JOURNAL OF CHEMICAL INFORMATION AND MODELING

Article

pubs.acs.org/jcim

Structure-Based Virtual Screening of MT_2 Melatonin Receptor: Influence of Template Choice and Structural Refinement

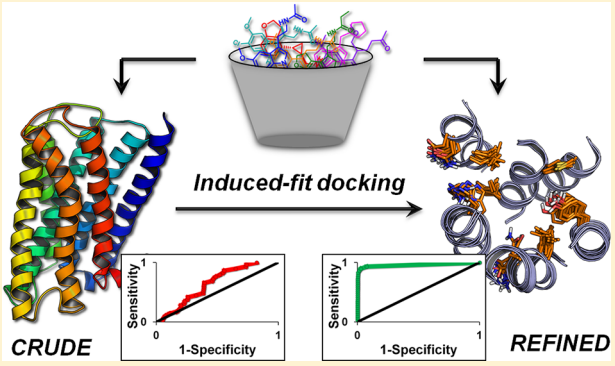
Daniele Pala,[†] Thijs Beuming,*[,‡] Woody Sherman,[‡] Alessio Lodola,[†] Silvia Rivara,*[,†] and Marco Mor[†]

[†]Dipartimento di Farmacia, Università degli Studi di Parma, Parco Area delle Scienze 27/A, I-43124 Parma, Italy

[‡]Schrödinger, Inc., 120 West 45th Street, New York, New York 10036, United States

Supporting Information

ABSTRACT: Developing GPCR homology models for structure-based virtual screening requires the choice of a suitable template and refinement of binding site residues. We explored this systematically for the MT_2 melatonin receptor, with the aim to build a receptor homology model that is optimized for the enrichment of active melatoninergic ligands. A set of 12 MT_2 melatonin receptor models was built using different GPCR X-ray structural templates and submitted to a virtual screening campaign on a set of compounds composed of 29 known melatonin receptor ligands and 2560 drug-like decoys. To evaluate the effect of including a priori information in receptor models, 12 representative melatonin receptor ligands were placed into the MT_2 receptor models in poses consistent with known mutagenesis data and with assessed pharmacophore models. The receptor structures were then adapted to the ligands by induced-fit docking. Most of the 144 ligand-adapted MT_2 receptor models showed significant improvements in screening enrichments compared to the unrefined homology models, with some template/refinement combinations giving excellent enrichment factors. The discriminating ability of the models was further tested on the 29 active ligands plus a set of 21 inactive or low-affinity compounds from the same chemical classes. Rotameric states of side chains for some residues, presumed to be involved in the binding process, were correlated with screening effectiveness, suggesting the existence of specific receptor conformations able to recognize active compounds. The top MT_2 receptor model was able to identify 24 of 29 active ligands among the first 2% of the screened database. This work provides insights into the use of refined GPCR homology models for virtual screening.



INTRODUCTION

Melatonin (*N*-acetyl-5-methoxytryptamine, Figure 1) is a tryptophan-derived hormone mainly secreted by the pineal gland following a circadian rhythm. Besides its well-established functions in the regulation of circadian rhythms,[1,2] melatonin has effects in many other physiological processes, such as immune system regulation, blood pressure control, bone formation, and modulation of neuronal firing.[3,4] A growing body of evidence has shown additional roles of melatonin in antioxidant, antitumor, and neuroprotectant processes[3,4] as well as the implication of melatonin signaling pathways in the events leading to the development of type-2 diabetes.[5] For this reason, melatonin administration has been proposed for the treatment of a variety of pathological conditions, such as sleep disturbances, cancer, neurodegenerative diseases, hypertension, depression, epilepsy, and headache.[6] Moreover, melatonin receptor ligands, which are either marketed or currently being tested in clinical trials, have been shown to be efficient in the treatment of circadian rhythm sleep disorders, jet lag, and major depression (Figure 1).[7]

In mammals, melatonin is the endogenous agonist of two G protein-coupled receptors (GPCRs), named MT_1 and MT_2,[8] exhibiting binding affinities in the nanomolar range. In addition to these two class A GPCRs, melatonin binds with lower affinity to the so-called MT_3 binding site, characterized as the hamster homologue of the human enzyme quinone reductase 2.[9] In humans, MT_1 and MT_2 receptors are mainly expressed in different areas of the central nervous system, such as the suprachiasmatic nucleus, hippocampus, nucleus accumbens, and substantia nigra. However, melatonin receptors have also been found in different peripheral tissues, such as retina, immune cells, coronary arteries, and epithelial cells.[3]

Thus far, experimental studies have shown that activation of the MT_1 receptor produces vasoconstriction, inhibits hormone secretion, and reduces neuronal firing of the suprachiasmatic nucleus, whereas activation of the MT_2 receptor induces splenocyte proliferation, vasodilatation, inhibits leukocyte rolling, and regulates oxytocin receptor gene expression.[8] However, despite the number of studies aimed at clarifying the patho/physiological role of melatonin receptor subtypes and the number of melatoninergic ligands reported in the literature and in patents,[7] the differential role of both MT_1 and MT_2 receptors has yet to be fully elucidated. To this aim and to

Received: January 8, 2013
Published: March 29, 2013

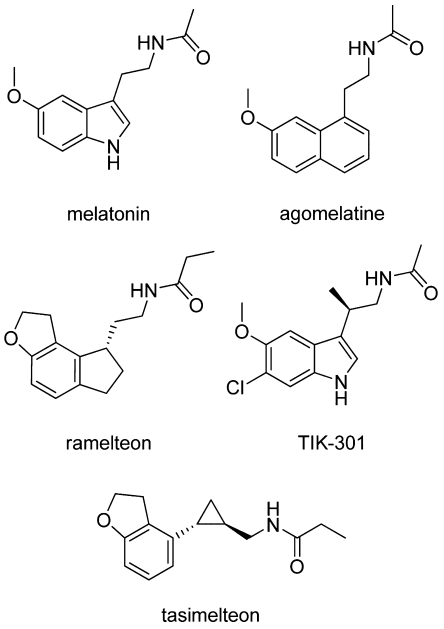


Figure 1. Melatonin and melatonin receptor agonists approved or in clinical trials.

develop novel therapeutic agents, there is intense research activity seeking novel subtype-selective melatoninergic compounds.

Since their cloning in the mid-1990s, a number of different MT_1 and MT_2 melatonin receptor three-dimensional (3D) models have been published.[10−22] Most of these receptor models were built starting from the structure of bovine rhodopsin, which was the only structural template available for GPCR modeling until 2007. Only recently have novel homology models of melatonin receptors built on the newly released X-ray structures of druggable GPCRs appeared in the literature.[23] However, none of these 3D models have been investigated by means of virtual screening (VS) campaigns, aimed at testing the ability of homology models to retrieve known active compounds or to identify new promising candidates in a data set of drug-like compounds, and melatoninergic ligands discovered so far had mostly been designed by ligand-based approaches.

Prediction of the 3D structures of melatonin receptors can be challenging because all available GPCR templates show a limited sequence identity (sID), i.e., lower than 30%, to the target MT_1 or MT_2 receptors within the transmembrane (TM) domains. Indeed, recent studies performed on GPCR homology models clearly showed that the reliability of a model structure is strongly related to the degree of sID between the template and the target sequence.[24−27] Particularly, while the availability of an experimentally solved GPCR template showing a sID > 30−35% to the target enabled the building of reliable receptor models, "remote-template" homology modeling procedures (i.e., when the sID between template and receptor decreases under 30%) often failed to produce accurate receptor models.[24,25,28] A further element that makes melatonin receptor modeling challenging is the lack of well-characterized ligand−receptor contacts. Indeed, site-directed mutagenesis studies performed on both MT_1 and MT_2 melatonin receptors do not provide a definite binding scheme for their lipophilic natural ligand. In addition, melatonin receptors belong to a branch in the phylogenetic tree of class A GPCRs distinct from those containing receptors for biogenic amines.[29] As a consequence, both melatonin receptors lack the pattern of key binding site residues that are known to anchor the polar substituents of endogenous amines (e.g., the conserved aspartic residue located in TM3).

Because modeling of melatonin receptors requires the use of remote structural templates, available ligand-based information, such as SARs and pharmacophore models, should be used as additional guidance to increase the quality of models. The inclusion of prior knowledge in the refinement of GPCRs homology models has been shown to be crucial to overcome the limitations due to the lack of a suitable template structure.[24,27] Indeed, the GPCR X-ray crystal structures of about 20 unique receptors deposited in the Protein Data Bank at the time of writing cover only the ∼2% of nearly 800 GPCRs identified in the human genome,[30] providing closely related templates only for a small subset of receptors. However, a significant number of therapeutically relevant GPCRs show moderate to low sID percentages with available X-ray crystal structures, making homology modeling a necessary option. In these cases, the exploitation of prior information (e.g., site-directed mutagenesis studies and biophysical data) during receptor model refinement can remarkably improve the ability to recognize active compounds or to reproduce known SARs.[27,31−34]

To evaluate the influence of template selection and of knowledge-driven structural refinement on the reliability of MT_2 receptor models, we developed a homology modeling approach in which the ability of different MT_2 receptor models to retrieve known active compounds among a set of drug-like decoys has been evaluated by means of VS campaigns (Figure 2). Initially, different MT_2 receptor structures were built

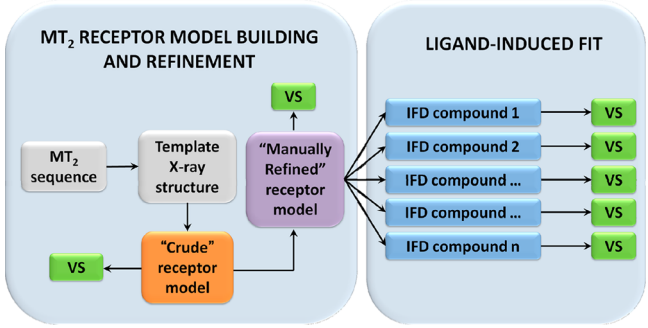


Figure 2. Schematic representation of the workflow applied to each template-based MT_2 receptor model. The initial "crude" MT_2 receptor structure (orange block) was submitted to an initial refinement stage conducted on binding site residues in the absence of a ligand in the binding site (purple block). Explorative VS calculations were performed on both "crude" and "manually refined" MT_2 receptor structures (green blocks). Induced-fit docking[35] (IFD) calculations (blue blocks) were performed on the "manually refined" MT_2 receptor structure with a series of potent representative melatoninergic ligands, which were used to adapt the binding site region of each MT_2 receptor structure prior to VS campaigns.

starting from available GPCR X-ray structures in either their active or inactive state and subsequently submitted to a refinement procedure aimed at optimizing the orientation of binding site residues, independent of any ligands in the binding site. Both "crude" and "manually refined" MT_2 receptor models were then submitted to VS campaigns using a subset of drug-like decoys and a small subset of active melatoninergic ligands.

dx.doi.org/10.1021/ci4000147 | J. Chem. Inf. Model. 2013, 53, 821−835

Table 1. Enrichment Factor Values Calculated in the First 2% of Database ($EF^{2\%}$) for "Crude" and "Manually Refined" MT_2 Receptor Structures

template	A_{2A}	A_{2AACT}[a]	β_2	β_{2ACT}[a]	β_1	β_{1ACT}[a]	D3	H_1	Rho	Rho_{ACT}[a]	CXCR4	MULTI[b]
crude	1.7[c]	1.7[c]	1.7[c]	3.4[c]	1.7[c]	0	0	0	0	0	0	5.2[c]
manually refined	6.9[c]	0	0	0	0	0	3.4[c]	0	0	0	5.2	1.7[c]

[a]When the template is either in its active conformation or is bound to an agonist, it is labeled as "ACT". [b]MULTI indicates a consensus model with the inactive GPCR crystal structures used as templates. [c]Active melatoninergic ligands were only docked outside the putative orthosteric binding site.

Moreover, to evaluate the effect of a ligand-guided adaptation of the binding site cavity, a subset of chemically diverse melatoninergic ligands were used as "shape inducers" to further refine the binding pocket before VS calculations. The ability of the resulting ligand-adapted MT_2 receptor models to recognize active compounds among a set of drug-like decoys, as well as to discriminate between known active and known inactive melatoninergic ligands, was evaluated by means of VS calculations.

RESULTS AND DISCUSSION

Initially, 11 different MT_2 receptor models were built starting from a wide range of available GPCR X-ray structures, including the β_1 and β_2 adrenergic, A_{2A} adenosine, D3 dopamine, H_1 histamine, and CXCR4 chemokine receptors, and rhodopsin. For those receptors that had been solved in different activation states, i.e., rhodopsin, β_1 and β_2 adrenoceptors, and A_{2A} adenosine receptor, the crystallographic structures of both states were used for homology modeling. An additional "consensus" MT_2 receptor model was built, starting from all available GPCR X-ray structures crystallized in an inactive state. In fact, recent studies performed on "remote-template" GPCR homology models highlighted the importance of incorporating different templates during model building to obtain reliable 3D receptor structures.[25,32,36] The program Prime[37] was used to build the starting 3D structure of the 12 MT_2 receptor models. These "crude" models were submitted to visual inspection, and side chains of residues thought to be important for ligand binding were manually rotated toward the binding site pocket (purple block in Figure 2). The location of the putative binding cavity was inferred from site-directed mutagenesis experiments, which had shown that a subset of residues located between TM4 and TM7, including $Asn175^{4.60}$, $His208^{5.46}$, $Asn268^{6.52}$, and $Tyr298^{7.43}$ (superscripts refer to Ballesteros–Weinstein numbering),[38] were involved in melatonin binding.[16,22,39] The putative binding region of the MT_2 melatonin receptor outlined by these studies corresponds well with the prototypic binding site cavity observed in structures of several class A GPCRs, which is lined by TM3, TM4, TM5, TM6, and TM7 and located in the extracellular portion of the helical bundle.[40]

The 12 "crude" and the 12 "manually refined" (i.e., those 12 receptor models in which side chains of residues important for ligand binding were directed toward the binding site cavity) MT_2 receptor models were submitted to energy minimization (with Prime) and to VS campaigns (with Glide[41,42]) using a data set composed of drug-like decoys enriched with a subset of active melatoninergic compounds. A set of 29 active MT_2 receptor ligands belonging to 12 different chemical classes was collected from the literature (Table S1, Supporting Information).[43−58] The compounds included both agonists and antagonists, covering most of the chemical scaffolds so far identified for active melatoninergic ligands.[7,59] The selected chemotypes are characterized by diverse chemical structures, yet retaining the key pharmacophoric features of the melatonin receptor ligands, i.e., an aromatic nucleus, usually carrying a methoxy group, which is connected to an amide group through a short linker. The drug-like decoys were extracted from the ChemBridge data set and submitted to a filtering procedure to select decoys similar to the active melatoninergic ligands in terms of physicochemical properties (see Experimental section and Figure S1, Supporting Information), in order to avoid artificial enrichments.[60] The final database used in VS calculations was composed of 29 active melatoninergic compounds and 2560 decoys.

VS calculations performed on "crude" and "manually refined" MT_2 receptor models gave low enrichment factors (Table 1). A visual inspection of docking results showed that, in most cases, both actives and decoys were docked outside the putative MT_2 binding pocket. In fact, side chains of residues delimiting the putative binding cavity prevented the accommodation of ligands in almost all initial models. The side chains blocking the binding site were often associated with differences in amino acid residues between the template and target at these positions. While these differences often reflect the propensity for different targets to recognize chemically distinct ligands, the lack of an identical template residue makes the placement difficult. This is the case of the MT_2 melatonin receptor, which shares limited sID percentages (19−29%) with available GPCRs within the TM domains and even lower in the putative binding site region (12−24%). The MT_2 receptor models described so far were often characterized by closely packed binding sites, which would sterically exclude the majority of ligands. Two exceptions to this trend were the models built from the CXCR4 receptor, in which all the active melatoninergic compounds were docked into the helical bundle, in the proximity of the putative binding site region. However, while in the initial CXCR4-based MT_2 receptor model none of the actives was retrieved in the first 2% of the database (i.e., among the 52 best-scored compounds), in the "manually refined" structure, three active ligands were ranked in the top 2% of the screened database, yielding an enrichment factor ($EF^{2\%}$) of 5.2. The ability of these CXCR4-based homology models to accommodate a small number of active compounds into the putative binding cavity is mainly due to the peculiar geometry and position of the binding cavity in the CXCR4 crystal structure. Indeed, compared to other GPCR structures, the binding pocket of the CXCR4 receptor is more open and located closer to the extracellular surface. Moreover, the extracellular loop (ECL) 2 of the CXCR4 structure is bent toward the tip of TM3 and ECL1,[61] favoring the accommodation of both active melatoninergic ligands and decoy molecules into the solvent-exposed binding cavity (Figure S2, Supporting Information). The ability of these models to accept the docking of different compounds in a large binding cavity has no relation with specific recognition of melatoninergic ligands.

dx.doi.org/10.1021/ci4000147 | J. Chem. Inf. Model. 2013, 53, 821−835

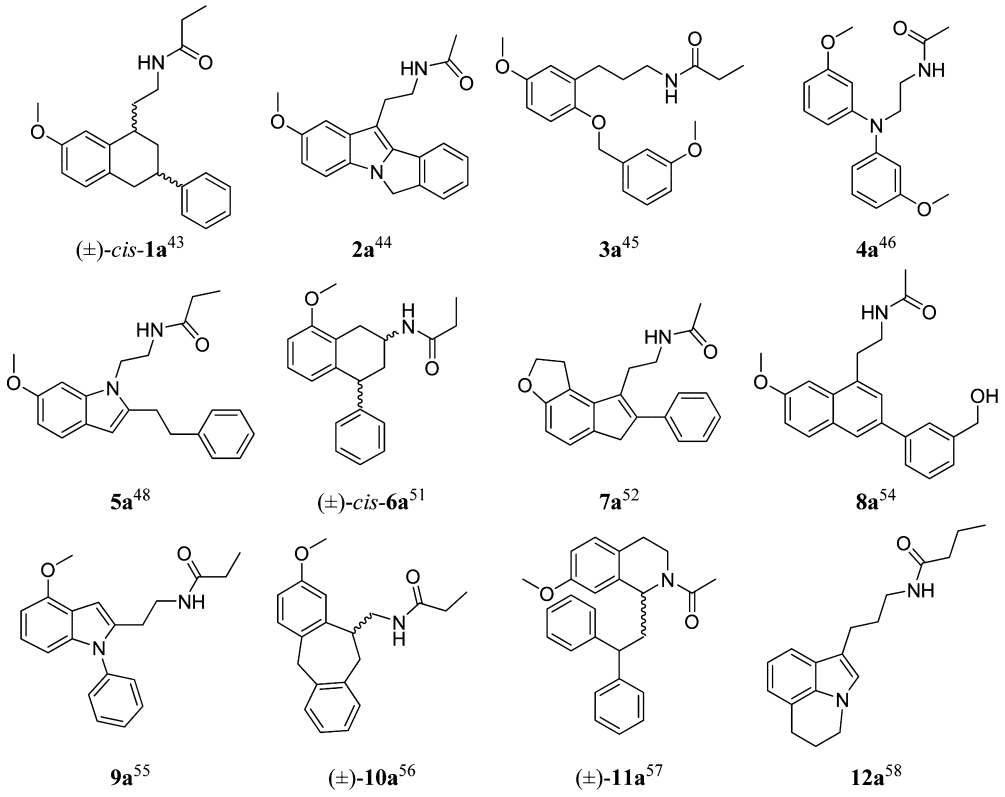


Figure 3. Twelve representative melatoninergic ligands selected for Induced-Fit Docking (IFD) calculations.

Table 2. $EF^{2\%}$ Values for MT_2 Receptor Models Built on Different Templates and Adapted around Representative Ligands 1a–12a through IFD Calculations

shape inducer		1a	2a	3a	4a	5a	6a	7a	8a	9a	10a	11a	12a	Mean
template	A_{2A}	20.6	25.8	18.9	25.8	18.9	20.6	18.9	22.3	6.9	3.4	8.6	6.9	16.5
	A_{2AACT}[a]	8.6	24.0	10.3	15.5	6.9	1.7	8.6	30.9	6.9	10.3	5.2	3.4	11.0
	β_2	15.5	18.9	24.0	13.7	22.3	13.7	12.0	8.6	17.2	20.6	1.7	1.7	14.2
	β_{2ACT}[a]	41.2	22.3	30.9	10.3	10.3	36.1	22.3	15.5	24.0	3.4	3.4	1.7	18.5
	β_1	29.2	17.2	10.3	32.6	20.6	13.7	18.9	13.7	13.7	13.7	8.6	1.7	16.2
	β_{1ACT}[a]	24.0	15.5	13.7	20.6	17.2	10.3	15.5	5.2	5.2	8.6	1.7	1.7	11.6
	D3	15.5	18.9	25.8	22.3	13.7	17.2	12.0	25.8	12.0	8.6	5.2	1.7	14.9
	H_1	18.9	12.0	8.6	8.6	17.2	13.7	25.8	15.5	12.0	6.9	1.7	3.4	12.0
	Rho	17.2	18.9	6.9	8.6	24.0	20.6	10.3	6.9	20.6	5.2	0.0	10.3	12.5
	Rho_{ACT}[a]	13.7	5.2	8.6	3.4	8.6	18.9	22.3	12.0	13.7	3.4	8.6	1.7	10.0
	CXCR4	0.0	5.2	3.4	12.0	10.3	6.9	1.7	10.3	8.6	5.2	12.0	3.4	6.6
	MULTI[b]	10.3	13.7	25.8	10.3	13.7	10.3	12.0	15.5	13.7	3.4	1.7	3.4	11.2
	Mean	17.9	16.5	15.6	15.3	15.3	15.3	15.0	15.2	12.9	7.7	4.9	3.4	

[a]When the template is either in its active conformation or is bound to an agonist, it is labeled as "ACT". [b]MULTI indicates a consensus model with the inactive GPCR crystal structures used as templates.

Because docking to the initial homology models produced poor poses and enrichment of the active melatoninergic ligands into the MT_2 receptor binding cavities, an alternative approach was applied employing potent ligands to reshape the binding site before docking calculations. In particular, the most potent melatoninergic ligand for each of the 12 chemical classes was selected (Figure 3) and docked into the 12 MT_2 receptor models. The Induced-Fit Docking (IFD) protocol was applied, allowing mutual adaptation of the putative binding region and the ligand. This process gave 144 different models employed for the following docking studies.

Top poses from standard IFD protocol with default settings were unsatisfactory, as determined by comparing to experimentally deduced interactions. Either poses would dock outside the putative MT_2 binding site, or would lack interactions with residues important for ligand binding, or would assume conformations not consistent with pharmacophore models previously developed for melatonin receptor ligands.[49,51,62] A new protocol was therefore established by explicitly taking into account information from mutagenesis studies and pharmacophore models in docking ligands into the binding site. To this end, a pharmacophore model was built superposing the structures of the 12 melatoninergic ligands

dx.doi.org/10.1021/ci4000147 | J. Chem. Inf. Model. 2013, 53, 821–835

represented in Figure 3 on the conformation of the most rigid compound, **6a**, reproducing the active conformation of melatonin as previously defined (Figure S3, Supporting Information).[62] Conformational analysis on the 12 melatonergic ligands was then performed using a mixed torsional/low-mode sampling approach (MCMM/LMOD, see Experimental section for details) and for each compound three diverse conformations giving good fitting to the pharmacophoric elements were selected, resulting in 36 conformations. Each conformation was manually placed into the 12 MT_2 receptor models using information retrieved from mutagenesis studies as guidance for initial ligand placement. In particular, when present, the methoxy group of the ligand was placed in proximity of Asn175$^{4.60}$, and the amide group was placed close to Tyr298$^{7.43}$. Indeed, site-directed mutagenesis had shown that Asn175$^{4.60}$ is likely to be involved in the stabilization of the 5-methoxy group of melatonin[39] and that Tyr298$^{7.43}$ plays a crucial role in ligand stabilization.[16] The ensemble of 432 (36 × 12) starting complexes was submitted to the structure refinement procedure implemented in the default IFD workflow, where the side chains of binding site residues are adapted around the docked ligand. The refined 432 MT_2 receptor complexes were then ranked according to a composite customized knowledge-based scoring function (IFD_{KB}, eq 1 in Experimental section), considering both the energy of ligand–receptor complex and its consistency with mutagenesis studies and pharmacophore models. Then, for each of the 12 template-based MT_2 receptor models, and for each representative ligand, the complex showing the best IFD_{KB} score was selected, which resulted in the set of 144 ligand-adapted MT_2 receptor structures with experimentally consistent ligand binding modes that could be employed for virtual screening.

These receptor models showed significant increases in VS performance, compared to both initial "crude" and "manually refined" MT_2 receptor models (Tables 1 and 2). However, screening performance strongly depended on the choice of both the template used for homology modeling and the ligand used for knowledge-driven binding site refinement. Table 2 reports the $EF^{2\%}$ values for the 144 ligand-adapted MT_2 receptor models. $EF^{1\%}$, $EF^{5\%}$, and $EF^{10\%}$ values, as well as other performance descriptors, are reported in Table S2 of the Supporting Information. Because high correlations between $EF^{2\%}$ values and other enrichment metrics (e.g., the Boltzmann-enhanced discrimination of receiver operating characteristic (BEDROC)[63] or the robust initial enrichment (RIE)[64]) were found (R^2= 0.93, 0.92, and 0.93 for $EF^{2\%}$ and BEDROC (α = 20.0), BEDROC (α = 160.9), and RIE, respectively), we decided to report and discuss the commonly used $EF^{2\%}$ metric. The row and column average values in Table 2 suggest that the influence of ligand-induced fit is larger than that of the choice of receptor template. In fact, Kruskal–Wallis analysis,[65] a non-parametric test based on the ranking of medians for non-normally distributed samples, indicates a significative difference among columns (different shape-inducing compounds) of Table 2 (χ^2 = 52.91, df = 11, $p = 1.86 \times 10^{-7}$) and a non-significant difference among rows (receptor templates: χ^2 = 19.21, df =11, p = 0.057). The distribution of $EF^{2\%}$ values among columns is strongly biased by the presence of two subsets of models, corresponding to those adapted around **11a** and **12a**, which clearly show the worst performances. When these two columns are removed, Kruskal–Wallis analysis does not detect a significant difference among columns anymore (χ^2 = 15.31, df =9, p = 0.083). On the contrary, in the reduced data table a significant difference among rows (templates) has emerged (χ^2 = 29.56, df =11, p = 0.002), mainly due to the limited $EF^{2\%}$ values showed by the CXCR4-based MT_2 receptor models. In fact, when the corresponding row was removed, Kruskal–Wallis analysis could not detect significant differences among the remaining 11 samples built from other templates (χ^2 = 16.43, df =10, p = 0.088).

The limited VS performances observed for some MT_2 receptor models could be explained from chemical or biological points of view. Compared to other compounds, the aromatic core of **12a** bearing the amide chain does not carry an additional aromatic substituent, which usually positively affects ligand potency (Figure 3).[45,46,53,66] Conversely, compound **11a** is characterized by an additional aromatic group not present in the other active compounds. The effects of these two compounds on reshaping of the binding cavity should therefore be rather different from those induced by other ligands. As for templates, CXCR4 represents a peculiar case, in which the ligand binding site is located in a solvent accessible region near the ECL2, whereas in the aminergic or purinergic class A GPCRs it is located deeper into the TM region.

It should be stressed, however, that performance differences within groups (either columns or rows in Table 2) were often larger compared to differences among group averages. Consequently, the lack of significant difference in non-parametric tests could be due to high within-group variability rather than to a real lack of differences among the models. In particular, a few combinations of templates and ligands gave outstanding enrichments, higher than expected from the sum of statistical contributions of row and column means. For example, looking at Table 2, it can be seen that a subset of MT_2 receptor models showed $EF^{2\%}$ values higher than 30, comparable or even better than those reported in other studies on GPCR homology models.[32,67−71] Most notably, the MT_2 receptor model built from the β_2 adrenergic receptor in its active conformation refined with **1a** correctly recognizes 24 active compounds in the first 2% of the ranked data set (i.e., in the first 52 best-ranked compounds according to the SP DockingScore), yielding the highest $EF^{2\%}$ value (41.2) observed among the 144 ligand-adapted receptor models. The accommodation of the partial agonist **1a** into the binding site of this model is shown in Figure 4.

Docked poses of active ligands into this model shared a common orientation in the receptor binding site (Figure S4,

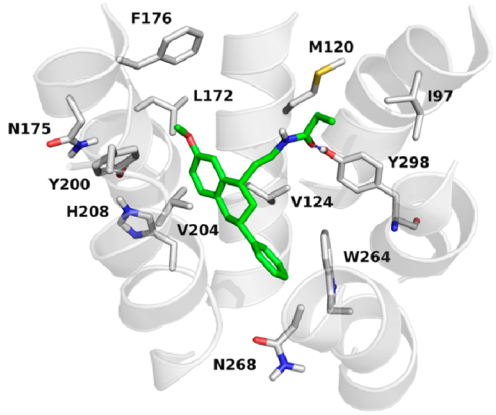


Figure 4. Best-ranked pose of **1a** into the β_{2ACT}-based MT_2 receptor model according to the IFD_{KB} score.

dx.doi.org/10.1021/ci4000147 | J. Chem. Inf. Model. 2013, 53, 821−835

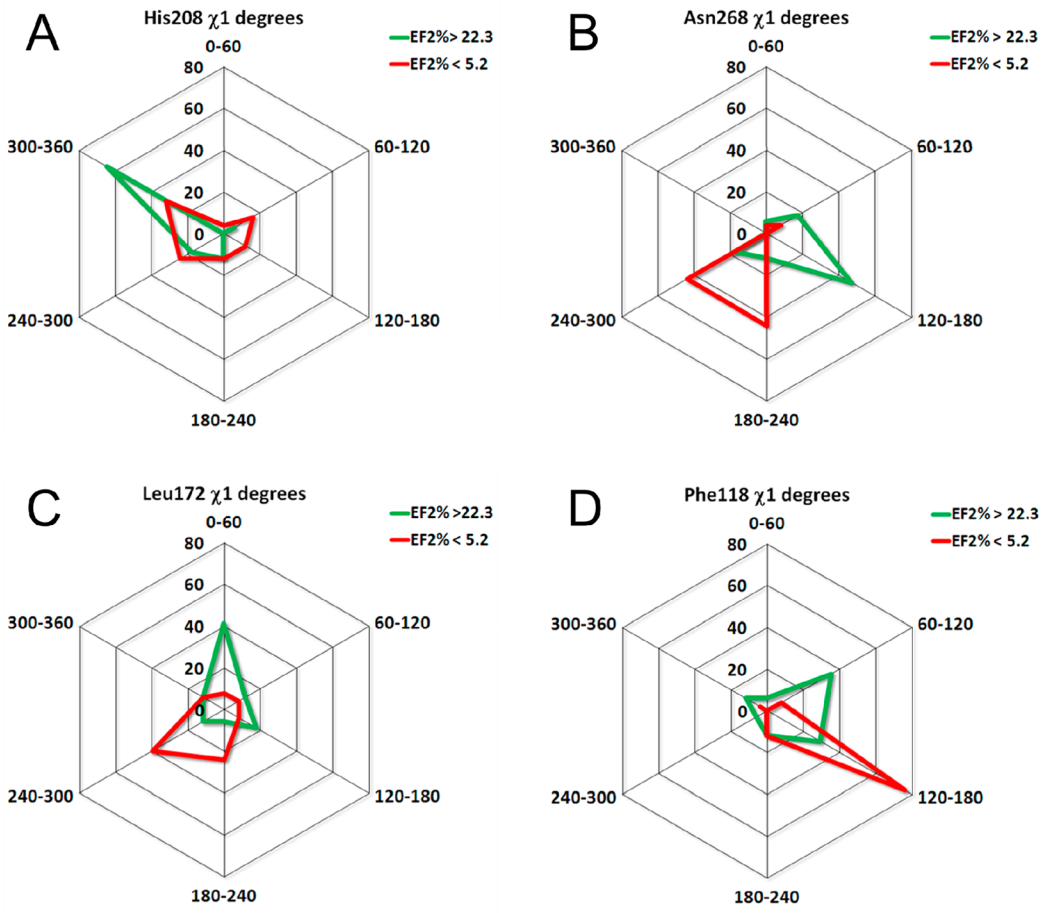


Figure 5. Radial histograms showing the distribution of the $\chi 1$ dihedral angles for (A) His208$^{5.46}$, (B) Asn268$^{6.52}$, (C) Leu172$^{4.57}$, and (D) Phe118$^{3.30}$. The range of $\chi 1$ was divided in bins of 60 degrees. The percentage of selected structures showing $\chi 1$ values within each bin are reported along the radial axis. Only MT_2 receptor models with high (green lines, for models showing $EF^{2\%}$ values greater than 22.3) and low (red lines, for models showing $EF^{2\%}$ values lower than 5.2) VS efficiency were considered.

Supporting Information), located within the helical bundle lined by TM3, TM4, TM5, TM6, and TM7. With few exceptions, the methoxy group (or a bioisoster) occupied a cavity formed by Gly121$^{3.33}$, Leu172$^{4.57}$, Phe176$^{4.61}$, and Tyr200$^{5.38}$ in the proximity of the extracellular terminus of TM4. The amide oxygen was often H-bonded to Tyr298$^{7.43}$ located on TM7, while the aromatic rings of melatoninergic compounds were sandwiched between residues located on TM3 and TM5. The additional aromatic group present in a subset of melatoninergic ligands (i.e., the 3-phenyl ring of **1a** and **8a**, the N-phenyl ring of **4a** and **9a**, the fused benzene rings of **2a** and **10a**, the phenyl-alkyl(oxy) substituents of **3a** and **5a**, and the 4-phenyl substituent of **6a**) was accommodated in the proximity of TM6 and TM7, in a pocket delimited by Val128$^{3.40}$, Phe209$^{5.47}$, Trp264$^{6.48}$, Leu267$^{6.51}$, and Asn268$^{6.52}$. The alkyl chain carrying the amide group formed extensive hydrophobic contacts with neighboring residues located on TM3 and on the tips of TM2 and TM7, including Ile97$^{2.61}$, Met120$^{3.32}$, and Tyr294$^{7.39}$.

Although the use of multiple templates has been shown to increase the accuracy of GPCR homology models, as well as their performances in docking studies,[25,32,36] the MT_2 receptor models based on a multiple template approach showed screening efficiencies comparable to those obtained for models built using single templates. This could be due to the limited sID percentage between the MT_2 receptor and the crystallized GPCRs used for homology modeling. Indeed, it is likely that the significant phylogenetic distance between available template structures and melatonin receptors mirrors remarkable differences also in secondary and tertiary structure elements, such as shifts and kinks of TM domains, as well as the geometry and the secondary structure content of loop regions. In this scenario, the combination of multiple templates, which all greatly differ from the target sequence in terms of amino acid composition and, likely, in 3D structural features, might not be sufficient to improve the accuracy of the MT_2 receptor model compared to a single template approach.

In order to find structural clues able to explain the different enrichment efficiencies observed among the 144 ligand-adapted MT_2 models, the conformations assumed by binding site residues were investigated in detail. A root-mean-square fluctuation value (RMSF) of side chain heavy atoms was calculated by aligning each ligand-adapted MT_2 receptor model on the average structure built from the ensemble of 144 MT_2 receptor models (see Figure S5, Supporting Information and Experimental section for details). RMSF values indicate that a subset of residues delimiting the binding site region is characterized by high fluctuations and, consequently, assumed different rotameric states among different receptor models. We were able to identify a set of residues for which rotameric preferences could be correlated with enrichment efficiencies. Figure 5 shows radial histograms that highlight some of the

dx.doi.org/10.1021/ci4000147 | J. Chem. Inf. Model. 2013, 53, 821−835

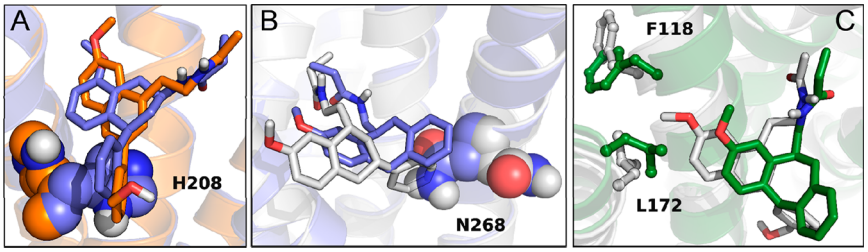


Figure 6. (A) Superposition of the A_{2AACT}-based MT_2 receptor models adapted with **6a** (violet carbons, His208$^{5.46}$ $\chi1$ = 200°) and **8a** (orange carbons, His208$^{5.46}$ $\chi1$ = 313°). His208$^{5.46}$ is shown as spheres, whereas ligand molecules are shown as sticks. TM6 was removed for clarity. The orientation colored orange is associated with high VS-efficiency. (B) Superposition of the β_{2ACT}-based MT_2 receptor models adapted with **1a** (light gray carbons, Asn268$^{6.52}$ $\chi1$ = 160°) and **10a** (violet carbons, Asn268$^{6.52}$ $\chi1$ = 292°). Asn268$^{6.52}$ is shown as spheres; ligand molecules are shown as sticks. TM5 was removed for clarity. The orientation colored light gray is associated with high VS-efficiency. (C) Superposition of the D3-based MT_2 receptor model adapted with **8a** (light gray carbons, Phe118$^{3.30}$ $\chi1$ = 77° and Leu172$^{4.57}$ $\chi1$ = 60°) and the multi-template MT_2 receptor model adapted with **10a** (dark green carbons, Phe118$^{3.30}$ $\chi1$ = 166° and Leu172$^{4.57}$ $\chi1$ = 276°). Phe118$^{3.30}$ and Leu172$^{4.57}$ are shown as ball-and-stick, whereas ligand molecules are shown as sticks. TM5 and the ECL2 were removed for clarity. The orientation colored light gray is associated with high VS-efficiency.

residues with rotameric preferences that correlate with improved enrichments.

For example, analysis of the His208$^{5.46}$ $\chi1$ dihedral angle among MT_2 receptor models with high enrichment efficiency values (defined as those with EF$^{2\%}$ values greater than 22.3) clearly showed a peak of the $\chi1$ distribution at values between 300 and 360 degrees. In this orientation, His208$^{5.46}$ points away from the helical bundle, bringing the imidazole ring in the proximity of TM4 and allowing a better accommodation of ligands into the binding cavity (Figure 6A). Similarly, the $\chi1$ dihedral angle distribution observed for Asn268$^{6.52}$ among VS-efficient MT_2 receptor models showed a peak between 120 and 180 degrees, corresponding to rotation of the side chain toward TM5. In this rotameric state, the asparagine side chain lies outside the binding site region, favoring the accommodation of melatoninergic compounds (Figure 6B). The $\chi1$ dihedral angle distributions of Leu172$^{4.57}$ for MT_2 receptor structures with high and low screening efficiencies showed two peaks at $0 < \chi1 < 60$ and $240 < \chi1 < 300$, respectively. Visual inspection of the best performing MT_2 receptor models showed that Leu172$^{4.57}$ was oriented toward TM3, forming extensive hydrophobic interactions with Phe118$^{3.30}$ and Leu122$^{3.34}$. In this orientation, Leu172$^{4.57}$ did not obstruct the binding pocket and, consequently, did not hinder the accommodation of ligands into the binding site cavity. On the contrary, in MT_2 receptors with the worst screening performances Leu172$^{4.57}$ side chain hindered the binding site region, impeding the accommodation of melatonin receptor ligands (Figure 6C). In the case of Phe118$^{3.30}$, only MT_2 receptors with low screening efficiency showed a clear peak of the $\chi1$ dihedral angle between 120 and 180 degrees, whereas VS-efficient MT_2 receptor models did not show a preferred value. A visual inspection of MT_2 receptor models characterized by low EF$^{2\%}$ values revealed that Phe118$^{3.30}$ pointed toward TM4, thus causing a rotation of the Leu172$^{4.57}$ side chain into the binding site region and indirectly hampering ligand docking (Figure 6C). For other binding site residues, smooth distributions characterized by multiple peaks were observed. It should be stressed that, for relevant residues, the best performing conformations were characterized by the displacement of side chains away from the binding site. Thus, steric compliance plays a major role in ligand recognition by these models, with lower or negligible contribution by specific polar interactions. This could be related to the neutral and lipophilic nature of melatonin receptor ligands, but it is not fully consistent with mutagenesis data, which suggest specific contributions to the binding of melatoninergic ligands, particularly for His208$^{5.46}$ and Asn268$^{6.52}$.

The previous sections demonstrate that with a suitable template/ligand combination (e.g., β_{2ACT} and **1a**) remarkable screening results can be obtained. Next, we performed a series of additional tests to critically evaluate some crucial aspects for the use of virtual screening in drug design. We first evaluated the ability of a ligand-based approach to distinguish between active and inactive molecules, using a shape screening (performed with Phase 3.3, see Experimental section for details). Following a recently published approach,[72] each molecule of the data set was flexibly aligned on a reference ligand (compounds **1a**−**12a**) in its pharmacophoric conformation, considering volume overlaps between pharmacophoric sites of the same type. This procedure retrieved a remarkable number of actives in the top 2% of the ranked data set. In particular, screening conducted on compound **1a** identified 26 actives out of 29 in the first 52 best-ranked compounds, yielding an EF$^{2\%}$ value of 44.6. Also compounds **2a**−**12a** showed remarkable values of EF$^{2\%}$, ranging from 18.9 for **12a** to 46.4 for **2a** (Table S3, Supporting Information). These were generally higher than mean values obtained for receptor models with each IFD ligand.

Next, we tested the ability of receptor models to discriminate between agonists and antagonists. Among the 29 active ligands, 16 were antagonists or partial agonists with intrinsic activity lower than 0.3; the other 13 were full agonists or partial agonists with high intrinsic activity. Table 3 reports the EF$^{2\%}$ values calculated for agonists and antagonists, considered as different subsets of "active" ligands. For each subset, antagonists or agonists, respectively, were considered as "inactive" decoys. Only receptor models built from crystal structures available both in their inactive and active states were considered; three full agonists (**2a**, **3a**, and **7a**) and three antagonists (**8a**, **10a**, and **11a**) were used as IFD ligands. Again, enrichment factors were dependent on the combination of single template and IFD ligand, with large deviations within the same group of combinations. Agonists were more efficient as IFD ligands than antagonists, while no significant difference was observed between templates. For agonists, β_2 receptor was a better template in its active form than in its inactive one (mean values: 38.3 ± 6.6 vs 23.0 ± 3.9, p = 0.01), but the same was not

dx.doi.org/10.1021/ci4000147 | J. Chem. Inf. Model. 2013, 53, 821−835

Table 3. $EF^{2\%}$ Values for a Subset of Agonist- and Antagonist-Adapted MT_2 Receptor Models

		agonists[a]			antagonists[b]		
	ligand used in IFD	2a	3a	7a	8a	10a	11a
template	A_{2AACT}[c]	26.8	11.5	3.8	24.9	15.6	6.2
	β_{2ACT}[c]	42.1	42.1	30.6	9.3	6.2	3.1
	β_{1ACT}[c]	19.1	19.1	19.1	6.2	9.3	3.1
	A_{2A}	34.5	30.6	26.8	15.6	3.1	15.6
	β_2	26.8	23.0	19.1	9.3	18.7	3.1
	β_1	23.0	11.5	23.0	12.4	18.7	6.2

[a]Values of $EF^{2\%}$, based on the number of melatoninergic agonists found in the top 2% of the ranked data set. [b]Values of $EF^{2\%}$, based on the number of melatoninergic antagonists found in the top 2% of the ranked data set. [c]When the template is in its active conformation or is bound to an agonist, it is labeled as "ACT".

observed for β_1 and A_{2A} templates. In fact, β_{1ACT} and A_{2AACT} templates accommodate agonists with conformations very similar to those assumed in antagonist-bound crystal structures, a likely consequence of crystallization conditions.[73,74] Differently, the β_{2ACT} template has a conformation quite dissimilar to that of the inactive form.[75]

We also evaluated the ability of each receptor model to discriminate between antagonists and agonists. Table 4 reports the number of agonists found in the top 13 positions of a subset composed of the 29 active ligands for receptor models already reported in Table 3. It was expected that models based on active-state templates and on agonist inducers would have ranked most of the 13 agonist ligands in the first 13 positions. However, no significant difference was observed among the different classes of templates and IFD ligands. Accordingly, modeling studies on the β_2 adrenergic and A_{2A} adenosine receptors had shown that a binding site of an activated GPCR or an agonist-bound receptor in its inactive state were able to preferentially recognize agonist ligands, compared to antagonists in computational approaches.[76] Only the combination of the β_2 receptor in its active state as a template and **2a** as an IFD ligand gave a model with remarkable discriminating ability, ranking 11 agonists in the first 13 positions.

Table 4. Number of Agonists Found in the Top 13 positions within the Set of 29 Melatoninergic Ligands, for a Subset of MT_2 Receptor Models

		agonists			antagonists		
	ligand used in IFD	2a	3a	7a	8a	10a	11a
template	A_{2AACT}[a]	6	6	6	7	7	7
	β_{2ACT}[a]	11	8	8	8	5	4
	β_{1ACT}[a]	6	8	7	8	6	6
	A_{2A}	7	9	9	8	7	5
	β_2	8	5	7	7	7	5
	β_1	6	7	7	7	4	6

[a]When the template is in its active conformation or is bound to an agonist, it is labeled as "ACT".

Finally, we tested the ability of the 144 ligand-adapted MT_2 receptor models to rank order active and inactive compounds belonging to the same chemical classes as the active compounds studied here, which is a challenging task in docking studies. A subset of 21 "inactive" melatonin receptor ligands from the 12 chemical classes of the 29 active compounds was retrieved from the literature (Table S4, Supporting Information).[21,43−45,47,50,52,55−58,77−81] Additional ligands were classified as inactive when their potency was at least 10-fold lower than that of active ligands from the same class. New screening runs were performed on the 50 compound set, and the number of false positives, i.e., the number of inactive compounds found in the top 29 positions, is reported in Table 5. The number of false positives varied among different ligand-adapted MT_2 receptor models, ranging from 6 to 17. However, there were no clear differences among average performances for rows and columns of Table 5, with mean values ranging from 8.8 to 11.6 false positives in the top 29 positions. Thus, these models (and particularly the aforementioned β_{2ACT}/**1a** model) retained some ability to reproduce basic structure−activity relationships within chemical classes of melatonin receptor ligands. A comparison between $EF^{2\%}$ in the active-decoys subset and these false positives outlined a general agreement between these two performances. Indeed, the five models with best $EF^{2\%}$ values also gave small numbers of false positives (6, 7, 9, 8, and 8 for β_{2ACT}/**1a**, β_{2ACT}/**6a**, β_1/**4a**, β_{2ACT}/**3a**, and A_{2AACT}/**8a**, respectively).

Table 5. Number of False Positives Calculated for 144 Ligand-Adapted MT_2 Receptor Models

	ligand used in IFD	1a	2a	3a	4a	5a	6a	7a	8a	9a	10a	11a	12a	Mean
template	A_{2A}	10	10	6	10	8	12	13	10	12	14	10	10	10.4
	A_{2AACT}[a]	8	9	8	12	10	10	8	8	8	12	12	9	9.5
	β_2	6	7	11	11	8	11	11	9	9	11	13	11	9.8
	β_{2ACT}[a]	6	11	8	10	10	7	9	11	10	14	10	17	10.3
	β_1	9	11	8	9	7	9	10	7	13	10	9	12	9.5
	β_{1ACT}[a]	8	8	10	9	8	10	6	9	9	10	12	11	9.2
	D3	8	9	8	8	12	11	11	8	9	11	12	11	9.8
	H_1	9	8	9	11	11	9	10	10	11	11	14	13	10.5
	Rho	9	10	11	8	9	9	12	11	7	12	15	9	10.2
	Rho_{ACT}[a]	10	13	11	11	12	10	9	11	9	12	9	13	10.8
	CXCR4	12	9	10	10	9	12	10	7	10	11	9	10	9.9
	MULTI[b]	11	8	8	12	10	7	9	9	8	10	14	13	9.9
	Mean	8.8	9.4	9.0	10.1	9.5	9.8	9.8	9.2	9.6	11.5	11.6	11.6	

[a]When the template is either in its active conformation or is bound to an agonist, it is labeled as "ACT". [b]MULTI indicates a consensus model with the inactive GPCR crystal structures used as templates.

dx.doi.org/10.1021/ci4000147 | J. Chem. Inf. Model. 2013, 53, 821−835

CONCLUSIONS

The present work was aimed at developing MT_2 receptor models for virtual screening. Significant challenges were associated with the low sequence identity of the target to crystal structure templates and with the need to refine the binding sites of the homology models. None of the initial MT_2 receptor models, irrespective of the structural template employed, gave screening results significantly better than random on a data set composed of 29 active melatoninergic ligands and 2560 decoys due to steric incompatibility of the binding site with the active ligands. Significant improvements in screening enrichments were achieved by a binding site refinement guided by knowledge of the experimental SARs. Representative melatoninergic ligands were docked using information from mutagenesis studies and pharmacophore models, followed by ligand-induced adaptation of the binding sites. In some cases, excellent enrichment factors were observed; in particular, the model built from the β_2 adrenoceptor in its active state and adapted around a 1-acylaminoethyl tetralin partial agonist (**1a**) found 24 of the 29 active compounds in the first 52 positions of the 2589 compounds set ($EF^{2\%}$ = 41.2 out of a maximum value of 50.0) and 22 in the first 26 positions ($EF^{1\%}$ = 75.5 out of a maximum value of 100.0). This model was also able to accurately discriminate active compounds from a small set of less potent melatoninergic derivatives from the same chemical classes. On the other hand, a ligand-based approach based on shape similarity with IFD ligands on-average outperformed virtual screening results based on receptor models. This could be due to the particular data set of "active" melatoninergic ligands used here, which all belong to chemical classes that are similar from both a topological and a physicochemical point of view. Moreover, a number of studies have shown that ligand-based approaches give better screening results with "homogenous" data sets compared to docking-based algorithms,[82,83] which are better able to separate gross ligand features.[60,84] That said, the highest performing model here showed experimentally consistent binding poses for a large range of docked actives, making this an attractive tool for pose-prediction of novel MT_2 ligands, something which is not possible with a shape based algorithm.

Screening performance was strongly affected by the choice of the template and ligand combination in a rather unpredictable way. Structural analysis revealed that in the best performing receptor models some residues previously reported to be involved in ligand-binding interactions adopted distinct conformations. Interestingly, these preferred conformations did not form expected polar interactions with ligand functional groups, illustrating the inherent uncertainty related to the use of comparative modeling and its application to virtual screening. Even from the best-performing models, it is not always possible to derive structural hypotheses about ligand−receptor interactions from VS results, which suggests that the models must be calibrated and tested to define their domain of applicability. Encouragingly, it was possible to obtain high enrichment values starting with all of the templates, suggesting that the domain of applicability for GPCR homology models can be greatly expanded with an appropriate binding site refinement protocol. However, the best ligand to use for induced-fit docking calculations varied between the templates, implying that further work is needed to apply this method in cases where limited information about active ligands is available. Nevertheless, reasonable performance was obtained in almost all cases where the modified induced-fit docking protocol that accounts for known ligand SAR was applied. In addition, the best models performed remarkably well and justify the extra effort required to generate refined homology models optimized for virtual screening.

EXPERIMENTAL SECTION

Molecular modeling was performed using the Schrödinger software suite. Homology models were built with Prime 3.0[37] and refined in Maestro 9.2.[85] Ligand molecules were prepared using LigPrep 2.5,[86] whereas docking studies were carried out with Glide 5.7 using the SP scoring function. Default settings were used, unless stated otherwise.

Homology Modeling. At the outset of this study, the following structures were available and were used as templates for homology modeling of the MT_2 receptor: the crystal structures of rhodopsin (PDB IDs: 1GZM[87] (chain A) and 2X72[88] (chain A)), β_2 adrenergic receptor (PDB IDs: 2RH1[89] and 3P0G[75]), β_1 adrenergic receptor (PDB IDs: 2VT4[90] (chain B) and 2Y02[73] (chain B)), A_{2A} adenosine receptor (PDB IDs: 3EML[91] and 2YDO[74]), CXCR4 chemokine receptor (PDB ID: 3OE6[61] (chain A)), H_1 histamine receptor (PDB ID: 3RZE[92]) and D3 dopamine receptor (PDB: 3PBL[93] (chain A)), which were retrieved from the Protein Data Bank. The amino acid sequence of the human MT_2 melatonin receptor was obtained from the Universal Protein Resource[94] (UniProt ID: P49286). Automated sequence alignments were performed with Prime and subsequently refined taking into account conserved sequence motifs among class A GPCRs (Figure S6, Supporting Information).[38,95] An additional MT_2 receptor model was built using all the GPCR X-ray structures crystallized in the presence of antagonists or inverse agonists (i.e., 1GZM, 2RH1, 2VT4, 3EML, 3OE6, 3RZE, and 3PBL), and the multiple sequence alignment is reported at the end of Figure S6 of the Supporting Information.

The 12 "crude" receptor models were processed using the Protein Preparation Wizard,[96] and protein C- and N-termini were capped with acetyl and methylamino groups, respectively. Hydrogen atoms were added to each structure, and the overall hydrogen bonding network was optimized by sampling the orientation of hydroxyl and thiol groups, together with the side chain amides of asparagine and glutamine residues and by adjusting the tautomerization states of histidine residues. This optimization procedure first identifies localized clusters of hydrogen bonding species and, for each cluster, all possible combinations of tautomeric forms, protonation states, and rotamers of thiol and hydroxyl groups and of terminal amide side chains of glutamine and asparagines are enumerated. For each cluster, combinations are then scored using multiple criteria, including (i) the deviation from an ideal hydrogen bond geometry, (ii) the presence of clashes between close polar hydrogens, and (iii) the penalties (or the rewards) derived from the mismatch (or the exact match) between the current protonation state of a polar residue and that identified as optimal on the basis of neighboring hydrogen bond partners. The best-scored combination is then selected for each cluster. At the end of this procedure, histidines were in their neutral form, lysines and arginines were protonated, and glutamates and aspartates were negatively charged. Histidines carried their hydrogen atom either on the δ or ε nitrogen to maximize the hydrogen bond interactions with neighboring residues. The 12 receptor models were in the same protonation state, while

dx.doi.org/10.1021/ci4000147 | J. Chem. Inf. Model. 2013, 53, 821−835

histidines could have different tautomeric forms, due to the optimization of hydrogen bonding networks. The 12 "crude" receptor structures were submitted to a manual structural refinement, in which the side chains of residues thought to be important for ligand binding, i.e., Asn175$^{4.60}$, His208$^{5.46}$, Asn268$^{6.52}$, and Tyr298$^{7.43}$, were reoriented, when necessary, toward the binding site region. Accordingly, neighboring side chains were adjusted to avoid major steric clashes, and the tautomeric forms of binding site residues were reevaluated with the Protein Preparation Wizard. The two subsets of 12 "crude" and 12 "manually refined" MT_2 receptor structures were then submitted to a Prime minimization using the OPLS2005 force field[97] to an RMSD of 0.01 Å and applying the Surface Generalized Born (SGB) continuum solvation model for water. The stereochemical quality of these 24 MT_2 receptor models was assessed through the Protein Report Tool implemented in Maestro 9.2. The receptor models showed more than 90% of residues within the allowed regions of the Ramachandran plot.

Database Preparation. A set of 50 melatonin receptor ligands belonging to 12 diverse chemical classes was retrieved from the literature. Compounds with p*K*i values higher than 8.0 and lower than 7.0 were defined as active and inactive compounds, respectively. When these criteria could not be met within a class, members were selected in order to maintain a difference of at least one logarithmic unit of binding affinity between active and inactive ligands (e.g., compounds **7a** and **7b** vs compound **7c**). Selection of compounds **11a**−**11d** was based on their pIC50 values. The final melatoninergic data set was composed of 29 active and 21 inactive melatonin receptor ligands (Tables S1 and S3, Supporting Information). Decoy molecules were selected carefully to avoid biasing screening results. An initial data set of 31500 decoys with a molecular weight lower than 550 was selected from the ChemBridge database. Decoys were subsequently submitted to a filtering procedure aimed at retaining molecules with similar properties (molecular weight, number of rotatable bonds, number of heavy atoms, number of hydrogen bond donor/acceptor groups) compared to the 29 active melatoninergic ligands (Figure S1, Supporting Information) using the ligparse utility in Maestro 9.2. The 50 melatonin receptor ligands and the 2631 decoys thus selected were prepared with LigPrep to identify and generate tautomers and stereoisomers. Charged species were removed from the final data set, yielding 2560 unique decoys and 50 unique melatonin receptor ligands.

Pharmacophore Model Building. Previous pharmacophore analysis and SAR studies on active melatonin receptor ligands[49,51,62,98] had led to the identification of an active conformation of melatonin, with the acetylaminoethyl side chain, in its fully extended conformation, perpendicular to the indole ring and the methyl of the 5-methoxy group directed toward position 4. Taking into consideration the pharmacophore hypothesis previously developed, 12 representative melatonin receptor ligands (i.e., the most potent compounds within each of the 12 chemical classes) were used to devise a new pharmacophore model. Because *cis*-(2*S*,4*S*)-**6b** has been recently found to be the most potent of **6b** stereoisomers,[51] the corresponding *cis*-(2*S*,4*S*)-**6a** (carrying the methoxy group) was used as the reference compound for ligand alignment and pharmacophore model building given its limited flexibility. The conformation of compound *cis*-(2*S*,4*S*)-**6a** selected as reference was the most abundant one previously identified by a combined NMR/stochastic dynamics simulation study, showing a die-quatorial arrangement of the phenyl group and of the amide side chain and an *anti* arrangement of the H_2-C_2-N-H dihedral.[51] This conformation is consistent with the requirements of the pharmacophore models cited at the beginning of the paragraph. Superposition of the 12 representative ligands on *cis*-(2*S*,4*S*)-**6a** was performed by a rigid fit procedure, superimposing the four atoms of the amide function when present, the centroid of the phenyl ring bearing the methoxy group, and the oxygen atom of the methoxy group when present. When no indications were given about compound stereochemistry (i.e., for compounds **10a** and **11a**) or when biological data were available for enantiomeric mixtures only (compound **1a**), the stereoisomer showing the best superposition was selected. A pharmacophore model was then created using the "create_hypoConsensus" utility implemented in Phase 3.3.[99] It performs an initial clustering on each type of pharmacophoric site found among all the prealigned ligands. Clusters are then included in the final model only if all the pharmacophoric sites enclosed in the cluster are within a user-specified distance. Finally, the centroid of each cluster is selected as the representative site in the final hypothesis. A 5-points pharmacophore model was obtained setting the clustering tolerance to 3 Å and the minimum number of ligands that must be included in a cluster to 10. It was composed of an aromatic core bearing a hydrogen bond acceptor group connected to an amide fragment through an alkyl linker (Figure S3, Supporting Information).

Conformational Search and Pharmacophore-Consistent Conformer Selection. The 12 representative compounds were submitted to a conformational search by means of the mixed torsional/low-mode sampling approach (MCMM/LMOD) implemented in Macromodel 9.9,[100] applying OPLS2005 as force field and the GBSA water solvation treatment. Conformers were minimized using the Polak−Ribiere conjugate gradient method with a convergence threshold of 0.05 kJ mol^{-1} Å^{-1}. During the conformational search, two minimum-energy conformations were considered as duplicates if the maximum distance observed between any pair of corresponding atoms after rigid-body superposition was lower than 1 Å. Moreover, all conformers showing energy greater than 21 kJ mol^{-1} (5 kcal mol^{-1}) above the minimum energy conformation were discarded. For each of the 12 representative melatoninergic ligands, the final subset of filtered conformers was superimposed to its initial ligand conformation used for pharmacophore model building using heavy atoms.

To evaluate the consistency of the obtained conformations with the pharmacophore model for melatoninergic ligands, a score-in-place procedure was applied to the conformations of each ligand using the 5-points pharmacophore built starting from the 12 representative compounds. Conformations were required to match a minimum of 4 sites with a distance matching tolerance of 5 Å. The only exception was represented by compound **11a**, whose scoring procedure was based on 3 sites. Indeed, the tetrahydroisoquinoline core of compound **11a** does not carry the methoxy group, and it also lacks the NH donor group of the amide functionality. Before the pharmacophore fitting procedure, an additional filtering stage was conducted on the conformations of compounds **2a**, **5a**, and **7a** to improve the quality of conformational selection through the exploitation of prior information known for these chemical classes. Indeed, these compounds are characterized by a melatonin-like structure, in which a 6 + 5 bicyclic nucleus is connected to an amide functionality through a linker of two methylene groups. In the putative bioactive conformation of

dx.doi.org/10.1021/ci4000147 | J. Chem. Inf. Model. 2013, 53, 821−835

melatonin, the amide side chain lays perpendicular to the indole ring in its fully extended conformation (dihedral angle τ: C2−C3−Cβ−Cα∼90°, dihedral angle θ: C3−Cβ−Cα−N∼180°, Figure S7, Supporting Information). Therefore, all the conformations of compounds **2a**, **5a**, and **7a** that did not show an extended geometry of the amide side chain were discarded before the fitting procedure on the common pharmacophore model.

All the pharmacophore-screened conformations were subsequently ranked according to their Fitness score, a measure of how well the aligned ligand conformer matches the pharmacophore points. The two best ranked conformers showing a root-mean-square deviation (RMSD) to the initial ligand conformation greater than 0.5 Å were selected. The RMSD criterion could not be applied to conformationally constrained derivatives **2a** and **6a** because their pharmacophore-consistent conformations showed RMSD values lower than 0.5 Å. For this reason, conformations of compounds **2a** and **6a** were selected only on the basis of their Fitness score. Thus, the initial ligand conformation and the two additional pharmacophore-consistent conformations for each of the 12 representative melatoninergic ligands (i.e., 36 structures) were passed to the next phase of induced-fit docking (see Knowledge-Based IFD Procedure).

Shape Screening. Shape screening was performed with the "phase_shape" utility implemented in Phase 3.3.[72] In brief, each ligand is treated as a set of pharmacophoric sites, identified according to the phase feature definitions, and represented as hard spheres. Included features are aromatic, hydrophobic, H-bond acceptor, H-bond donor, negative ionic, and positive ionic. Once features are created, each molecule of the database is flexibly aligned to the reference one using the pharmacophore-derived spheres as aligning points. Volume overlap scores are then computed between sites of the same type. A single shape similarity value, proportional to the total volume overlap, is finally calculated for each ligand molecule. Compounds **1a**−**12a** were kept fixed in the conformation used to build the 5-points pharmacophore model. A conformational search was applied to each molecule of the database during the screening process, and the *trans* geometry was imposed to amide fragments. All the other parameters were set to default values.

Induced-Fit Docking (IFD) Calculations. *Automated IFD Procedure.* Initial IFD calculations of the 12 reference compounds in Figure 3 were conducted on the 12 "manually refined" MT_2 receptor models applying default settings. In particular, during the softened-potential docking runs, van der Waals radii of protein and ligand nonpolar atoms were scaled by a factor of 0.5, and Glide grids were centered on Met120$^{3.32}$, which is positioned in the middle of the putative binding region.

Because none of the ligand docking poses was either accommodated into the putative binding site or was consistent with site-directed mutagenesis or pharmacophore models, the initial IFD protocol was slightly modified. Thus, during the initial softened-potential docking run, side chains of His208$^{5.46}$, Phe209$^{5.47}$, and Asn268$^{6.52}$ were temporarily mutated to alanines to create additional space within the binding site region. van der Waals radii scaling factors applied to ligand and receptor nonpolar atoms were set to 0.5 and 0.7, respectively. Unfortunately, also this procedure did not produce docking poses consistent with prior information.

Knowledge-Based IFD Procedure. Given the unsatisfactory results obtained with automated IFD procedures, we decided to develop a customized knowledge-based refinement of the MT_2 receptor models, in which pharmacophore-consistent ligand poses were used as "shape inducers" to adapt the binding site region of each MT_2 receptor model. Thus, the 36 pharmacophore-consistent conformations obtained for the 12 representative melatoninergic compounds were manually inserted into the putative binding site region of each of the 12 "manually refined" MT_2 receptor structures (i.e., those 12 receptor models in which side chains of residues important for ligand binding were directed toward the binding site cavity). Initial ligand placement was guided by information retrieved from mutagenesis studies. In particular, the amide oxygen was placed in the proximity of Tyr298$^{7.43}$ located on TM7 and the methoxy group of active melatoninergic ligands was accommodated near Asn175$^{4.60}$ located on TM4. During the initial ligand placement, side chains of His208$^{5.46}$, Asn268$^{6.52}$, and Tyr294$^{7.39}$ were temporarily mutated to alanines. The initial set of 36 ligand−receptor complexes were refined using the Prime stage of the IFD workflow. Once residues that were previously mutated to alanine were reintroduced, residues within a shell of 5 Å around the ligand pose were refined by a side chain conformational search. The optimization of side chain positions was repeated five times to favor a better side chain conformation prediction, and it was followed by an energy minimization of the residues and the ligand molecule. To retain the pharmacophore-consistent ligand conformation in the final Glide stage of the IFD workflow, each ligand structure obtained at the end of the protein structure refinement was not redocked into the induced-fit receptor structure, but it was only energetically optimized (refined) in the field of the receptor and subsequently scored using default Glide settings.

Scoring of Ligand−Receptor Complexes. A knowledge-based scoring function was developed to recognize energetically favorable ligand−receptor complexes consistent with both mutagenesis data and ligand-based information and was used to select the best ligand-induced receptor structures among the 3 conformations of each of the 12 representative melatoninergic ligands. The knowledge-based IFDScore (IFD_{KB}) was defined as

$$IFD_{KB} = \Sigma(IFDScore - SIFts - Fitness + \cos\theta) \quad (1)$$

The IFDScore was taken directly from the IFD calculations. Structural interaction fingerprints (SIFts)[101] were calculated to evaluate the molecular interactions between the ligand and the residues supposed to be involved in ligand binding because of mutagenesis studies (Met120$^{3.32}$, Asn175$^{4.60}$, Val204$^{5.42}$, His208$^{5.46}$, Asn268$^{6.52}$, Leu272$^{6.56}$, Ala275$^{6.59}$, Val291$^{7.36}$, Leu295$^{7.40}$, and Tyr298$^{7.43}$). SIFts were calculated as the number of contacts between ligand and the relevant residues cited above. A contact was considered if the minimum distance between ligand-residue heavy atoms was less than 4 Å. To assess the consistency of ligand conformation with the pharmacophore model, the ligand structure was extracted from its ligand−receptor complex and superimposed to the five pharmacophoric points using Phase. The Fitness score calculated with Phase is proportional to the RMSD between the ligand molecule and the pharmacophore model. In the putative bioactive conformation of melatonin, the torsion angle C3−Cβ−Cα−N (Figure S7, Supporting Information) had been set to 180°. Because compounds **2a**, **5a**, and **7a** include a substructure having the topology of melatonin, the cosine of

dx.doi.org/10.1021/ci4000147 | J. Chem. Inf. Model. 2013, 53, 821−835

their corresponding torsion angles (θ) was included in IFD_{KB} to favor the conformations having θ at 180 degrees. The IFD_{KB} scoring function has been implemented as a custom python script available upon request.

Virtual Screenings (VS). Each ligand−receptor complex obtained at the end of the knowledge-guided IFD protocol was ranked according to the IFD_{KB} score. The best complex for each of the 12 representative "shape inducers" and for each of the 12 template-based MT_2 receptor models was selected for VS campaigns, yielding a final ensemble of 144 ligand-adapted MT_2 receptor models. These ligand−receptor complexes were submitted to a Prime energy minimization, using the OPLS2005 force field to an RMSD of 0.01 Å, applying the SGB continuum solvation model for water.

The 12 "crude", 12 "manually refined", and 144 ligand-adapted MT_2 receptor models were then mutually superimposed using the Structalign utility implemented in Prime. Docking grids were centered in an arbitrary position comprising the putative binding region, setting enclosing and bounding boxes to 30 Å and 10 Å on each side, respectively. One pose per ligand was collected during each VS run. The resulting ensemble of docking poses was ranked according to the Docking Score, which incorporates the ionization penalties calculated with Epik[102] to the main Glide Score term.

Enrichment factors were calculated at 1%, 2%, 5%, and 10% of the total database screened, using the following equation

$$EF^{x\%} = (act/n)/(ACT/N)$$

where *act* is the number of actives retrieved in the first *n* positions of the database and ACT is the total number of actives included in the database of N compounds. All the other enrichment metrics were computed with the enrichment.py script available from the Schrödinger Script Center (www.schrodinger.com/scriptcenter).

Root-Mean-Square Fluctuation (RMSF) Calculation. The 144 ligand-adapted MT_2 receptor models were mutually superimposed using the Structalign utility implemented in Prime. To compare MT_2 receptor models of the same amino acid length, residues before $Ser44^{1.36}$ and after Arg317 were deleted. The average_structure.py script (www.schrodinger.com/scriptcenter) was then used to build an average structure. The 144 MT_2 models were then realigned to the TM Cα carbons of the average structure, and for each model, the RMSD value of side chain heavy atoms was calculated for all binding site residues using the rmsd_by_residue.py script (www.schrodinger.com/scriptcenter). Root-mean-square fluctuation (RMSF) was defined as the mean RMSD value obtained for each residue side chain among the 144 MT_2 receptor models.

Kruskal−Wallis Analysis. Kruskal−Wallis analysis[65] was performed with the statistical program R.[103]

■ ASSOCIATED CONTENT

Supporting Information

Molecular properties of melatoninergic ligands and decoy molecules; docking results for the "manually refined" CXCR4-based MT_2 receptor model and for the β_{2ACT}-based MT_2 structure adapted with **1a**; representation of the 5-points pharmacophore model built from the 12 representative melatoninergic ligands; RMSF values of binding site residues; sequence alignments between the human MT_2 melatonin receptor and template structures; molecular structure and MT_2 binding affinity for the 50 melatoninergic ligands; 3D structures of the 144 ligand-adapted MT_2 receptor models in Maestro and pdb formats; $EF^{2\%}$ values obtained from the shape screening on compounds **1a**−**12a**; and enrichment metrics for the 144 ligand-adapted MT_2 receptor structures. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*Phone: +39 0521 905062 (S.R.), +1 212 295 5800 (T.B.). Fax: +39 0521 905006 (S.R.), +1 212 295 5801 (T.B.). E-mail: silvia.rivara@unipr.it (S.R.), thijs.beuming@schrodinger.com (T.B.).

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

Daniele Pala thanks Prof. Harel Weinstein for his kind hospitality during the time spent in his lab.

■ ABBREVIATIONS

BEDROC, Boltzmann-enhanced discrimination of receiver operating characteristic; EF, enrichment factor; GBSA, generalized born surface area; IFD, induced-fit docking; IFD_{KB}, knowledge-based induced-fit docking score; MT_1, melatonin receptor subtype 1; MT_2, melatonin receptor subtype 2; RIE, robust initial enrichment; RMSD, root-mean-square deviation; RMSF, root-mean-square fluctuation; SGB; surface generalized born; sID, sequence identity; SIFts, structural interaction fingerprints; SP, standard precision; TM, transmembrane domains; VS, virtual screening

■ REFERENCES

(1) Arendt, J. Melatonin: Characteristics, concerns, and prospects. *J. Biol. Rhythms* **2005**, *20*, 291−303.

(2) Pévet, P.; Bothorel, B.; Slotten, H.; Saboureau, M. The chronobiotic properties of melatonin. *Cell Tissue Res.* **2002**, *309*, 183−191.

(3) Pandi-Perumal, S. R.; Trakht, I.; Srinivasan, V.; Spence, D. W.; Maestroni, G. J. M.; Zisapel, N.; Cardinali, D. P. Physiological effects of melatonin: Role of melatonin receptors and signal transduction pathways. *Prog. Neurobiol.* **2008**, *85*, 335−353.

(4) Pandi-Perumal, S. R.; Srinivasan, V.; Maestroni, G. J. M.; Cardinali, D. P.; Poeggeler, B.; Hardeland, R. Melatonin: Nature's most versatile biological signal? *FEBS J.* **2006**, *273*, 2813−2838.

(5) Lyssenko, V.; Nagorny, C. L. F.; Erdos, M. R.; Wierup, N.; Jonsson, A.; Spégel, P.; Bugliani, M.; Saxena, R.; Fex, M.; Pulizzi, N.; Isomaa, B.; Tuomi, T.; Nilsson, P.; Kuusisto, J.; Tuomilehto, J.; Boehnke, M.; Altshuler, D.; Sundler, F.; Eriksson, J. G.; Jackson, A. U.; Laakso, M.; Marchetti, P.; Watanabe, R. M.; Mulder, H.; Groop, L. Common variant in MTNR1B associated with increased risk of type 2 diabetes and impaired early insulin secretion. *Nat. Genet.* **2009**, *41*, 82−88.

(6) Sánchez-Barceló, E. J.; Mediavilla, M. D.; Tan, D. X.; Reiter, R. J. Clinical uses of melatonin: Evaluation of human trials. *Curr. Med. Chem.* **2010**, *17*, 2070−2095.

(7) Mor, M.; Rivara, S.; Pala, D.; Bedini, A.; Spadoni, G.; Tarzia, G. Recent advances in the development of melatonin MT_1 and MT_2 receptor agonists. *Expert Opin. Ther. Pat.* **2010**, *20*, 1059−1077.

(8) Dubocovich, M. L.; Delagrange, P.; Krause, D. N.; Sugden, D.; Cardinali, D. P.; Olcese, J. International union of basic and clinical pharmacology. LXXV. Nomenclature, classification, and pharmacology of G protein-coupled melatonin receptors. *Pharmacol. Rev.* **2010**, *62*, 343−380.

(9) Nosjean, O.; Ferro, M.; Coge, F.; Beauverger, P.; Henlin, J. M.; Lefoulon, F.; Fauchere, J. L.; Delagrange, P.; Canet, E.; Boutin, J. A.

Identification of the melatonin-binding site MT_3 as the quinone reductase 2. *J. Biol. Chem.* **2000**, *275*, 31311−31317.
(10) Mazna, P.; Grycova, L.; Balik, A.; Zemkova, H.; Friedlova, E.; Obsilova, V.; Obsil, T.; Teisinger, J. The role of proline residues in the structure and function of human MT2 melatonin receptor. *J. Pineal Res.* **2008**, *45*, 361−372.
(11) Farce, A.; Chugunov, A. O.; Logé, C.; Sabaouni, A.; Yous, S.; Dilly, S.; Renault, N.; Vergoten, G.; Efremov, R. G.; Lesieur, D.; Chavatte, P. Homology modeling of MT_1 and MT_2 receptors. *Eur. J. Med. Chem.* **2008**, *43*, 1926−1944.
(12) Voronkov, A. E.; Ivanov, A. A.; Baskin, I. I.; Palyulin, V. A.; Zefirov, N. S. Molecular modeling study of the mechanism of ligand binding to human melatonin receptors. *Dokl. Biochem. Biophys.* **2005**, *403*, 284−288.
(13) Grol, C. J.; Jansen, J. M. The high affinity melatonin binding site probed with conformationally restricted ligands–II. Homology modeling of the receptor. *Bioorg. Med. Chem.* **1996**, *4*, 1333−1339.
(14) Chugunov, A. O.; Farce, A.; Chavatte, P.; Efremov, R. G. Differences in binding sites of two melatonin receptors help to explain their selectivity to some melatonin analogs: A molecular modeling study. *J. Biomol. Struct. Dyn.* **2006**, *24*, 91−107.
(15) Kokkola, T.; Salo, O. M. H.; Poso, A.; Laitinen, J. T. The functional role of cysteines adjacent to the NRY motif of the human MT1 melatonin receptor. *J. Pineal Res.* **2005**, *39*, 1−11.
(16) Mazna, P.; Obsilova, V.; Jelinkova, I.; Balik, A.; Berka, K.; Sovova, Z.; Ettrich, R.; Svoboda, P.; Obsil, T.; Teisinger, J. Molecular modeling of human MT2 melatonin receptor: the role of Val204, Leu272 and Tyr298 in ligand binding. *J. Neurochem.* **2004**, *91*, 836−842.
(17) Uchikawa, O.; Fukatsu, K.; Tokunoh, R.; Kawada, M.; Matsumoto, K.; Imai, Y.; Hinuma, S.; Kato, K.; Nishikawa, H.; Hirai, K.; Miyamoto, M.; Ohkawa, S. Synthesis of a novel series of tricyclic indan derivatives as melatonin receptor agonists. *J. Med. Chem.* **2002**, *45*, 4222−4239.
(18) Navajas, C.; Kokkola, T.; Poso, A.; Honka, N.; Gynther, J.; Laitinen, J. T. A rhodopsin-based model for melatonin recognition at its G protein-coupled receptor. *Eur. J. Pharmacol.* **1996**, *304*, 173−183.
(19) Sugden, D.; Chong, N. W.; Lewis, D. F. Structural requirements at the melatonin receptor. *Br. J. Pharmacol.* **1995**, *114*, 618−623.
(20) Ivanov, A. A.; Voronkov, A. E.; Baskin, I. I.; Palyulin, V. A.; Zefirov, N. S. The study of the mechanism of binding of human ML1A melatonin receptor ligands using molecular modeling. *Dokl. Biochem. Biophys.* **2004**, *394*, 49−52.
(21) Rivara, S.; Lorenzi, S.; Mor, M.; Plazzi, P. V.; Spadoni, G.; Bedini, A.; Tarzia, G. Analysis of structure-activity relationships for MT_2 selective antagonists by melatonin MT_1 and MT_2 receptor models. *J. Med. Chem.* **2005**, *48*, 4049−4060.
(22) Mazna, P.; Berka, K.; Jelinkova, I.; Balik, A.; Svoboda, P.; Obsilova, V.; Obsil, T.; Teisinger, J. Ligand binding to the human MT2 melatonin receptor: the role of residues in transmembrane domains 3, 6, and 7. *Biochem. Biophys. Res. Commun.* **2005**, *332*, 726−734.
(23) Rivara, S.; Pala, D.; Lodola, A.; Mor, M.; Lucini, V.; Dugnani, S.; Scaglione, F.; Bedini, A.; Lucarini, S.; Tarzia, G.; Spadoni, G. MT_1-selective melatonin receptor ligands: Synthesis, pharmacological evaluation, and molecular dynamics investigation of N-{[(3-O-substituted)anilino]alkyl}amides. *ChemMedChem* **2012**, *7*, 1954−1964.
(24) 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.
(25) Michino, M.; Abola, E.; Brooks, C. L., 3rd; Dixon, J. S.; Moult, J.; Stevens, R. C. Community-wide assessment of GPCR structure modelling and ligand docking: GPCR Dock 2008. *Nat. Rev. Drug Discovery* **2009**, *8*, 455−463.
(26) Beuming, T.; Sherman, W. Current assessment of docking into GPCR crystal structures and homology models: Successes, challenges, and guidelines. *J. Chem. Inf. Model.* **2012**, *52*, 3263−3277.
(27) Katritch, V.; Rueda, M.; Lam, P. C.-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.
(28) Forrest, L. R.; Tang, C. L.; Honig, B. On the accuracy of homology modeling and sequence alignment methods applied to membrane proteins. *Biophys. J.* **2006**, *91*, 508−517.
(29) Vassilatis, D. K.; Hohmann, J. G.; Zeng, H.; Li, F.; Ranchalis, J. E.; Mortrud, M. T.; Brown, A.; Rodriguez, S. S.; Weller, J. R.; Wright, A. C.; Bergmann, J. E.; Gaitanaris, G. A. The G protein-coupled receptor repertoires of human and mouse. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 4903−4908.
(30) Fredriksson, R.; Lagerström, M. C.; Lundin, L.-G.; Schiöth, H. B. The G-protein-coupled receptors in the human genome form five main families. Phylogenetic analysis, paralogon groups, and fingerprints. *Mol. Pharmacol.* **2003**, *63*, 1256−1272.
(31) Ivanov, A. A.; Barak, D.; Jacobson, K. A. 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.
(32) Kneissl, B.; Leonhardt, B.; Hildebrandt, A.; Tautermann, C. S. Revisiting automated G-protein coupled receptor modeling: the benefit of additional template structures for a neurokinin-1 receptor model. *J. Med. Chem.* **2009**, *52*, 3166−3173.
(33) Evers, A.; Klebe, G. Successful virtual screening for a submicromolar antagonist of the neurokinin-1 receptor based on a ligand-supported homology model. *J. Med. Chem.* **2004**, *47*, 5381−5392.
(34) Kołaczkowski, M.; Bucki, A.; Feder, M.; Pawłowski, M. Ligand-optimized homology models of D_1 and D_2 dopamine receptors: Application for virtual screening. *J. Chem. Inf. Model.* **2013**, *53*, 638−648.
(35) 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.
(36) Mobarec, J. C.; Sanchez, R.; Filizola, M. Modern homology modeling of G-protein coupled receptors: Which structural template to use? *J. Med. Chem.* **2009**, *52*, 5207−5216.
(37) *Prime*, version 3.0; Schrödinger, LLC: New York, 2011.
(38) Ballesteros, J.; Weinstein, H. Integrated methods for the construction of three-dimensional models of structure−function relations in G protein-coupled receptors. *Methods Neurosci.* **1995**, *25*, 366−428.
(39) Gerdin, M. J.; Mseeh, F.; Dubocovich, M. L. Mutagenesis studies of the human MT_2 melatonin receptor. *Biochem. Pharmacol.* **2003**, *66*, 315−320.
(40) Jacobson, K. A.; Costanzi, S. New insights for drug design from the X-ray crystallographic structures of G-protein-coupled receptors. *Mol. Pharmacol.* **2012**, *82*, 361−371.
(41) *Glide*, version 5.7; Schrödinger, LLC: New York, 2011.
(42) 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.
(43) Durieux, S.; Chanu, A.; Bochu, C.; Audinot, V.; Coumailleau, S.; Boutin, J. A.; Delagrange, P.; Caignard, D. H.; Bennejean, C.; Renard, P.; Lesieur, D.; Berthelot, P.; Yous, S. Design and synthesis of 3-phenyltetrahydronaphthalenic derivatives as new selective MT_2 melatoninergic ligands. Part II. *Bioorg. Med. Chem.* **2009**, *17*, 2963−2974.
(44) Faust, R.; Garratt, P. J.; Jones, R.; Yeh, L. K.; Tsotinis, A.; Panoussopoulou, M.; Calogeropoulou, T.; Teh, M. T.; Sugden, D. Mapping the melatonin receptor. 6. Melatonin agonists and antagonists derived from 6H-isoindolo[2,1-*a*]indoles, 5,6-dihydroindolo[2,1-*a*]isoquinolines, and 6,7-dihydro-5H-benzo[*c*]-azepino[2,1-*a*]indoles. *J. Med. Chem.* **2000**, *43*, 1050−1061.
(45) Hu, Y.; Ho, M. K. C.; Chan, K. H.; New, D. C.; Wong, Y. H. Synthesis of substituted *N*-[3-(3-methoxyphenyl)propyl] amides as

highly potent MT_2-selective melatonin ligands. *Bioorg. Med. Chem. Lett.* **2010**, *20*, 2582−2585.
(46) Rivara, S.; Lodola, A.; Mor, M.; Bedini, A.; Spadoni, G.; Lucini, V.; Pannacci, M.; Fraschini, F.; Scaglione, F.; Sanchez, R. O.; Gobbi, G.; Tarzia, G. *N*-(substituted-anilinoethyl)amides: Design, synthesis, and pharmacological characterization of a new class of melatonin receptor ligands. *J. Med. Chem.* **2007**, *50*, 6618−6626.
(47) Bedini, A.; Spadoni, G.; Gatti, G.; Lucarini, S.; Tarzia, G.; Rivara, S.; Lorenzi, S.; Lodola, A.; Mor, M.; Lucini, V.; Pannacci, M.; Scaglione, F. Design and synthesis of *N*-(3,3-diphenylpropenyl)-alkanamides as a novel class of high-affinity MT_2-selective melatonin receptor ligands. *J. Med. Chem.* **2006**, *49*, 7393−7403.
(48) Mor, M.; Spadoni, G.; Di Giacomo, B.; Diamantini, G.; Bedini, A.; Tarzia, G.; Plazzi, P. V.; Rivara, S.; Nonno, R.; Lucini, V.; Pannacci, M.; Fraschini, F.; Stankov, B. M. Synthesis, pharmacological characterization and QSAR studies on 2-substituted indole melatonin receptor ligands. *Bioorg. Med. Chem.* **2001**, *9*, 1045−1057.
(49) Rivara, S.; Mor, M.; Silva, C.; Zuliani, V.; Vacondio, F.; Spadoni, G.; Bedini, A.; Tarzia, G.; Lucini, V.; Pannacci, M.; Fraschini, F.; Plazzi, P. V. Three-dimensional quantitative structure-activity relationship studies on selected MT_1 and MT_2 melatonin receptor ligands: requirements for subtype selectivity and intrinsic activity modulation. *J. Med. Chem.* **2003**, *46*, 1429−1439.
(50) Zlotos, D. P.; Attia, M. I.; Julius, J.; Sethi, S.; Witt-Enderby, P. A. 2-[(2,3-dihydro-1*H*-indol-1-yl)methyl]melatonin analogues: a novel class of MT_2-selective melatonin receptor antagonists. *J. Med. Chem.* **2009**, *52*, 826−833.
(51) Bedini, A.; Lucarini, S.; Spadoni, G.; Tarzia, G.; Scaglione, F.; Dugnani, S.; Pannacci, M.; Lucini, V.; Carmi, C.; Pala, D.; Rivara, S.; Mor, M. Toward the definition of stereochemical requirements for MT_2-selective antagonists and partial agonists by studying 4-phenyl-2-propionamidotetralin derivatives. *J. Med. Chem.* **2011**, *54*, 8362−8372.
(52) Koike, T.; Takai, T.; Hoashi, Y.; Nakayama, M.; Kosugi, Y.; Nakashima, M.; Yoshikubo, S.; Hirai, K.; Uchikawa, O. Synthesis of a novel series of tricyclic dihydrofuran derivatives: Discovery of 8,9-dihydrofuro[3,2-*c*]pyrazolo[1,5-*a*]pyridines as melatonin receptor (MT_1/MT_2) ligands. *J. Med. Chem.* **2011**, *54*, 4207−4218.
(53) Koike, T.; Hoashi, Y.; Takai, T.; Nakayama, M.; Yukuhiro, N.; Ishikawa, T.; Hirai, K.; Uchikawa, O. 1,6-Dihydro-2*H*-indeno[5,4-*b*]furan derivatives: Design, synthesis, and pharmacological characterization of a novel class of highly potent MT_2-selective agonists. *J. Med. Chem.* **2011**, *54*, 3436−3444.
(54) Poissonnier-Durieux, S.; Ettaoussi, M.; Pérès, B.; Boutin, J. A.; Audinot, V.; Bennejean, C.; Delagrange, P.; Caignard, D. H.; Renard, P.; Berthelot, P.; Lesieur, D.; Yous, S. Synthesis of 3-phenyl-naphthalenic derivatives as new selective MT_2 melatoninergic ligands. *Bioorg. Med. Chem.* **2008**, *16*, 8339−8348.
(55) Spadoni, G.; Balsamini, C.; Diamantini, G.; Tontini, A.; Tarzia, G.; Mor, M.; Rivara, S.; Plazzi, P. V.; Nonno, R.; Lucini, V.; Pannacci, M.; Fraschini, F.; Stankov, B. M. 2-N-acylaminoalkylindoles: Design and quantitative structure−activity relationship studies leading to MT2-selective melatonin antagonists. *J. Med. Chem.* **2001**, *44*, 2900−2912.
(56) Lucini, V.; Pannacci, M.; Scaglione, F.; Fraschini, F.; Rivara, S.; Mor, M.; Bordi, F.; Plazzi, P. V.; Spadoni, G.; Bedini, A.; Piersanti, G.; Diamantini, G.; Tarzia, G. Tricyclic alkylamides as melatonin receptor ligands with antagonist or inverse agonist activity. *J. Med. Chem.* **2004**, *47*, 4202−4212.
(57) Karageorge, G. N.; Bertenshaw, S.; Iben, L.; Xu, C.; Sarbin, N.; Gentile, A.; Dubowchik, G. M. Tetrahydroisoquinoline derivatives as melatonin MT_2 receptor antagonists. *Bioorg. Med. Chem. Lett.* **2004**, *14*, 5881−5884.
(58) Tsotinis, A.; Panoussopoulou, M.; Eleutheriades, A.; Davidson, K.; Sugden, D. Design, synthesis and melatoninergic activity of new unsubstituted and β,β'-difunctionalised 2,3-dihydro-1*H*-pyrrolo[3,2,1-*ij*]quinolin-6-alkanamides. *Eur. J. Med. Chem.* **2007**, *42*, 1004−1013.
(59) Zlotos, D. P. Recent progress in the development of agonists and antagonists for melatonin receptors. *Curr. Med. Chem.* **2012**, *19*, 3532−3549.
(60) Verdonk, M. L.; Berdini, V.; Hartshorn, M. J.; Mooij, W. T. M.; Murray, C. W.; Taylor, R. D.; Watson, P. Virtual screening using protein−ligand docking: Avoiding artificial enrichment. *J. Chem. Inf. Comput. Sci.* **2004**, *44*, 793−806.
(61) Wu, B.; Chien, E. Y. T.; Mol, C. D.; Fenalti, G.; Liu, W.; Katritch, V.; Abagyan, R.; Brooun, A.; Wells, P.; Bi, F. C.; Hamel, D. J.; Kuhn, P.; Handel, T. M.; Cherezov, V.; Stevens, R. C. Structures of the CXCR4 chemokine GPCR with small-molecule and cyclic peptide antagonists. *Science* **2010**, *330*, 1066−1071.
(62) Rivara, S.; Diamantini, G.; Di Giacomo, B.; Lamba, D.; Gatti, G.; Lucini, V.; Pannacci, M.; Mor, M.; Spadoni, G.; Tarzia, G. Reassessing the melatonin pharmacophore−enantiomeric resolution, pharmacological activity, structure analysis, and molecular modeling of a constrained chiral melatonin analogue. *Bioorg. Med. Chem.* **2006**, *14*, 3383−3391.
(63) 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.
(64) Sheridan, R. P.; Singh, S. B.; Fluder, E. M.; Kearsley, S. K. Protocols for bridging the peptide to nonpeptide gap in topological similarity searches. *J. Chem. Inf. Comput. Sci.* **2001**, *41*, 1395−1406.
(65) Kruskal, W. H.; Wallis, W. A. Use of ranks in one-criterion variance analysis. *J. Am. Stat. Assoc.* **1952**, *47*, 583−621.
(66) Sugden, D.; Pickering, H.; Teh, M. T.; Garratt, P. J. Melatonin receptor pharmacology: Toward subtype specificity. *Biol. Cell* **1997**, *89*, 531−537.
(67) McRobb, F. M.; Capuano, B.; Crosby, I. T.; Chalmers, D. K.; Yuriev, E. Homology modeling and docking evaluation of aminergic G protein-coupled receptors. *J. Chem. Inf. Model.* **2010**, *50*, 626−637.
(68) Vilar, S.; Karpiak, J.; Berk, B.; Costanzi, S. In silico analysis of the binding of agonists and blockers to the β_2-adrenergic receptor. *J. Mol. Graph. Modell.* **2011**, *29*, 809−817.
(69) Katritch, V.; Jaakola, V.-P.; Lane, J. R.; Lin, J.; Ijzerman, A. P.; Yeager, M.; Kufareva, I.; Stevens, R. C.; Abagyan, R. Structure-based discovery of novel chemotypes for adenosine A_{2A} receptor antagonists. *J. Med. Chem.* **2010**, *53*, 1799−1809.
(70) Evers, A.; Klabunde, T. Structure-based drug discovery using GPCR homology modeling: Successful virtual screening for antagonists of the alpha1A adrenergic receptor. *J. Med. Chem.* **2005**, *48*, 1088−1097.
(71) Evers, A.; Hessler, G.; Matter, H.; Klabunde, T. Virtual screening of biogenic amine-binding G-protein coupled receptors: Comparative evaluation of protein- and ligand-based virtual screening protocols. *J. Med. Chem.* **2005**, *48*, 5448−5465.
(72) Sastry, G. M.; Dixon, S. L.; Sherman, W. Rapid shape-based ligand alignment and virtual screening method based on atom/feature-pair similarities and volume overlap scoring. *J. Chem. Inf. Model.* **2011**, *51*, 2455−2466.
(73) Warne, T.; Moukhametzianov, R.; Baker, J. G.; Nehmé, R.; Edwards, P. C.; Leslie, A. G. W.; Schertler, G. F. X.; Tate, C. G. The structural basis for agonist and partial agonist action on a β_1-adrenergic receptor. *Nature* **2011**, *469*, 241−244.
(74) Lebon, G.; Warne, T.; Edwards, P. C.; Bennett, K.; Langmead, C. J.; Leslie, A. G. W.; Tate, C. G. Agonist-bound adenosine A_{2A} receptor structures reveal common features of GPCR activation. *Nature* **2011**, *474*, 521−525.
(75) Rasmussen, S. G. F.; Choi, H.-J.; Fung, J. J.; Pardon, E.; Casarosa, P.; Chae, P. S.; DeVree, B. T.; Rosenbaum, D. M.; Thian, F. S.; Kobilka, T. S.; Schnapp, A.; Konetzki, I.; Sunahara, R. K.; Gellman, S. H.; Pautsch, A.; Steyaert, J.; Weis, W. I.; Kobilka, B. K. Structure of a nanobody-stabilized active state of the β_2 adrenoceptor. *Nature* **2011**, *469*, 175−180.
(76) Katritch, V.; Abagyan, R. GPCR agonist binding revealed by modeling and crystallography. *Trends Pharmacol. Sci.* **2011**, *32*, 637−643.
(77) Rivara, S.; Vacondio, F.; Fioni, A.; Silva, C.; Carmi, C.; Mor, M.; Lucini, V.; Pannacci, M.; Caronno, A.; Scaglione, F.; Gobbi, G.; Spadoni, G.; Bedini, A.; Orlando, P.; Lucarini, S.; Tarzia, G. *N*-(Anilinoethyl)amides: design and synthesis of metabolically stable,

selective melatonin receptor ligands. *ChemMedChem* **2009**, *4*, 1746−1755.
(78) Dubocovich, M. L.; Masana, M. I.; Iacob, S.; Sauri, D. M. Melatonin receptor antagonists that differentiate between the human Mel_{1a} and Mel_{1b} recombinant subtypes are used to assess the pharmacological profile of the rabbit retina ML_1 presynaptic heteroreceptor. *Naunyn Schmiedebergs Arch. Pharmacol.* **1997**, *355*, 365−375.
(79) Spadoni, G.; Mor, M.; Tarzia, G. Structure-affinity relationships of indole-based melatonin analogs. *Biol. Signals Recept.* **1999**, *8*, 15−23.
(80) Depreux, P.; Lesieur, D.; Mansour, H. A.; Morgan, P.; Howell, H. E.; Renard, P.; Caignard, D. H.; Pfeiffer, B.; Delagrange, P.; Guardiola, B. Synthesis and structure−activity relationships of novel naphthalenic and bioisosteric related amidic derivatives as melatonin receptor ligands. *J. Med. Chem.* **1994**, *37*, 3231−3239.
(81) Spadoni, G.; Bedini, A.; Diamantini, G.; Tarzia, G.; Rivara, S.; Lorenzi, S.; Lodola, A.; Mor, M.; Lucini, V.; Pannacci, M.; Caronno, A.; Fraschini, F. Synthesis, enantiomeric resolution, and structure−activity relationship study of a series of 10,11-dihydro-5*H*-dibenzo-[a,d]cycloheptene MT_2 receptor antagonists. *ChemMedChem* **2007**, *2*, 1741−1749.
(82) Krüger, D. M.; Evers, A. Comparison of structure- and ligand-based virtual screening protocols considering hit list complementarity and enrichment factors. *ChemMedChem* **2010**, *5*, 148−158.
(83) Hu, G.; Kuang, G.; Xiao, W.; Li, W.; Liu, G.; Tang, Y. Performance evaluation of 2D fingerprint and 3D shape similarity methods in virtual screening. *J. Chem. Inf. Model.* **2012**, *52*, 1103−1113.
(84) Huang, N.; Shoichet, B. K.; Irwin, J. J. Benchmarking sets for molecular docking. *J. Med. Chem.* **2006**, *49*, 6789−6801.
(85) *Maestro*, version 9.2; Schrödinger, LLC: New York, 2011.
(86) *LigPrep*, version 2.5; Schrödinger, LLC: New York, 2011.
(87) Li, J.; Edwards, P. C.; Burghammer, M.; Villa, C.; Schertler, G. F. X. Structure of bovine rhodopsin in a trigonal crystal form. *J. Mol. Biol.* **2004**, *343*, 1409−1438.
(88) Standfuss, J.; Edwards, P. C.; D'Antona, A.; Fransen, M.; Xie, G.; Oprian, D. D.; Schertler, G. F. X. The structural basis of agonist-induced activation in constitutively active rhodopsin. *Nature* **2011**, *471*, 656−660.
(89) 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 β_2-adrenergic G protein-coupled receptor. *Science* **2007**, *318*, 1258−1265.
(90) Warne, T.; Serrano-Vega, M. J.; Baker, J. G.; Moukhametzianov, R.; Edwards, P. C.; Henderson, R.; Leslie, A. G. W.; Tate, C. G.; Schertler, G. F. X. Structure of a β_1-adrenergic G-protein-coupled receptor. *Nature* **2008**, *454*, 486−491.
(91) Jaakola, V.-P.; Griffith, M. T.; Hanson, M. A.; Cherezov, V.; Chien, E. Y. T.; Lane, J. R.; Ijzerman, A. P.; Stevens, R. C. The 2.6 angstrom crystal structure of a human A_{2A} adenosine receptor bound to an antagonist. *Science* **2008**, *322*, 1211−1217.
(92) Shimamura, T.; Shiroishi, M.; Weyand, S.; Tsujimoto, H.; Winter, G.; Katritch, V.; Abagyan, R.; Cherezov, V.; Liu, W.; Han, G. W.; Kobayashi, T.; Stevens, R. C.; Iwata, S. Structure of the human histamine H_1 receptor complex with doxepin. *Nature* **2011**, *475*, 65−70.
(93) 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.
(94) The UniProt Consortium.. Reorganizing the protein space at the Universal Protein Resource (UniProt). *Nucleic Acids Res.* **2012**, *40*, D71−D75.
(95) Mirzadegan, T.; Benkö, G.; Filipek, S.; Palczewski, K. Sequence analyses of G-protein-coupled receptors: similarities to rhodopsin. *Biochemistry* **2003**, *42*, 4310−4310.
(96) *Schrödinger Suite 2011 Protein Preparation Wizard, Epik*, version 2.2; Schrödinger, LLC, New York, 2011; *Impact*, version 5.7; Schrödinger, LLC: New York, 2011; *Prime*, version 3.0; Schrödinger, LLC: New York, 2011.
(97) Kaminski, G. A.; Friesner, R. A.; Tirado-Rives, J.; Jorgensen, W. L. Evaluation and reparametrization of the OPLS-AA force field for proteins via comparison with accurate quantum chemical calculations on peptides. *J. Phys. Chem. B* **2001**, *105*, 6474−6487.
(98) Spadoni, G.; Balsamini, C.; Diamantini, G.; Di Giacomo, B.; Tarzia, G.; Mor, M.; Plazzi, P. V.; Rivara, S.; Lucini, V.; Nonno, R.; Pannacci, M.; Fraschini, F.; Stankov, B. M. Conformationally restrained melatonin analogues: synthesis, binding affinity for the melatonin receptor, evaluation of the biological activity, and molecular modeling study. *J. Med. Chem.* **1997**, *40*, 1990−2002.
(99) *Phase*, version 3.3; Schrödinger, LLC: New York, 2011.
(100) *MacroModel*, version 9.9; Schrödinger, LLC: New York, 2011.
(101) Deng, Z.; Chuaqui, C.; Singh, J. Structural interaction fingerprint (SIFt): a novel method for analyzing three-dimensional protein−ligand binding interactions. *J. Med. Chem.* **2004**, *47*, 337−344.
(102) *Epik*, version 2.2; Schrödinger, LLC: New York, 2011.
(103) R Development Core Team. *R: A Language and Environment for Statistical Computing*, 2012. http://www.R-project.org (accessed April 4, 2013).