands were recovered within the top 0.72%, 2.25%, and 9.55% of the screened database, with model 1, model 2, and model 2-F290g+, respectively. Moreover, in the top 2.5% of the screened database, the yield was 65.0%, 53.33%, and 30.0% with model 1, model 2, and model 2-F290g+, respectively (see Table 3). The same trend was evident from the analysis of the EF (see Figure 9).

4. Discussion

As mentioned, rhodopsin has been for years the only receptor with an experimentally elucidated 3D structure. However, recent breakthroughs led to the solution of the crystal structure of a few additional receptors, including the β_2 -AR, while more structures are expected to be solved in the near future. In this context, our analysis of the applicability of the crystal structure of a GPCR to virtual screening is indeed very timely. The positive results that we obtained in this case study encourage the application of computer-aided techniques to the rational identification of GPCR ligands, and are very much in line with the outcome of actual virtual screening campaigns based on the crystal structures of the β_2 adrenergic and the adenosine A_{2A} receptors [22-24]. Notably, our virtual screening experiments did not necessitate the consideration of explicit water molecules in the binding pocket. However, for other systems, the inclusion of one or more water molecules crucial for the mediation of receptor ligand-interactions may be essential [25].

Crystal structures, however, are only static representations of proteins that, instead, are inherently flexible. Neglecting protein flexibility by using one rigid representation of the target has been shown to adversely affect virtual screening performance [26-28]. Receptor ensemble docking, a method where docking is performed on several distinct structures and the results are subsequently combined, has been successfully used to account for protein flexibility in virtual screening campaigns for ligands of protein kinases [18,20,29,30], GPCRs [19], and other protein targets [28,31]. Here, we implemented receptor ensemble docking to mimic the flexibility of the β_2 -AR by generating and using more than one binding site conformation, and observed significantly improved virtual screening yields, in agreement with what was found for other receptors.

Moreover, all the known crystal structures of GPCRs, including that of the β_2 -AR, have been obtained in complex with antagonists or inverse agonists, thus raising possible concerns on their applicability to the identification of agonists. According to mutagenesis and crystallographic data, blockers and agonists of the β_2 -AR establish an ionic interaction Asp113 in TM3, through their protonated amino group, and a hydrogen-bond with Asn312 in TM7, through the β -hydroxyl group [32,33]. As shown by the crystal structure, the inverse agonist Carazolol also establishes a direct interaction with Ser203 in TM5. The catechol hydroxyl groups of the agonists are also supposed to interact with this residue as well as two other serines located in TM5, namely Ser204 and Ser207 [34]. However, such interactions cannot be detected when docking of agonists to a rigid representation of the receptor crystallized in its ground state. For these interactions to occur, a conformational change in the protein is necessary. Modeling studies suggested that agonists triggered a counterclockwise rotation of TM5 [35,36], leading to the formation of hydrogen bonds with the three above-mentioned Ser residues. In agreement with the need for a conformational change, our virtual screening based on the crystal structure generally prioritized blockers over agonists, although being clearly able to distinguish the latter from the decoys (Figure 2). In this study, we defined a method that successfully inverted this tendency, by utilizing a model optimized around an agonist (see Figure 5). Notably, De Graaf and Rognan have shown that a crystal structure modified to reflect the early conformational events in β_2 -AR activation led to the selective retrieval of full and partial agonists, in a screen that combined docking and molecular interaction fingerprints [37]. Similarly, Abagyan *et al.* have improved the retrieval of agonists by changing the conformation of TM5, a domain which is thought to be related with the activation of the receptor [38]. However, the procedure utilized here for the generation of the agonist optimized receptor does not necessitate prior knowledge of the activation process. It just requires docking an agonist at the crystal structure and subjecting the resulting complex to automatic optimization according to the InducedFit protocol, as implemented in the Schrodinger package [11,12,17]. Confirming these theoretical findings, a recent virtual screening campaign for ligands of the free fatty receptor 1 (FFA1), based on a model optimized in complex with a potent agonist [39-41], prioritized the retrieval of activators over blockers [42]. It is worth noting that the protocol that we used for the generation of the isoproterenol-adapted conformation of the β_2 -AR accounts only for local changes within the ligand-binding pocket, and is not intended to study the activation process. In fact, GPCR activation is a complex phenomenon that may imply relatively large conformational changes as well alterations of dimerization interfaces [32,33,43]. We have been able to follow the initial phases of the activation process by means of a modified Induced Fit protocol that granted flexibility to entire domains of the receptor (data still unpublished). Alternatively, computational hypotheses on the activation process have been inferred through molecular dynamics and coarse-grain simulations [44-47].

Lastly, we demonstrated that docking-based virtual screening techniques are not only applicable to crystal structures but also to homology models. This point is of extreme importance because it extends their area of applicability to the vast majority of receptors, for which crystal structures are not available. Not surprisingly, the three homology models that we tested performed according to their level of accuracy. Notably, the most accurate model, although obviously outperformed by the crystal structure, yielded excellent results. The conformation of the second extracellular loop, a domain heavily involved in ligand binding and receptor activation [46,48], affected substantially the accuracy of the binding site, and thus the virtual screening, which showed significantly improved performances when the loop was built *de novo*. Thus, accurately modeling the extracellular loops appears as one of the most crucial aspects to ensure the applicability of virtual screening to GPCR models – the use of mixed computational and experimental techniques in which the helical bundle is modeled *in silico* while the extracellular regions are modeled through NMR may offer an attractive solution to this challenging problem [49]. It is worth noting that with the HTVS

scoring function, which is the fastest but also the least sophisticated of those that we applied, model 1 and model 2 did not perform outstandingly, although yielding results by far better than a random selection (represented by the dotted diagonal in Figure 6). This observation suggests that, when conducting virtual screening campaigns based on GPCR homology models, it might be convenient to use directly a scoring function with a higher level of accuracy, rather than subjecting the database to a preliminary screen with the HTVS scoring function.

5. Conclusion

This study shows the applicability of the β_2 -AR crystal structure to docking-based virtual screening, defines a method to optimize the results by mimicking the flexibility of the protein through the use of generated alternative conformations of the binding pocket, defines a method to guide the screening towards the discovery of agonists, and provides evidence that also *in silico* models are powerful tools on which to base the virtual screenings. Our results definitely encourage the application of computer-aided techniques to the discovery of novel GPCR ligands. Moreover, the practical applicability of these computer-based drug discovery techniques promises to considerably expand and bloom in the near future, as experimental structures are expected to be produced and become available for an ever greater number of GPCRs.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This research was supported, in part, by the Intramural Research Program of the NIH, NIDDK.

References

1. Muller G. Towards 3D structures of G protein-coupled receptors: A multidisciplinary approach. *Curr. Med. Chem* 2000;7:861–888. [PubMed: 10911020]
2. Costanzi S, Siegel J, Tikhonova IG, Jacobson KA. Rhodopsin and the Others: A Historical Perspective on Structural Studies of G Protein-Coupled Receptors. *Curr. Pharm. Design* 2009;15:3994–4002.
3. Hanson MA, Stevens RC. Discovery of New GPCR Biology: One Receptor Structure at a Time. *Structure* 2009;17:8–14. [PubMed: 19141277]
4. Cherezov V, Rosenbaum DM, Hanson MA, Rasmussen SGF, Thian FS, Kobilka TS, Choi HJ, Kuhn P, Weis WI, Kobilka BK, Stevens RC. High-resolution crystal structure of an engineered human beta(2)-adrenergic G protein-coupled receptor. *Science* 2007;318:1258–1265. [PubMed: 17962520]
5. Rosenbaum DM, Cherezov V, Hanson MA, Rasmussen SGF, Thian FS, Kobilka TS, Choi HJ, Yao XJ, Weis WI, Stevens RC, Kobilka BK. GPCR engineering yields high-resolution structural insights into beta(2)-adrenergic receptor function. *Science* 2007;318:1266–1273. [PubMed: 17962519]
6. Costanzi S. On the applicability of GPCR homology models to computer-aided drug discovery: A comparison between *in silico* and crystal structures of the beta(2)-adrenergic receptor. *J. Med. Chem* 2008;51:2907–2914. [PubMed: 18442228]
7. Vilar S, Karpik 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. [PubMed: 19569204]
8. Irwin JJ, Shoichet BK. ZINC - A free database of commercially available compounds for virtual screening. *J. Chem. Inf. Model* 2005;45:177–182. [PubMed: 15667143]
9. <http://zinc.docking.org/>
10. MOE, version 2008.10. Chemical Computing Group, Inc.; 2008. www.chemcomp.com

11. Maestro, version 8.5. Schrödinger, LLC; New York, NY: 2008.
12. Glide, version 5.0. Schrödinger, LLC; New York, NY: 2008.
13. Cavasotto CN, Orry AJW, Murgolo NJ, Czarniecki MF, Kocsi SA, Hawes BE, O'Neill KA, Hine H, Burton MS, Voigt JH, Abagyan RA, Bayne ML, Monsma FJ. 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. [PubMed: 18198821]
14. Diaz P, Phatak SS, Xu J, Astruc-Diaz F, Cavasotto CN, Naguib M. 6-Methoxy-N-alkyl Isatin Acylhydrazone Derivatives as a Novel Series of Potent Selective Cannabinoid Receptor 2 Inverse Agonists: Design, Synthesis, and Binding Mode Prediction. *J. Med. Chem* 2009;52:433–444. [PubMed: 19115816]
15. Diaz P, Phatak SS, Xu JJ, Fronczek FR, Astruc-Diaz F, Thompson CM, Cavasotto CN, Naguib M. 2,3-Dihydro-1-Benzofuran Derivatives as a Series of Potent Selective Cannabinoid Receptor 2 Agonists: Design, Synthesis, and Binding Mode Prediction through Ligand-Steered Modeling. *Chemmedchem* 2009;4:1615–1629. [PubMed: 19637157]
16. ICM, Molsoft, LLC.; www.molsoft.com
17. Prime, version 2.0. Schrödinger, LLC; New York, NY: 2008.
18. Cavasotto CN, Abagyan RA. Protein flexibility in ligand docking and virtual screening to protein kinases. *J. Mol. Biol* 2004;337:209–225. [PubMed: 15001363]
19. Engel S, Skoumbourdis AP, Childress J, Neumann S, Deschamps JR, Thomas CJ, Colson AO, Costanzi S, Gershengorn MC. A virtual screen for diverse ligands: Discovery of selective G protein-coupled receptor antagonists. *J. Am. Chem. Soc* 2008;130:5115–5123. [PubMed: 18357984]
20. Cavasotto CN, Kovacs JA, Abagyan RA. Representing receptor flexibility in ligand docking through relevant normal modes. *J. Am. Chem. Soc* 2005;127:9632–9640. [PubMed: 15984891]
21. Sherman W, Day T, Jacobson MP, Friesner RA, Farid R. Novel procedure for modeling ligand/receptor induced fit effects. *J. Med. Chem* 2006;49:534–553. [PubMed: 16420040]
22. Kolb P, Rosenbaum DM, Irwin JJ, Fung JJ, Kobilka BK, Shoichet BK. Structure-based discovery of beta(2)-adrenergic receptor ligands. *Proc. Natl. Acad. Sci. USA* 2009;106:6843–6848. [PubMed: 19342484]
23. 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]
24. Carlsson J, Yoo L, Gao Z-G, Irwin JJ, Shoichet BK, Jacobson KA. Structure-based discovery of A_{2A} adenosine receptors ligands. *J. Med. Chem.* 2010 DOI: 10.1021/jm100240h.
25. Ivanov AA, Barak D, Jacobson KA. Evaluation of homology modeling of G-protein-coupled receptors in light of the A(2A) adenosine receptor crystallographic structure. *J. Med. Chem* 2009;52:3284–3292. [PubMed: 19402631]
26. Cavasotto CN, Singh N. Docking and high throughput docking: Successes and the challenge of protein flexibility. *Curr. Comput-Aided Drug Des* 2008;4:221–234.
27. Cozzini P, Kellogg GE, Spyraakis F, Abraham DJ, Costantino G, Emerson A, Fanelli F, Gohlke H, Kuhn LA, Morris GM, Orozco M, Pertinhez TA, Rizzi M, Sottriffer CA. Target Flexibility: An Emerging Consideration in Drug Discovery and Design. *J. Med. Chem* 2008;51:6237–6255. [PubMed: 18785728]
28. B-Rao C, Subramanian J, Sharma SD. Managing protein flexibility in docking and its applications. *Drug Discov. Today* 2009;14:394–400. [PubMed: 19185058]
29. Sperandio O, Mouawad L, Pinto E, Villoutreix BO, Perahia D, Miteva MA. How to choose relevant multiple receptor conformations for virtual screening: a test case of Cdk2 and normal mode analysis. *Eur. Biophys. J.* 2010 DOI 10.1007/s00249-010-0592-0.
30. Kovacs JA, Cavasotto CN, Abagyan R. Conformational sampling of protein flexibility in generalized coordinates: Application to ligand docking. *J. Comput. Theor. Nanosci* 2005;2:354–361.
31. Ferrari AM, Wei BQQ, Costantino L, Shoichet BK. Soft docking and multiple receptor conformations in virtual screening. *J. Med. Chem* 2004;47:5076–5084. [PubMed: 15456251]

32. Strader CD, Fong TM, Tota MR, Underwood D, Dixon RAF. Structure and Function of G-Protein-Coupled Receptors. *Annu. Rev. Biochem* 1994;63:101–132. [PubMed: 7979235]
33. Audet M, Bouvier M. Insights into signaling from the beta(2)-adrenergic receptor structure. *Nat. Chem. Biol* 2008;4:397–403. [PubMed: 18560432]
34. Liapakis G, Ballesteros JA, Papachristou S, Chan WC, Chen X, Javitch JA. The forgotten serine - A critical role for Ser-203(5.42) in ligand binding to and activation of the beta(2)-adrenergic receptor. *J. Biol. Chem* 2000;275:37779–37788. [PubMed: 10964911]
35. Freddolino PL, Kalani MYS, Vaidehi N, Floriano WB, Hall SE, Trabanino RJ, Kam VWT, Goddard WA. Predicted 3D structure for the human beta 2 adrenergic receptor and its binding site for agonists and antagonists. *Proc. Natl. Acad. Sci. USA* 2004;101:2736–2741. [PubMed: 14981238]
36. Bhattacharya S, Hall SE, Li H, Vaidehi N. Ligand-stabilized conformational states of human beta(2) adrenergic receptor: Insight into G-protein-coupled receptor activation. *Biophys. J* 2008;94:2027–2042. [PubMed: 18065472]
37. de Graaf C, Rognan D. Selective structure-based virtual screening for full and partial agonists of the beta 2 adrenergic receptor. *J. Med. Chem* 2008;51:4978–4985. [PubMed: 18680279]
38. Reynolds KA, Katritch V, Abagyan R. Identifying conformational changes of the beta(2) adrenoceptor that enable accurate prediction of ligand/receptor interactions and screening for GPCR modulators. *J. Comput. Aided Mol. Des* 2009;23:273–288. [PubMed: 19148767]
39. Tikhonova IG, Sum CS, Neumann S, Thomas CJ, Raaka BM, Costanzi S, Gershengorn MC. Bidirectional, iterative approach to the structural delineation of the functional “Chemoprint” in GPR40 for agonist recognition. *J. Med. Chem* 2007;50:2981–2989. [PubMed: 17552505]
40. Sum CS, Tikhonova IG, Neumann S, Engel S, Raaka BM, Costanzi S, Gershengorn MC. Identification of residues important for agonist recognition and activation in GPR40. *J. Biol. Chem* 2007;282:29248–29255. [PubMed: 17699519]
41. Costanzi S, Neumann S, Gershengorn MC. Seven transmembrane-spanning receptors for free fatty acids as therapeutic targets for diabetes mellitus: Pharmacological, phylogenetic, and drug discovery aspects. *J. Biol. Chem* 2008;283:16269–16273. [PubMed: 18385136]
42. Tikhonova IG, Sum CS, Neumann S, Engel S, Raaka BM, Costanzi S, Gershengorn MC. Discovery of novel Agonists and antagonists of the free fatty acid receptor 1 (FFAR1) using virtual screening. *J. Med. Chem* 2008;51:625–633. [PubMed: 18193825]
43. Guo W, Shi L, Filizola M, Weinstein H, Javitch JA. Crosstalk in G protein-coupled receptors: changes at the transmembrane homodimer interface determine activation. *Proc. Natl. Acad. Sci. U S A* 2005;102:17495–17500. [PubMed: 16301531]
44. Tikhonova IG, Best RB, Engel S, Gershengorn MC, Hummer G, Costanzi S. Atomistic insights into rhodopsin activation from a dynamic model. *J. Am. Chem. Soc* 2008;130:10141–10149. [PubMed: 18620390]
45. Deflorian F, Engel S, Colson AO, Raaka BM, Gershengorn MC, Costanzi S. Understanding the structural and functional differences between mouse thyrotropin-releasing hormone receptors 1 and 2. *Proteins* 2008;71:783–794. [PubMed: 17979196]
46. Costanzi S, Joshi BV, Maddileti S, Mamedova L, Gonzalez-Moa MJ, Marquez VE, Harden TK, Jacobson KA. Human P2Y(6) receptor: Molecular modeling leads to the rational design of a novel agonist based on a unique conformational preference. *J. Med. Chem* 2005;48:8108–8111. [PubMed: 16366591]
47. Bhattacharya S, Vaidehi N. Computational mapping of the conformational transitions in agonist selective pathways of a G-protein coupled receptor. *J. Am. Chem. Soc* 2010;132:5205–5214. [PubMed: 20235532]
48. Sum CS, Tikhonova IG, Costanzi S, Gershengorn MC. Two Arginine-Glutamate Ionic Locks Near the Extracellular Surface of FFAR1 Gate Receptor Activation. *J. Biol. Chem* 2009;284:3529–3536. [PubMed: 19068482]
49. Tikhonova IG, Costanzi S. Unraveling the structure and function of G protein-coupled receptors through NMR spectroscopy. *Curr. Pharm. Des* 2009;15:4003–4016. [PubMed: 20028318]

Figure 1.

Flowchart representation of the different procedures presented in this study. within various s of top scoring compounds in the screening based on the β_2 -AR crystal structure, applying the London dG and XP scoring functions one the poses selected with the SP docking.

Figure 2.

Percentage yield of ligand retrieval plotted versus percentage of screened compounds, sorted according to their docking rank, for the β_2 -AR crystal structure using the HTVS, SP, SP-XP and SP-London dG scoring functions. Red dots represent blockers and green dots represent agonists.

NIH-PA Author Manuscript

Figure 4.

Alternative conformations of the β_2 -AR in complex with three representative ligands (red crystal, yellow carvedilol protein, orange carvedilol ligand, light green ritodrine protein, dark green ritodrine ligand, cyan carazolol protein, blue carazolol ligand). The ribbon representation of the backbone, identical for all the structures since the optimizations did not allow movements of the backbone, is colored with a smooth transition from red to blue moving from the N-terminus to the C-terminus.

Figure 5.

Comparison of the virtual screening results obtained when using the β_2 -AR crystal structure alone (blue) and when combining it with three alternative conformations according to the receptor ensemble docking approach (red). The percentage yield of ligand retrieval is plotted versus percentage of screened compounds, sorted according to their docking rank, using the HTVS, SP, SP-XP and SP-London dG scoring functions.

Figure 6.

Alternative protein conformation optimized with the full agonist isoproterenol (red: receptor in the crystal structure; pink: carazolol; green: isoproterenol-adapted receptor; cyan: isoproterenol). The backbone of the crystal structure is shown in ribbon representation, and is colored with a smooth transition from red to blue moving from the N-terminus to the C-terminus. The backbone of the isoproterenol-adapted receptor is shown as a thin white line.

Figure 7.

Percentage yield of ligand retrieval plotted versus percentage of screened compounds, sorted according to their docking rank, for the isoproterenol-adjusted structure of the β_2 -AR using the HTVS, SP, SP-XP and SP-London dG scoring functions. Red dots represent blockers and green dots represent agonists. A significant migration of the agonists towards the top scoring positions can be noticed when comparing these results with those shown in Figure 2.

Figure 8.

Percentage yield of ligand retrieval plotted versus percentage of screened compounds, sorted according to their docking rank, for the β_2 -AR crystal structure and three homology models using the HTVS (red), SP (green), SP-XP (yellow) and SP-London dG (blue) scoring functions.

Figure 9.

Enrichment of true ligands concentration within the top scoring 10 to 500 compounds when compared to their concentration in the whole database, for the screenings based on the three homology models. Values expressed as enrichment factors (EF). Numerical values can be found in Table S2 of the supplementary data.

Table 1

Percentage yield of ligand retrieval and enrichment factor (EF) within various percentages of top scoring compounds in the screening based on the β_2 -AR crystal structure, applying the London dG and XP scoring functions one the poses selected with the SP docking.

SP-London dG			SP-XP		
Screened Database (%)	Yield (%)	EF	Screened Database (%)	Yield (%)	EF
0.10	35.00	34 9.16	0.10	30.00	29 9.28
0.20	46.67	23 2.79	0.20	41.67	20 7.85
0.50	73.33	14 6.83	0.50	55.00	11 0.13
1.00	85.00	84.95	1.00	61.67	61.63
2.50	91.67	36.66	2.50	76.67	30.66
5.00	96.67	19.34	5.00	81.67	16.34
10.00	96.67	9.67	10.00	88.33	8.83

Table 2

Blockers and agonists retrieved within the top scoring compounds using the crystal structure and the isoproterenol-adjusted structure.

		HTVS		SP	
		Crystal structure	Isoproterenol-adjusted structure	Crystal structure	Isoproterenol-adjusted structure
TOP-50	Blockers	2	0	6	1
	Agonists	0	11	0	14
	Total	2	11	6	15
TOP-100	Blockers	4	1	7	2
	Agonists	0	14	2	18
	Total	4	15	9	20
TOP-250	Blockers	8	2	13	2
	Agonists	1	17	2	20
	Total	9	19	15	22
TOP-500	Blockers	8	2	17	3
	Agonists	5	18	4	21
	Total	13	20	21	24
		SP-XP		SP-London dG	
		Crystal structure	Isoproterenol-adjusted structure	Crystal structure	Isoproterenol-adjusted structure
TOP-50	Blockers	13	1	16	1
	Agonists	4	11	5	2
	Total	17	12	21	3
TOP-100	Blockers	16	1	17	1
	Agonists	8	16	10	6
	Total	24	17	27	7
TOP-250	Blockers	18	3	24	3
	Agonists	13	17	18	11
	Total	31	20	42	14
TOP-500	Blockers	18	7	28	6
	Agonists	18	20	23	17
	Total	36	27	51	23

Table 3

Percentage yield of ligand retrieval and enrichment factor (EF) within various percentages of top scoring compounds in the screening based on the β_2 -AR homology models, applying the London dG and XP scoring functions one the poses selected with the SP docking.

Model 1					
SP-London dG			SP-XP		
Screened Database (%)	Yield (%)	EF	Screened Database (%)	Yield (%)	EF
0.10	3.33	33.22	0.10	3.33	33.22
0.20	5.00	24.94	0.20	5.00	24.94
0.50	11.67	23.37	0.50	13.33	26.69
1.00	21.67	21.66	1.00	20.00	19.99
2.50	30.00	12.00	2.50	30.00	12.00
5.00	41.67	8.33	5.00	33.33	6.67
10.00	51.67	5.17	10.00	41.67	4.17

Model 2					
SP-London dG			SP-XP		
Screened Database (%)	Yield (%)	EF	Screened Database (%)	Yield (%)	EF
0.10	5.00	49.88	0.10	10.00	99.76
0.20	8.33	41.55	0.20	11.67	58.21
0.50	16.67	33.38	0.50	16.67	33.38
1.00	40.00	39.98	1.00	21.67	21.66
2.50	53.33	21.33	2.50	23.33	9.33
5.00	60.00	12.00	5.00	31.67	6.33
10.00	73.33	7.33	10.00	40.00	4.00

Model 2-F290g+					
SP-London dG			SP-XP		
Screened Database (%)	Yield (%)	EF	Screened Database (%)	Yield (%)	EF
0.10	20.00	19.952	0.10	11.67	11.642
0.20	25.00	12.470	0.20	20.00	99.76
0.50	45.00	90.11	0.50	25.00	50.06
1.00	56.67	56.64	1.00	26.67	26.65
2.50	65.00	25.99	2.50	30.00	12.00

Model 2-F290g+						
SP-London dG			SP-XP			
Screened Database (%)	Yield (%)	EF	Screened Database (%)	Yield (%)	EF	EF
5.00	75.00	15.00	5.00	43.33	8.67	8.67
10.00	86.67	8.67	10.00	56.67	5.67	5.67