Protein-Based Virtual Screening of Chemical Databases. II. Are Homology Models of G-Protein Coupled Receptors Suitable Targets?

Caterina Bissantz,[1,2] **Philippe Bernard,**[1] **Marcel Hibert,**[1] **and Didier Rognan**[1*]

[1]*Laboratoire de Pharmacochimie de la Communication Cellulaire, UMR CNRS 7081, 74 route du Rhin, B.P. 24 , F-67401 Illkirch, France.*
[2]*Department of Applied Biosciences, ETH Zürich, Winterthurerstrasse 190, CH-8057 Zürich, Switzerland*

ABSTRACT The aim of the current study is to investigate whether homology models of G-Protein-Coupled Receptors (GPCRs) that are based on bovine rhodopsin are reliable enough to be used for virtual screening of chemical databases. Starting from the recently described 2.8 Å-resolution X-ray structure of bovine rhodopsin, homology models of an "antagonist-bound" form of three human GPCRs (dopamine D3 receptor, muscarinic M1 receptor, vasopressin V1a receptor) were constructed. The homology models were used to screen three-dimensional databases using three different docking programs (Dock, FlexX, Gold) in combination with seven scoring functions (ChemScore, Dock, FlexX, Fresno, Gold, Pmf, Score). Rhodopsin-based homology models turned out to be suitable, indeed, for virtual screening since known antagonists seeded in the test databases could be distinguished from randomly chosen molecules. However, such models are not accurate enough for retrieving known agonists. To generate receptor models better suited for agonist screening, we developed a new knowledge- and pharmacophore-based modeling procedure that might partly simulate the conformational changes occurring in the active site during receptor activation. Receptor coordinates generated by this new procedure are now suitable for agonist screening. We thus propose two alternative strategies for the virtual screening of GPCR ligands, relying on a different set of receptor coordinates (antagonist-bound and agonist-bound states). Proteins 2003; 50:5–25. © 2002 Wiley-Liss, Inc.

Key words: docking; scoring; structure-based ligand design; homology modeling; GPCRs

INTRODUCTION

High-throughput screening of chemical libraries is a well-established method for finding new lead compounds in drug discovery.[1] However, as the available databases get larger and larger, the costs of such screenings rise whereas the hit rates decrease. A possibility to avoid these problems is not to screen the whole database experimentally but only a small subset, which should be enriched in compounds that are likely to bind to the target. This preselection can be done by virtual screening (VS), a computational method to select the most promising compounds from an electronic database for experimental screening.[2] Virtual screening can be carried out by searching databases for molecules fitting either a known pharmacophore[3] or a three-dimensional (3-D) structure of the macromolecular target.[4] Pharmacophore-based screening has been extensively used over the last decade and has proven to be successful in many cases.[5] It has the advantage of being applicable to ligands for which the target 3-D structure is unknown, but it requires several 3- or 4-point pharmacophoric descriptions.[6] Protein-based virtual screening should be more efficient than the pharmacophore-based method since the protein environment of the ligand is taken into account. However, it still suffers from docking/scoring inaccuracies,[7,8] and it requires knowledge of the 3-D structure of the target. Therefore, with few exceptions,[9,10] it has only been applied to targets for which a high-resolution X-ray structure is known.[11] However, with the sequencing of the human genome, computational chemists will have to face an overwhelming number of potential targets for which no or very few experimental 3-D information is available. Therefore, it will be very important in the near future to be able to use not only X-ray or NMR structures, but also protein models for protein-based virtual screening of chemical libraries.

This is notably true for G-Protein coupled receptors (GPCRs) that represent one of the most important families of pharmaceutical targets.[12] Thus, there is high interest in developing new lead compounds for most human GPCRs. However, difficulties in obtaining significant amounts of pure and active recombinant GPCRs have rendered the determination of high-resolution GPCR X-ray structures very challenging. For long, GPCR models[13,14] have been

based on low-resolution maps of either bacteriorhodopsin[15] or bovine rhodopsin.[16,17] These models were useful for studying the functional architecture of GPCRs, but were not reliable enough for structure-based ligand design.[13,18] This situation hopefully changed with the recent determination of a high-resolution X-ray structure of bovine rhodopsin.[19] Based on this X-ray structure, it is possible to build refined models of other GPCRs, which might be close enough to their true 3-D structures for protein-based ligand design.

In the current study, we present two consecutive studies in which we investigated whether rhodopsin-based GPCR homology models are reliable enough for carrying out virtual screening of chemical libraries focused on either antagonists or agonist ligands of test GPCRs. We first constructed "antagonist-bound" molecular models of three human GPCRs (dopamine D3 receptor, acetylcholine muscarinic M1 receptor, vasopressin V1a receptor) for which a large amount of information is available.[20–22] The three models were evaluated in terms of their ability to identify known antagonists seeded into a database of randomly chosen "drug-like" compounds. A previously described virtual screening procedure[23] combining three different docking algorithms (Dock,[24] Gold,[25] FlexX[26]) in association with seven scoring functions (Dock,[24] Gold,[25] FlexX,[26] PMF,[27] ChemScore,[28] Fresno,[29] Score[30]) was used. Consensus scoring[31] was then applied to generate small subsets (hit lists) comprising only the top scorers common to two or three scoring lists. In the second part of this study, different models of an "agonist-bound: form of three human GPCRs (dopamine D3, β2-adrenergic, and δ-opioid receptors) were constructed and used for identifying true agonists embedded in test databases.

COMPUTATIONAL METHODS
Preparation of 3-D Databases

The Advanced Chemical Directory (ACD v.2000-1)[32] was filtered in order to eliminate unwanted compounds such as chemically reactive molecules, inorganic compounds, and molecules with unsuitable molecular weights (lower than 250, higher than 600) using an in-house SLN (Sybyl Line Notation)[33] script. These boundaries were applied to retrieve most of the "drug-like" compounds of the starting database and to avoid molecular weight biases when comparing known ligands with randomly chosen compounds. We have previously shown that larger molecules tend to be overestimated by free energy scoring functions as they can statistically provide more interactions to the target.[23] The randomly chosen compounds should, therefore, have a molecular weight not below that of the known ligands seeded into the test databases.

Out of the 75,000 remaining molecules, 990 were randomly chosen. The most likely protonation state of ionizable moieties (amines, amidines, carboxylic acids, etc.) present in this test database was then assigned using another SLN script. For each of the 990 compounds, three-dimensional coordinates and Gasteiger atomic charges[34] were then generated using Concord 4.0.[35] Final coordinates were stored in a multi mol2 file (a TRIPOS file format[33] for storing coordinates of several molecules in a single file). For each of the three receptors used for antagonist screening (D3, M1, and V1a receptors), a set of 10 known antagonists of the respective target was prepared (Fig. 1). For each of the receptors used for agonist screening (D3, β2, δ receptors), a set of 10 known full agonists was prepared (Fig. 2). The 10 reference ligands were chosen to be as different as possible in order to span the broadest chemical diversity for a given target. In order to keep our docking results as unbiased as possible, we did not include any of the ligands used for receptor refinement (see the following) in these test sets. Starting from Isis/Draw[36] 2D structures, a 3-D structure of each known ligand was generated with Gasteiger charges using the Concord conversion program. These new structures were then added to the random database to create six final libraries of 1,000 molecules.

Alignment of Amino Acid Sequences

The amino acid sequences of the five target receptors were retrieved from the Swiss-Prot database (accession numbers: human dopamine D3 receptor, P35462; human muscarinic acetylcholine M1 receptor, P11229; human vasopressin V1a receptor, P37288; human β2-adrenergic receptor, P07550; human δ-opioid receptor, P41143) and aligned to the sequence of bovine rhodopsin (accession number: P02699) using the ClustalW multiple alignment program.[37] A slow pairwise alignment using BLOSUM matrix series[38] (the matrix of the series used for the alignment is decided by ClustalW itself based on the similarity of the amino acid sequences to align and is thus not specified by the user) and a gap opening penalty of 15.0 were chosen for aligning the amino acid sequences to the sequence of bovine rhodopsin in two steps: (1) from the N-terminus to the first 5 residues of the third intracellular loop I3, (2) from the last 5 residues of the I3 loop to the C-terminus. Because the disulfide bridge occurring between the third transmembrane segment (TM III) and the second extracellular loop (E2) in the structure of bovine rhodopsin is conserved in all five test receptors, we manually adjusted the alignment of the extracellular loop E2 to align the respective cysteines.

Preparation of Starting Protein Coordinates

The 3-D models of the five test receptors were constructed by mutating the side chains of the amino acids in rhodopsin to the respective side chains in M1, D3, V1a, β2, and δ-opioid receptors. Standard geometries for the mutated side chains were given by the BIOPOLYMER module of SYBYL.[33] Whenever possible, the side chain torsional angles were kept to the values occurring in bovine rhodopsin. Otherwise, a short scanning of side chain angles was performed to remove steric clashes between the mutated side chain and the other amino acids. The third intracellular loop between helices 5 and 6, which shows a high degree of variability, was not included in any of the three models. This loop is responsible for G protein coupling but is not involved in direct interactions with the ligand.[39] We, therefore, assume that omitting this loop should not

influence our docking results. Insertions/deletions occurred only in loops but not in secondary structure elements (α-helix, β-sheet). The insertions/deletions in the loops were achieved through a simple knowledge-based loop search procedure using the LOOPSEARCH module of the SYBYL package.[33] In this procedure, a set of 1,478 high-resolution X-ray structures was searched for loops of similar length and similar distance between the Cα atoms of the residues delimiting the loop window. The loop showing the highest sequence identity and the lowest rms deviations was then selected for insertion in the model. Special caution had to be given to the second extracellular (E2) loop, which has been described in bovine rhodopsin to fold back over transmembrane helices,[19] and, therefore, limits the size of the active site. Hence, amino acids of this loop could be involved in direct interactions with the ligands. A driving force to this peculiar fold of the E2 loop might be the presence of a disulfide bridge between cysteines in TM III and E2. Since this covalent link is conserved in all receptors modeled in the current study, the E2 loop was modeled using a rhodopsin-like constrained geometry around the E2-TMIII disulfide bridge. After the heavy atoms were modeled, all hydrogen atoms were added, and the protein coordinates were then minimized with AMBER[40] using the AMBER95 force field.[41] The minimizations were carried out by 1,000 steps of steepest descent followed by conjugate gradient minimization until the rms gradient of the potential energy was less than 0.05 kcal/mol.Å. A twin cut-off (10.0, 15.0 Å) was used to calculate non-bonded electrostatic interactions at every minimization step, and the non-bonded pair-list was updated every 25 steps. A distance-dependent ($\epsilon = 4r$) dielectric function was used.

From here on, we will differentiate between protein coordinates generated by minimization with a receptor antagonist and coordinates generated with a bound agonist. We assume that minimization with an antagonist gives a set of protein coordinates that represents the receptor in its "antagonist-bound" state (ground state), whereas minimization with an agonist is assumed to result in a representation of an "agonist-bound" state (activated state) of the receptor.

Modeling "antagonist-bound" GPCR Models

To obtain an antagonist-bound model of the D3, M1, and V1a receptors, the initial models of the receptors were refined by the following procedure: A known antagonist was first manually docked into each active site according to experimental binding data (D3: BP-897, M1: pirenzepine, V1a: conivaptan; Fig. 1).[42–47] The starting 3-D conformations of the ligands were generated with Concord.[35] The D3 and M1 ligand were orientated into its respective active site so that a protonated nitrogen atom was located 3 Å away from AspIII:11. Afterwards, the orientation of the ligand as well as torsional angles were manually adjusted in order to reproduce expected interactions with receptor side chains.[42–47] For the hydrophobic V1a antagonists, there is no directed interaction known compared to the salt bridge of the D3 and M1 ligands with

their respective receptor. We, therefore, orientated the V1a antagonist conivaptan so that its shape optimally fits in that of the binding pocket.

The resulting protein-ligand complexes were then refined by minimization using the above-described AMBER parameters. Removing the ligand atoms from the minimized complexes finally yielded for each of the three receptor one set of coordinates for an antagonist-bound structure.

Modeling Agonist-Bound GPCR Models

3-D models of the D3, β2, and δ-opioid receptors were generated for future agonist screening. These three receptors were selected as test cases since several full agonists are already known for each of these GPCRs. Different agonist-bound GPCR models were set up as we speculate that the "activated state" of GPCRs is conformationally more flexible than the antagonist-bound ground state. Two agonist-bound models were built for each receptor using the same refinement procedure as for the antagonist-bound models but now using full agonists rather than antagonists for receptor refinement. The agonists used for the refinement (Fig. 2) were apomorphine and pergolide (D3 receptor), epinephrine, and nylidrine (β2 receptor), SNC-80, and TAN67(δ-opioid receptor) .

An alternative activated-state model was generated by substituting the single ligand-biased receptor refinement protocol with the following procedure: (1) rotation of TM VI by 30° anticlockwise around its axis (when viewed from the extracellular side), (2) manual docking of one known full agonist (D3: apomorphine, β2: epinephrine, δ-opioid: SNC-80; Fig. 2)into the active site according to known experimental data,[42–45,48–55] (3) definition of a "pharmacophore fingerprint" by superimposing several structurally unrelated full agonists (Fig. 2) onto the reference ligand docked in step (2) using standard parameters of the FlexS program,[56] (4) if necessary, manual adjustment of the FlexS alignments to position the ligands within the active site according to known experimental data, (5) multi-ligands biased receptor minimization with AMBER 5.0[40] as previously described, but keeping the ligands coordinates fixed at the alignment-derived positions ("belly" option of AMBER). The starting 3-D coordinates of the ligands used in this procedure were generated with Concord.

Dock4.01 Docking

A solvent-accessible surface of each receptor active site was generated using a 1.4-Å probe radius and used to generate a set of overlapping spheres. To compute interaction energies, a 3-D grid of 0.35 Å resolution was centered on the active sites. Energy scoring grids were obtained by using an all atom model and a distance-dependent dielectric function ($\epsilon = 4r$) with a 10-Å cut-off. Amber95 atomic charges[41] were assigned to all protein atoms. Database molecules were then docked into the protein active site by matching sphere centers with ligand atoms. A flexible docking of all molecules (peripheral search and torsion drive) with subsequent
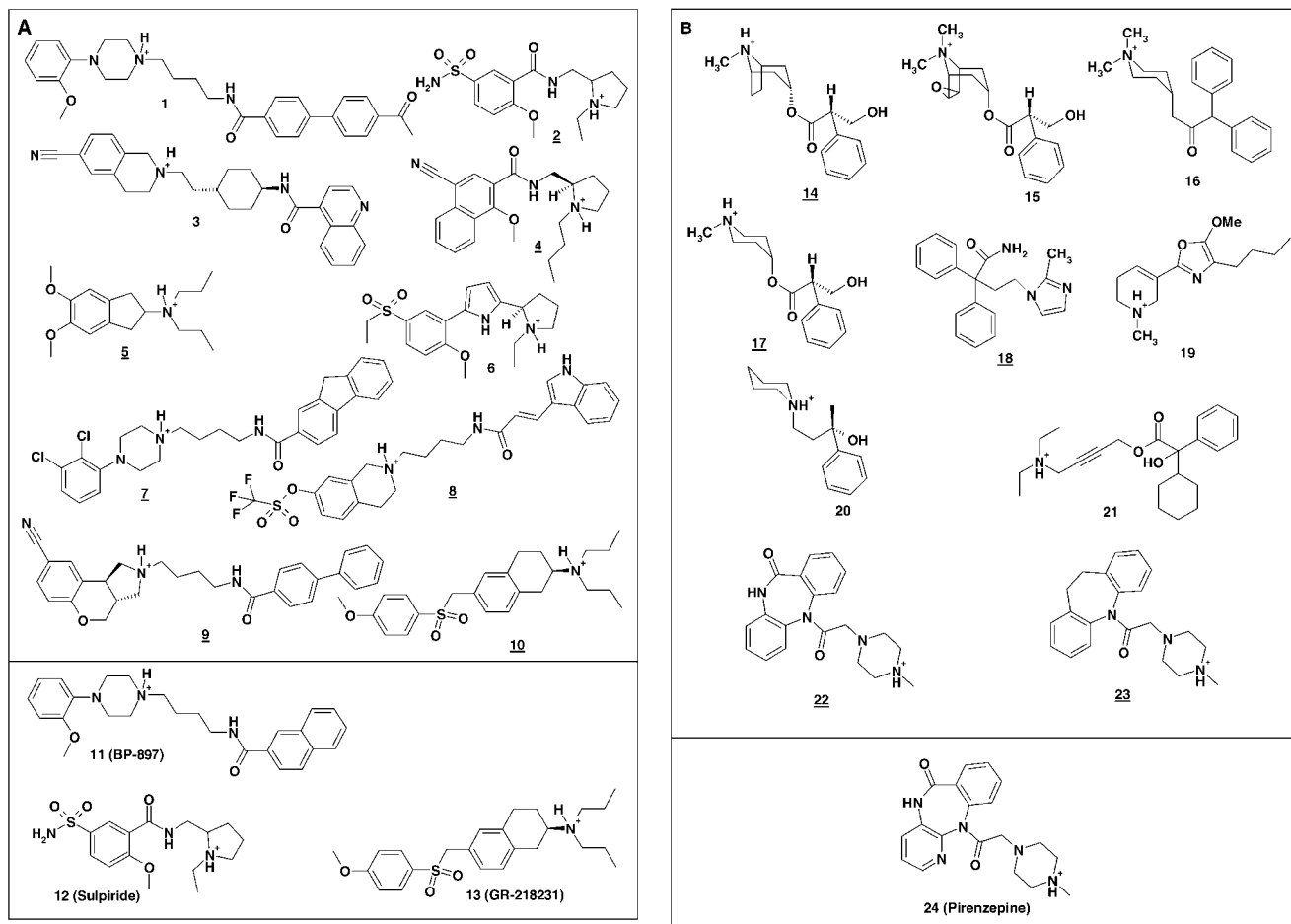
Fig. 1. Structures of receptor antagonists used in the test databases (top boxes) and in the minimization protocol (bottom boxes). **A:** Dopamine D3 antagonists. **B:** Muscarinic M1 antagonists. **C:** Vasopressin V1a antagonists. Pirenzepine and conivaptan were used for refining the M1 and V1a antagonist-bound receptor models. For the dopamine D3 receptor, three independent minimizations were carried out using BP-897, sulpiride, and GR-218231, thus leading to three sets of coordinates for the antagonist-bound form of the D3 receptor. Underlined numbers represent true antagonists predicted as virtual hits by the best virtual screening strategy.

minimization was performed as follows: (1) automatic selection and matching of an anchor fragment within a maximum of 100 orientations, (2) iterative growing of the ligand using at least 30 conformations (peripheral seeds) for seeding the next growing stage with assignment of energy-favored torsion angles, (3) simultaneous relaxation of the base fragments as well as of all peripheral segments and final relaxation of the entire molecule. Orientations/conformations were relaxed in 100 cycles of 100 simplex minimization steps to a convergence of 0.1 kcal/mol. The top solution corresponding to the best Dock energy or chemical score of each ligand was then stored in a single multi mol2 file.

FlexX1.8 Docking

Standard parameters of the FlexX1.9 program as implemented in the 6.62 release of the SYBYL package were used for iterative growing and subsequent scoring of FlexX poses. Only the top solution was retained and further stored in a single mol2 file.

Gold1.1 Docking

For each of the 10 independent genetic algorithm (GA) runs, a maximum number of 1,000 GA operations was performed on a single population of 50 individuals. Operator weights for crossover, mutation, and migration were set to 100, 100, and 0, respectively. To allow poor non-bonded contacts at the start of each GA run, the maximum distance between hydrogen donors and fitting points was set to 5 Å, and non-bonded van der Waals energies were cut off at a value equal to k_{ij} (well depth of the van der Waals energy for the atom pair i,j). To further speed up the calculation, the GA docking was stopped when the top three solutions were within 1.5 Å rmsd. If this criterion is met, we can assume that these top solutions represent a reproducible pose for the ligand.

It should be mentioned that all docking tests were applied using fast virtual screening parameters (pace of approximately 60–90 sec/ligand on a standard UNIX workstation) to closely match database screening condi-
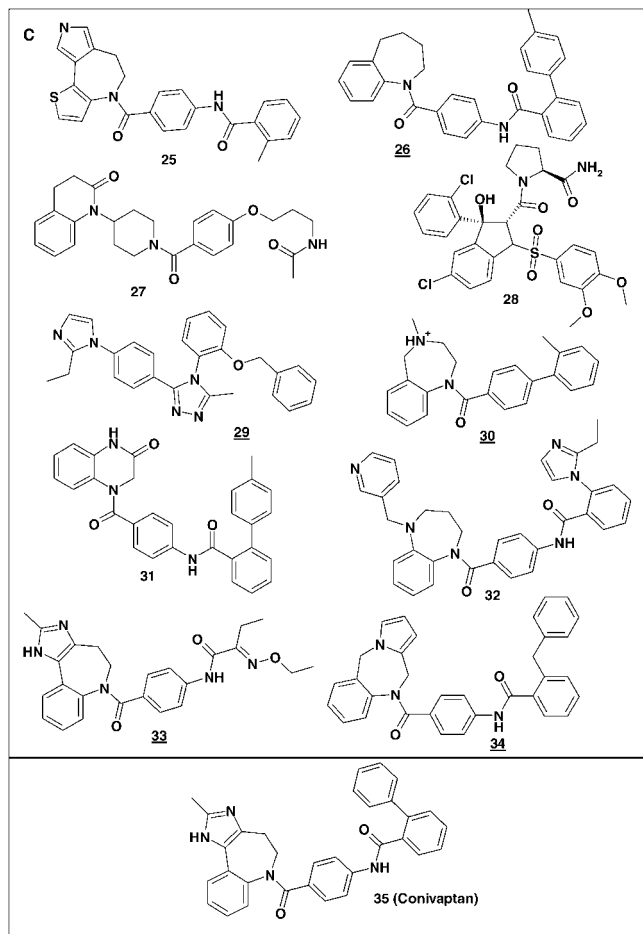
Figure 1. (Continued)

tions. This strategy explains why the top-ranked pose only was selected for further rescoring.

Consensus Scoring: Definition of a Hit List

Docked poses were rescored using the CScore™ module of SYBYL6.62 that comprises the following scoring functions: ChemScore, Dock, FlexX, Gold, and Pmf. It should be noted that the original Dock4.0 and Gold scores as implemented in the respective docking program significantly differ from the respective scores calculated by CScore™. The scoring function of Cscore corresponding to the native scoring function of the applied docking algorithm (Dock, FlexX, Gold) was always discarded. Two additional scoring functions, Fresno[29] and Score,[30] were used as part of in-house SPL (Sybyl Programming Language)[33] scripts. For rescoring, no relaxation of the bound ligands was performed. Rescoring the whole database with all scoring functions took about 30 min on a standard UNIX workstation.

Having rescored the docked poses with these seven scoring functions, we then carried out the "consensus scoring." For the D3, V1a, and β2 receptor screenings, we defined the top 15% of the individual ranking lists as top scorers. For the δ-opioid receptors, the top 20% scorers

were selected. Last, the top 25% scorers were retrieved when screening the M1 receptor A pairwise comparison of the lists of top scorers (21 combinations) yielded the consensus lists (from here on also called "hit lists") that consequently contain only those compounds that were ranked among the top scorers from the two compared scoring functions. We, furthermore, generated the consensus lists comparing the top scorers of three single ranking lists (giving 35 additional consensus lists).

The reason why we had to adjust the definition of "top scorers" from one GPCR to another resides in the expected variation of accuracy of our homology models. The number of top scorers used for consensus scoring represents for each receptor the best possible compromise between the total size of the virtual hit lists and the number of true ligands it contains (see Quantitative Descriptors of the Hit Lists).

Quantitative Descriptors of the Hit Lists

Two properties of every generated hit list were computed: the hit rate and the yield. The hit rate (or purity) describes the accuracy of the virtual screening and is defined as the percentage of known true ligands in the hit list (Eqn. 1). A random screening of the full database would thus have given a hit rate of 1% (10 true hits out of 1,000 molecules). The yield is defined as the percentage of true hits retrieved by our virtual screening protocol (Eqn. 2). It describes the sensitivity of the screening.

$$\text{Hit Rate} = (t/l) \times 100 \qquad (1)$$

$$\text{Yield} = (t/T) \times 100 \qquad (2)$$

t = number of true hits in the hit list
T = number of true hits in the full database
l = number of compounds in the hit list

RESULTS
Building Starting GPCR Models

The amino acid sequences of the target receptors were aligned to the sequence of the bovine rhodopsin template (Fig. 3), excluding the third intracellular loop, which shows too high variability in amino acid composition and length. The sequence identity between the transmembrane (TM) domains of bovine rhodopsin and the human dopamine D3, muscarinic M1, Vasopressin V1a, β2-adrenergic, and δ-opioid receptor is 29, 21, 21, 22, and 23%, respectively. The alignment is in agreement with known structural features of GPCRs.[57] Amino acids known to be highly conserved in the family of the (rhodopsin-like) GPCRs (AsnI:21, AspII:13, ArgIII:29, TrpIV:11, ProV:16, ProVI:21, NPxxY in TM VII; A:n, position n of the transmembrane helix A) are aligned against each other. The disulfide bridge in bovine rhodopsin is conserved in all test receptors, and the respective cysteines are also aligned to each other. It should be emphasized that we did not model the loops exactly with the deletions/insertions as indicated in the alignment since there are obviously several ways of aligning these very variable sequences. We rather carried out a loop search to make sure that our loops have
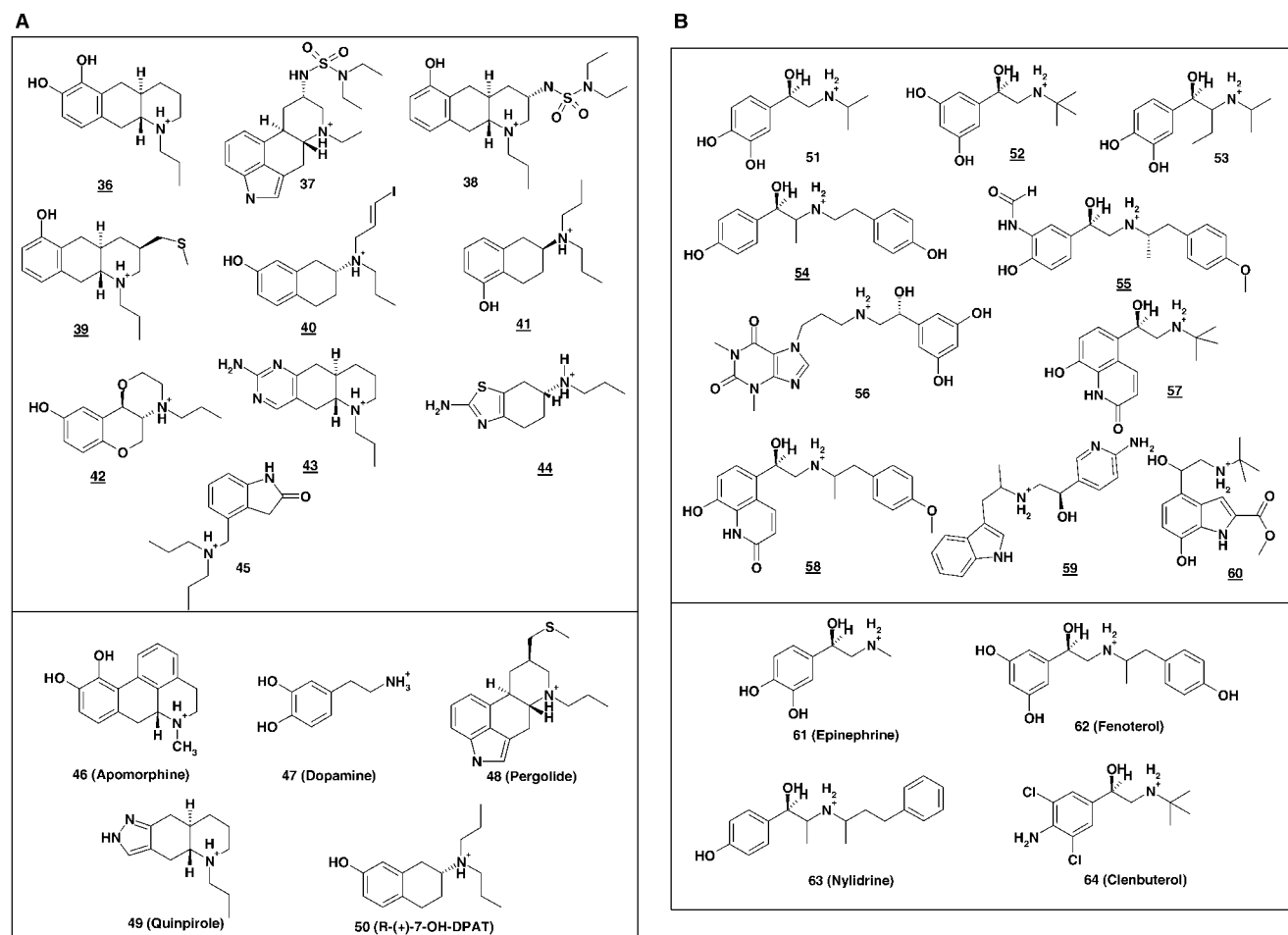
Fig. 2. Structures of receptor agonists used in the test databases (top boxes) and in the minimization protocol (bottom boxes). **A:** Dopamine D3 agonists. **B:** β2-adrenergic agonists. **C:** δ-opioid agonists. Underlined numbers represent true agonists predicted as virtual hits by the best virtual screening. Single-ligand minimized receptor models were obtained with apomorphine and pergolide (D3 receptor), epinephrine and nylidrine (β2 receptor), and SNC-80 and TAN67 (δ-opioid receptor). Pharmacophore-based receptor models were obtained by first docking one reference agonist (apomorphine, epinephrine, and SNC-80) into the TM cavity and then superimposing other ligands (bottom boxes) with FlexS onto the receptor-bound reference agonist.

conformations similar to loops in X-ray structures that have similar length and similar distance between the Cα atoms of the residues delimiting the loop.

Based on this alignment, we mutated the side chains of the amino acids in the X-ray structure of bovine rhodopsin to the respective side chains in the new receptor (see Computational Methods). Deletions occurring in the E2 loop of the three GPCRs (see Fig. 3) could be accommodated easily without modifying the respective orientations of the seven α-helices. Minimizing the resulting structures gave for each receptor one initial model. Any amino acid side chain directed towards the inner part of the transmembrane ligand binding domain was selected as a potential anchor and thus selected as part of a putative binding site (Fig. 3).

Building Antagonist-Bound Models of D3, M1, and V1a Receptors

The X-ray structure of bovine rhodopsin corresponds to its ground state[19] in which retinal is covalently bound

(antagonist-bound conformation). Homology models constructed from this structure are not directly suitable for docking purpose for two main reasons: (1) depending on the type of bound ligand (agonist, antagonist), GPCRs are known to adopt different conformational states,[58–60] (2) retinal is a rather flat molecule, and thus the modeled binding sites are too narrow—especially since the E2 loop in bovine rhodopsin folds deeply into the center of the receptor to completely enclose retinal—for most of the known GPCR ligands to fit in. We, therefore, have to "expand" the active site in our models somewhat by pushing the E2 loop out of the TM cavity center. Hence, preliminary attempts to automatically dock known ligands into the starting models directly obtained from rhodopsin (minimized in absence of any ligand) usually failed whatever the docking tool used (data not shown).

Thus, we feel it is necessary to refine the receptor model by minimization in complex with a known ligand to achieve expansion of the TM cavity. Minimization with an antagonist is supposed to give a model close to an antago-
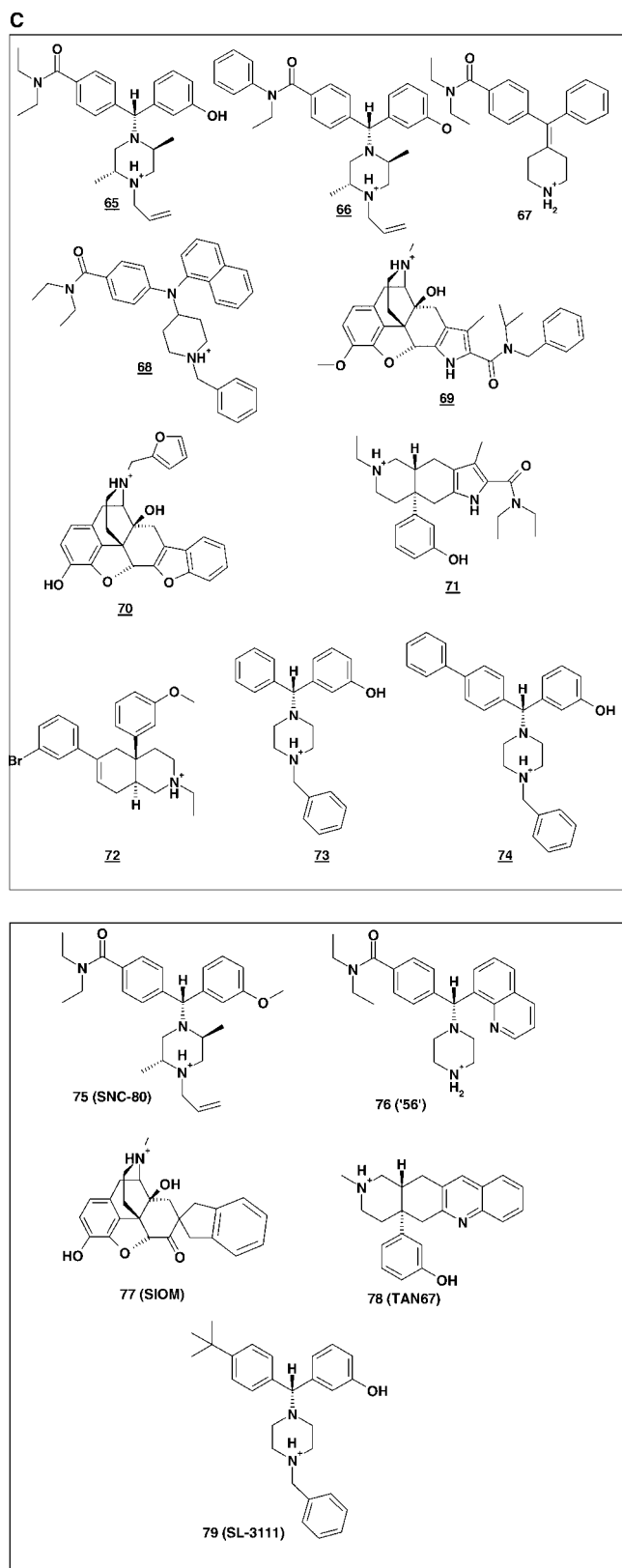
Figure 2.   (Continued)

nist-bound state (ground state) of the receptor, whereas minimization with a full agonist should lead to a model closer to the agonist-bound state (activated state). For this purpose, we have chosen ligands for which a large amount of experimental data (location of the binding site) was available.

To get 3-D coordinates of the antagonist-bound state of the D3, M1, and V1a receptors, we chose BP-897 (D3 receptor), pirenzepine (M1 receptor), and conivaptan (V1a receptor), docked them according to known experimental data into their respective receptors,[42−47] and energy minimized the complexes. These reference ligands were chosen in the light of their conformational freedom (as rigid as possible) and of known experimental data to guide us in the initial docking. The D3 antagonist BP-897 has been oriented so that its protonated nitrogen forms a salt bridge with AspIII:11. Additionally, the amide group of BP-897 interacts with ThrVII:7. The basic nitrogen of the muscarinic antagonist pirenzepine also forms a salt bridge with AspIII:11. The lactam group interacts with AsnVI:23, and its aromatic nitrogen with ThrV:8. The V1a antagonist conivaptan is orientated so that the benzazepine moiety lies in the pocket between TM III, V, and VI, whereas the diphenyl group occupies the pocket between TM I, II, and VII. It should be noted that the three docking modes reported herein are compatible with all known experimental data.[42−47] Removing the bound ligand from the energy minimized complex (refined as previously described for the starting models) gave a set of protein coordinates (inactive state) for each receptor.

Building Agonist-Bound Models of the D3, β2, and δ-Opioid Receptors

Two agonist-bound models were built for each GPCR using the same strategy as that used for the antagonist-bound models. The two dopamine agonists apomorphine (S1 model) and pergolide (S2 model) were manually docked into the active site with their protonated nitrogens within salt bridge distance (3 Å) from AspIII:11. Additionally, the two phenolic groups of apomorphine and the indolic nitrogen atom of pergolide H-bond to three serines on TM V (SerV:8, SerV:9, SerV:12). For generating the β2-models, epinephrine (S1 model) and nylidrine (S2 model) were used as reference agonists. They were also docked with the basic nitrogens forming a salt bridge with AspIII:11, and the phenolic groups interacting with the three serines on TM V. δ-opioid receptor models were generated by docking SNC-80 (S1 model) and TAN67 (S2 model) into the active site, with the protonated amine establishing a salt bridge with AspIII:11, the phenol rings lying in the pocket between TMs III,V,VI, whereas the diethylamide (SNC-80) and the quinoline ring (TAN67) were embedded in the pocket between TMs VI and VII. All complexes were then refined using the above-described AMBER refinement protocol.

For each of the three receptors, we furthermore generated one activated-state model using a multi-ligand based modeling procedure (M models). This protocol requires first defining a pharmacophoric fit of few reference ago-
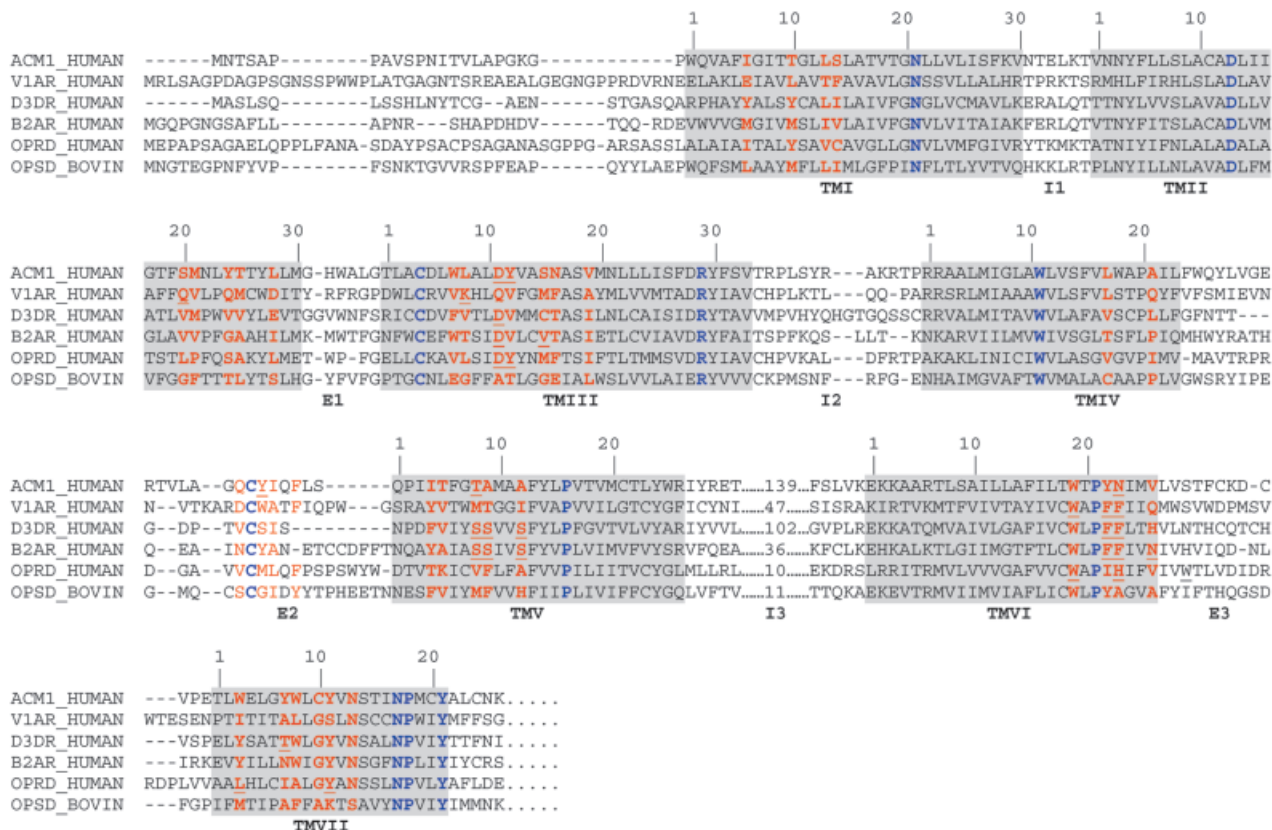
```
                    1        10        20        30        1        10
ACM1_HUMAN   ------MNTSAP---------PAVSPNITVLAPGKG------------PWQVAFIGITTGLLSLATVTGNLLVLISFKVNTELKTVNNYFLLSLACADLII
V1AR_HUMAN   MRLSAGPDAGPSGNSSPWWPLATGAGNTSREAEALGEGNGPPRDVRNEELAKLEIAVLAVTFAVAVLGNSSVLLALHRTPRKTSRMHLFIRHLSLADLAV
D3DR_HUMAN   ------MASLSQ---------LSSHLNYTCG--AEN------STGASQARPHAYYALSYCALILAIVFGNGLVCMAVLKERALQTTTNYLVVSLAVADLLV
B2AR_HUMAN   MGQPGNGSAFLL--------APNR---SHAPDHDV------TQQ-RDEVWVVGMGIVMSLIVLAIVFGNVLVITAIAKFERLQTVTNYFITSLACADLVM
OPRD_HUMAN   MEPAPSAGAELQPPLFANA-SDAYPSACPSAGANASGPPG-ARSASSLALAIAITTALYSAVCAVGLLGNVLVMFGIVRYTKMKTATNIYIFNLALADALA
OPSD_BOVIN   MNGTEGPNFYVP--------FSNKTGVVRSPFEAP------QYYLAEPWQFSMLAAYMFLLIMLGFPINFLTLYVTVQHKKLRTPLNYILLNLAVADLFM
                                                         TMI                     I1        TMII

              20        30        1        10        20        30        1        10        20
ACM1_HUMAN   GTFSMNLYTTYLLMG-HWALGTLACDLWLALDYVASNASVMNLLLISFDRYFSVTRPLSYR---AKRTPRRAALMIGLAWLVSFVLWAPAILFWQYLVGE
V1AR_HUMAN   AFFQVLPQMCWDITY-RFRGPDWLCRVVKHLQVFGMFASAYMLVVMTADRYIAVCHPLKTL---QQ-PARRSRLMIAAAWVLSFVLSTPQYFVFSMIEVN
D3DR_HUMAN   ATLVMPWVVYLEVTGGVWNFSRICCDVFVTLDVMMCTASILNLCAISIDRYTAVVMPVHYQHGTGQSSCRRVALMITAVWVLAFAVSCPLLFGFNTT---
B2AR_HUMAN   GLAVVPFGAAHILMK-MWTFGNFWCEFWTSIDVLCVTASIETLCVIAVDRYFAITSPFKQS--LLT--KNKARVIILMVWIVSGLTSFLPIQMHWYRATH
OPRD_HUMAN   TSTLPFQSAKYLMET-WP-FGELLCKAVLSIDYYNMFTSIFTLTMMSVDRYIAVCHPVKAL---DFRTPAKAKLINICIWVLASGVGVPIMV-MAVTRPR
OPSD_BOVIN   VFGGFTTTLYTSLHG-YFVFGPTGCNLEGFFATLGGEIALWSLVVLAIERYVVVCKPMSNF---RFG-ENHAIMGVAFTWVMALACAAPPLVGWSRYIPE
                     E1              TMIII                       I2                TMIV

                1        10        20                    1        10        20
ACM1_HUMAN   RTVLA--GQCYIQFLS------QPIITFGTAMAAFYLPVTVMCTLYWRIYRET......139...FSLVKEKKAARTLSAILLAFILTWTPYNIMVLVSTFCKD-C
V1AR_HUMAN   N--VTKARDCWATFIQPW---GSRAYVTWMTGGIFVAPVVILGTCYGFICYNI......47.....SISRAKIRTVKMTFVIVTAYIVCWAPFFIIQMWSVWDPMSV
D3DR_HUMAN   G--DP--TVCSIS---------NPDFVIYSSVVSFYLPFGVTVLVYARIYVVL......102...GVPLREKKATQMVAIVLGAFIVCWLPFFLTHVLNTHCQTCH
B2AR_HUMAN   Q--EA--INCYAN-ETCCDFFTNQAYAIASSIVSFYVPLVIMVFVYSRVFQEA......36.....KFCLKEHKALKTLGIIMGTFTLCWLPFFIVNIVHVIQD-NL
OPRD_HUMAN   D--GA--VVCMLQFPSPSWYW-DTVTKICVFLFAFVVPILIITVCYGLMLLRL......10.....EKDRSLRRITRMVLVVVGAFVVCWAPIHIFVIWTLVDIDR
OPSD_BOVIN   G--MQ--CSCGIDYYTPHEETNNESFVIYMFVVHFIIPLIVIFFCYGQLVFTV......11.....TTQKAEKEVTRMVIIMVIAFLICWLPYAGVAFYIFTHQGSD
                 E2            TMV                I3            TMVI              E3

                1        10        20
ACM1_HUMAN   ---VPETLWELGYWLCYVNSTINPMCYALCNK.....
V1AR_HUMAN   WTESENPTITITALLGSLNSCCNPWIYMFFSG.....
D3DR_HUMAN   ---VSPELYSATTWLGYVNSALNPVIYTTFNI.....
B2AR_HUMAN   ---IRKEVYILLNWIGYVNSGFNPLIYIYCRS.....
OPRD_HUMAN   RDPLVVAALHLCIALGYANSSLNPVLYAFLDE.....
OPSD_BOVIN   ---FGPIFMTIPAFFAKTSAVYNPVIYIMMNK.....
                  TMVII
```

Fig. 3. Amino acid sequence alignment of 5 GPCRs (ACM1_HUMAN: human acetylcholine muscarinic M1, D3DR_HUMAN: dopamine D3, V1AR_HUMAN: human vasopressin V1a, OPRD_HUMAN: human δ-opioid receptor, B2AR_HUMAN: human δ-opioid receptor) with bovine rhodopsin (OPSD_BOVIN). Transmembrane domains (TM I to TM VII) are boxed. E1-3 and I1-3 indicate the positions of extracellular and intracellular loops, respectively. Residues colored in red delimit a putative binding site for GPCR ligands. Amino acids indicated in blue are conserved in the GPCR family. For facilitating the comparison of different receptors, transmembrane helical sequences (TM I to TM VII), assigned as in the X-ray structure of bovine rhodopsin,[29] are numbered from the N- to the C-terminal end. Numbers inserted in the I3 loop describe the number of residues omitted in the current study. The residues referenced in the text are underlined. The numbering of residues referenced in the text is as follows: A:n indicates position n of the transmembrane helix A.

nists. For defining a dopamine D3 agonist pharmacophore, 5 full agonists [Fig. 2(A)] were selected: two catecholamines (dopamine, apomorphine), one ergoline (pergolide), one pyrazoloquinoline (quinpirole), and one 2-aminotetraline derivative (7-OH-DPAT). The protonated amine common to all D3 ligands is supposed to form a salt bridge with the highly conserved AspIII:11, and the catechol moiety of the catecholamines or its bioisostere in other ligand classes are expected to H-bond to three serine side chains on helix 5 (SerV:8, SerV:9, SerV:12).[43] The phenylalanine residues PheVI:22 and PheVI:23 also seem to be important for agonist binding, probably via π-π interactions with the aromatic catechol ring or its bioisostere.[45] We thus manually docked a rigid molecule first (apomorphine) as previously described and used this structural template for superimposing the other full dopamine D3 receptor agonists by fitting three chemical features (protonated amine, aromatic ring, H-bond donor moiety) with the FlexS program.[56] The FlexS-aligned conformations were slightly modified in order to correctly orientate the N-alkyl substituents and the thioether chain of pergolide within the active site. In the proposed alignment, protonated amines and H-bond forming groups of the aromatic rings are superim-

posed [Fig. 4(A)]. The agonists are orientated in the active site so that the protonated amines form salt bridges with AspIII:11; the para-hydroxyl group of apomorphine and dopamine H-bonds to SerV:12, the meta-hydroxyl substituent of apomorphine, dopamine, and 7-OH-DPAT to SerV:8 and SerV:9. The indolic nitrogen atom in the ergoline derivatives binds to SerV:12. Since TM VI was previously rotated by 30°, the aromatic ring of PheVI:22 lies now in a plane parallel to that of the aromatic rings of the agonists and can thus form π-π interactions with the ligands. This π-π interaction cannot be formed when docking the ligands directly in the rhodopsin-derived model as it was done for generating the models S1 and S2.

Like the D3 receptor ligands, all β2 agonists have a protonated nitrogen [Fig. 2(B)] expected to form a salt bridge with AspIII:11,[49] and the catechol moiety (or its bioisostere) is probably H-bonded to three serines on helix 5 (SerV:8, SerV:9, SerV:12).[48,52] Additionally, the β2 agonists possess a β-hydroxyl group expected to H-bond to AsnVI:26.[51] The phenylalanine residues PheVI:22 and PheVI:23 have also been proposed to be involved in ligand binding.[50] After manually docking epinephrine according to the expected interactions, other β2 full agonists were
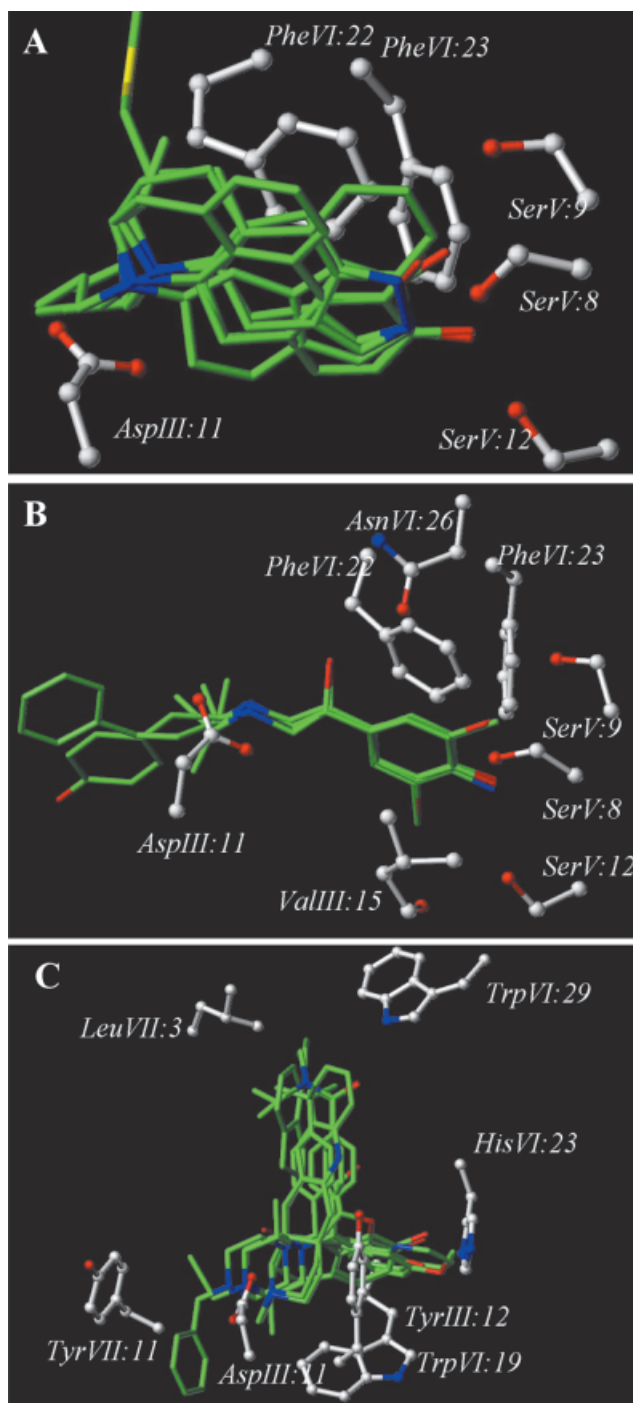
Fig. 4. Receptor-based alignment of the ligands used for receptor refinement of model M of the dopamine D3 receptor (**A**), the β2-adrenergic receptor (**B**), and the δ-opioid receptor (**C**). Ligands are shown as capped sticks, whereas amino acid side chains of the receptor are indicated in ball-and-stick representation. The color coding is the following: carbon (agonists, green; receptor, white); oxygen, red; nitrogen, blue; sulfur, yellow.

superimposed onto the reference agonist using FlexS [Fig. 2(B)]. This alignment is straightforward due to the common phenyl-2-aminoethanol core of the β2 agonists [Figs. 2(B), 4(B)]. We only had to adjust the conformations of the

large N-aralkyl substituents in fenoterol and nylidrine so that they do not bump into protein atoms. The ligands are orientated in the active site so that the protonated amines form salt bridges with AspIII:11 [Fig. 4(B)]. The para-hydroxyl group of epinephrine and nylidrine H-bonds to the side chain of SerV:12. The meta-hydroxyl of epinephrine and one of the two phenolic moieties of fenoterol H-bond to the side chains of SerV:8 and SerV:9, the other hydroxyl group of fenoterol to either the side chain of SerV:12 or to the backbone carbonyl of ValIII:15. The amino group of clenbuterol is proposed to interact with SerV:12. All aromatic catechol moieties can again form a π-π interaction with PheVI:22. Last, the large hydrophobic residues on the protonated amines of fenoterol and nylidrine occupy the pocket between TMs I, II, and VII [Fig. 4(B)]. As for the D3 receptor, the interactions between the ligands and the important residues on TM VI (PheVI:22, AsnVI :26) cannot be formed in the single ligand-biased models (S1 and S2 models).

Last, a pharmacophoric alignment was defined for δ-opioid receptor agonists from 5 compounds [Fig. 2(C)]: two compounds of the SNC series (SNC-80 and 56[61]), one opiate (SIOM) and one simplified opiate (TAN67), as well as one true non-peptide[62] compound (SL-3111). The δ-agonists are regarded as consisting of two parts: the "address" and the "message" part.[63] The address part is generally a large hydrophobic residue that is supposed to bind in the pocket between TMs VI and VII; important residues hereby are TrpVI:29 and LeuVII:3.[53,55] The message part is expected to bind in the central cavity between TMs III, V, and VI; important residues are AspIII:11, TyrIII:12, TrpVI:19, and TyrVII:11.[54] FlexS was again used to superimpose the agonists onto the manually docked reference ligand (SNC-80). The alignment of the SNC derivative 56 onto SNC-80 is straightforward. The two compounds of the SNC series are orientated in the active site so that the protonated amine can form a salt bridge with AspIII:11; the N,N-diethylamide (address-part), which is orientated out-of-plane with respect to the benzene ring, lies in the pocket between TMs VI and VII, with the carbonyl pointing to TrpVI:29 and one of the two ethyl groups interacting with LeuVII:3 [Fig. 4(C)]. The methoxyphenyl and quinoline moiety, respectively, fit in the central pocket between TMs III, V, and VI, with the methoxy group of SNC-80 pointing away from HisVI:23. Though structurally very similar, the true non-peptide compounds are expected to bind in a different mode than the SNC derivatives.[62] Whereas O-methylation is tolerated in the SNC-series, it results in a loss of activity of the true non-peptides. We, therefore, adjusted the FlexS fitted solution of SL-3111 so that its phenolic OH group can form a H-bond to HisVI:23. Besides this small modification of the automated alignment, SL-3111 is docked similarly to the SNC compounds, with the protonated amine superimposed onto the amine of SNC-80 and the t-butyl group onto the N,N-diethylamide. The N-benzyl substituent of SL-3111 occupies the pocket between TMs II and VII. The FlexS solution for SIOM was not accepted because the address part was superimposed onto the phenoxy ring of SNC-80, thus lying

in the central cavity rather than in the pocket between TMs VI and VII. We, therefore, manually superimposed SIOM onto the other agonists. The address part of the opiate SIOM overlaps now with the N,N-diethylamide of SNC-80, its phenolic OH with the one of SL-3111, thus interacting with HisVI:23 [Fig. 4(C)]. The FlexS alignment of the simplified opiate TAN67 to the reference agonist SNC-80 was modified in order to maximize overlap with the opiate derivative (SIOM) so that its perhydroisoquinoline overlaps the corresponding part of SIOM, thus positioning the protonated nitrogens, the phenolic hydroxyl moities, and the address parts in the same regions of space.

Again, energy minimization (previously described AMBER protocol) and removal of the bound ligands afforded for each receptor three independent sets of receptor coordinates (S1, S2, M models) in its supposed activated state.

Virtual Screening of Receptor Antagonists

Screening the antagonist D3 receptor model with FlexX and Gold, hit rates up to 5, 14, and 30% were obtained after single, double, and triple scoring, respectively [Fig. 5(A)]. Only the consensus hit lists out of Dock poses did not contain any true antagonist. The best compromise between hit rate and yield is achieved when combining FlexX as docking tool with PMF, Gold, and ChemScore for consensus scoring (hit rate of 18.92%, yield of 70%). This consensus list of 37 molecules contains 7 true ligands (compounds **2, 4–5, 7–10**) [Fig. 1(A)]. Regarding the docked conformations, we can observe a clear difference between the conformations proposed by Dock and those generated by Gold and FlexX. Whereas Dock was unable to place any of the 10 chosen D3 antagonists into the transmembrane cavity, the conformations proposed by FlexX and Gold are mostly in agreement with the expected interactions: the basic nitrogens form salt bridges with AspIII:11, and the amide oxygen found in many D3 antagonists H-bonds to ThrVII:7 [Fig. 6(A)].

When screening the M1 receptor, poorer hit rates (2–2.5% for Dock and Gold poses, almost 10% for FlexX poses) were achieved [Fig. 5(B)]. The best consensus hit list (FlexX as docking program in combination with FlexX, Gold, and PMF as scoring functions) comprised 53 molecules. It contains 5 of the 10 known ligands (compounds **14, 17, 18, 22, 23**) [Fig. 1(B)]. Interestingly, in contrast to both the retinal binding mode[19] and the previous dopamine D3 antagonist binding mode, the M1 antagonist binding site seems to be restricted to the pocket between TMs III, V, and VI. Whereas the D3 receptor cavity is wide enough to allow complete occupation of the binding site from one extreme (TM I) to the other (TM V in our orientation) [Fig. 6(A)], the M1 binding site is almost separated in two distinct cavities [Fig. 6(B)] by two tyrosine side chains (TyrIII:12, TyrE2:17), one of which (TyrIII:12) is part of TM3, whereas the other one (TyrE2:17) belongs to the second extracellular loop (E2) whose rhodopsin-like fold still remains putative and has to be experimentally verified.
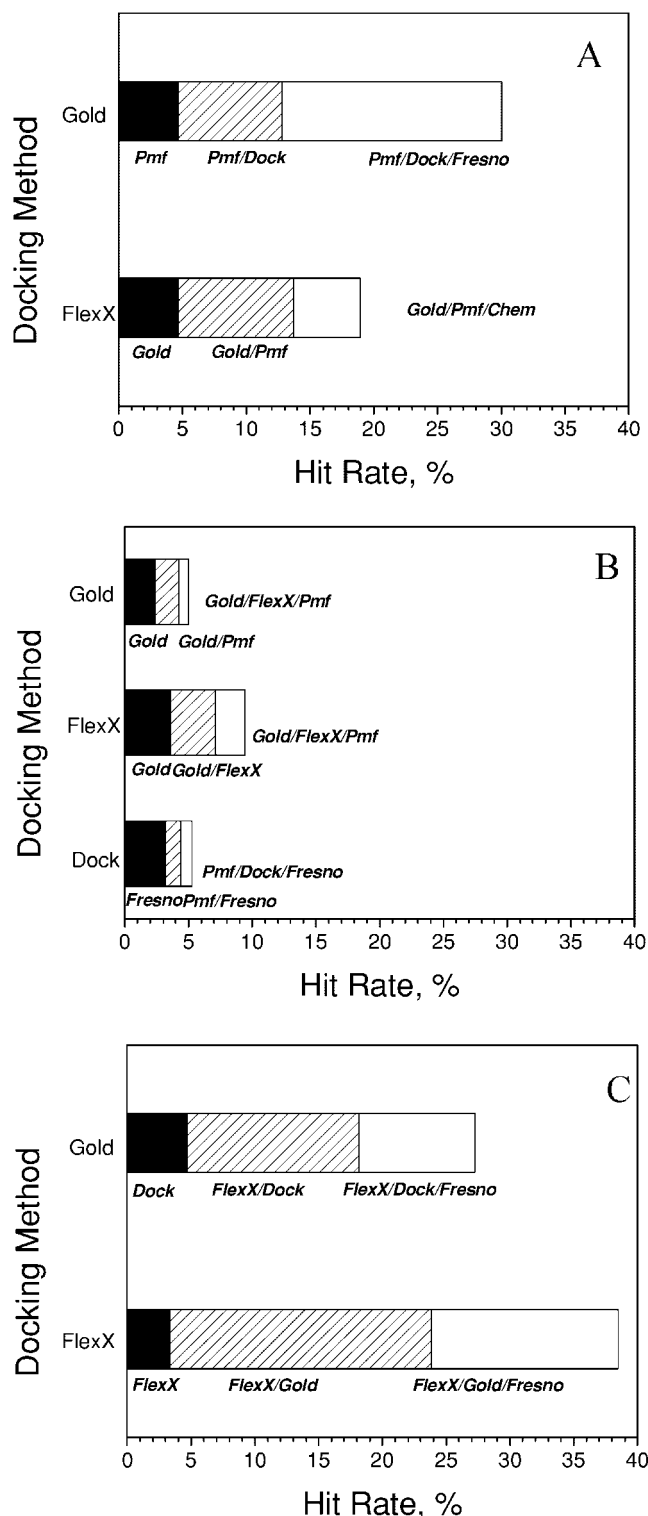

Fig. 5. Virtual screening of receptor antagonists. Hit rates among the top scorers after single (dark bars) or consensus scoring (double scoring: hashed bars, triple scoring: white bars). The top 15% scorers of the single ranking lists were used for the D3 and V1a receptor whereas the top 25% scorers were retrieved for the M1 receptor screening. Only the best docking/scoring combinations are shown here. **A:** D3 receptor. **B:** M1 receptor. **C:** V1a receptor.
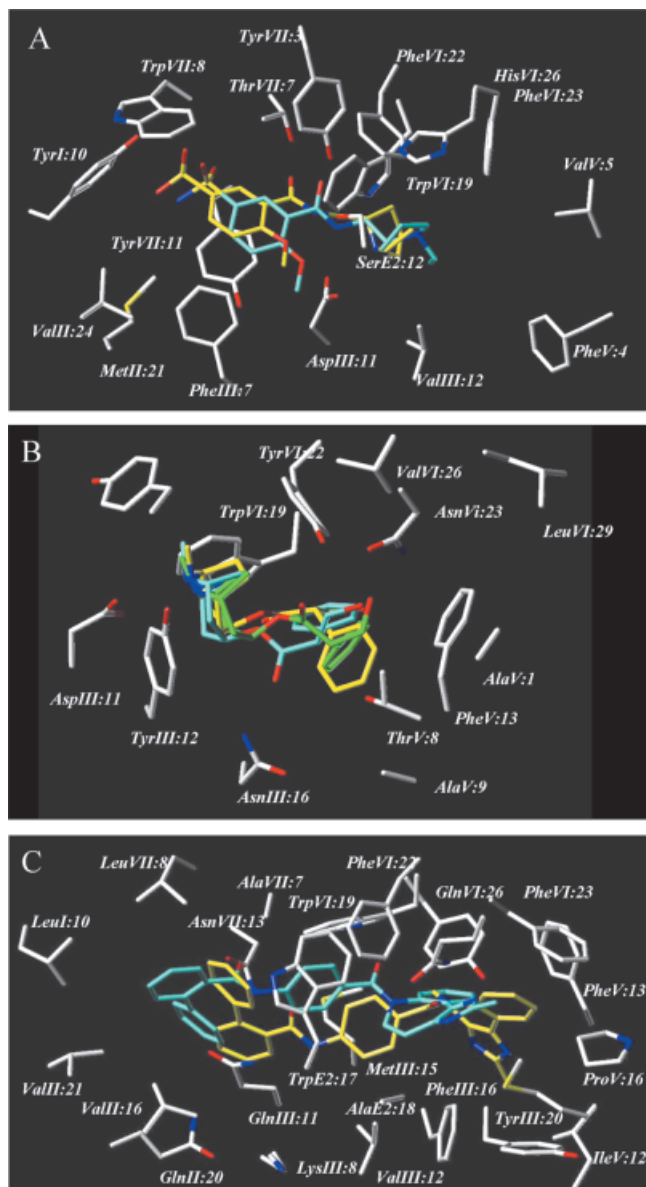
Fig. 6. Proposed binding mode for (**A**) dopamine D3 receptor antagonists (exemplified by sulpiride, compound **2**), (**B**) muscarinic M1 receptor antagonists (exemplified by atropine, compound **14**), (**C**) V1a receptor antagonists (exemplified by compound **34**). Dock, FlexX, and Gold poses are displayed by green, yellow, and cyan carbon atoms, respectively. Receptor side chains are displayed by white carbon atoms.

Screening the V1a receptor, FlexX, and Gold led to hit rates of up to 40% after triple scoring [Fig. 5(C)]. The proposed docked orientations of V1a receptor antagonists are compatible with site-directing mutagenesis data,[47] identifying aromatic residues at TM VI (TrpVI:19, PheVI: 22, PheVI:23) and hydrophilic amino acids at TM II (GlnII:20) and TM III (LysIII:8) as part of the "antagonist-binding site" [see docked orientation of compound **34** in Fig. 6(C)]. Five ligands (compounds **26, 29, 30, 33, 34**) [Fig. 1(C)] were retrieved in the best consensus hit list (FlexX as docking program in combination with FlexX, Gold, and

Fresno as scoring functions), which comprises only a total number of 13 molecules.

Virtual Screening of Receptor Agonists

It should be emphasized again that none of the full agonists used to assist receptor refinement in the different protocols (S1, S2 or M models) was included in the set of 10 true agonists seeded in the test 1,000 compounds database. For clarity, we present herein for each screening only the best consensus combination (Fig. 7), giving the optimal compromise between hit rate and yield for the computed consensus hit lists.

Screening the D3 receptor models, the pharmacophore-based model M proved to be superior to both single ligand-biased receptor models for enriching hit lists in true D3 agonists [Fig. 7(A)] whatever the docking tool used. Only the knowledge- and pharmacophore-based model M afforded hit lists with both good hit rates and good yields. It is important to notice that 60–70% of the known ligands were included in the model M-derived hit list [Fig. 2(A)], thus spanning a wide chemical diversity. The better screening results of the pharmacophore-based model can be explained looking at the docked conformations of the ligands. Only this model allows an automated docking of the 10 true agonists as expected with the protonated nitrogen atom facing AspIII:11, the catechol (or its bioisosteric counterpart) H-bonding to the three serines of the TMV (SerV:8, SerV:9, SerV:12), and the phenyl ring developing p-p stacking interactions with PheVI:22 [compare Fig. 8(A–C)]. The conformations achieved with the other models do not form these important interactions.

A similar result was found when comparing the different β2 adrenergic receptor models. After rescoring and consensus scoring analysis, the best results were again obtained with the multi-ligand-based model M. Starting from either FlexX or Dock poses, resulting hit rate and yield in true agonists were significantly higher when the pharmacophore-based model of the receptor was used [Fig. 7(B)]. As for the D3 receptor screening, all agonist classes compounds were represented in the optimal consensus hit list [Fig. 2(B)]. Only model M of the β2 receptor is compatible with a proper docking of all known β2 full agonists with the protonated nitrogen atom of the agonists forming a salt bridge with AspIII:11, the catechol moiety or its bioisostere H-bonded to the serines on TM V (SerV:8, SerV:9 and SerV:12), the β-OH groups interacting with AsnVI:26, and the aromatic moiety establishing π-π interactions with PheVI:22 [compare Fig. 9(A–C)].

In our third test case, the δ-opioid receptor, the docking, scoring, and consensus scoring was carried out as already explained for the D3 and β2 receptors with the exception that here the top 20% scorers were selected for consensus scoring. Comparing the consensus hit lists obtained with the different δ-opioid receptor models [Fig. 7(C)], we still observe a clear difference between the knowledge- and multi-ligand-based model M and the single-ligand minimized models S1 and S2. The best result was again obtained with model M. Rotation of helix 6 in the pharmacophore model results in a much better interaction of two

crucial residues (TrpVI:19, HisVI:23) with the bi-phenyl moiety found in most δ-opioid agonists [Fig. 10(A, B)]. As already observed for the D3 and β2 receptors, screening against model M enables us to retrieve true agonists representing different structural classes [Fig. 2(C)]. Notably, 90% of the known agonists were included in the consensus hit list of FlexX.

Strikingly, although Gold was able to propose reliable poses for most GPCR agonists, none of the 10 true agonists was present in any of the final consensus lists for the three receptors used in our validation study.
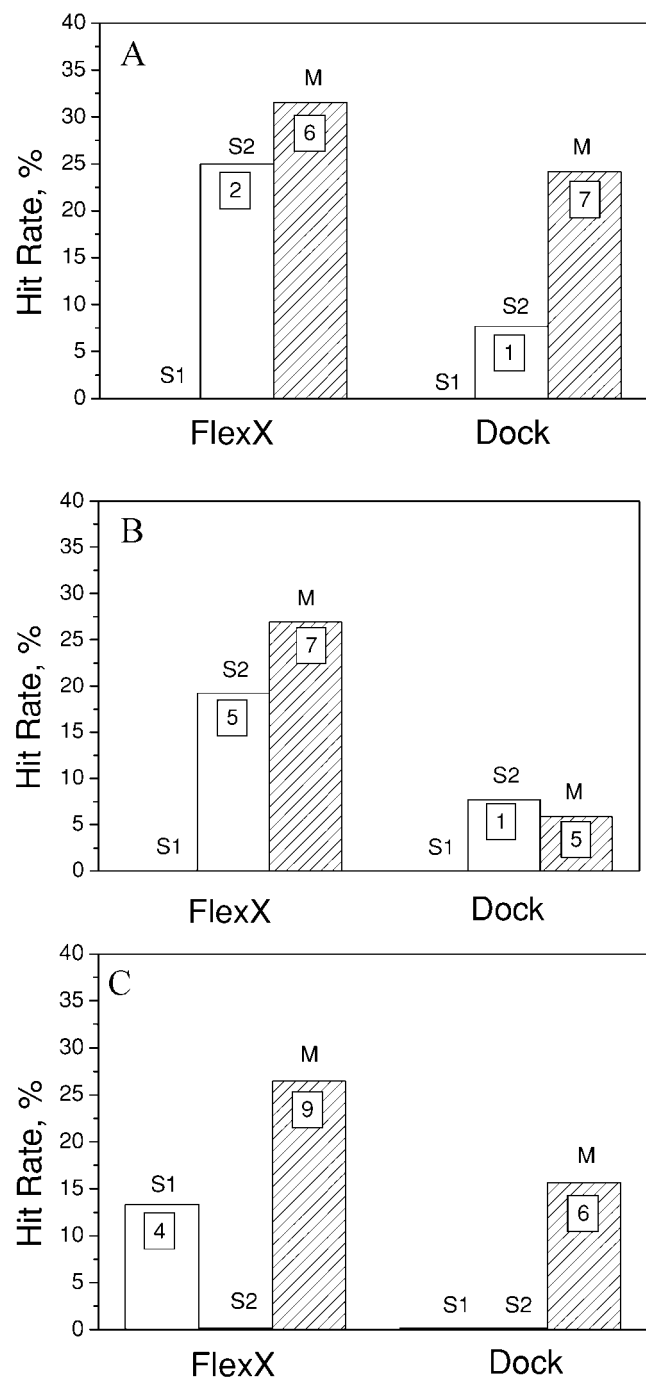
DISCUSSION
Rhodopsin-Based Homology Modeling of GPCRs

Although the sequence identity between the TMs of the modeled receptors and the template (bovine rhodopsin) is only between 21–29%, we believe that the receptors modeled in the current study share the same fold as the template since the modeled GPCRs include several conserved motifs (sequences of highly conserved amino acids) showing that, despite the low sequence identity when taking the whole TM sequence in account, the structurally and functionally important amino acids are highly conserved or substituted by amino acids of high similarity (e.g., DRY on TMIII, NPxxY on TMVI). GPCR models based on a template with an identity of 20–30% can thus be expected to be of higher accuracy than when modeling other type of proteins based on a template with such a low-sequence identity. However, we have to clearly distinguish between the transmembrane domains and loop regions. The above-stated reasoning holds true only for the TMs. We can assume that our models are of relatively high accuracy in the TMs, but at this moment we are not able to achieve a fine modeling of the loop regions as well as the N- and C-terminal parts. These extramembrane parts are much too flexible and divergent in amino acid sequence from the rhodopsin template. Since the active site of the investigated receptors is located in the transmembrane core, for our purpose the loops—generating models that are suitable for virtual screening—are not very important. The only exception is the second extracellular E2 loop, which folds back into the transmembrane core and thus completely encloses retinal in the rhodopsin X-ray structure.[19] The question arises whether this particular fold is conserved in other GPCRs. At this point, it is very difficult to give a definitive answer to this question. We can only imagine that a driving force to this fold is the disulfide bridge between cysteines in TM III and E2. Since this covalent link is conserved in our test receptors, we decided to keep this special fold unchanged in our models.

Another important point is that our GPCR models are static though proteins are in reality more or less flexible. This might lead to further inaccuracies especially for

Fig. 7. Virtual screening of receptor agonists. Consensus scoring lists have been defined out of the top 15% scorers (D3 and β2 receptor) or the top 20% scorers (δ-opioid screening) of single ranking lists. For each screening of the three receptor models (S1, S2, and M), the best consensus combination is shown. The boxed numbers in the bars indicate the number of true agonists present in the respective consensus list. Gold results are not shown as consensus scoring of Gold poses did not allow the retrieval of a single true agonist for any of the three receptors investigated herein. **A:** D3 receptor screening. FlexX poses have been scored with FlexX, Gold, and ChemScore, Dock poses with Dock, FlexX, and Gold The consensus lists from the S1-model screenings did not include any of the known agonists, thus giving hit rates of 0. **B:** β2 receptor screening. FlexX poses have been scored with FlexX, Gold, and Fresno, Dock poses scored with Dock, FlexX, and Gold. The consensus lists from the S1-model screenings did not include any of the known agonists, thus giving hit rates of 0. **C:** δ-opioid receptor screening. FlexX poses have been scored with FlexX, Dock, and Fresno, Dock poses scored with Gold, PMF, and Fresno. The consensus lists from the FlexX screening against S1 and the Dock screenings against S1 and S2-model did not include any of the known agonists, thus giving hit rates of 0.
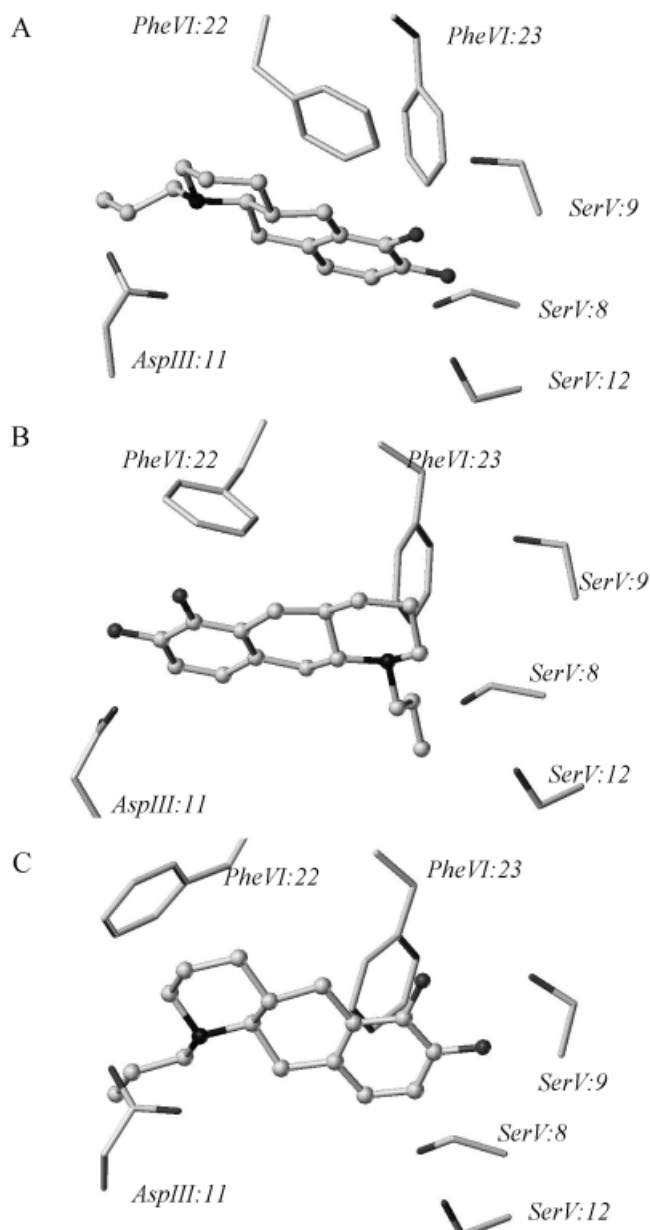
Fig. 8. Illustration of FlexX poses of the D3 receptor agonist 36 bound to different dopamine D3 receptor models. **A:** Model M. **B:** Model S1. **C:** Model S2. The ligands are represented as ball-and-sticks, receptor side chains as capped sticks. Carbon atoms are colored in light gray, heteroatoms in black.
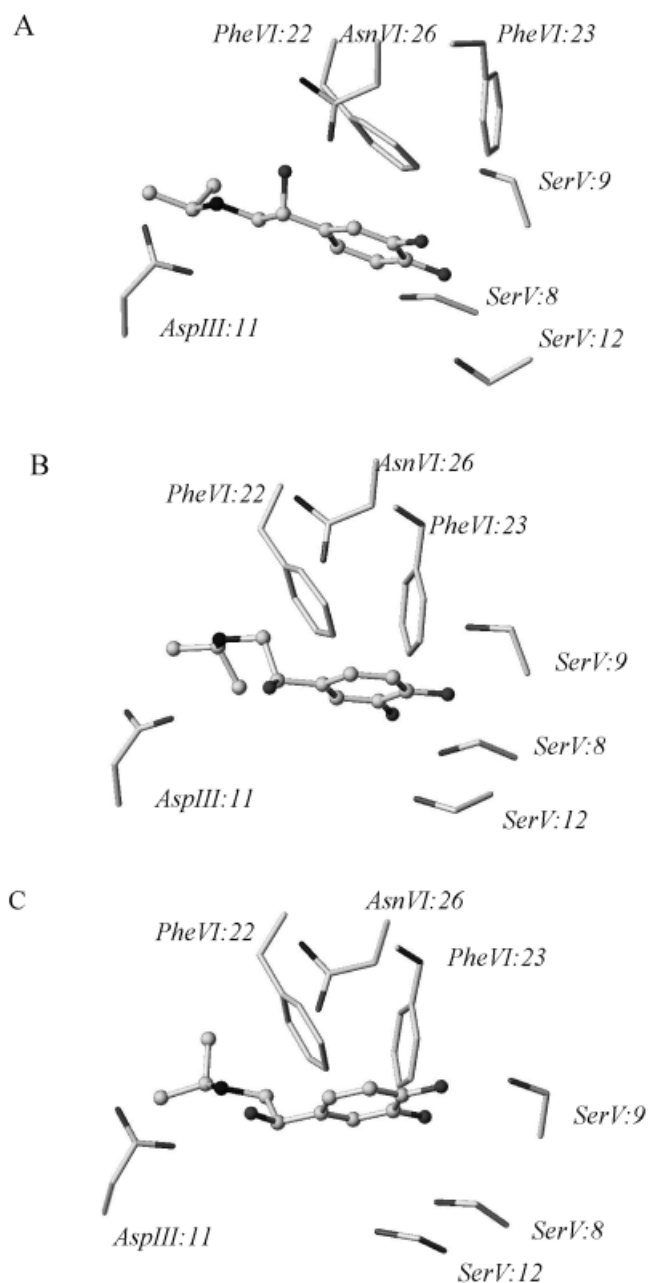

Fig. 9. Illustration of FlexX poses of the β2 receptor agonist 51 bound to different β2 adrenergic receptor models. **A:** Model M. **B:** Model S1. **C:** Model S2. The ligands are represented as ball-and-sticks, receptor side chains as capped sticks. Carbon atoms are colored in light gray, heteroatoms in black.

activated GPCR models since the active site of the activated state of GPCRs probably has a higher conformational flexibility than in the ground state. We, therefore, cannot expect that all ligands will be optimally docked to our static models.

Setting-Up Test Ligand Databases

The test databases screened against these receptors contained 990 randomly selected compounds from the ACD. We assumed that these compounds are not ligands of our test receptors, although no experimental data verify this assumption. Thus, random compounds that are included in final consensus lists should be regarded as "false positives." Assuming that a protonated nitrogen and an aromatic moiety are necessary features for binding, 15 out of the 990 randomly chosen compounds that possess these two features have a higher probability to be a potential true hit than the other random compounds. Regarding the final consensus lists, however, we could not observe that these 15 compounds are more often included in the consensus lists than the other compounds.

GPCR Models Are Reliable Enough for an Unbiased Screening of Receptor Antagonists

Inactive states of three human GPCRs were built by minimizing the rhodopsin-derived homology models with a known antagonist and challenged for their ability to accommodate 10 known antagonists seeded in a database of 990 randomly selected "drug-like" molecules.

Screening against the D3 receptor model, hit rates up to 30% were achieved. Thus, the herein presented D3 receptor model leads to hit rates similar to that already obtained when screening is performed against high-resolution X-ray structures.[23,31] These high hit rates can be explained by the fact that both FlexX and Gold docking algorithms were able to generate poses for most D3 receptor antagonists that are in good agreement with known experimental data. The rationale to test several docking programs is here justified by the poor performance of the third docking tool (Dock), which was not able to reliably dock any of the 10 chosen antagonists into the TM cavity. This failure of Dock is rather unexpected. Since the orientation of base fragments in Dock is based on shape complementarity between the ligand and the receptor,[24] we would have expected Dock to perform rather well for the very hydrophobic dopamine D3 antagonists. In the cases where the screening was successful, it is important to notice the chemical diversity of the true D3 antagonists included in the consensus hit lists. Using FlexX as docking program and rescoring with PMF, Gold, and ChemScore, the consensus hit list contains 7 out of the 10 starting true antagonists. Ligands structurally unrelated to the antagonist (the arylpiperazine BP-897) used for refining the D3 receptor minimization are indeed retrieved as virtual hits, such as the 2-aminotetraline derivative **10**, the tetrahydroisoquinoline **8**, the benzopyranopyrrolidine **9**, or chemically diverse compounds (**2, 4, 5**). Thus, the present inactive state of the dopamine D3 model seems to be suitable for a virtual screening program aimed at discovering new D3 antagonist lead structures.

Much poorer hit rates (approximately 10%) have been obtained by screening the muscarinic M1 antagonist model. Although the best computed hit rate is still 10-fold higher than that given by random screening, it is clear that virtual screening is less efficient for detecting M1 antagonists than dopamine D3 antagonists [compare Fig. 5(A,B)]. The most likely reason is that the current M1 receptor model is of lower quality than the previously described D3 receptor model. One explanation for this observation might be the different predicted binding mode of M1 and D3 receptor antagonists. D3 antagonists, like retinal bound to rhodospin,[19] are proposed to occupy the whole binding site from TMs I to VII. In contrast, the M1 antagonist binding site seems restricted to the narrow cavity between TMs III, V, and VI. Therefore, the X-ray structure of bovine rhodopsin might be a better template for the D3 receptor than for the M1 receptor. Since the M1 model seems of limited accuracy, screening results are somewhat influenced by the ligand used in the minimization step. The best consensus list (FlexX as docking program in combination with FlexX, Gold, and PMF as scoring functions) contains 5 of the 10 known ligands (compounds **14, 17, 18, 22, 23**) [Fig. 1(B)], two of which (**22,23**) are structurally very close to the ligand used for the minimization (pirenzepine). The other three compounds belong, however, to different structural classes than pirenzepine.

In the last screening test targeting the V1a vasopressin receptor, consensus scoring of FlexX and Gold poses led to remarkable hit rates of up to 40% after triple scoring [Fig. 5(C)]. The quality of the V1a receptor model, from the virtual screening point of view, is thus similar to that of high-resolution X-ray structures.[23] This remarkable feature can be explained by the ability of FlexX and Gold to place the true reference antagonists in agreement with site-directed mutagenesis data.[18,47] Five ligands (compounds **26, 29, 30, 33, 34**) [Fig. 1(C)] were retrieved in the best consensus hit list (FlexX poses scored with FlexX, Gold, and Fresno). Except for the SR-49059 compound (**28**), all known vasopressin V1a antagonists are structurally similar. Within this limited diversity, it is, however, possible to retrieve antagonist with different chemical features like a pyrrolobenzodiazepine (compound **34**), an oxime (compound **33**), or the triazole **29**.

The presented screening results against three different GPCRs (two monoamine and a peptide receptor) suggest that the herein described strategy for modeling GPCR inactive states is suitable for virtual screening. Importantly, ligands structurally different from the antagonist used for the receptor refinement were present in the final hit lists generated by consensus scoring. Thus, we feel that the ligand-based minimization protocol obviously does not strongly bias the results of the virtual screening toward a user-defined antagonist. To ascertain this issue, the influence of the ligand-based minimization procedure on the composition of final hit lists was studied by building two further inactive state models of one GPCR (dopamine D3 receptor separately minimized with sulpiride and GR-218231, respectively) and challenging the two new models for virtual screening of a 1,000-compound database comprising 10 true D3 dopamine antagonists. As mentioned earlier, we took care that the ligands used for receptor refinement were not included in the test database. Thus, sulpiride was replaced with BP-897 [Fig. 1(A)] in the database docked to the sulpiride-biased receptor model (model 2) whereas GR-218231 was replaced by BP-897 in the database docked to the GR-218231-biased receptor model (model 3). For the sake of clarity, we used only the previously reported best screening strategy for dopamine D3 antagonists (FlexX docking, consensus scoring of the 15% top scorers with Pmf, Gold, and ChemScore) [see Fig. 5(A)].

The sulpiride-based receptor model 2 is very similar to the previously used model 1 (the root-mean square deviation of heavy atoms of the active site is 0.3 Å only), differing mainly only in the conformation of TyrI:10 [Fig. 11(A)]. The GR-218231-based model 3 differs more from the initial model (rmsd of 0.8 Å from model 1), mainly in the positions of the side chains forming the pocket between
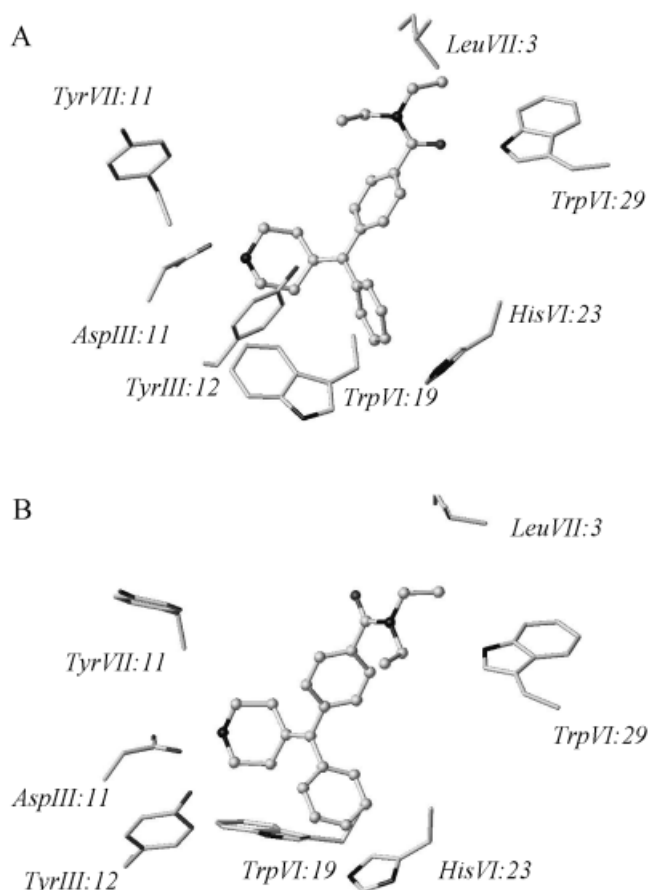
Fig. 10. Illustration of FlexX poses of the Δ-opioid receptor agonist 67 bound to different δ-opioid receptor models. **A:** Model M. **B:** Model S1. FlexX did not find any docking solution for model S2. The ligands are represented as ball-and-sticks, receptor side chains as capped sticks. Carbon atoms are colored in light gray, heteroatoms in black.
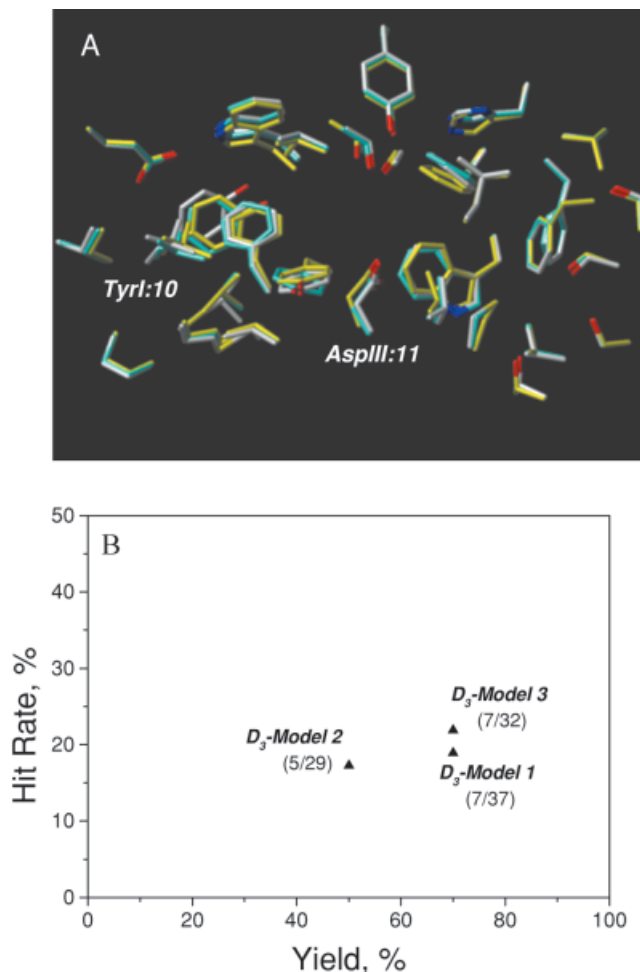

Fig. 11. Influence of the D3 receptor coordinates in the virtual screening results. The rhodopsin-based receptor model was independently minimized with three ligands belonging to three different chemical classes (phenylpiperazine: BP-897; orthomethoxybenzamide: sulpiride; 2-aminotetraline GR-218231) to lead to three independent receptor models (Model1: BP-897-biased model; Model 2: sulpiride-biased model; Model3: GR-218231-biased model). For each screening, the ligand used for minimizing the receptor was discarded from the set of 10 known antagonists seeded in the 990-compounds random library. **A:** Binding site of three human D3 receptor models (model 1, white; model 2, cyan; model 3, yellow). **B:** Hit rate and yield of the final hit lists, calculated from the previously-defined optimal protocol (consensus scoring of FlexX poses by Pmf, Gold and ChemScore and selection of the top 15% scorers common to the three lists) [recall Fig. 5(B)]. Numbers in brackets indicate the number of true hits and the total number of compounds in the final hit list, respectively.

TMs I, II, and VII [Fig. 11(A)]. This observation shows that the 3-D coordinates of the present models are not significantly biased by the ligand used for expanding the active site, explaining that we are able to retrieve ligands of different structure than the one used for the receptor refinement. Moreover, the final hit lists obtained by consensus scoring are consequently very similar in terms of hit rate, yield, and total size of the full hit list [Fig. 11(B)]. Five known dopamine D3 antagonists were included in the consensus list obtained when screening against "model 2" (compounds **4, 7, 9, 10, 11**), 7 in the consensus list produced by screening against "model 3" (compounds **1, 3, 4, 6, 7, 8, 11**). The ligands retrieved in the hit lists vary slightly in the different consensus lists, showing that relatively small differences in receptor side chain conformations can nevertheless influence virtual screening outputs. However, none of the models shows biasing in the sense that minimization with a specific ligand results in retrieving only ligands of the same structure and thus size and shape, for example, both the large arylpiperazine 7 as well as the smaller-sized nafadotride (compound **4**), were included in all three hit lists.

Pharmacophore-Biased Receptor Refinement Generates Suitable Active State GPCR Models

In the second part of this work, we used the previous modeling strategy—now minimizing with a known agonist docked to the active site—to generate for each of three test receptors (dopamine D3, β2-adrenergic, and δ-opioid receptors) two independent models of an agonist-bound form of the receptor. However, we have to consider that the template (bovine rhodopsin) has been crystallized in its ground state that conformationally differs from the activated state.[19] GPCRs models based on this X-ray struc-

ture, are therefore expected to be closer to their inactive form than to their activated, agonist-bound state. We, therefore, anticipated that a slight minimization with one known agonist docked to the active site might not be enough to generate correct models of an activated state of a GPCR.

Until now, there is only a crude picture of the conformational changes that occur during receptor activation. Recent studies based on electron paramagnetic resonance and fluorescence spectroscopy[64] suggest an outward movement of the cytoplasmic end of TMs III and VI,[65,66] as well as an anticlockwise rotation of TM VI around its helical axis when viewed from the extracellular side. Other helices probably adjust their positions upon activation as well. We developed a specific knowledge- and multi-ligand-based modeling procedure that better mirrors GPCR activation than single-ligand minimization. This procedure notably includes as the first step manually rotating TM VI by 30° anticlockwise around its helical axis. To simulate all the other changes that might occur in the active site, we then performed as the second step of the new procedure a multi-ligand-based minimization. A receptor-restricted "pharmacophore" was first defined by superimposing several different full agonists onto a previously docked reference ligand. The resulting pharmacophore is thus not only defined by the chemical structures of the agonists, but also takes into account the protein environment as well as known experimental data. It represents all the important structural features of the different agonist classes of the target receptor. Following minimization of the receptor in the presence of several different ligands avoids biasing the receptor 3-D structure toward a specific agonist structure. We assume that the resulting receptor conformation rather reflects a compromise between possible activated states. This is important to notice since minimization with several ligands at once results in a rather strong change of the active site.

Comparing the virtual screening results obtained with the different models generated by either single-ligand (models S1 and S2) or multi-ligand-based receptor refinement (models M) clearly shows that the pharmacophore-based refinement protocol gives receptor models that are much better suited for agonist screening (Fig. 7). This observation can be explained by differences observed in the conformations of the ligands docked to the different models.

Most of the dopamine D3 receptor agonists were correctly docked into the active site of model M. In this model, the side chains of the amino acids in the active site are orientated so that the D3 agonists can develop important specific interactions with the receptor, including a salt bridge with AspIII:11, H-bonds to three serines on TM V (SerV:8, SerV:9, SerV:12) and π-π interactions with PheVI:22. Therefore, the docked solutions get high scores from the scoring functions and are thus among the top scorers of the ranking lists used for consensus scoring. Using model M, we not only achieved high hit rates and yields, but the retrieved agonists also show structural diversity, an important feature for virtual screening. At least one compound

of each chemical class represented in the test set has been identified as a potential hit in the consensus lists. Notably, we were able to retrieve agonists whose structures were not represented in the pharmacophore definition. Thus, the knowledge- and pharmacophore-based modeling procedure seems to have properly simulated the conformational rearrangements in the active site during receptor activation. This is obviously not the case for the single-ligand minimization. Though the distances between AspIII:11 and the serines on TM V in the models S1 and S2 are similar to the respective distances in model M, the orientations of the side chains delimiting the active site are not reliable for docking D3 agonists (Fig. 8). In both model S1 and S2, neither PheVI:22 nor PheVI:23 can form a π-π interaction with the aromatic moiety of the agonists, showing the importance of the rotation of TM VI in the new procedure. It has, however, to be mentioned that we do not intend to suggest here that reorientation of PheVI:22 is the mechanism of receptor activation. Such a conclusion cannot be drawn on the basis of the current model. The docking solutions found for the models S1 and S2 are, therefore, partially or totally in disagreement with the expected interactions, resulting in low scores for the known agonists and explaining the low hit rates and yields of the corresponding consensus lists.

In the case of the β2-adrenergic receptor, FlexX could successfully dock only one agonist to the active site of the model S1, and Dock was unable to place any of the 10 known β2 agonists into this active site. A reason for this failure might be that epinephrine, the ligand used for refining model S1, has no large N-alkyl substituent. The pocket between TMs II and VII is thus too small to accommodate the larger hydrophobic residues of other ligands. Using nylidrine for receptor minimization (model S2), the ligands could be placed into the binding site because this pocket is now larger. But since nylidrine lacks the meta hydroxyl group on the catechol ring, the positions of the serines on TM V cannot be correctly adjusted. Since the position of TM VI and thus of the phenylalanines PheVI:22 and PheVI:23 were not optimized either, the docked conformations show the expected interactions only partially (Fig. 9). Nevertheless, we could achieve a rather good hit rate and yield with the FlexX consensus list of the S2 screening. However, 60% of the retrieved known agonists (compounds **53–55**) belong to the same chemical class as the compound used for the receptor refinement (nylidrine) showing that in this case the model is obviously biased toward this agonist class. These results of the S1 and S2 screenings suggest that we need a pharmacophore to represent all the important pharmacophoric features of the β2-agonists. Only model M represents a correct model of the activated state in which critical amino acids are properly orientated for the majority of known agonists. Accordingly, in contrast to the conformations obtained for the models S1 and S2, the docked conformations for model M are in agreement with the expected interactions. Therefore, the docked solutions for model M get high scores from the scoring functions and are among the top scorers of the ranking lists used for consensus scoring. As for the D3

receptor agonist screening, we could even retrieve known agonists whose structures were not represented in the initial pharmacophore definition, as for example the 2-oxoquinoline derivatives **57, 58**.

When screening against the δ-opioid receptor models, neither FlexX nor Dock was able to dock any of the agonists to the active site of model S2. In this model, HisVI:23 points into the central cavity between TMs III, V, and VI, thus reducing the size of this pocket. Moreover, as TAN67, which has been used for the receptor refinement, has only a small methyl group on the protonated nitrogen, the TyrVII:11 and LeuII:20 side chains come very close to the protonated amine, closing the pocket between TMs II and VII and thus not leaving enough space to accommodate larger substituents present in other δ-agonists. Screening against model S1, FlexX was able to dock six true ligands to the active site. Four of them were included in the consensus list, but all four belong to the same chemical class as the ligand used for prior minimization (SNC-80). Dock was not able to place any of the agonists in the active site of model S1. As for the D3 and β2 receptors, model M of the δ-opioid receptor gave again with both FlexX and Dock, the best result in terms of docking conformations, hit rate, and yield of the final consensus lists, as well as diversity of the agonists present in the consensus lists. Ninety percent of the known agonists were included in the FlexX consensus list, thus representing all classes of the test sets and showing that this model is not biased toward a specific agonist class.

Strikingly, rescoring Gold poses did not allow the retrievement of a single true agonist in any of the test cases investigated herein. This observation is rather surprising as Gold has been recently demonstrated to be the most accurate docking tool of the three programs used in the current study.[67] One possible explanation of the Gold failure comes from the automated positioning of polar hydrogen atoms of the protein target to optimize intermolecular hydrogen bonds during the docking process. However, it is not possible to save different sets of protein coordinates for each ligand. Thus, the unique set of original receptor coordinates that has to be used for rescoring Gold poses might not be optimal for some ligands. As polar interactions play a more important role for agonists than antagonists, the consequence of the latter Gold feature is much more dramatic for receptor agonists than for receptor antagonists. In any case, this result justifies our screening strategy[23] that systematically challenges several docking/scoring combinations.

Agonist vs. Antagonist Screening: A Matter of Receptor Flexibility?

We have shown that single-ligand minimization is enough to correctly model an antagonist-bound form of GPCRs, but not to generate an reliable agonist-bound model. However, the activated-state models generated by the more complex knowledge- and multi-ligand based procedure are indeed suitable targets for virtual screening. This difference in antagonist and agonist-state modeling can be explained by the conformational changes taking

TABLE I. Root-Mean Square Deviations in Å (Backbone TM Atoms) of the Different Dopamine D3 Receptor Models From the X-Ray Structure of Bovine Rhodopsin

Antagonist model			Agonist model		
Model 1[a]	Model 2[b]	Model 3[c]	S1 model[d]	S2 model[e]	M model[f]
0.85	0.80	0.81	0.74	1.06	1.26

[a]Receptor refinement with docked BP-897.
[b]Receptor refinement with docked sulpiride.
[c]Receptor refinement with docked GR-218231.
[d]Receptor refinement with docked apomorphine.
[e]Receptor refinement with docked pergolide.
[f]Pharmacophore-based multi ligands-biased receptor model (see Computational Methods).

place upon receptor activation. But a further reason might lie in the fact that a GPCR is more flexible in its activated conformation than in the ground state. There is evidence for several different agonist-bound states, but only for one ground state.[68] We have seen that single-ligand minimization with an antagonist gives an antagonist-bound state model suitable to retrieve antagonists of different structural classes by virtual screening, supporting the idea that all antagonists stabilize a very similar ground state. However, single-ligand minimization with a known full agonist results in a model that is either not useful at all in virtual screening or biased toward the ligand structure used to assist receptor refinement (see FlexX screenings of the β2-receptor S2 model and of the δ-opioid receptor S1 model). The modeled active site is, in these two cases, clearly biased toward one structural class, suggesting that structurally unrelated agonists stabilize slightly different activated states of the same receptor. To give a better idea how much the models had changed after the different minimization strategies, the rmsd of the different models from the rhodopsin X-ray template are exemplified in Table I for the D3 receptor. As expected, the antagonist models show the smallest rmsd, whereas the pharmacophore-based minimization strategy modifies the backbone atoms the most, especially in TM VI, which was manually rotated during receptor refinement [Fig. 12(A)]. The TMIII of this model also significantly deviates from the template in one helix turn. This observation might be of biological importance since helix 3 is also suspected to play a major role in the receptor activation. Figure 12(B,C) shows the antagonist and agonist models, respectively, before and after receptor refinement. Again, the differences in the pharmacophore-based agonist model are much larger than in the antagonist model.

Crossdocking Experiments

Two of the three test GPCRs selected for agonist screening (the dopamine D3 and β2-adrenergic receptor) are catecholamine receptors and bind structurally related agonists. To address the selectivity of virtual screening on related targets, the 10 known D3 agonists were docked to model M of the β2 receptor whereas the 10 β2-adrenergic receptor agonists were docked to the dopamine D3 receptor model M. The best consensus scoring schemes previously
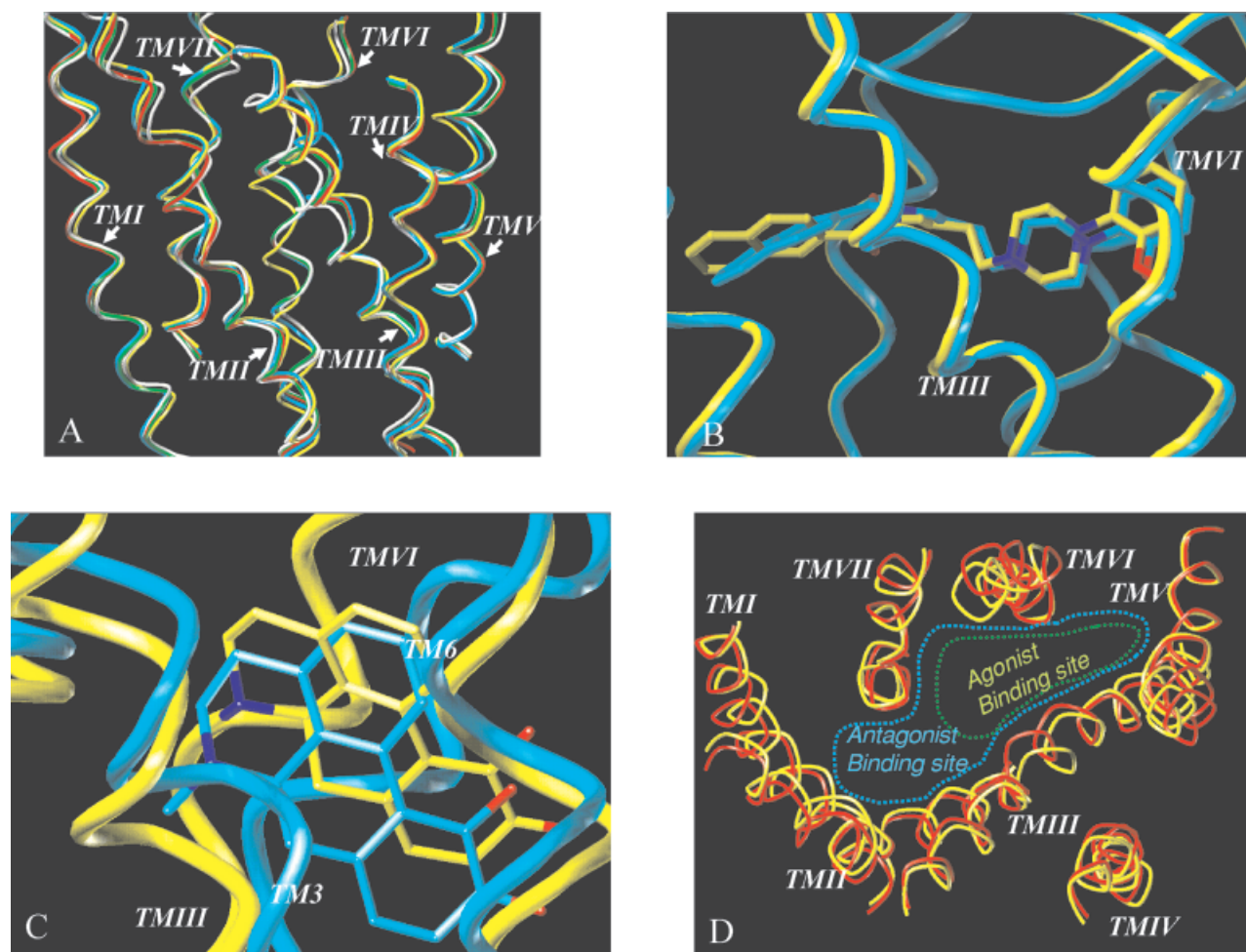
Fig. 12. Influence of the different refinement protocols on the overall GPCR structure exemplified by the dopamine D3 receptor. **A:** Ribbon view of backbone atoms of the seven transmembrane (TM) helices (bovine rhodopsin, white; D3 receptor antagonist model 1, red; D3 agonist pharmacophore-based model M, yellow; D3 agonist model S1, green; D3 agonist model S2, cyan). **B:** Comparison of the D3 antagonist model before (yellow) and after (cyan) receptor minimization. Ribbons indicate the backbone atoms of both structures. The docked ligand BP-897 is displayed by sticks. **C:** Comparison of the D3 pharmacophore-based agonist model before (yellow) and after (cyan) receptor minimization. Ribbons indicate the backbone of both structures. For clarity, only one of the 5 superimposed ligands(apomorphine) is shown as sticks. **D:** Ribbon view of the backbone atoms of the D3 antagonist "model 1" (red) and the D3 agonist pharmacophore-based model M (yellow). Please note that helix 6 was manually rotated in the pharmacophore-based model to fit known pharmacophoric requirements. The usual locations of the antagonist and agonist binding sites in GPCRs are indicated by dashed lines. Generally, GPCR antagonists occupy the complete transmembrane cavity (from the pocket between TMs III, V, and VI to the pocket between TMs II and VII), whereas the agonist binding site is often restricted to the pocket between TMs III, V, VI. However, there are exceptions to this rule, as for example β2-agonists that also partially occupy the pocket between TMs II and VII.

identified for both receptors [Fig. 7(A,B)] were then applied to calculate the enrichment in true and "wrong" ligands. Though the positions of the AspIII:11 and the serines on helix 5 are very similar in both receptors, neither FlexX nor Dock was capable to find reliable poses upon docking the D3 agonists to the β2 receptor. A comparison of the D3 and β2 models shows that D3 agonists have to change their binding mode when binding to the β2 receptor to avoid bumping into the second extracellular loop E2. The E2 loop is 8 residues longer in the β2 than in the D3 receptor (Fig. 3), thus decreasing the overall size of the β2 binding pocket. Since the D3 agonists cannot form the expected interactions with the conserved aspartate (TM III) and the serines on TM V, they are all badly scored whatever the scoring function used. Consequently, none of the true D3 agonists would have been

included in the consensus hit lists when screening for β2 receptor agonists (Table II).

Docking the β2 agonists to the active site of the D3 receptor gives better solutions, at least when using FlexX as docking algorithm. The small catecholamines **51–53** as well as compound **57** can be docked as expected (salt bridge between the basic nitrogen and AspIII:11, H-bonds of the catechol to the serines on TM V). Fitting the larger β2 agonists (**52, 54–56, 58–60**) in the D3 active site is more difficult because the pocket between TM II and VII is not as opened as in the β2 model. However, possible poses are nevertheless found in which the large N-aralkyl substituents are directed towards TMs VI and VII. However, only two β2-agonists (compounds **58, 59**) [Fig. 2(B)] would have been selected as potential hits by screening the D3 receptor model. In conclusion, these "cross-docking" experi-

TABLE II. Percentage of Known Ligands Identified as True Hits After Cross-Docking β2 Agonists to the D3 Receptor and D3 Agonists to the β2 Receptor[†]

Ligands	Receptor (model M)	Docking method	Hit rate (%)
D3 agonists	D3	Dock	24.1
		FlexX	31.6
	β2	Dock	0.0
		FlexX	0.0
β2 agonists	D3	Dock	0.0
		FlexX	7.7
	β2	Dock	5.8
		FlexX	26.9

[†]Hit rates were calculated from the consensus list derived from the top 15% scorers of single hit lists (D3 receptor: FlexX poses scored by FlexX, Gold, and ChemScore, Dock poses scored by Dock, FlexX, and Gold; β2 receptor: FlexX poses scored by FlexX, Gold, and Fresno, Dock poses scored by Dock, FlexX, and Gold). In both dockings, the putative best model (model M) was used for each receptor. Gold was not used as docking tool since rescoring Gold poses for GPCR agonists was previously proven to be inefficient.

ments show that virtual screening against the new agonist-bound states of related GPCRs is selective enough to distinguish not only true ligands from randomly chosen drug-like molecules but also true hits from chemically related inactive compounds.

CONCLUSIONS

Antagonist-bound state models of three human G-Protein coupled receptors have proven to be indeed suitable for virtual screening of GPCR antagonists. For the D3 and the V1a receptor, obtained hit rates are 20- to 40-fold higher than what can be obtained by random screening. True antagonists structurally unrelated to the one used for receptor refinement were predicted to be virtual hits, suggesting that the antagonist used in the minimization step does not significantly bias the virtual screening. Hence, the topology of the transmembrane cavity remains relatively similar after using different ligands for expanding the binding site, thus explaining that (1) hit lists contain ligands belonging to chemical series different from that used for receptor refinement and (2) the outcome of the virtual screening in terms of hit rate, yield, and structures of the virtual hits is rather similar whatever ligand used for expanding the binding site.

However, challenging single-ligand biased receptor models for retrieving known full agonists was inefficient. A logical explanation for this observation resides in the fact that all GPCR models have been derived from the inactive state of bovine rhodospin, which is closer to an "antagonist-bound state" than to an "agonist-bound state" of the target GPCR, as well as in the higher flexibility of the active site in the activated state than in the ground state. Minimizing with a known ligand in the active site is obviously sufficient to expand the active site but not for simulating the conformational changes occurring in the receptor activation process. A pharmacophore-based receptor refinement method was proposed to generate agonist-bound state models of GPCRs still using the X-ray structure of bovine

rhodopsin as template. Applied to three human GPCRs, such receptor models are indeed precise enough for discriminating known agonists from randomly chosen "drug-like" molecules. Most importantly, we were able to retrieve in the virtual hit lists true agonists whose chemical structures had not previously been used for generating the pharmacophore and refining the receptor model. This confirms that the current GPCR activated-state models are not significantly biased toward the agonists used in the minimization step. This very important feature should enable us to retrieve new lead structures that could be totally unrelated to any known GPCR ligand.

We, thus, suggest two different strategies for virtual screening of GPCRs, depending on the type of compounds that are searched for: (1) searching for antagonists requires only minimization of the rhodopsin-derived homology model with a single true ligand; (2) searching for full agonists demands a much more complex procedure, still starting from rhodopsin, in which a proper activated state model has to be first derived by a pharmacophore-based receptor refinement.

REFERENCES

1. Mander T. Beyond uHTS: ridiculously HTS? Drug Discov Today 2000;5:223–225.
2. Walters WP, Stahl MT, Murcko MA. Virtual screening: an overview. Drug Discov Today 1998;3:160–178.
3. Pickett SD, McLay IM, Clark DE. Enhancing the hit-to-lead properties of lead optimization libraries. J Chem Inf Comput Sci 2000;2:263–272.
4. Gohlke H, Klebe G. Statistical potentials and scoring functions applied to protein-ligand binding. Curr Opin Struct Biol 2001;11:231–235.
5. Mason JS, Hermsmeier MA. Diversity assessment. Curr Opin Chem Biol 1999;3:342–349.
6. Mason JS, Morize I, Menard PR, Cheney DL, Hulme C, Labaudiniere RF. New 4-point pharmacophore method for molecular similarity and diversity applications: overview of the method and applications, including a novel approach to the design of combinatorial libraries containing privileged substructures. J Med Chem 1999;42:3251–3264.
7. Dixon JS. Evaluation of the CASP2 docking section. Proteins 1997;(Suppl 1):198–204.
8. Tame JRH. Scoring functions: A view from the bench. J Comput-Aided Mol Design 1999;13:99–108.
9. Marhefka CA, Moore BM, Bishop C, Kirkovsky L, Mukherjee A, Dalton JT, Miller DD. Homology modeling using multiple molecular dynamics simulations and docking studies of the human androgen receptor ligand binding domain bound to testosterone and nonsteroidal ligands. J Med Chem 2001;44:1729–1740.
10. Schafferhans A, Klebe G. Docking ligands onto binding site representations derived from proteins built by homology modelling. J Mol Biol 2001;307:407–427.
11. Schneider, G., Böhm, H.J. Virtual screening and fast automated docking methods. Drug Discov. Today 2002;7:64–70.
12. Marchese A, George SR, Kolakowski LF, Lynch KR, O'Dowd BF. Novel GPCRs and their endogenous ligands: expanding the bound-

aries of physiology and pharmacology. Trends Pharmacol Sci 1999;20:370–375.

13. Hibert MF, Trumpp-Kallmeyer S, Bruinvels A, Hoflack J. Three-dimensional models of neurotransmitter G-binding protein-coupled receptors. Mol Pharmacol 1991;40:8–15.

14. Bikker JA, Trumpp-Kallmeyer S, Humblet C. G-Protein coupled receptors: models, mutagenesis, and drug design. J Med Chem 1998;41:2911–2927.

15. Henderson R, Baldwin JM, Ceska TA, Zemlin F, Beckmann E, Downing KH. Model for the structure of bacteriorhodopsin based on high-resolution electron cryo-microscopy. J Mol Biol 1990;213:899–929.

16. Schertler GF, Villa C, Henderson R. Projection structure of rhodopsin. Nature 1993;362:770–772.

17. Baldwin JM, Schertler GF, Unger VM. An alpha-carbon template for the transmembrane helices in the rhodopsin family of G-protein-coupled receptors. J Mol Biol 1997;272:144–164.

18. Mouillac B, Chini B, Balestre MN, Elands J, Trumpp-Kallmeyer S Hoflack J, Hibert M, Jard S, Barberis C. The binding site of neuropeptide vasopressin V1a receptor. Evidence for a major localization within transmembrane regions. J Biol Chem 1995;270:25771–25777.

19. Palczewski K, Kumasaka T, Hori T, Behnke CA, Motoshima H, Fox BA, Trong IL, Teller DC, Okada T, Stenkamp RE, Yamamoto M, Miyano M. Crystal structure of Rhodopsin: a G protein-coupled receptor. Science 2000;289:739–745.

20. Sokoloff P, Schwartz JC. Novel dopamine receptors half a decade later. Trends Pharmacol Sci 1995;16:270–275.

21. Felder CC, Bymaster FP, Ward J, DeLapp N. Therapeutic opportunities for muscarinic receptors in the central nervous system. J Med Chem 2000;43:4333–4353.

22. Hibert M, Hoflack J, Trumpp-Kallmeyer S, Mouillac B, Chini B, Mahe E, Cotte N, Jard S, Manning M, Barberis C. Functional architecture of vasopressin/oxytocin receptors. J Recept Signal Transduct Res 1999;19:589–596.

23. Bissantz C, Folkers G, Rognan D. Protein-based virtual screening of chemical databases. 1. Evaluation of different docking/scoring combinations. J Med Chem 2000;43:4759–4767.

24. Ewing TJA, Kuntz ID. Critical evaluation of search algorithms for automated molecular docking and database screening. J Comput Chem 1997;18:1175–1189.

25. Jones G, Wilett P, Glen RC, Leach AR, Taylor R. Development and validation of a genetic algorithm for flexible docking. J Mol Biol 1997;267:727–748.

26. Rarey M, Kramer B, Lengauer T, Klebe G. A fast flexible docking method using an incremental construction algorithm. J Mol Biol 1996;261:470–489.

27. Muegge I, Martin YC. A general and fast scoring function for protein-ligand interactions: a simplified potential approach. J Med Chem 1999;42:791–804.

28. Eldridge M, Murray CW, Auton TA., Paolini GV, Lee RP. Empirical scoring functions: I. The development of a fast empirical scoring function to estimate the binding affinity of ligands in receptor complexes. J Comput-Aided Mol Des 1997;11:425–445.

29. Rognan D, Laumoeller SL, Holm A, Buus S, Tschinke V. Predicting binding affinities of protein ligands from three-dimensional coordinates: application to peptide binding to class I major histocompatibility proteins. J Med Chem 1999;42:4650–4658.

30. Wang R, Liu L, Lai L, Tang Y. SCORE: a new empirical method for estimating the binding affinity of a protein-ligand complex. J Mol Model 1998;4:379–384.

31. Charifson PS, Corkery JJ, Murcko MA, Walters WP. Consensus scoring: A method for obtaining improved hit rates from docking databases of three-dimensional structures into proteins. J Med Chem 1999;42:5100–5109.

32. MDL Information Systems, Inc., San Leandro, CA 94577.

33. SYBYL 6.62, Tripos Inc., St. Louis, MO 63144.

34. Gasteiger J, Marsili M. Iterative partial equalization of orbital electronegativity: a rapid access to atomic charges . Tetrahedron 1980;36:3219–3228.

35. Concord 4.0 is part of the SYBYL software distribution (http://www.tripos.com).

36. Isis/Draw is distributed from MDL Information Systems (http://www.mdli.com).

37. Thompson JD, Higgins DG, Gibson TJ. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res 1994;22:4673–4680.

38. Henikoff S, Henikoff J. G. Amino acid substitution matrices from protein blocks Proc Natl Acad Sci USA 1992;89:10915–10919.

39. Strader CD, Dixon RA, Cheung AH., Candelore MR, Blake AD, Sigal IS. Mutations that uncouple the beta-adrenergic receptor from Gs and increase agonist affinity. J Biol Chem 1987;262:16439–16443.

40. Case DA, Pearlman DA, Caldwell JW, Cheatham III TE, Ross WS, Simmerling CL, Darden TA, Merz KM, Stanton RV, Cheng AL, Vincent JJ, Crowley M, Ferguson, DM, Radmer RJ, Seibel GL, Singh,UC, Weiner PK, Kollman PA. AMBER 5.0. University of California, San Francisco; 1997.

41. Cornel WD, Cieplak P, Bayly CI, Gould IR, Merz JKM, Ferguson DM, Spellmeyer DM, Fox T, Caldwell JW, Kollman PA. A second generation force field for the simulation of proteins, nucleic acids and organic molecules. J Am Chem Soc 1995;117:5179–5197.

42. Alberts GL, Pregenzer JF, Im WB. Identification of transmembrane regions critical for ligand binding to the human D3 dopamine receptor using various D3/D1 transmembrane chimeras. Mol Pharmacol 1998;54:379–88.

43. Mansour A, Meng F, Meador-Woodruff JH, Taylor LP, Civelli O, Akil H. Site-directed mutagenesis of the human dopamine D2 receptor. Eur J Pharmacol 1992;227:205–214.

44. Lundstrom K, Turpin P, Large C, Robertson G, Thomas P, Lewell XQ. Mapping of dopamine D3 receptor binding site by pharmacological characterization of mutants expressed in CHO cells with the Semliki Forest virus system. J Recept Signal Transduct Res 1998;18:133–150.

45. Cho W, Taylor LP, Mansour A, Akil H. Hydrophobic residues of the D2 dopamine receptor are important for binding and signal transduction. J Neurochem 1995;65:2105–2115.

46. Bourdon H, Trumpp-Kallmeyer S, Schreuder H, Hoflack J, Hibert M, Wermuth CG. Modelling of the binding site of the human m1 muscarinic receptor: Experimental validation and refinement. J Comput-Aided Mol Design 1997;11:317–332.

47. Cotte N, Balestre MN, Aumelas A, Mahe E, Phalipou S, Morin D, Hibert M, Manning M, Durroux T, Barberis C, Mouillac B. Conserved aromatic residues in the transmembrane region VI of the V1a vasopressin receptor differentiate agonist vs. antagonist ligand binding. Eur J Biochem 2000;267:4253–4263.

48. Strader CD, Candelore MR, Hill WS, Sigal IS, Dixon RAF. Identification of two serine residues involved in agonist activation of the β2-adrenergic receptor. J Biol Chem 1989;264:13572–13578.

49. Strader CD, Gaffney T, Sugg EE, Candelore MR, Keys R, Patchett AA, Dixon RAF. Allele-specific activation of genetically engineered receptors. J Biol Chem 1991;266:5–8.

50. Strader CD, Fong TM, Tota MR, Underwood D, Dixon RA. Structure and function of G protein-coupled receptors. Annu Rev Biochem 1994;63:101–132.

51. Wieland K, Zuurmond HM, Krasel C, Ijzerman AP, Lohse MJ. Involvement of Asn-293 in stereospecific agonist recognition and in activation of the β2-adrenergic receptor. Pharmacology 1996;93:9276–9281.

52. Liapaski G, Ballesteros JA, Papachristou S, Chan WC, Chen X, Javitch JA. The Forgotten Serine. J Biol Chem 2000;275:37779–37788.

53. Valiquette M, Vu HK, Yue SY, Waehlestedt C, Walker P. Involvement of Trp-284, Val-296, and Val-297 of the human δ-opioid receptor in binding of δ-selective ligands. J Biol Chem 1996;271:18789–18796.

54. Befort K, Tabbara L, Kling D, Maigret B, Kiefer BL. Role of aromatic transmembrane residues of the δ-opioid receptor in ligand recognition. J Biol Chem 1996;271:10161–10168.

55. Pepin M, Yue SY, Roberts E, Wahlestedt C, Walker P. Novel "restoration of function" mutagenesis strategy to identify amino acids of the δ-opioid receptor involved in ligand binding. J Biol Chem 1997;272:9260–9267.

56. Lemmen C, Lengauer T, Klebe G. FlexS: a method for fast flexible ligand superimposition. J Med Chem 1998;41:4502–4520.

57. Schwartz TW. Locating ligand-binding sites in 7 TM receptors by protein engineering. Cur Opin Biotechnol 1994;5:434–444.

58. Gether U, Kobilka BK. G protein-coupled receptors. II. Mechanism of agonist activation. J Biol Chem 1998;273:17979–17982.

59. Hulme EC, Lu ZL, Ward SDC, Allman K, Curtis CA. The

conformational switch in 7-transmembrane receptors: the muscarinic receptor paradigm. Eur J Pharmacol 1999;375: 247–260.

60. Ghanouni P, Steenhuis JJ, Farrens DL, Kobilka BK. Agonist-induced conformational changes in the G-protein-coupling domain of the β2 adrenergic receptor. Proc Natl Acd Sci USA 2001;98:5997–6002.

61. Plobeck N, Delorme D, Wei Z, Yang H, Zhou F, Schwarz P, Gawell L, Gagnon H, Pelcman B, Schmidt R, Yue SY, Walpole C, Brown W, Zhou E, Labarre M, Payza K, St-Onge S, Kamassah A, Morin P, Projean D, Ducharme J, Roberts E. New Diarylmethylpiperazines as potent and selective nonpeptidic δ opioid receptor agonists with increased in vitro metabolic stability. J Med Chem 2000;43:3878–3894.

62. Liao S, Alfaro Lopez J, Shenderovich MD, Hosohata K, Lin J, Li X, Stropova D, Davis P, Jernigan KA, Porreca F, Yamamura HI, Hruby VJ. De novo design, synthesis, and biological activities of high-affinity and selective non-peptide agonists of the δ-opioid receptor. J Med Chem 1998;41:4767–4776.

63. Portoghese PS. From models to molecules: opioid receptor dimers, bivalent ligands, and selective opioid receptor probes. J Med Chem 2001;44:2259–2269.

64. Farrens DL, Altenbach C, Yang K, Hubell WL, Khorana HG. Requirement of rigid-body motion of transmembrane helices for light activation of rhodopsin. Science 1996;274:768–770.

65. Dunham TD, Farrens DL. Conformational changes in Rhodopsin. J Biol Chem 1999;274:1683–1690.

66. Jensen AD, Guarnieri F, Rasmussen SGF, Asmar F, Ballesteros JA, Gether U. Agonist-induced conformational changes at the cytoplasmatic side of transmembrane segment 6 in the β2 adrenergic receptor mapped by site-selective fluorescent labeling. J Biol Chem 2001;276:9279–9290.

67. Paul N, Rognan D. ConsDock: A new program for the consensus analysis of protein-ligand interactions. Proteins 2002;47:521–533.

68. Ghanouni P, Gryczynski Z, Steenhuis JJ, Lee TW, Farrens DL, Lakowicz JR, Kobilka BK. Functionally different agonists induce distinct conformations in the G protein coupling domain of the β2 adrenergic receptor. J Biol Chem 2001;276:24433–24436.