JOURNAL OF CHEMICAL INFORMATION AND MODELING

Article

pubs.acs.org/jcim

Ligand-Optimized Homology Models of D_1 and D_2 Dopamine Receptors: Application for Virtual Screening

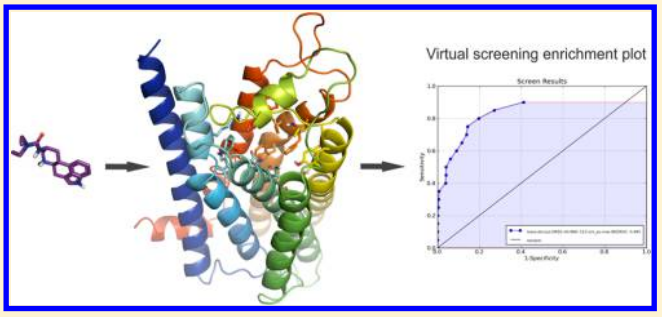
Marcin Kołaczkowski,*[,†,‡] Adam Bucki,[†] Marcin Feder,[‡] and Maciej Pawłowski[†]

[†]Department of Pharmaceutical Chemistry, Jagiellonian University Collegium Medicum, 9 Medyczna Street, 30-688 Kraków, Poland

[‡]R&D Department, Adamed Ltd., Pienków 149, 05-152 Czosnów, Poland

Supporting Information

ABSTRACT: Recent breakthroughs in crystallographic studies of G protein-coupled receptors (GPCRs), together with continuous progress in molecular modeling methods, have opened new perspectives for structure-based drug discovery. A crucial enhancement in this area was development of induced fit docking procedures that allow optimization of binding pocket conformation guided by the features of its active ligands. In the course of our research program aimed at discovery of novel antipsychotic agents, our attention focused on dopaminergic D_2 and D_1 receptors (D_2R and D_1R). Thus, we decided to investigate whether the availability of a novel structure of the closely related D_3 receptor and application of induced fit docking procedures for binding pocket refinement would permit the building of models of D_2R and D_1R that facilitate a successful virtual screening (VS). Here, we provide an in-depth description of the modeling procedure and the discussion of the results of a VS benchmark we performed to compare efficiency of the ligand-optimized receptors in comparison with the regular homology models. We observed that application of the ligand-optimized models significantly improved the VS performance both in terms of BEDROC (0.325 vs 0.182 for D_1R and 0.383 vs 0.301 for D_2R) as well as EF1% (20 vs 11 for D_1R and 18 vs 10 for D_2R). In contrast, no improvement was observed for the performance of a D_2R model built on the D_3R template, when compared with that derived from the structure of the previously published and more evolutionary distant β_2 adrenergic receptor. The comparison of results for receptors built according to various protocols and templates revealed that the most significant factor for the receptor performance was a proper selection of "tool ligand" used in induced fit docking procedure. Taken together, our results suggest that the described homology modeling procedure could be a viable tool for structure-based GPCR ligand design, even for the targets for which only a relatively distant structural template is available.



INTRODUCTION

Despite an increasing number of crystallized proteins being actual or putative drug targets, a majority of them, especially membrane proteins, still remain experimentally unresolved. Therefore homology modeling is the method of choice for obtaining models suitable for structure-based drug discovery process.[1] Recently several new crystal structures of monoaminergic G protein-coupled receptors (GPCRs) were published and some novel modeling techniques were introduced, creating opportunities for improvements in homology modeling-based approaches.[2−4] Nevertheless, issues such as the choice of an appropriate template, binding site refinement, or selection of a receptor conformation that is best suited for virtual screening (VS) represent an ongoing challenge.[5]

In the course of research devoted to homology modeling of GPCRs, we described the successful application of a ligand-guided model selection for serotonin 5-HT_{1A} and 5-HT_7 receptors.[6,7] This methodology could be called "inverse virtual screening" because the procedure was based on a selection of a proper model by docking active, possibly rigid ligands to an abundant conformational ensemble of the receptor, created at an initial modeling stage.[8,9] The advantage of this approach was inclusion of some degree of receptor flexibility in the process of automated docking as well as exploiting the structure of the ligand for selection of the best binding pocket conformation. However, the potential drawback of this methodology was a lack of ligand context at the stage of model preparation, resulting from random generation of an initial set of conformations that was subsequently selected based on docking procedure. This demanded a thorough search of the receptor's conformational space, resulting also in conformations that may not be relevant to binding site optimization. On the other hand, a limited number of conformations taken for ligand-based selection might not guarantee a sufficiently full conformational search of the most important regions of the receptor.

Induced fit docking (IFD), introduced by Schrödinger, Inc., is a method that combines flexible ligand docking with receptor structure prediction and side chain refinement. The workflow

Received: August 30, 2012
Published: February 12, 2013

   dx.doi.org/10.1021/ci300413h | J. Chem. Inf. Model. 2013, 53, 638−648

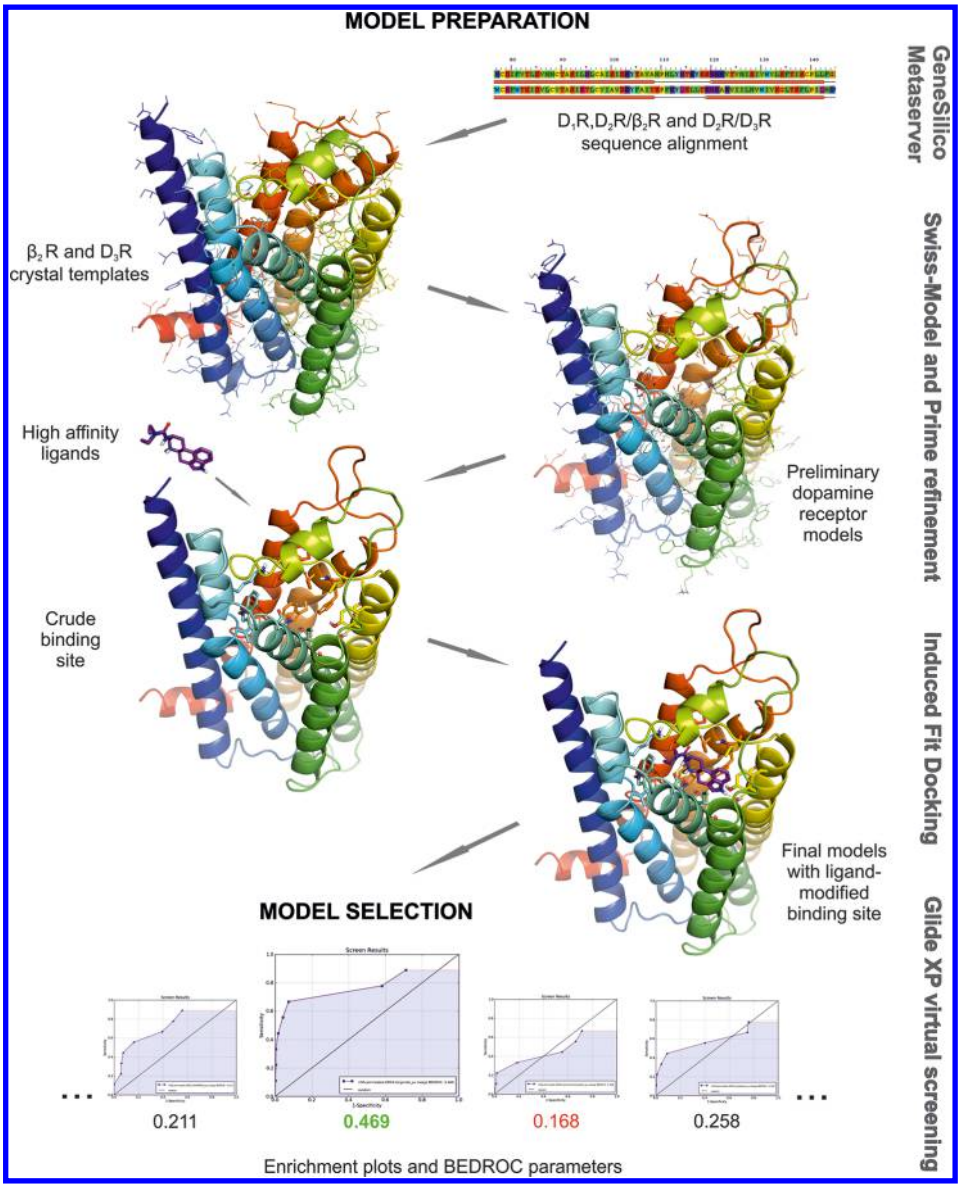


Figure 1. Homology model preparation and selection based on ligand-optimized modeling procedure.

consists of alternating automated docking of the ligands to the initial models of receptors, using a Glide multistep algorithm, followed by refinement of the obtained complexes with Prime, which enables ligand-induced receptor relaxation.[10,11] It can be also considered as a method of ligand-based protein structure optimization and seems to be very useful in homology modeling of GPCRs.[12] Considering this fact, we have recently developed a ligand-steered, homology modeling approach, utilizing the IFD technique, and successfully applied it to the modeling of serotoninergic receptors.[13]

In the present project devoted to discovery of novel ligands as potential antipsychotic drugs, our attention has been focused on dopaminergic D_2 and D_1 receptors (D_2R and D_1R). Antagonism of the former is traditionally associated with antipsychotic activity, while agonism of the latter has potential for treatment of poorly controlled cognitive deficits in schizophrenia and other mental disorders.[14−16] In order to rationally design new ligands with desired profiles of D_2 and D_1 receptor activity, we built and validated their homology models using their closest, structurally characterized homologues: adrenergic β_2R and dopaminergic D_3R as templates followed by binding pocket optimization using a novel, ligand-optimized modeling approach. In this study, we aimed to address several questions: (1) Is induced-fit docking based binding pocket refinement suitable for optimizing homology models of monoaminergic GPCRs? (2) How does it perform comparing to "inverse virtual screening" of conformational ensembles, the procedure described by us previously? (3) What is the impact of template choice in homology modeling of those receptors for the results of virtual screening? (4) How does model performance depend on the selection of the ligand taken for induced-fit procedure? And finally, (5) How should the best receptor for virtual screening of a large database be efficiently preselected?

Herein, we present detailed descriptions of the procedures used for a model preparation and selection as well as a series of virtual screening (VS) experiments that were carried out to answer the above-mentioned questions.

dx.doi.org/10.1021/ci300413h | J. Chem. Inf. Model. 2013, 53, 638−648

METHODOLOGY

Overview. The global fold and steric arrangement of the main secondary structure elements is highly conserved in GPCRs.[17] Therefore, we presumed that the state-of-the-art homology modeling allows us to assemble reasonable "scaffolds" for dopamine D_1 and D_2 receptor models. On the other hand, the detailed structural arrangement of side chains in the ligand binding pocket may differ significantly, so the quality of homology models in this area is likely to be inadequate for VS purposes. We therefore decided to use known dopaminergic receptor ligands and reconstruct binding site conformations in a way resembling reconstruction of a lock on a key.

A multistep procedure of human dopamine D_1 and D_2 receptor modeling was therefore performed. The SwissModel server was involved in initial homology model building, while an induced fit docking (IFD) protocol introduced by Schrödinger, Inc. was used as a tool for subsequent binding site optimization (Figure 1).

For comparison, some reference models were generated utilizing the "inverse virtual screening" method that was previously applied by us[6,7] as well as some other authors[8,9] for successful GPCR modeling. VS trials were performed as a tool to verify and compare quality of the models and to select those that differentiate the affinity of tested compounds the most efficiently.

Homology Modeling. The models used in main experiments were built on the template of β_2 adrenergic receptor (β_2R) crystal structure (Protein Data Bank ID 2RH1). Reasonable sequence similarity between dopamine and adrenergic β_2 receptors, which belong to the same GPCR subfamily, justified the choice of the latter structure as an initial state-of-the-art template for homology modeling.[18] However, the first crystal structure of a dopaminergic receptor, the D_3 receptor (D_3R), was recently resolved and published (PDB ID 3PBL).[19] The latter seemed to be a natural choice of template for D_1R and D_2R modeling, but, in fact, β_2R was found to possess both higher similarity in the binding site region (63% vs 56%) as well as overall (33% vs 32%) sequence identity to D_1R. Moreover, both proteins exhibit the same length of the ECL2 region, the evolutionary variable loop capable of interacting with ligand (5 residues instead of 4 in D_3R), and thus, it had an advantage over the recent D_3R template, which was therefore not considered for D_1R modeling. In the case of D_2R, we maintained the initial models based on β_2R template, since the β_2R crystal structure is the most extensively used template for D_2R homology modeling. Nevertheless, due to a very high sequence similarity in the binding site region between D_2R and D_3R—97% vs 56% in case of β_2R (overall 48% vs 26%) as well as the same length of ECL2 fragment forming ligand binding site (4 residues)—we also prepared a group of D_2R models based on D_3R crystal structure, for comparison with those built using β_2R template.

We used the complete crystal structures containing the heptahelical bundle and all the originally occurring loops, the only exception being the artificial fragments replacing the third intracellular loop (ICL3) in the crystallized receptors, which were removed. Sequence alignment was performed via GeneSilico Metaserver[20] that is a gateway to multiple fold recognition servers (for sequence alignments see Supporting Information Figure S1). The results proposed by a profile−profile comparison tool, *hhsearch*, were manually selected as the starting alignments between the sequences of D_1 and D_2 dopamine receptors (UniProt accession numbers P21728 and P14416, respectively) and respective templates. The alignments were manually adjusted in Swiss Pdb Viewer and submitted to the SwissModel server.[21] We constrained a part of the second extracellular loop (ECL2) secondary structure of dopamine D_1 receptor model to an α-helix as it is present in the crystal structure of the β_2-adrenoreceptor. This decision was based on the length of the ECL2 loop (about 30 amino acids, even longer than in β_2R) and the fact that α-helix is recognized as the most compact way to pack amino acids. Moreover, secondary structure prediction algorithms, like pssfinder, ssp, sspal, and nnssp suggested that an α-helix organization is appropriate. Using the SwissModel model building algorithm, the cores of the new models were determined by average backbone atom positions of the template structure, while insertions and deletions were constructed using an ensemble of fragments compatible with the neighboring stems. The reconstruction of the side chains was based on the conformation of corresponding residues in the template. Ultimately, a scoring function was applied to assess the most likely conformation, which underwent energy minimization using the GROMOS96.[22−26] After removal of the artificial fragments, short loops were created by joining Leu 230 and Lys 263 of β_2R or Asn 224 and Arg 360 of D_3R to avoid modeling of the long ICL3, which is distant from the binding site and therefore beyond the scope of our experiments.

Further improvements of the models were applied using software implemented in Schrödinger Suite 2009. All the operations engaging Schrödinger's tools were executed with default settings, unless stated otherwise. The homology models were validated by processing in Protein Preparation Wizard, which consisted of optimizing the hydrogen-bonding network and restrained minimization of the whole system with the OPLS_2005 force field. The least conserved fragments of the models, i.e. extracellular loops, among which ECL 2 is known to be involved in ligand binding, were additionally refined using the Prime Refinement tool.

Ligand-Based Binding Site Optimization. Binding pockets of the resulting homology models were optimized using modified induced fit docking workflow.[10,11] This method of ligand-based structure optimization is very useful in homology modeling of GPCRs, since it combines flexible ligand docking with protein structure prediction and side chains refinement. The workflow consists of alternating automatic docking of the ligands to the initial models of receptors using Glide algorithm and refinement of the resulting complexes with Prime.[27,28] Glide docks ligands through a multistep algorithm. The first phase includes approximate positioning and scoring, followed by torsion angle flexible optimization of the poses (OPLS-AA). A Monte Carlo sampling of pose conformation refines the best of them, and finally, GlideScore discriminates docking poses.[29,30] Receptors prepared using this procedure are more likely to bind ligands in their correct conformation. A modification that we used in the present study consisted of appending another cycle of optimization of the receptor structure and docking the ligand using Glide XP (extra precision). The modified procedure allows the models to improve in terms of ligand binding mode and obtained GlideScores.

Two groups of bioactive compounds were selected for the purpose of ligand-steered binding site optimization. Some of them are characterized by high affinity for D_1 (pK_i = 7.2−9.7, Figure 2) and the others for D_2 dopamine receptors (pK_i =

dx.doi.org/10.1021/ci300413h | J. Chem. Inf. Model. 2013, 53, 638−648

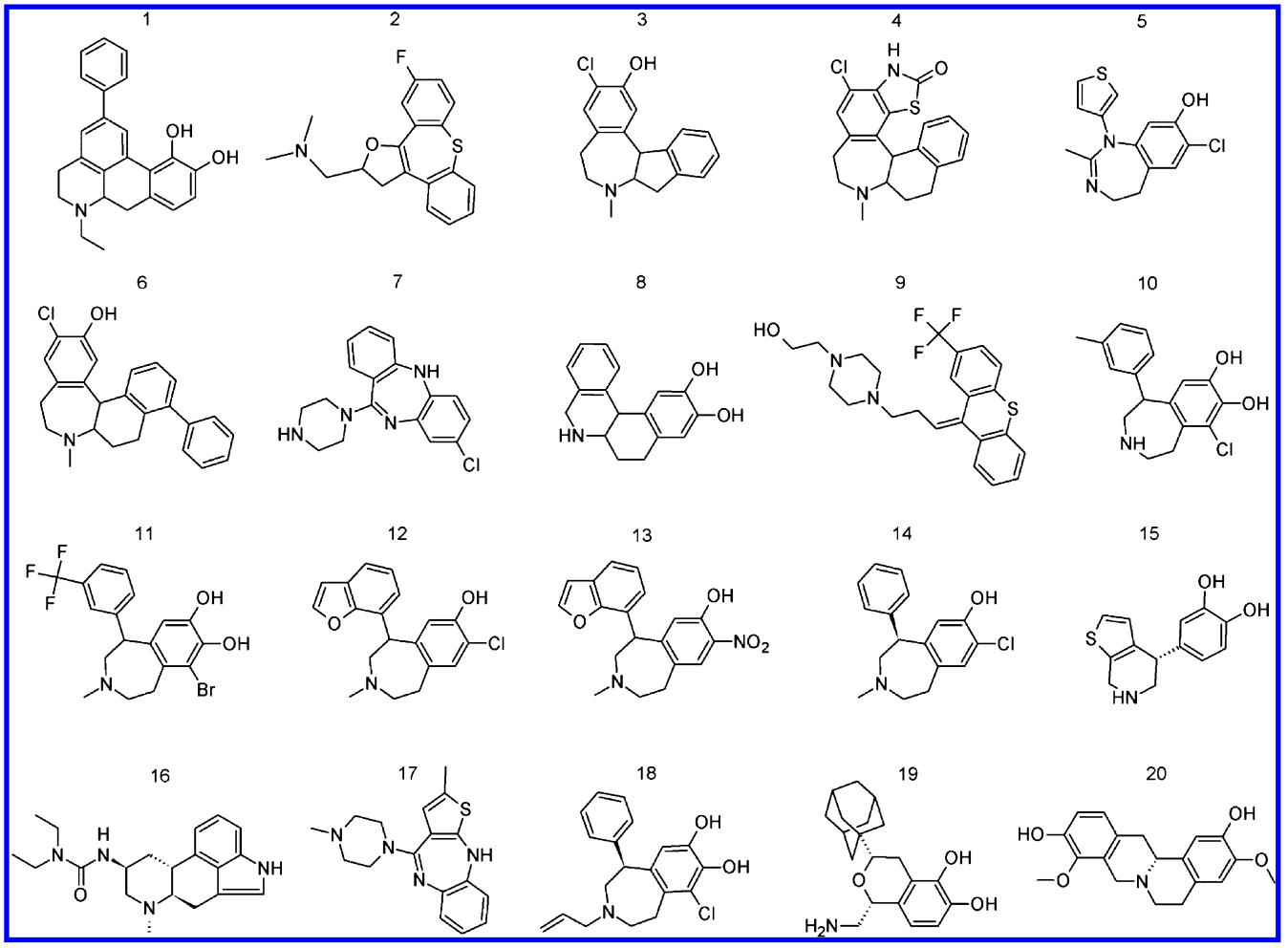


Figure 2. D_1 dopamine receptor ligands.

7.1−10.6, Figure 3). Binding affinities of those compounds were derived from PDSP K_i database (http://pdsp.med.unc.edu/pdsp.php).[31]

2D structures of the ligands were sketched or sourced from ChEMBL database (https://www.ebi.ac.uk/chembldb) and then converted to 3D and optimized using LigPrep. If available, ring conformations for cyclic ligands were taken from their crystal structures deposited in the Cambridge Crystalographic Data Centre (CCDC) database and retained in the process of flexible docking.[32] The outcome of induced fit docking of the ligands to the crude receptor models prepared on β_2R template were 187 and 158 complexes of the D_1 and D_2 receptors, respectively. An additional 158 complexes were obtained for the D_2R models based on D_3R template. The top-scored complexes were inspected visually to check the compliance with common binding mode for monoaminergic receptor ligands. That procedure resulted in 20 models of both D_1R and D_2R (the best one for each ligand used for model optimization) that served as molecular targets in further docking studies. Numbers of ligands used in the optimization step (Figures 2 and 3) were assigned to the resulting receptor models.

In each case of Glide docking, general parameters were set to defaults, with the exception of docking precision set to XP (extra precision) and flexible docking option retaining original conformations of amide bonds. H-bond constraints were applied on Asp3.32, since its interaction with the protonated nitrogen atom of the ligand is considered crucial for the monoaminergic GPCRs.[33] The centroid of a grid box (22 Å × 22 Å × 22 Å) was located on Asp3.32. In the ligand preparation stage, OPLS_2001 force field was applied. The same force field is integrated in Glide calculations by default, and it successfully optimized the structures we processed. In the VS studies, only the best-scored poses of each compound were taken into consideration during ranking preparation.

Generation of Reference Models. Reference homology models of D_1 and D_2 dopamine receptors were generated using previously developed method.[5,6] The only methodological differences consisted in the use of β_2R template, instead of bovine rhodopsin, and the choice of software for automated docking, which was Glide XP instead of FlexX . Briefly, in the initial step, a wide variety of model structures (100 in this case) that cover the conformational space of the binding site, were generated using Modeller[34−36] and numbered in the range 1−100. The proper models for main VS experiment were then selected by comparison of BEDROC parameters derived from VS trials on database 3 (1000 decoys) merged with 20 ligands that were drawn for VS studies with minor databases (see the Virtual Screening paragraph below), resulting in 20 final models for each receptor subtype.

Generation of one crude model using the described method is quicker than in a ligand steered approach, which requires more attention during model building. Nevertheless screening of 100 models demands huge computational time, whereas the IFD-based procedure yields already ranked models. Hence, the

dx.doi.org/10.1021/ci300413h | J. Chem. Inf. Model. 2013, 53, 638−648

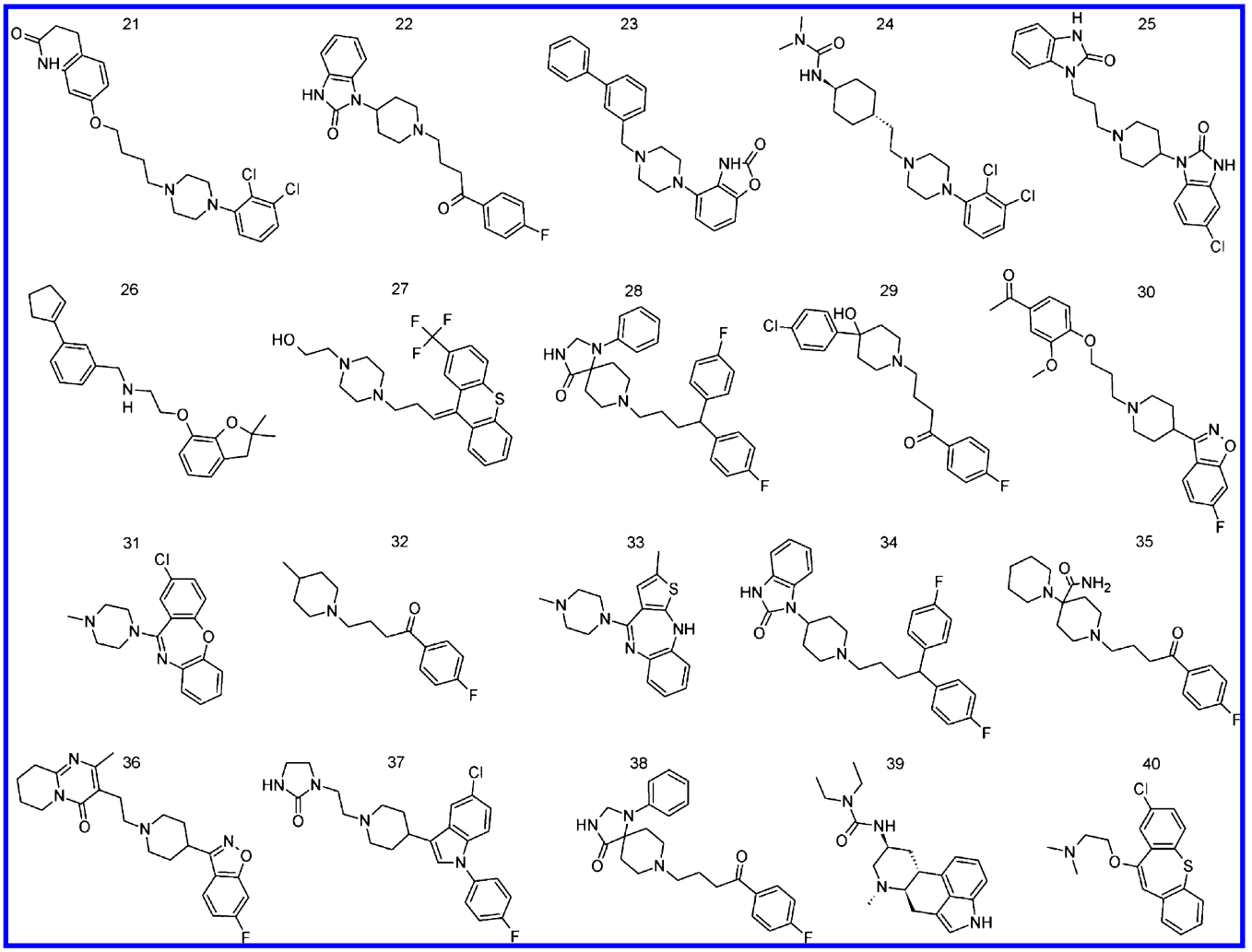


Figure 3. D_2 dopamine receptor ligands.

overall analytical time required to select 20 final models by either the reference or the ligand-steered methods is similar.

Virtual Screening. Four databases were utilized for virtual screening trials. The main database (1) was prepared based on ZINC's[37] usual CNS permeable subset. We started with downloading 344 596 chemical entities, characterized by properties, which favor blood−brain barrier permeability (polar surface area (PSA) between 0 and 60, molecular weight 150−400, and predicted xlogP between 1.5 and 2.7) and prepared for docking studies (generated 3D conformation, protonation variants in pH 5.75−8.25, and tautomers of the compounds). In order to exclude compounds considered to be unable to bind with momoaminergic receptors, the database was filtered with a simple pharmacophore model, representative for monoaminergic receptor ligands. To prepare this model, we selected several known antipsychotic drugs and found common pharmacophore features employing Phase.[38,39] The two most important were taken into account: positively charged group and aromatic ring, situated in a distance of 6 ± 1 Å from each other. The final version of the main database (1) consisted of 17 196 compounds and is further referred to as the 17K database. To initially verify performance of the prepared models and to select the most suitable one for principal virtual screening of a larger, real-life database 1, we utilized three minor 1000-compound databases. Database 2 is a commonly accessible database, structurally not related to the main database 1, consisting of compounds preprepared by Schrödinger, Inc. for testing of docking procedures.[29,40] The other state-of-the-art approach is to generate the databases containing property-matched decoys[41] in the context of active ligands. Such databases were generated separately for D_1R and D_2R ligands (database 3A and 3B, respectively). The databases were generated throughout the DUD-E online tool (http://dude.docking.org). As an input, 20 active ligands of D_1R and D_2R which were added to minor 1K databases in VSs (see below) were uploaded. The method uses molecular properties and fingerprint-derived Tanimoto coefficient to exclude potential ligands among the decoys. It resulted in two databases containing unique 1000 property-matched decoys each. We prepared yet another database (3), by selecting a random subset of 1000 compounds out of the main 17K database 1. This database is further referred to as the 1K database.

For virtual screening trials, the databases were mixed with high-affinity ligands of D_1R or D_2R. The ligands were acquired from ChEMBL database. Two sets of 340 ligands having lowest K_i values for D_1R or D_2R were mixed with 17K database 1. The randomly chosen 20 compounds from the above-mentioned group of 340 ligands were included in VS tests utilizing smaller databases 2−4. The obtained databases were docked to receptor models using Glide XP. For each VS result, enrichment factors and BEDROC (Boltzmann-Enhanced Discrimination of Receiver-Operating Characteristic)[42] param-

dx.doi.org/10.1021/ci300413h | J. Chem. Inf. Model. 2013, 53, 638−648

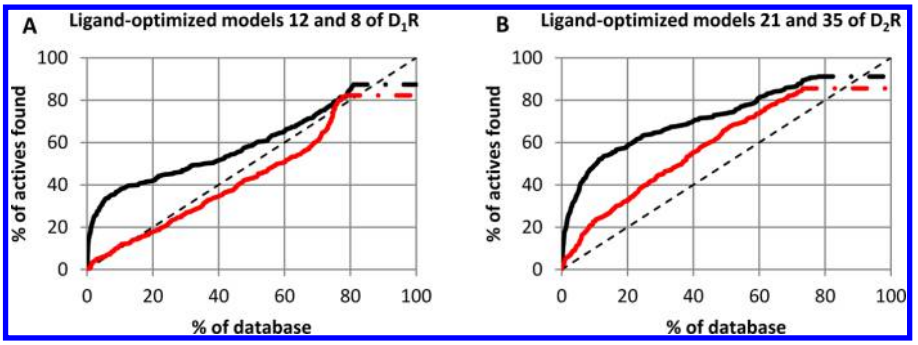


Figure 4. Enrichment factor plots showing VS performance of ligand-optimized model (A) **12** (black) and **8** (red) of D_1R; (B) **21** (black) and **35** (red) of D_2R (solid line), compared to random choice (dotted line).

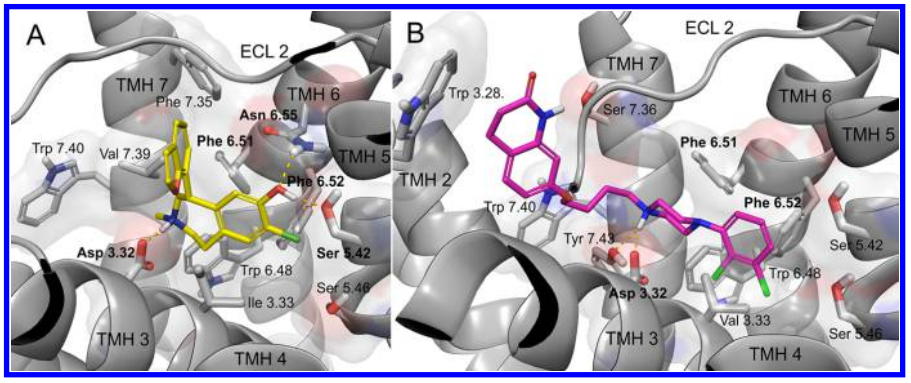


Figure 5. Binding modes of tool ligands in ligand-optimized model **12** of D_1R (A) and **21** of D_2R (B). Residues in the binding site are presented as thick sticks with transparent surfaces. Dotted lines represent H-bonds with polar residues: (TMH) transmembrane helix; (ECL) extracellular loop.

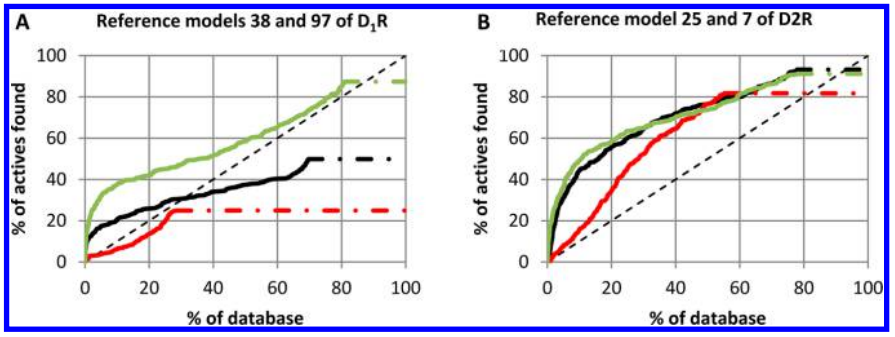


Figure 6. Enrichment factor plots showing VS performance of reference model (A) **38** (black) and **97** (red) of D_1R; (B) **25** (black) and **7** (red) of D_2R (solid line), compared to random choice (dotted lines). Performance of the best ligand-optimized models was shown for comparison (green).

eters were calculated and enrichment plots were generated. Enrichment factor (EF) measures increases in the number of active compounds found within a defined fraction of the ordered list relative to the random distribution, while BEDROC is called "virtual screening usefulness scale", combining improvement over random scale evaluation with the notion of early recognition.

The reference receptor models evaluation was performed on 17K database 1, mixed with the same 340 ligands that were depicted above. All the docking parameters were kept unchanged. That enabled direct comparison of models constructed through ligand-steered and "inverse virtual screening" methods.

To determine whether associating multiple model screening would bring an improvement to the results, outcomes of three ligand-steered models characterized by the highest BEDROC values were combined. Their GlideScore results were merged, and only the best score for each ligand was stored. Glide_sort application was used to manage the data.

RESULTS

Comparison of Models Built Using Different Methods.

The comparison was based on virtual screening results from trials engaging the 17K database, using novel and reference models. Final scores of the VS trials varied widely, depending on the methodology used and the ligand taken for the model building stage. In the case of IFD-based models, BEDROC values ranged from 0.067 to 0.325 and from 0.139 to 0.383 for D_1R and D_2R, respectively. Best-performing models were prepared based on the following ligands (BEDROC and EF 1% shown respectively in brackets): **12** (0.325; 17) for D_1R and **21** (0.383; 18) for D_2R. The best model of D_2R was characterized by slightly higher BEDROC values and EF 1% comparing to D_1R. The score of the worst-performing D_2R model based on ligand **35** (0.139; 4,7) was also higher than the analogous D_1R

dx.doi.org/10.1021/ci300413h | J. Chem. Inf. Model. 2013, 53, 638−648

model based on ligand **8** (0.067; 0,3). Enrichment plots for the above-mentioned models are presented on Figure 4, while their complexes with tool ligands are shown on Figure 5. All the VS results for IFD-based models are presented in Supporting Information Table S1.

Models prepared with previously used "inverse virtual screening" methodology gave the following results: for D_1R models, BEDROC values ranged between 0.046 and 0.182, while for D_2R models between 0.091 and 0.301. The best outcomes were obtained for model **38** (0.182; 11) for D_1R and **25** (0.301; 7.1) for D_2R. The worst results were obtained for D_1R model **97** (0.046; 1.8) and D_2R model **7** (0.091; 0.6) . Enrichment plots for the above-mentioned models are presented in Figure 6. All the VS results for reference models are presented in Supporting Information Table S2.

The present results show high performance of the IFD-based methodology in preparing models suitable for virtual screening, which was reflected in both high BEDROC and EF 1% values. This is especially significant in view of the fact that the database of compounds that are assumed to be inactive was preprocessed with a pharmacophore filter, creating stringent conditions requiring a distinction to be made between active and inactive ligands of a relatively similar structure. Although the visual examination of Figure 6 or comparison of BEDROC values, especially in the case of D2R, do not give the impression of marked differences in performance between the methods, the analysis of EF1% (17 vs 11 and 18 vs 7.1) reveals significant superiority of ligand-optimized models.

Comparison of Models Built Using Different Templates. In the case of D_2R models obtained using the IFD-based method on the template of D_3R, the best performing model was prepared with ligand **38** and was characterized by slightly lower BEDROC (0.369) and EF1% (16) values compared to the best model built on β_2R template. The worst performing model was prepared with ligand **29**, and the BEDROC value was equal (0.139) to the worst β_2R based model result, though having a lower EF1%—2.7 (Figure 7 and Supporting Information Table S3). The lower BEDROC value for the best performing model results from the fact that less active compounds were successfully docked to this model (284 vs 310), making it relatively less sensitive than its β_2R-based counterpart.

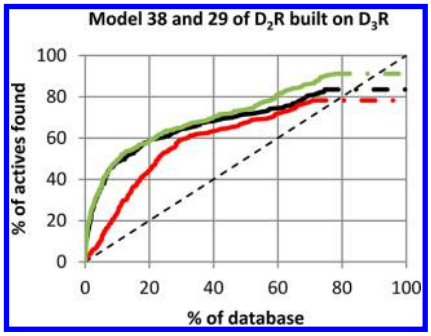


Figure 7. Enrichment factor plots showing VS performance of model **38** (black) and **29** (red) of D_2R built on D_3R template (solid line), compared to random choice (dotted lines). Performance of the best ligand-optimized D_2R model built on β_2R template was shown for comparison (green).

These observations indicate that despite much higher sequence similarity, especially in the binding site region, D_2R models obtained using IFD-based method on the template of D_3R crystal structure did not allow for an improvement in VS performance compared with β_2R-based models.

Combination of Homology Models. The performance of ligand-steered homology models constructed on β_2R template seems superior to that of the other models presented above. Nevertheless, it would be desirable to make models that are more selective and precise. We therefore tested if a combination of homology models would improve the quality of results. The docking outcomes of three D_2R models were combined: 21 (0.383; 18), 36 (0.360; 13), and 26 (0.332; 16)—BEDROC and EF1% values given in parentheses. The obtained score was 0.361; 17, which was higher than the average for those three receptors, but less than the result achieved by the highest-ranked model, by itself. Therefore, it seems that use of such a combination of multiple models does not improve the performance of the VS.

Comparison of Smaller Databases for Selection of the Best Models. The large size of most virtual databases limits the number of models that can be used for their screening. This, in turn, requires a robust method for selection of a suitable model. One strategy is to perform VS trials on a database of decoy compounds enriched with known actives and to analyze essential parameters describing usefulness of the model, like EF or BEDROC. However, comprehensive screening of a real, large database performed on tens of models would consume a great deal of computational time. Therefore, smaller databases may be used for preliminary VS trials in order to select models that would perform well in a large scale.[43] We used freely accessible database 2 of 1000 drug-like decoys provided by Schrödinger, Inc. and compared VS performance of the models selected on its base with those selected using the 17K database 1. The results for database 2, represented by BEDROC values (Supporting Information Table S4), corresponded to some extent with those obtained for database 1, effecting in correlation coefficient (R^2) reaching 0.57 for D_1R and 0.68 for D_2R models (Figure 8). Similar correlations were obtained for property-matched decoy databases (3A and 3B—Supporting Information Table S5), reaching 0.54 for D_1R and 0.75 for D_2R (Figure 9).

Nevertheless, such a level of correlation does not guarantee that models which performed best in the small database, would necessarily be equally efficient in the large one. We therefore proposed another way of using the smaller database for selection of the best model using a newly prepared 1K database 4, consisting of a randomly selected part of the 17K database 1, and therefore being similar to it. VS parameter values (Supporting Information Table S6) collected for database 4 correlated much better with those of database 1 than did those obtained with databases 2 or 3A/3B, resulting in R^2 coefficients equaling 0.80 (D_1R) and 0.91 (D_2R) (Figure 10).

Similarity of the databases to the active compounds was determined using Qiksim module of Schrödinger Suite that utilizes Tanimoto coefficient and Euclidean distance. Database entities as well as the active compounds were ranked according to their similarity to the virtual probe molecule obtained by computing the averages of the descriptors for the actives (for details see QikProp 3.0 User Manual from www.schrodinger.com). Enrichment factors were computed in order to show how the active compounds are distinguishable by similarity with particular database entities (Figure 11). Database 4, as anticipated, displayed the same, substantial similarity level to the actives as database 1, which was expressed in modest EF values. On the other hand databases 2 and 3A/3B differed from

dx.doi.org/10.1021/ci300413h | J. Chem. Inf. Model. 2013, 53, 638−648

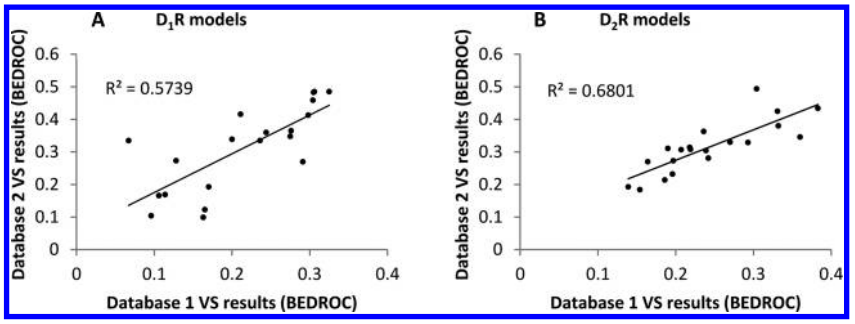


Figure 8. Correlation of BEDROC values obtained from VS of databases 1 and 2 using D_1R (A) and D_2R (B) models.

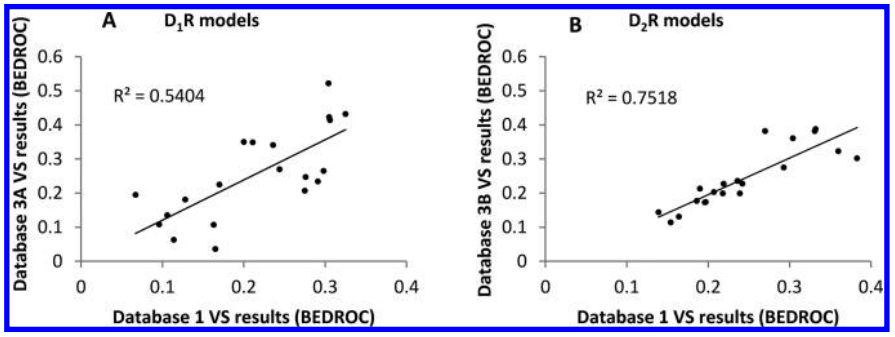


Figure 9. Correlation of BEDROC values obtained from VS of databases 1 and 3A/3B using D_1R (A) and D_2R (B) models.

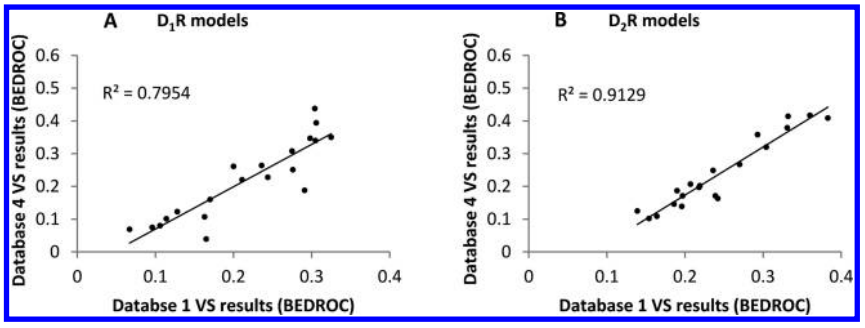


Figure 10. Correlation of BEDROC values obtained from VS of databases 1 and 4 using D_1R (A) and D_2R (B) models.

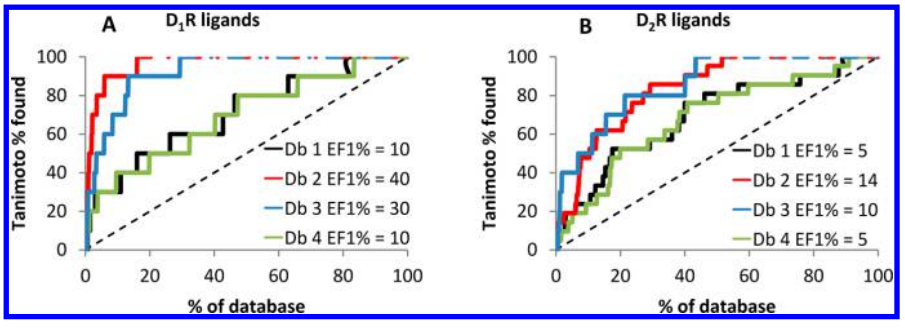


Figure 11. Percentage of the D_1R (A) and D_2R (B) active compounds found by similarity vs percentage of molecules processed for database 1 (black), 2 (red), 3 (blue), and 4 (green) (solid lines), compared to the random choice (dotted lines).

them significantly, being much less similar and resulting in a very high enrichment factor, based on similarity comparison only.

It is noteworthy that Schrödinger's database 2 as well as property-matched decoys databases 3A/3B are much more structurally varied than databases 1 and 4, prepared using the pharmacophore filter for monoamine receptors. It makes Schrödinger's and property-matched decoys databases entities theoretically easier to distinguish from actives, which may promote some receptor models that are not capable of distinguishing similar ligands (preselected using a pharmacophore model) and therefore makes it relatively less relevant for model qualification in the case of such a biased larger database. It should be emphasized that the EF1% values determined from similarity comparison using databases 2 and 3A/3B, are similar to or even higher than the EF1% values determined in docking experiments with those databases. On the other hand, in the case of databases 1 and 4, the EF1% values resulting from

dx.doi.org/10.1021/ci300413h | J. Chem. Inf. Model. 2013, 53, 638−648

similarity comparison are significantly lower than that achievable from docking to the best model (Table 1). These observations support the validity of using databases such as 4 in proper model selection.

Table 1. EF1% Obtained Using Similarity Comparison and Docking to Ligand-Optimized Models

receptor	database	similarity EF1%	docking EF1%
D_1	1	10	20
	2	40	31
	3A/3B	30	31
	4	10	31
D_2	1	5	18
	2	14	15
	3A/3B	10	20
	4	5	20

DISCUSSION AND CONCLUSIONS

In the present paper we show an example of ligand steered, IFD-based homology modeling of dopaminergic D_1 and D_2 receptors, using different experimental conditions and their performance in virtual screening trials. We successfully addressed several questions identified in the Introduction. In particular, application of the novel methodology resulted in receptor models that consistently reproduced ligand binding modes and demonstrated good performance in VS experiments, which was reflected in high BEDROC (0.3−0.4) and EF1% (18−20) values. It should be stressed that those results were obtained for the decoy database that was prefiltered using a pharmacophore model for monoaminergic ligands, and therefore, the VS task was significantly more difficult than for more diversified databases. The results obtained using a publicly available, unbiased database from Schrödinger, Inc., or property-matched decoys database were even higher (BEDROC > 0.5; EF1% > 30) which provides additional support for the quality of those models.

The reference models, built using previously described "inverse virtual screening" methodology, performed worse, resulting in lower BEDROC (0.2−0.3) and EF1% values (10−11) and underlining the superiority of the currently used approach. This is most probably due to the favorable ligand effect at the stage of model building, which is absent from previously used methodology. It focuses on the conformational variations of the most significant areas of the receptor, especially the binding site, and drives them toward improved ligand binding, while sparing unnecessary changes.

The main template used in this study for modeling of both receptors, using different methods, was the commonly employed β_2R crystal structure. Nevertheless, since a crystal structure of D_3R has recently been described and this receptor displays a very high similarity with D_2R, especially in a binding site region (sequence homology 97%), we compared results obtained using D_2R models built on D_3R template with those modeled on β_2R. The VS performance of models based on both templates was basically equal, with slightly lower maximal BEDROC (0.369 vs 0.383) and maximal EF1% (16 vs 18) of D_3R based models. Those results are somewhat surprising for us, given the much higher binding site similarity of D_3R template, which did not result in improvement of VS parameters. On the other hand, we observed that models of D_2R based on D_3R were more useful in a detailed analysis of binding modes and allowed for drawing more consistent conclusions on differences in shape and specific interaction points of the binding site, with highest importance for D_2/D_1 receptor selectivity (data in publication).

One of the cornerstones of the induced-fit docking based homology modeling is the use of a ligand at the stage of receptor preparation. Our results suggest that the inclusion of a proper tool ligand in a group used for binding site optimization is of fundamental importance for a model quality, especially in terms of efficacy in VS trials. It should be noted that performance of models in VS trials depended much more crucially on the ligand used in the IFD procedure than on a choice of template for homology modeling or a method of model preparation. This is described by much higher deviations of BEDROC and EF1% between receptors based on different ligands, within given template or building methodology (e.g., 0.383; 18 vs 0.139; 4.7), than between best-performing receptors obtained on different templates (e.g., 0.383; 18 vs 0.369; 16) or using different methods (e.g., 0.383; 18 vs 0.315; 11). Our findings suggest that, despite the profound impact of a tool ligand structure on binding site formation, the resulting models do not seem to favor only similar compounds. In fact, scoring values of tested compounds do not correlate with Tanimoto coefficients based on topological torsions fingerprints describing their similarity to tool ligand, which testifies to the objectivity of the method (see Figure S2 in the Supporting Information).

Low conformational flexibility of tool ligands theoretically favors correct transferring of electronic and steric features to newly generated models, and thus, such ligands should be the most valuable. In fact, the ligands that contributed to creation of the best models were not necessarily the most rigid ones, and they differed structurally from each other depending on methodology or template used for receptor modeling as well as type of database used in VS. Lack of a universal recipe for a tool ligand makes it necessary to use numerous and diversified high-affinity compounds for binding site tuning, to increase a probability of obtaining efficacious models. This in turn requires a robust method of model selection.

We performed a comparison of efficacy of three smaller test databases in selection of the best model for VS of larger database. This procedure proved that the properly selected smaller database, here extracted from the original one, may be efficient in identifying the most promising models for VS studies, considerably saving computational time. The extracted 1k database proved to be most suitable, showing the superiority of a "dedicated" test database that is similar to the large one.

ASSOCIATED CONTENT

Supporting Information

Figure S1: multiple sequence alignment between β_2R, D_1R, D_2R, and D_3R. Table S1: database 1 VS results for ligand-optimized D_1R and D_2R models. Table S2: database 1 VS results for reference D_1R and D_2R models. Table S3: database 1 VS results for D_2R models built on D_3R template. Table S4: database 2 VS results for ligand-optimized D_1R and D_2R models. Table S5: database 3A/3B VS results for ligand-optimized D_1R and D_2R models. Table S6: database 4 VS results for ligand-optimized D_1R and D_2R models. Figure S2: similarity of the active ligands to the tool ligand (expressed as Tanimoto coefficient) vs GlideScores derived from docking of the active ligands to the model optimized based on the tool ligand.

This information is available free of charge via the Internet at http://pubs.acs.org

AUTHOR INFORMATION

Corresponding Author

*E-mail: marcin.kolaczkowski@uj.edu.pl.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

Part of this work was presented at the international conference "Joint Meeting on Medicinal Chemistry", Budapest 2009. Glide, induced fit docking, LigPrep, Phase, Prime, and Protein Preparation Wizard were implemented in Schrödinger Suite 2009, which was licensed for Adamed Ltd.

ABBREVIATIONS

BEDROC, Boltzmann-enhanced discrimination of receiver operating characteristic; β_2R, β_2 adrenergic receptor; CCDC, Cambridge Crystalographic Data Centre; D_1R, D_1 dopamine receptor; D_2R, D_2 dopamine receptor; D_3R, D_3 dopamine receptor; ECL, extracellular loop; EF, enrichment factor; GPCR, G protein coupled receptor; IFD, induced-fit docking; TMH, transmembrane helix; VS, virtual screening

REFERENCES

(1) Tang, H.; Wang, X. S.; Hsieh, J.-H.; Tropsha, A. Do crystal structures obviate the need for theoretical models of GPCRs for structure-based virtual screening? *Proteins* **2012**, *80*, 1503−1521.

(2) Katritch, V.; Cherezov, V.; Stevens, R. C. Diversity and modularity of G protein-coupled receptor structures. *Trends Pharmacol. Sci.* **2012**, *33*, 17−27.

(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) Katritch, V.; Rueda, M.; Lam, PC-H; Yeager, M.; Abagyan, R. GPCR 3D homology models for ligand screening: lessons learned from blind predictions of adenosine A2a receptor complex. *Proteins* **2010**, *78*, 197−211.

(5) Kufareva, I.; Rueda, M.; Katritch, V.; Stevens, R. C.; Abagyan, R. Status of GPCR modeling and docking as reflected by community-wide GPCR Dock 2010 assessment. *Structure* **2011**, *19*, 1108−1126.

(6) Kołaczkowski, M.; Nowak, M.; Pawłowski, M.; Bojarski, A. J. Receptor-based pharmacophores for serotonin 5-HT7R antagonists-implications to selectivity. *J. Med. Chem.* **2006**, *49*, 6732−6741.

(7) Nowak, M.; Kołaczkowski, M.; Pawłowski, M.; Bojarski, A. J. Homology modeling of the serotonin 5-HT1A receptor using automated docking of bioactive compounds with defined geometry. *J. Med. Chem.* **2006**, *49*, 205−214.

(8) Evers, A.; Gohlke, H.; Klebe, G. Ligand-supported homology modelling of protein binding-sites using knowledge-based potentials. *J. Mol. Biol.* **2003**, *334*, 327−345.

(9) 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 D3 receptor homology model and crystal structure. *Nat. Chem. Biol.* **2011**, *7*, 769−778.

(10) Sherman, W.; Beard, H. S.; Farid, R. Use of an induced fit receptor structure in virtual screening. *Chem. Biol. Drug Des.* **2006**, *67*, 83−84.

(11) Sherman, W.; Day, T.; Jacobson, M. P.; Friesner, R. A.; Farid, R. Novel procedure for modeling ligand/receptor induced fit effects. *J. Med. Chem.* **2006**, *49*, 534−553.

(12) 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.

(13) Kowalski, P.; Jaśkowska, J.; Bojarski, A. J.; Duszyńska, B.; Bucki, A.; Kołaczkowski, M. Evaluation of 1-arylpiperazine derivative of hydroxybenzamides as 5-HT 1A and 5-HT7 serotonin receptor ligands: An experimental and molecular modeling approach. *J. Heterocyclic Chem.* **2011**, *48*, 192−198.

(14) Newman-Tancredi, A.; Kleven, M. S. Comparative pharmacology of antipsychotics possessing combined dopamine D2 and serotonin 5-HT1A receptor properties. *Psychopharmacology (Berlin)* **2011**, *216*, 451−473.

(15) Rieckmann, A.; Karlsson, S.; Karlsson, P.; Brehmer, Y.; Fischer, H.; Farde, L.; Nyberg, L.; Bäckman, L. Dopamine D1 Receptor Associations Within and Between Dopaminergic Pathways in Younger and Elderly Adults: Links to Cognitive Performance. *Cereb. Cortex* **2011**, *21*, 2023−2032.

(16) Goldman-Rakic, P. S.; Castner, S. A.; Svensson, T. H.; Siever, L. J.; Williams, G. V. Targeting the dopamine D1 receptor in schizophrenia: insights for cognitive dysfunction. *Psychopharmacology (Berlin)* **2004**, *174*, 3−16.

(17) Congreve, M.; Langmead, C.; Marshall, F. H. The use of GPCR structures in drug design. *Adv. Pharmacol.* **2011**, *62*, 1−36.

(18) Cherezov, V.; Rosenbaum, D. M.; Hanson, M. A.; Rasmussen, S. G. F.; 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.

(19) Chien, E. Y. T.; 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* **2010**, *330*, 1091−1095.

(20) Kurowski, M. A.; Bujnicki, J. M. GeneSilico protein structure prediction meta-server. *Nucleic Acids Res.* **2003**, *31*, 3305−3307.

(21) Dolan, M. A.; Noah, J. W.; Hurt, D. Comparison of common homology modeling algorithms: application of user-defined alignments. *Methods Mol. Biol.* **2012**, *857*, 399−414.

(22) Arnold, K.; Bordoli, L.; Kopp, J.; Schwede, T. The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. *Bioinformatics* **2006**, *22*, 195−201.

(23) Guex, N.; Peitsch, M. C. SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. *Electrophoresis* **1997**, *18*, 2714−2723.

(24) Kiefer, F.; Arnold, K.; Künzli, M.; Bordoli, L.; Schwede, T. The SWISS-MODEL Repository and associated resources. *Nucleic Acids Res.* **2009**, *37*, D387−392.

(25) Peitsch, M. C. Protein modeling by E-mail. *Nat. Biotechnol.* **1995**, *13*, 658−660.

(26) Schwede, T.; Kopp, J.; Guex, N.; Peitsch, M. C. SWISS-MODEL: An automated protein homology-modeling server. *Nucleic Acids Re.s* **2003**, *31*, 3381−3385.

(27) Jacobson, M. P.; Friesner, R. A.; Xiang, Z.; Honig, B. On the role of the crystal environment in determining protein side-chain conformations. *J. Mol. Biol.* **2002**, *320*, 597−608.

(28) Jacobson, M. P.; Pincus, D. L.; Rapp, C. S.; Day, T. J. F.; Honig, B.; Shaw, D. E.; Friesner, R. A. A hierarchical approach to all-atom protein loop prediction. *Proteins* **2004**, *55*, 351−367.

(29) 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.

(30) Friesner, R. A.; Murphy, R. B.; Repasky, M. P.; Frye, L. L.; Greenwood, J. R.; Halgren, T. A.; Sanschagrin, P. C.; Mainz, D. T. Extra precision glide: docking and scoring incorporating a model of hydrophobic enclosure for protein-ligand complexes. *J. Med. Chem.* **2006**, *49*, 6177−6196.

(31) Roth, B.; Lopez, E.; Patel, S.; Kroeze, W. K. The multiplicity of serotonin receptors: Uselessly diverse molecules or an embarrassment of riches. *Neuroscientist* **2000**, *6*, 252−261.

(32) Allen, F. H. The Cambridge Structural Database: a quarter of a million crystal structures and rising. *Acta Crystallogr., B* **2002**, *58*, 380−388.
(33) Strader, C. D.; Sigal, I. S.; Candelore, M. R.; Rands, E.; Hill, W. S.; Dixon, R. A. Conserved aspartic acid residues 79 and 113 of the beta-adrenergic receptor have different roles in receptor function. *J. Biol. Chem.* **1988**, *263*, 10267−10271.
(34) Eswar, N.; Webb, B.; Marti-Renom, M. A.; Madhusudhan, M. S.; Eramian, D.; Shen, M.-Y.; Pieper, U.; Sali, A. Comparative protein structure modeling using Modeller. *Current Protocols in Bioinformatics*; Wiley: New York, **2006**; Chapter 5, Unit 5.6.
(35) Fiser, A.; Do, R. K.; Sali, A. Modeling of loops in protein structures. *Protein Sci.* **2000**, *9*, 1753−1773.
(36) Sali, A.; Blundell, T. L. Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol.* **1993**, *234*, 779−815.
(37) 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.
(38) Dixon, S. L.; Smondyrev, A. M.; Knoll, E. H.; Rao, S. N.; Shaw, D. E.; Friesner, R. A. PHASE: a new engine for pharmacophore perception, 3D QSAR model development, and 3D database screening: 1. Methodology and preliminary results. *J. Comput. Aided Mol. Des.* **2006**, *20*, 647−671.
(39) Dixon, S. L.; Smondyrev, A. M.; Rao, S. N. PHASE: a novel approach to pharmacophore modeling and 3D database searching. *Chem. Biol. Drug Des.* **2006**, *67*, 370−372.
(40) Halgren, T. A.; Murphy, R. B.; Friesner, R. A.; Beard, H. S.; Frye, L. L.; Pollard, W. T.; Banks, J. L. Glide: A New Approach for Rapid, Accurate Docking and Scoring. 2. Enrichment Factors in Database Screening. *J. Med. Chem.* **2004**, *47*, 1750−1759.
(41) Mysinger, M. M.; Carchia, M.; Irwin, J. J.; Shoichet, B. K. Directory of useful decoys, enhanced (DUD-E): better ligands and decoys for better benchmarking. *J. Med. Chem.* **2012**, *55*, 6582−6594.
(42) Truchon, J.-F.; Bayly, C. I. Evaluating virtual screening methods: good and bad metrics for the "early recognition" problem. *J. Chem. Inf. Model.* **2007**, *47*, 488−508.
(43) Cavasotto, C. N.; Orry, A. J. W.; 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. 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.