# Modeling the Similarity and Divergence of Dopamine $D_2$-like Receptors and Identification of Validated Ligand−Receptor Complexes

Frank Boeckler,[†] Harald Lanig,[‡] and Peter Gmeiner*[,†]

*Department of Medicinal Chemistry, Emil Fischer Center, Friedrich-Alexander University, Schuhstrasse 19, D-91052 Erlangen, Germany, and Computer-Chemie-Centrum, Friedrich-Alexander University, Nägelsbachstrasse 25, D-91052 Erlangen, Germany*

*Received May 27, 2004*

Focusing on the similarity and divergence of GPCR subtypes and their ligand interactions, we generated dopamine $D_2$, $D_3$, and $D_4$ receptor models based on the rhodopsin crystal structure and refined these with an extensive MM/MD protocol. After validation by diagnostic experimental data, subtype-specific relative positions of TM1, 2, 6, and 7 and bending angles of TM7 were found. To sample the conformational space of the complex, we performed simulated-annealing runs of the receptor protein with the sub-nanomolar antagonist spiperone. Docking a representative set of ligands, we were able to identify one superior model for each subtype when excellent correlations between predicted energies of binding and experimental affinities ($r^2$ = 0.72 for $D_2$, 0.91 for $D_3$ and 0.77 for $D_4$) could be observed. Further analysis revealed general ligand interactions with ASP3.32 and aromatic residues in TM6/7 and individual key interactions with TM1 and TM2 residues of the $D_3$ and $D_4$ receptor models, respectively.

## Introduction

It has been well-known since the early 1990s that the essential neurotransmitter dopamine mediates its physiological effects through interaction with two receptor subfamilies, the $D_1$- and $D_2$-like, consisting of five different subtypes.[1] Especially the $D_2$-like receptors $D_2$, $D_3$, and $D_4$, which share the coupling to $G_{i/0}$ proteins and can inhibit adenylyl cyclase,[2] have been associated with a variety of neuropathological diseases, brain disorders, and aspects of drug addiction as well as normal personality traits. As members of the class A rhodopsin like G protein-coupled receptors, they are basically composed of seven transmembrane helices (TM1−7) connected by three intracellular (IL1−3) and extracellular loops (EL1−3). According to the crystal structure of bovine rhodopsin,[3] another short helix (H8), which is directly connected to TM7, is expected to be found at the C-terminal end. The sequence lengths of the subtypes are quite different and depend largely on the length of the huge intracellular loop 3 (IL3), for which several splice variants have been reported by molecular-genetic studies on the different subtypes.

An important determinant of subtype individuality is the neuroanatomical region-specific expression. Because of its broad distribution, the $D_2$ receptor is important for mediating the effects of dopamine to control movement, certain aspects of behavior in the brain, and prolactin secretion from the anterior pituitary gland. Hence, it plays an established role in the therapy of schizophrenia and Parkinson's disease. $D_2$ is well-known as the primary pharmacological target for the classical antipsychotics, such as haloperidol, and is responsible for the extrapyramidal side effects of these compounds, mediated through $D_2$ receptors in the nigrostriatal system, as well as the endocrine side effects, which are mediated by $D_2$ receptors in the tubero-infundibulary system. Because of these problems, more attention has been drawn to the $D_3$ and $D_4$ receptors, which are known to be more selectively localized in limbic structures for $D_3$ and in neocortical and putatively limbic structures for $D_4$. Thus, $D_3$ has been ascribed therapeutic value for the treatment of disorders such as schizophrenia,[4] Parkinson's disease,[5] drug-induced dyskinesia,[6] and cocaine addiction.[7] Because of the high $D_4$ affinity of the exceptional atypical antipsychotic clozapine and its preference for $D_4$ over $D_2$ receptors, the $D_4$ subtype is considered to be an important target for the treatment of positive and negative symptoms of schizophrenia.[8] Unfortunately, this could not be corroborated by clinical trials, yet. Moreover, the $D_4$ receptor and, in particular, $D_4$ polymorphism have been suggested to be involved in the pathogenesis of other neuropathological diseases such as attention-deficit hyperactivity disorder (ADHD), obsessive-compulsive disorder (OCD) with tics, Tourette's syndrome, and drug abuse (excluding alcoholism).[8] Hence, it may be a primary target for the therapy of these diseases. Recently, treatment of erectile dysfunction was also discovered as a novel possible field of application for $D_4$ (partial) agonists.[9]

Considering all of these pharmacological aspects, it is clear that much effort has been devoted to obtaining highly potent and subtype-selective agonists, partial agonists and antagonists.[8,10] In this context, we recently reported the rationally based design of the highly selective complete $D_4$ receptor antagonist FAUC 213,[11] which exhibits characteristics of atypical antipsychotics.[12] In addition, we demonstrated an interactive drug discovery process, leading to the superpotent and highly selective $D_3$ partial agonist FAUC 346 and antagonist FAUC 365.[13] For another selective $D_3$ partial agonist

* To whom correspondence should be addressed. Phone: +49(9131)8529383. Fax: +49(9131)8522585. E-mail: gmeiner@pharmazie.uni-erlangen.de.
[†] Department of Medicinal Chemistry.
[‡] Computer-Chemie-Centrum.

10.1021/jm049612a CCC: $30.25 
Published on Web 01/08/2005

**Table 1.** Percent Identity and Similarity of the Complete Aligned Sequence and the Structurally Conserved Regions Only (SCR)

| | total % sequence identity/similarity | | | SCR % sequence identity/similarity | | |
|---|---|---|---|---|---|---|
| | D2DR[a] | D3DR[b] | D4DR[c] | D2DR[a] | D3DR[b] | D4DR[c] |
| D3DR[b] | 50/64 | | | 78/88 | | |
| D4DR[c] | 32/48 | 35/49 | | 50/71 | 53/72 | |
| OPSD[d] | 15/32 | 18/32 | 13/29 | 27/59 | 28/53 | 26/55 |

[a] $hD_{2\ long}$ (Swiss-Prot accession number: P14416-1). [b] $hD_3$ (P35462-1). [c] $hD_{4.7}$ (P21917). [d] OPSD_BOVIN (P02699).

of this series (FAUC 329), we were able to show an attenuation of 1-methyl-4-phenyl-1,2,3,6-tetrahydropyridine (MPTP) neurotoxicity, indicating a protective effect against dopamine depletion, predominantly in the nucleus accumbens of mice.[14]

With respect to the recent progress in the design of such highly selective compounds, we were intrigued by the question of how to explain such potency and subtype selectivity on a molecular level. Extending a number of three-dimensional GPCR models,[15,16] Livingstone et al.[17] reported a pioneering comparison of the rat $D_2$-like receptors based on the model of bacteriorhodopsin. Since 2000, when the crystal structure of bovine rhodopsin became available,[3] several modeling investigations dealing with discrete dopamine receptors have been published.[18−22] Hence, it should be noted that, to our knowledge, this is the first extensive study to systematically compare models of all three human $D_2$-like receptors based on the more recent structural template. Using a comparative modeling strategy, we first focused on the common characteristics and structural divergences of the $D_2$-like receptors. Thus, in this study we report the construction and refinement of $D_2$, $D_3$, and $D_4$ receptor models, followed by intensive analysis of their structural features. Subsequently, we applied a combined simulated-annealing and ligand-docking approach in order to identify significant complexes for each subtype. These were validated by the correlation between the estimated binding energy and the experimental affinities of all docked ligands.

## Results and Discussion

**1. Sequence Alignment and Degree of Conservation.** So far, X-ray crystallographic or nuclear magnetic resonance structural data are available for the dopamine or any of the aminergic G protein-coupled receptors. Thus, Palczewski et al.'s crystal structure (PDB entry: 1F88)[3] of bovine rhodopsin (BovR) has been widely used as the state-of-the-art template in homology modeling of GPCRs. Although BovR has only low sequence identity to other GPCRs, the specific arrangement of the transmembrane helices seems to be conserved among family A receptors.[23] This can be explained by the common observation that the decrease in model accuracy of sequences with less than 30% identity is mainly due to a rapid increase in alignment errors.[24] However, elaborate alignment studies[25,26] have yielded patterns of high conservation, which help to guide subsequent alignments and, therefore, strongly reduce the risk of large errors. Moreover, although the $D_2$-like receptors share only 13% to 18% sequence identity with BovR (Table 1), these percentages are almost doubled for the structurally conserved regions (SCR),[27] which scarcely differ from the transmembrane helices. In addition, when extending the comparison criteria from identity to similarity, as defined by the use of a Gonnet250 similarity matrix,[28] 29% to 32% similarity with BovR is found for the total sequence and even 53% to 59% for the SCRs. On the basis of experiences with the dopamine $D_2$,[18] muscarinic $M_1$[29] and histamine $H_4$ receptors,[30] the structure of rhodopsin seems to be a suitable template for such amine receptors, whereas it has been shown to be inappropriate for the $CCK_1$ (peptide) receptor.[31]

The sequence alignment between bovine rhodopsin and the $D_2$-like receptors, as shown in Figure 1, was generated using ClustalX.[32] We obtained an initial alignment by employing a Gonnet 250 protein weight matrix[28] with a standard gap-opening parameter of 10.0 and extension parameter of 0.1. Afterward, we checked and, where necessary, corrected this alignment with GeneDoc[33] in order to reflect the known alignment features of class A GPCRs, such as the highly conserved positions and gap-free transmembrane regions. According to the work of Ballesteros and Weinstein,[25] each transmembrane helix contains one most conserved position, which is assigned the reference number "helix#.50" (see Figure 1). These residues were used to verify and adjust the alignment, complemented by $CYS^{3.25}$, which is conserved because of the essential disulfide bond between TM3 and EL2, $ASP/GLU^{3.49}$, $TYR/PHE^{3.51}$ and $GLU^{6.30}$, which are conserved as part of the arginine-cage motif in TM3/6, as well as $PHE^{6.44}$, $CYS^{6.47}$, $TRP^{6.48}$, $PHE/TYR^{6.51}$, $ASN^{7.49}$, $ASN^{7.57}$, $PHE^{7.60}$ and $ARG^{7.61}$, which have been suggested to be important for the fold of TM6/7 and partially involved in the activation mechanism of the receptors. Parts of the sequence comprising the transmembrane domains and a few relevant adjacent residues (framed by a blue border in Figure 1) were later regarded as the structurally conserved regions (SCRs) and, thus, no gaps were allowed within them. In addition to this knowledge-based refinement process, we inspected the final alignment for consistency with data from the literature[18,22] and found it to be in good agreement with previously published alignments. As summarized in Table 1, sequence similarities and identities between the dopamine receptors and toward bovine rhodopsin were calculated on the basis of the corrected alignment using BioEdit[34] and, again, using a Gonnet 250 matrix. A comparison among dopamine receptors reveals a higher degree of identity/similarity between $D_2$ and $D_3$ (88% similarity in SCRs) than found for $D_2/D_4$ (71% similarity in SCRs) or $D_3/D_4$ (72% similarity in SCRs).

**2. Construction of the Homology Models.** Using the adjusted alignment, homology models of the three $D_2$-like subtypes were constructed within the Composer[35] suite of programs implemented in Sybyl 6.9.[36] For each of the three crystal structures of bovine rhodopsin (1F88, 1HZX, 1L9H), we used chain A as a template to built 3D coordinates for the SCR backbone of the corresponding dopamine receptor sequence. Initial side chain orientations were determined using the template residue conformation, as well as a database of secondary structure-dependent canonical residue conformations, which is incorporated in the Composer package. The loops connecting the SCRs were modeled

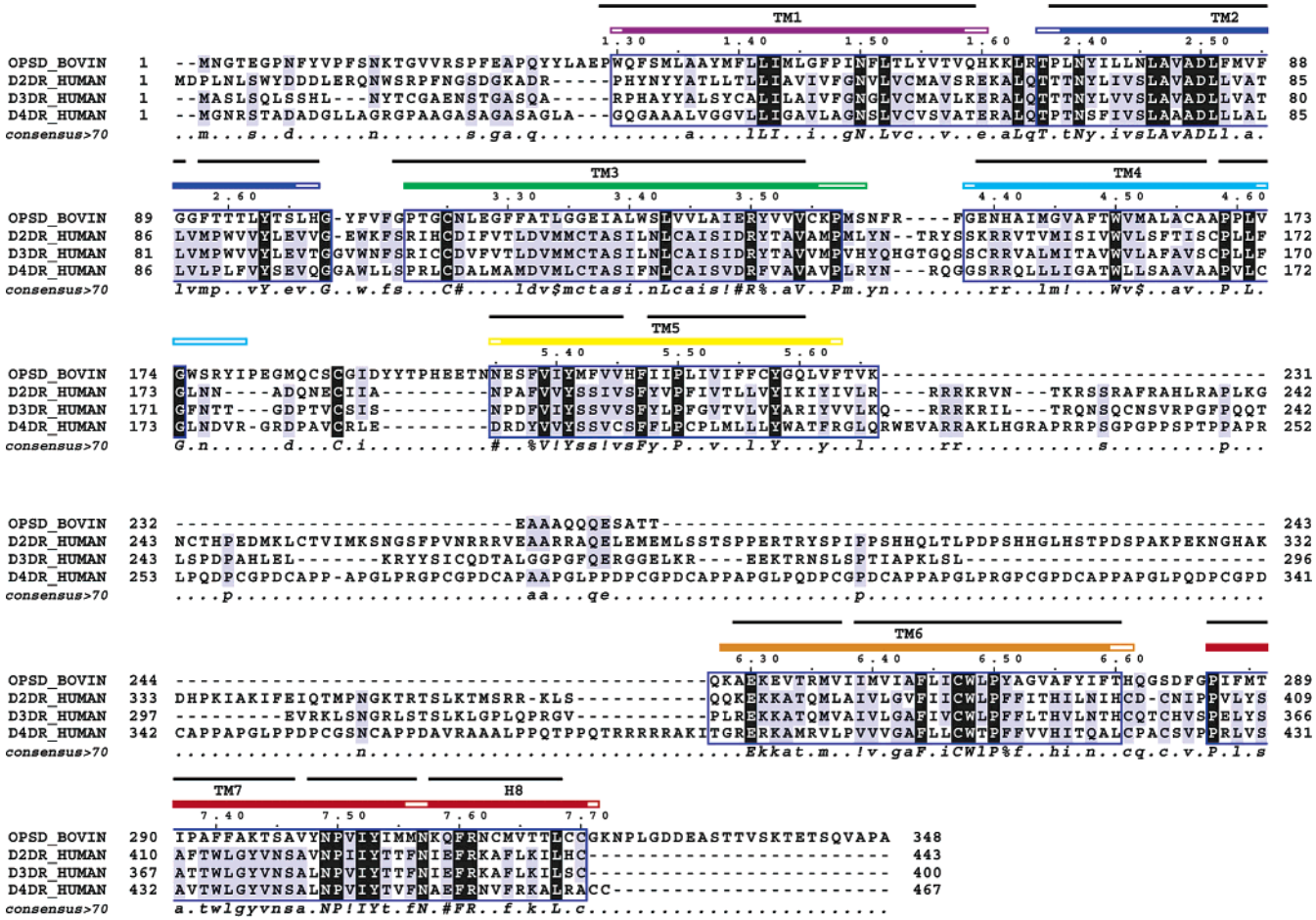



**Figure 1.** Multiple sequence alignment of bovine rhodopsin (OPSD_BOVIN) and the $D_2$, $D_3$, and $D_4$ receptors created with ClustalX.[32] Only marginal manual improvements were necessary using the alignment editor GeneDoc.[33] A consensus sequence was generated by applying standard criteria from MULTALIGN:[74] uppercase is identity, lowercase is consensus level > 0.7, ! is any one of the amino acid groups IV, $ is any one of LM, % is any one of FY, # is any one of NDQEBZ. Residues contributing to this consensus are drawn as white letters on black background, when strictly conserved throughout all 4 sequences, or shaded in gray when conserved in terms of any other MULTALIGN rules. The transmembrane parts TM1 to TM7 and the adjacent helix H8 are denoted above the sequences. The black bars on top of these labels indicate the helical areas found in the crystal structure of rhodopsin (1F88), whereas the colored bars below reflect helical structures found in all final models of the dopamine receptors. The parts of the bars without filling stand for areas that are found to be helical in some dopamine receptor models only. Beneath, the Ballesteros and Weinstein numbering scheme[25] is given with the dots indicating each fifth and the ticks indicating each tenth position. This scheme is used throughout this work, when positions of specific interest are labeled or discussed. Parts of the sequence enclosed by a blue border are reckoned as structurally conserved regions (SCRs) in the following modeling process.

with the Composer Loop Builder algorithm. All loops of identical length within a database of known structures were scored for their sequence homology toward the target loop and their RMS fit to the anchor regions. Loop fragments with steric overlaps with conserved regions or that were incompatible with the disulfide bonding geometry were rejected. In general, the loop structure with the best combination of high homology and minimal RMS fit was chosen. As the accuracy of this method's prediction is known to decrease with increasing loop length, it was found impossible to create an appropriate model for the huge IL3, which consists of 147 residues in $D_{2L}$, 104 residues in $D_3$, and 165 residues in $D_{4.7}$. Moreover, the size of the original IL3 in bovine rhodopsin is only 12 residues. Consequently, it cannot provide a structural basis for loop modeling. Interchanging this loop between the $D_2$/$D_3$ sequences in $D_2$/$D_3$ chimeras was shown to have no effect on ligand binding.[37,38] Hence, the IL3 has been suggested not to be crucial for ligand binding, which also justifies omitting this loop. Probably because of the absence of a second anchor region, the 26 to 31 amino acid long N-terminus also failed to yield a satisfactory model. Since there is no evidence from the literature for an involvement of this region in ligand binding, we decided to omit it, too. Consequently, the number of residues considered for all further steps of refinement and simulation was decreased from 443 to 265 (−40%) for $D_2$, from 400 to 270 (−33%) for $D_3$, and from 467 to 271 (−42%) for $D_4$. Thus, the simplification of the system leads to a reduction of computational load and more efficient use of resources, especially with respect to the extensive MD and docking investigations performed in this study. However, it is clear that the exclusion of the N-terminus and the IL3 potentially represents a limitation of our approach. However, Varady et al.[22] recently demonstrated that omission of even more loops still allows for the construction and simulation of a meaningful $D_3$ model, which was used successfully by the authors for structure-based virtual screening. As we can assume that IL3 stabilizes the packing of TM5 and TM6 to some extent, even though it is quite large, we used a

**Table 2.** RMSD (Å) of the SCR Backbone in the Homology Models after Construction with Composer and in the Final Models after the Molecular Dynamics Refinement Protocol

| | homology models | | | minimized models | | | final models | | |
|---|---|---|---|---|---|---|---|---|---|
| | 1F88[a] | D2DR[b] | D3DR[c] | 1F88[a] | D2DR[b] | D3DR[c] | 1F88[a] | D2DR[b] | D3DR[c] |
| D2DR[b] | 0.202 | | | 0.602 | | | 2.536 | | |
| D3DR[c] | 0.193 | 0.171 | | 0.445 | 0.533 | | 2.668 | 2.903 | |
| D4DR[d] | 0.196 | 0.191 | 0.200 | 0.437 | 0.603 | 0.387 | 2.528 | 2.517 | 3.123 |

[a] Crystal structure of bovine rhodopsin (template). [b] Homology/minimized model of the human $D_{2L}$ receptor/final, selected model $D_2^{4280}$. [c] Homology/minimized model of the human $D_3$ receptor/final, selected model $D_3^{3340}$. [d] Homology/minimized model of the human $D_{4.7}$ receptor/final, selected model $D_4^{4870}$.

relatively loose distance restraint between the unconnected ends of TM5 and TM6 as a precaution during the subsequent MD simulations.

Structural comparison among all dopamine receptor models and with the template structure of bovine rhodopsin (crystal structure: 1F88) reveals that, directly after the homology modeling process, all models are still closely related to each other as well as to the template structure. RMSD values measured for the SCR backbone range from 0.17 to 0.20 Å, as shown in detail in Table 2.

**3. Refinement of the Model Structures.** Before energetic refinements were started, a coarse relaxation of strain in the side chain groups was accomplished by using the Sybyl Fix_Sidechain command. The rotatable side-chain bonds were scanned in increments of 3° from the starting geometries for positions with no close van der Waals (vdW) contacts. The "hardness" of the van der Waals spheres was adjusted during this scan by scaling the vdW radii with a default factor of 0.9. On the basis of previous studies, $\chi_1$ of TRP$^{6.48}$ was modified from gauche$^+$ to trans to allow more favorable $\pi$-stacking interactions. The orientation of the amide groups in ASN and GLN residues was additionally evaluated with Sybyl Fix_Asn_Gln to achieve an optimal hydrogen-bonding network.

All further steps of the structural refinement were conducted with Amber5[39] using the Parm94 force field[40] and employing a distance-dependent dielectric constant of $\epsilon = 4r$ together with a nonbonded cutoff of 12 Å. To avoid spurious changes in the general fold and helix packing due to some still unfavorable electrostatic interactions or steric clashes, we applied positional restraints to the SCR-backbone atoms, which were released stepwise during the following four iteration cycles of minimization. The first three iterations consisted of a combination of 1000 steps of steepest descent and up to 9000 steps of conjugate gradient, while in the last iteration 1000 steps of steepest descent and up to 14000 steps of conjugate gradient were performed to ensure that the models converged to a local minimum. The harmonic force constant of 100.0 kcal·mol$^{-1}$·Å$^{-2}$ used for the initial positional restraint was modified to 10.0 kcal·mol$^{-1}$·Å$^{-2}$ for the second iteration, to 5.0 kcal·mol$^{-1}$·Å$^{-2}$ for the third, and to 1.0 kcal·mol$^{-1}$·Å$^{-2}$ for the last. Residue-based RMSDs were visually inspected with DeepView. We found in all subtype models a core area of minor RMSD values comprising most of the binding site, while the RMSD values increase in the direction of the borders of the restrained structurally conserved regions. Additionally, the RMSD values are slightly increasing toward the cytoplasmatic side of the receptor.

For a similar modeling task, Strahs and Weinstein[41] have demonstrated that initially structurally closely related models in principle have the ability to evolve into alternative structures during MD simulations. Moreover, they concluded that this ability is necessarily absent from GPCR models only constructed by minimization techniques and, thus, abstaining from MD simulation would yield a misleading conformity of structures, neglecting the available conformational space. We thus subjected the minimized models to molecular dynamics simulations with a total duration of 6000 ps each. By applying Shake bond-length restraints to bonds involving hydrogens, we were able to set the integration step size to 2 fs. The nonbonded interactions were updated every 25 steps. The receptor models were gradually heated to a simulation temperature of 310 K within 100 ps, preventing strong and unnatural deformations of the models within this long initial simulation phase. The simulation temperature was controlled by coupling to an external heat bath (coupling constant: 0.2 ps) for the rest of the 6 ns. When MD simulations are performed in the gas phase, disregard of the explicit macroscopic environment necessitates the use of a set of restraints, replacing the natural stabilizing effects of the membrane bilayer on helix length and packing. According to Sybyl standard parameters, hydrogen bonds, which are found to stabilize the helical conformation, were determined in the backbone of the minimized models. For the first 2.0 ns, semiharmonic restraints with a force constant of 10.0 kcal·mol$^{-1}$·Å$^{-2}$ were applied to these hydrogen bonds in order to preserve the actual interactions between the amide hydrogens and the carbonyl oxygens. The harmonic part of the restraint was set to change into a linear gradient below 0.8 Å and above 2.8 Å H-bond distance. After 2.0 ns, the force constant was gradually reduced to 5.0 kcal·mol$^{-1}$·Å$^{-2}$ over a time period of 100 ps, allowing the system to relax. According to preliminary studies about the optimal simulation setup (data not shown), we decided to refrain from decreasing the force constant below this value. In comparison to others, we suggest that this type of restraint is primarily geared to maintain the original structural information contained in the crystal structure of bovine rhodopsin, while, at the same time, giving the models freedom to evolve according to the structural characteristics of the subtypes. During all simulations, a "soft" (0.5 kcal·mol$^{-1}$·Å$^{-2}$) and a "hard" (10.0 kcal·mol$^{-1}$·Å$^{-2}$) semiharmonic distance restraint was additionally applied to both intracellular ends of TM5 and TM6, replacing the effect of the omitted large IL3. The "soft" restraint was set up to rise in parabolic shape up to a deviation of 5.0 Å from its current value and become linear afterward, having the slope of the parabola at 5.0 Å. The "hard" restraint, in contrast, was set up to equal 0.0 kcal·mol$^{-1}$·Å$^{-2}$ within a deviation of 5.0 Å, rising in a parabolic shape

between 5.0 and 10.0 Å and becoming linear with the slope of the parabola at 10.0 Å, afterward. The average wall clock time for a 1 ns simulation extrapolated to a single MIPS R14000-CPU (Origin3400 system) was about 35 h.

**4. Analysis of the Trajectories.** Two-dimensional root-mean-square deviation (2D-RMSD) plots (see Figure S1A in the Supporting Information) were generated using the Amber module Ptraj. In combination with plots of the potential energy (Figure S1B), convergence of the trajectories was monitored. Furthermore, these plots were used to guide clustering of the trajectories into homologous segments. For all three simulations, a stabilization of the potential energy is achieved after 1.5 to 2.5 ns. Fluctuations of the potential energy are, of course, apparent throughout the trajectory, allowing for the transition of the models into other energetically comparable conformational clusters. However, no significant trend to decrease the energy is observable from the smoothed plots. The number of clusters was chosen to comprise as many homologous structures of the trajectory as possible and, simultaneously, not to exceed an averaged RMSD value of 1.0 Å. In fact, the averaged RMSDs range between 0.89 Å to 1.02 Å for comparing the backbone atoms of the structurally conserved regions only and between 1.59 Å to 1.82 Å for comparison of all SCR atoms (see Table S1 in the Supporting Information). A weak tendency for a decrease of these averaged values toward the end of the trajectory can be found in the $D_2$ and $D_4$ simulations. Both the low standard deviations of about 0.12 Å for the backbone and 0.19 Å for the all atom comparison and the moderate maximal RMSD values (1.59−1.82 Å for the backbone and 2.14−2.66 Å for the all atom comparison) indicate that appropriate cluster dimensions have been selected.

The large number of cluster members (between 560 and 1720 snapshots taken each ps) clearly shows the structural stability of the trajectory within the simulation areas. Thus, we extracted a representative structure from each cluster by averaging over all conformations and performing stepwise optimizations with gradually reduced backbone restraints. At the first three stages, the averaged models were minimized for 100 steps using the steepest descent method, followed by up to 5000 steps of conjugate gradient, while at the final stage of optimization up to 1000 steps of steepest descent, followed by up to 10000 steps of conjugate gradient optimization, were performed. The convergence criterion for the energy gradient was set to 0.05 kcal·mol$^{-1}$·Å$^{-1}$ for the first two stages and afterward lowered to 0.01 kcal·mol$^{-1}$·Å$^{-1}$. The positional restraint on the SCR backbone was initially set to 1000.0 kcal·mol$^{-1}$·Å$^{-2}$ and afterward modified to 10.0 and 0.1 kcal·mol$^{-1}$·Å$^{-2}$. Only at the final stage was a completely restraint free optimization enabling full freedom of protein flexibility used in order to obtain unbiased and comparable energies. This careful, successive refinement protocol is essential, because averaging over a huge number of conformations is likely to result in deformations of the protein side chains. The maximal number of optimization steps at each stage was proven to be large enough that every minimization reached convergence as judged by the gradient criterion. On this basis, comparing final potential-energy values of structures representing different clusters of the trajectory to each other is significant. For $D_2$ and $D_4$, again, a clear tendency for more favorable energies is found toward the end of the trajectories, whereas for $D_3$, which reaches a constant energy level faster, the cluster representatives $D_3^{910}$ and $D_3^{3340}$ turned out to be most favorable at each stage of optimization. (For details of the optimization, see Table S2 in the Supporting Information.)

**Table 3.** Comparison[a] between Averaged, Minimized Cluster Representatives (Values in Å)

(a) Representative Structures of the $D_2$ Clusters

| | 1F88[b] | $D_2^{220}$ | $D_2^{1570}$ | $D_2^{2770}$ |
|---|---|---|---|---|
| $D_2^{220}$ | 2.434 | | | |
| $D_2^{1570}$ | 2.602 | 0.842 | | |
| $D_2^{2770}$ | 2.431 | 0.756 | 0.894 | |
| $D_2^{4280}$ | 2.536 | 1.013 | 1.071 | 0.960 |

(b) Representative Structures of the $D_3$ Clusters

| | 1F88[b] | $D_3^{910}$ | $D_3^{2630}$ | $D_3^{3340}$ | $D_3^{3990}$ |
|---|---|---|---|---|---|
| $D_3^{910}$ | 2.722 | | | | |
| $D_3^{2630}$ | 2.737 | 0.504 | | | |
| $D_3^{3340}$ | 2.668 | 0.675 | 0.630 | | |
| $D_3^{3990}$ | 2.729 | 0.573 | 0.501 | 0.623 | |
| $D_3^{5240}$ | 2.796 | 0.635 | 0.605 | 0.740 | 0.504 |

(c) Representative Structures of the $D_4$ Clusters

| | 1F88[b] | $D_4^{300}$ | $D_4^{1270}$ | $D_4^{2350}$ | $D_4^{3820}$ |
|---|---|---|---|---|---|
| $D_4^{300}$ | 2.640 | | | | |
| $D_4^{1270}$ | 2.682 | 0.663 | | | |
| $D_4^{2350}$ | 2.612 | 1.145 | 0.873 | | |
| $D_4^{3820}$ | 2.589 | 1.100 | 0.954 | 0.737 | |
| $D_4^{4870}$ | 2.528 | 1.112 | 0.981 | 0.717 | 0.362 |

[a] With respect to the RMSD of the SCR-backbone atoms. [b] Chain A of the crystal structure of bovine rhodopsin.

When comparing all minimized structures of a trajectory to each other and to chain A of the crystal structure of bovine rhodopsin, the calculation of the SCR backbone RMSD, given in Table 3, reveals a close relationship among the model structures (0.36−1.15 Å). In contrast, the RMSDs of the model structures versus the template crystal structure range between 2.43 and 2.80 Å. This divergence of all models from their common origin, together with the "convergence" of each model to a set of closely related conformational alternatives, indicates that, in the early parts of the simulation, a sequence-dependent evolvement of the model structures occurs, leading to distinct subtype-specific conformational clusters. This observation corresponds to findings of Strahs and Weinstein,[41] who described a characteristic development of model structures merely influenced by their "intrinsic energy requirements". Although the structural evolvement is achieved rapidly in our investigations, the slow transitions between alternative clusters demand a long-term simulation of the models. We have to assume that further transitions will probably occur with extended simulation time. Nevertheless, with 4 to 5 conformational clusters already found, we regard the selected duration of 6 ns as a reasonable balance between the elaborateness of the investigation and the efficient use of computing resources. Analyzing the data given in Table 3 in more detail, we find a slightly enlarged average RMSD between structures of the $D_3$ trajectory and the template (2.73 Å), in comparison to structures of the $D_4$ (2.61 Å) or $D_2$ trajectories (2.50 Å).

Interestingly, this order becomes inverted with respect to the average RSMD between the representative structures of the different conformational clusters. Here, $D_3$ shows a notably lower average RMSD of 0.60 Å than $D_4$ (0.86 Å) or $D_2$ (0.92 Å).

In order to check the entire refinement approach, we investigated bovine rhodopsin under identical conditions. Visualization of the 3D structures clearly indicated that BovR retained its folding very well (see Supporting Information: Figure S2) without noticeable deviations from the crystal structure. Only at the cytoplasmatic end of TM5 and along helix 8 substantial differences can be seen. A total RMSD of 1.8 Å for the structurally conserved regions (1.4 Å without the cytoplasmatic end of TM5 and H8) reflects that only minor topology changes have occurred. Moreover, BovR proved to be most stable around the binding region.

**5. Validation and Selection of the Model Alternatives.** First, all alternative subtype models were analyzed extensively using the What If[42] module What_Check.[43] This analysis consisted of checks for standard parameters such as bond lengths, angles, torsions, and, in particular, $\phi$- and $\psi$-angles, as well as examinations of the rotamer probability, the packing quality, and hydrogen-bonding features. In general, no major problems were detected. However, all indications of uncommon properties were carefully considered. Most of these were found to refer to loop substructures or, in the case of unsatisfied hydrogen-bond donors or acceptors, may be caused by the omission of water molecules in the binding pocket or at the intra- and extracellular sides of the receptor models.

Taking the energy of the optimized models and the plausibility of the What_Check results into account, the cluster representatives $D_2^{4280}$, $D_3^{3340}$, and $D_4^{4870}$ were chosen as the most convincing subtype models. Hence, all further investigations relate to these structures or, when addressing dynamic aspects, the term "model trajectory" will be used to indicate that only those parts of the entire simulation used to derive the selected models were considered. In order to give a more detailed characterization of these subtype models, the corresponding Ramachandran plots prepared with Rampage[44] are provided as Supporting Information (Figure S3). Acting on suggestions obtained by What_Check, we optimized the hydrogen-bonding network of the remaining three model structures with the What If command Hb2net.

Besides the validation adapted from intrinsic structural characteristics, we additionally employed experimental data retrieved from the literature to test our final models. The extensive SCAM (substituted-cysteine accessibility method) studies of Javitch and co-workers[45−47] provide very useful information to validate our $D_2$ model. In order to identify residues that participate in forming the binding-site crevice, they mutated single consecutive residues in TM 1 to 7 of the $D_2$ receptor into CYS and determined their accessibility from the effects of sulfhydryl-specific methanethiosulfonates on ligand binding. As a water-accessible CYS is far more reactive with these reagents than a CYS facing the lipid bilayer or protein interior, this method is considered very useful to discriminate packing and orientation of the seven transmembrane helices. As depicted in Figure 2, the

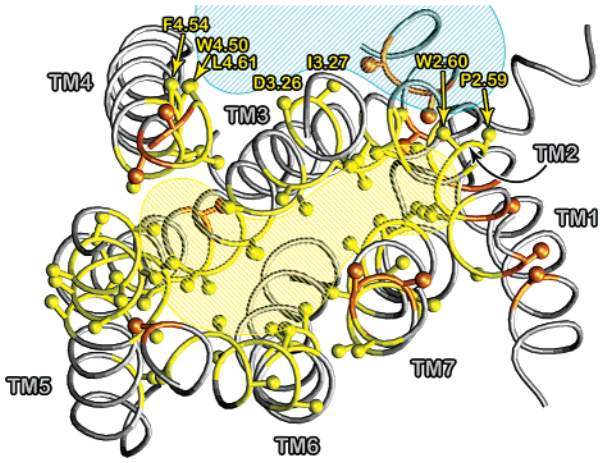



**Figure 2.** Illustrations of the model validation process based on direct and indirect experimental data. Projection of SCAM data from the literature on the structure of the $D_2$ model. C$\beta$ and its bond toward Cα are shown in ball-and-stick representation for residues found to be accessible after cysteine substitution. Residues colored in yellow have been described to be both accessible and protected, while orange residues are accessible, but not protected in the presence of a ligand. The yellow-shaded area indicates the binding pocket, and the cyan-shaded area marks the putative dimerization interface. Only exceptionally oriented residues that point to this interface are labeled. The figure was prepared with VMD.[75]

major part of the accessible residues (colored in yellow and orange) faces the central cleft, in good agreement with the experimental data. Only a few are pointing outward, and most of these are close to the extracellular ends, where two effects become relevant that can possibly explain this exceptional orientation. First, due to an attenuated stabilization of the helical secondary structure, increased dynamic flexibility is typically observable at the ends of the transmembrane regions, leading to broadly accessible segments. Second, the ends of the helices are probably not surrounded by the hydrophobic phospholipid chains, but by hydrophilic and water-associating phospholipid headgroups, which also cause increased accessibility of the adjacent, outward-facing residues. Another well-established exception is the highly conserved TRP$^{4.50}$, which is directed toward the lipid membrane together with PHE$^{4.54}$ (one helix turn above) and LEU$^{4.61}$ (three turns above), although all are found to be accessible. Recently, experimental evidence by cross-linking of CYS$^{6.58}$ has been reported that implicates this TM4-"back face" with homomeric $D_2$ receptor dimerization.[48] In addition, the residues PRO$^{2.59}$, TRP$^{2.60}$, ASP$^{3.26}$, and ILE$^{3.27}$ may also participate in this dimer interface, as all of them point in the right direction (meaning the putative region of interaction indicated in the cryo-electron microscopy structure of squid rhodopsin)[48,49] and were found to be accessible. A further phenomenon that has been discussed in the literature[18,50] is the appearance of 10 consecutive accessible residues in position 5.38 to 5.47 of the TM5, suggesting a generally enhanced flexibility of the extracellular pre-PRO$^{5.50}$ part of the helix or a coherence between different conformational states of TM5 and different functional states of the receptor. Very recently, Shi and Javitch[51] were able to show that even in the EL2, which connects TM4 and TM5 and is known to interact with retinal in the crystal structures of bovine

**Table 4.** Crossing Angles of Adjacent Helix Parts in the Model Trajectories[a]

| helix interaction | $D_2^{4280}$ model | $D_3^{3340}$ model | $D_4^{4870}$ model | bov rhodopsin |
|---|---|---|---|---|
| TM1−$TM2_{IC}$ | 137.5 ± 1.6 | 143.0 ± 1.3 | 146.2 ± 1.7 | 145.7 ± 0.3 |
| $TM2_{EC}$−TM3 | 129.4 ± 3.4 | 145.6 ± 3.6 | 144.6 ± 3.2 | 138.2 ± 2.0 |
| TM3−TM4 | 147.2 ± 1.2 | 146.6 ± 2.4 | 152.0 ± 1.6 | 146.1 ± 0.2 |
| TM4−$TM5_{EC}$ | 151.3 ± 2.3 | 152.9 ± 2.0 | 155.2 ± 2.2 | 158.5 ± 0.8 |
| TM3−$TM5_{IC}$ | 24.7 ± 2.5 | 17.6 ± 1.9 | 20.3 ± 2.5 | 23.6 ± 1.3 |
| $TM5_{IC}$−$TM6_{IC}$ | 140.5 ± 2.4 | 153.3 ± 2.6 | 146.7 ± 2.4 | 146.6 ± 0.4 |
| TM3−$TM6_{IC}$ | 141.3 ± 1.4 | 147.4 ± 1.5 | 154.0 ± 1.3 | 148.9 ± 1.0 |
| $TM5_{EC}$−$TM6_{EC}$ | 150.2 ± 2.8 | 160.4 ± 2.6 | 160.4 ± 2.8 | 148.9 ± 1.8 |
| $TM6_{EC}$−$TM7_{EC}$ | 148.0 ± 2.9 | 133.6 ± 3.2 | 140.1 ± 2.7 | 144.5 ± 0.9 |
| $TM6_{IC}$−$TM7_{IC}$ | 163.9 ± 3.4 | 143.1 ± 3.6 | 158.7 ± 4.6 | 157.6 ± 0.8 |
| $TM7_{IC}$−H8 | 73.3 ± 4.7 | 80.7 ± 4.5 | 114.8 ± 5.4 | 93.9 ± 0.8 |
| TM1−$TM7_{IC}$ | 39.4 ± 3.6 | 41.4 ± 4.1 | 39.4 ± 4.2 | 43.1 ± 0.8 |
| TM1−H8 | 80.9 ± 3.8 | 86.7 ± 3.9 | 100.4 ± 4.3 | 90.6 ± 0.6 |

[a] Interhelical crossing angles computed based on the Kahn algorithm,[52] which is used to determine the helical axes. TM2, TM5, TM6, and TM7 are regarded as two parts separated by the kink-inducing PRO residues in positions 2.59, 5.50, 6.50, and 7.50 and designated with the indices IC (intracellular) or EC (extracellular). Dihedral angles between the direction vectors of two consecutive or adjacent helices were measured for each 10 ps snapshot of the model trajectories. Values are presented as average ± standard deviation of 172 structures for $D_2$, 64 structures for $D_3$, and 113 structures for $D_4$. The averages and standard deviations given for bovine rhodopsin are obtained by comparable calculations applied on the three crystal structures: 1F88, 1HZX, 1L9H.

rhodopsin, the mutants I184C and N186C are highly reactive to various sulfhydryl reagents and I184C is significantly protected against this reaction by the bound ligand YM-09151-2. Like the major part of the SCAM data, these results are consistent with our model, where both residues are directed toward the binding site crevice.

**6. General Structural Characteristics of the Models.** By measuring the inter- and intrahelical angles, we quantified the global topology of the receptor models. We determined the interhelical crossing angles as dihedrals between the direction vectors of the helix axes, which were calculated using the Kahn algorithm.[52] Because of their characteristically kinked structures, TM2, TM5, TM6, and TM7 were split into extracellular (EC) and intracellular (IC) parts at the PRO residues in positions 2.59, 5.50, 6.50, and 7.50. In general, most of the angles between consecutive or adjacent helices deviate by 15° to 55° from an antiparallel orientation of 180° (Table 4). Narrow distributions of the crossing angles with standard deviations below 3° are most frequently found in the model trajectories. Only distinct substructures, such as the vertical helix 8 and the extracellular part of TM2, show enhanced flexibility, as suggested by atomic fluctuation results. Comparing the model structures to bovine rhodopsin, both the relationship of the models to their common origin and their subtype-specific development become obvious. While some interactions, such as TM3-TM4, TM4-$TM5_{EC}$, and TM1-$TM7_{IC}$, only change marginally, some differ only in one subtype ($TM2_{EC}$-TM3, $TM6_{EC}$-$TM7_{EC}$, and $TM6_{IC}$-$TM7_{IC}$) or have diverged generally in all models ($TM7_{IC}$-H8, TM1-H8). Especially the helices in TM6 and TM7 exhibit the largest variations between the dopamine receptors.

Because of the above-mentioned prolines in position 2.59, 5.50, 6.50, and 7.50, we analyzed not only interhelical crossing angles but especially intrahelical bend angles at the characteristic kinks in more detail. It is

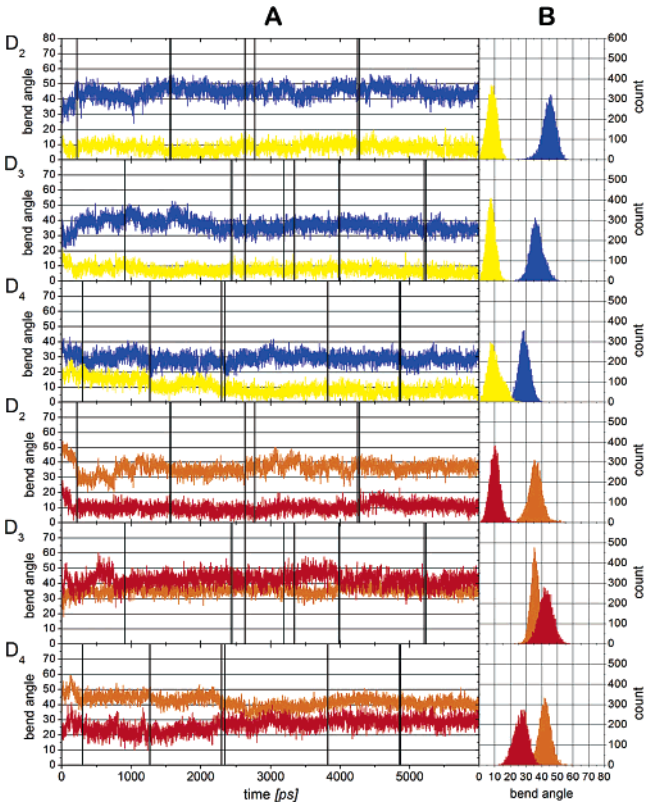



**Figure 3.** Variation of the bend angles of proline kinks in TM2 (blue), TM5 (yellow), TM6 (orange), and TM7 (red). The angles are defined according to the Prokink protocol of Visier et al.[76] and calculated using its implementation in Simulaid.[55] (A) Time evolution of the bend angle during the entire 6 ns of the simulation. With a sampling rate of one frame per ps, in total 6000 structures were analyzed for each receptor subtype. Vertical lines indicate borders of the trajectory clusters. (B) Histograms, showing the kink angle distribution over the entire trajectory, are plotted using a bin size of 0.5°.

well-known from the literature that PRO-containing regions are likely to cause structural deformations of regular α-helices, acting as molecular hinges, swivels, or switches.[53] As their side chain is covalently bound to the amide nitrogen of the backbone, kinks in the α-helices avoid steric clashes between the pyrrolidine ring and the backbone. In addition, the typical hydrogen bond between the amide nitrogen and the carbonyl oxygen in position (i-4) is lost.[54] Hence, we calculated the bend angles for TM2, TM5, TM6, and TM7 using the Prokink method of Visier et al.[26] as implemented in Simulaid.[55] In this procedure the helix is split into pre-proline and post-proline parts, with helix axes assigned to both by the Kahn method.[52] The bend angle is defined as the angle of deviation from a collinear orientation of the axes, as found in a normal α-helix. The graphs summarizing the evolution of this angle over the entire simulation (Figure 3) exhibit a higher flexibility during the first 500 ps, which is probably a consequence of the heating process (0 to 310 K during the first 100 ps). Afterward, most kinks are relatively stable, showing only slight or transient shifts, such as TM2 and TM6 in the $D_4$ trajectory after approximately 2.5 ns. Accordingly, the histograms reflect normal distributions of almost all kink angles with standard deviations between ±2.8° and ±4.6°. Focusing on the model structures and the related conformations within

their cluster, interestingly, very similar angles are found for all subtypes in TM5 ($D_2$, 8.2° ± 3.2°; $D_3$, 8.2° ± 2.6°; $D_4$, 7.6° ± 2.9°) and TM6 ($D_2$, 36.7° ± 2.9°; $D_3$, 34.2° ± 2.8°; $D_4$, 40.9° ± 2.7°). Additionally, these kink angles are fairly close to those found in rhodopsin (TM5, 13°; TM6, 37.6°). A connection between the extent of the PRO-kink in TM6 and receptor activation in aminergic GPCRs[56−58] was proposed in several recent publications. This proposal is also supported by experimental data on the water-accessible face of TM6 in the $D_2$ receptor.[59] Thus, it has been suggested that an explicitly kinked conformation is associated with the inactive state of the receptor, whereas a straightening of TM6 and a resulting movement of its intracellular part away from TM3 can account for receptor activation. According to this argumentation, it is noteworthy that all our subtype models show a similarly kinked TM6 and, therefore, should reflect the inactive receptor state. In contrast to this consistency, the kink in TM2 varies moderately between $D_2$ (45.2° ± 3.8°), $D_3$ (36.6° ± 3.0°), and $D_4$ (28.2° ± 3.1°) and even stronger differences are found for the bend angle in TM7 ($D_2$, 11.8° ± 3.1°; $D_3$, 46.5° ± 4.2°; $D_4$, 29.1° ± 3.3°). As the sequence identities and similarities in these areas are rather high, the differences in TM7 are quite surprising. However, two interactions in the interface between TM6 and TM7 may be responsible for this. The residues in positions 6.54 and 7.38 are proximate in all structures, but the van der Waals interactions and spatial requirements for THR and PHE ($D_2$), THR and THR ($D_3$), or VAL and VAL ($D_4$) necessarily distinguish between the subtypes. Another set of adjacent residues are VAL$^{6.43}$ and VAL$^{7.48}$ in the $D_2$ receptor, for which the arrangement of interaction is reversed in $D_3$ and $D_4$ (ALA$^{6.43}$ and LEU$^{7.48}$). For comparison, the kink angles of BovR as calculated by the Kahn method[52] are 32.4° for TM2, 13.0° for TM5, 37.6° for TM6, and 35.6° for TM7.

Comparing the topology of the other model structures pairwise to the $D_2$ receptor model (Figure 4), the most consistent elements found are TM4 and TM5, as well as the intracellular parts of TM3 and TM6. The extracellular portion of TM3 is slightly shifted toward the membrane in $D_3$ or toward the crevice in $D_4$. For both $D_3$ and $D_4$, the extracellular portion of TM6 is displaced in the direction of the receptor center, while in $D_4$, the whole TM6 is relocated downward to the cytoplasma by more than half a turn. In $D_3$, TM1 and TM2 are simultaneously tilted more upright relative to the other helices, thereby markedly increasing the accessibility of TM1 from the binding pocket. In contrast, TM1 and TM2 in the $D_4$ model are only moderately moved outward. Corresponding to the huge discrepancy of the kink angles, TM7 in the $D_3$ receptor is deeply bent into the central crevice and, hence, the intracellular end with helix 8 attached to it is somewhat shifted upward to the receptor center. In the $D_4$ model, the extracellular end of TM7 is more strongly bent outward, with otherwise only marginal repositioning compared to $D_2$. As shown in Table 2, the overall RMSD of the SCR backbone increased from about 0.2 Å after the homology modeling to 2.5−3.1 Å after the MD refinement. While there are some varieties in the comparison among the $D_2$-like receptors, the RMSD of the different subtypes compared to their common template, the crystal structure of rhodopsin, is uniformly 2.5 to 2.6 Å.

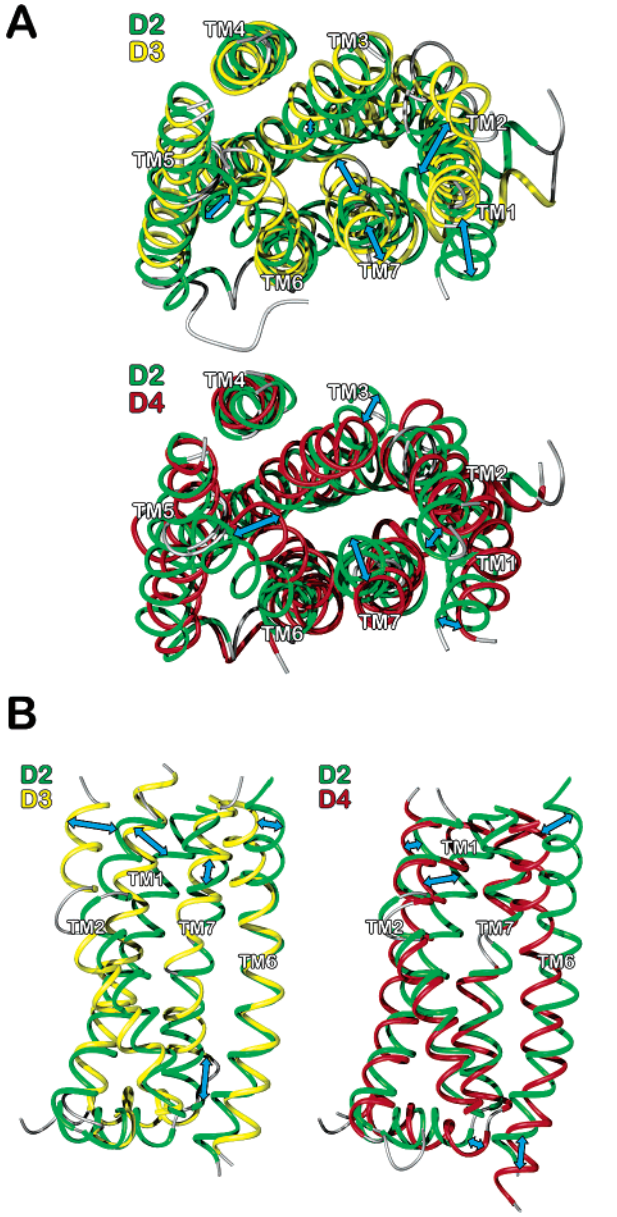



**Figure 4.** Pairwise superimposition of the final model structures of the $D_2$, $D_3$, and $D_4$ receptors. Shifts of major interest are indicated with blue double-headed arrows. All transmembrane helices are shown from an extracellular viewpoint in panel A, and only TM1, TM2, TM6, and TM7 are shown from a viewpoint in the membrane in panel B. For easier comparison, the $D_2$ receptor is always shown consistently in both pictures of panels A and B. The figure was prepared with MOLMOL.[77]

**7. Generation and Simulation of Ligand−Receptor Complexes.** In order to model and understand characteristic interactions of typical antagonists at the $D_2$, $D_3$, or $D_4$ receptor and to make our models applicable for computer-aided drug design, we used an extensive strategy of (1) exploring the ligand-bound conformational space of the receptors, (2) docking a diverse set of ligands, and (3) validating the obtained alternative models by their correlation to experimental binding affinities. This strategy was inspired by the work of other groups who have described the usefulness of ligand-assisted homology models for drug development.

**Table 5.** Receptor Binding and Selectivity Ratios for the Diverse Set of Dopamine Ligands Used for Validation Docking

| compound | | $K_i$ values[a] (nM) | | | ratio of $K_i$ values | | |
|---|---|---|---|---|---|---|---|
| no. | name | $hD2_{long}$ | hD3 | hD4 | $D2_{long}$/D3 | $D2_{long}$/D4 | D3/D4 |
| **1** | spiperone | 0.12 | 0.25 | 0.33 | 0.48 | 0.36 | 0.76 |
| **2** | FAUC 113 | 3200[b] | 5000[b] | 3.6[b] | 0.64 | 890 | 1400 |
| **3** | FAUC 213 | 3400[b] | 5300[b] | 2.2[b] | 0.64 | 1500 | 2400 |
| **4** | FAUC 365 | 3600[c] | 0.50[c] | 340[c] | 7200 | 11 | 0.0015 |
| **5** | | 31000[b] | 10000[b] | 1300[b] | 3.1 | 24 | 7.7 |
| **6** | haloperidol | 1.3 | 10 | 7.3 | 0.13 | 0.18 | 1.4 |
| **7** | | 110[c] | 1.1[c] | 30[c] | 100 | 3.7 | 0.037 |
| **8** | | 290[c] | 360[c] | 0.67[c] | 0.81 | 430 | 540 |
| **9** | sulpiride | 120 | 120 | 2100 | 1.0 | 0.057 | 0.057 |

[a] All affinities were obtained under sodium-free conditions. [b] Data corresponding to Löber et al. (2001).[11] [c] Data corresponding to Bettinetti et al. (2002).[13]

Bissantz et al.[19] constructed models of the dopamine $D_3$, muscarinic $M_1$, and vasopressin V1a receptors in their "antagonist-bound" form by minimizing the rhodopsin-derived homology models in the presence of a docked antagonist. These models were tested for their ability to discriminate known antagonists from randomly selected "druglike" compounds and, indeed, proved to be suitable for virtual screening purposes. With Mobile, Evers et al. describe a novel approach to integrate information about bioactive ligands at the stage of homology modeling, which they successfully applied to model factor Xa and aldose reductase.[60] Additionally, Evers and Klebe most recently also reported the use of this method for GPCRs, as they were able to obtain a homology model of the $NK_1$ receptor which was demonstrated to be capable of identifying putative lead compounds with submicromolar affinity by virtual screening.[61]

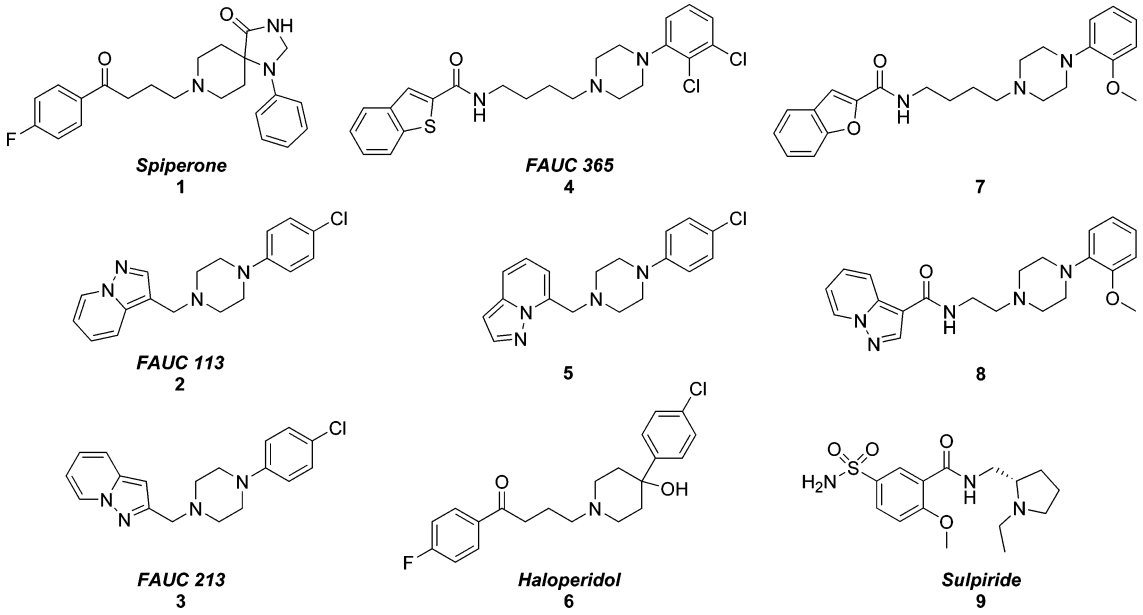
Within our outlined strategy, we manually inserted spiperone (**1**) on the basis of mutagenesis data into the binding area of each receptor. We suggest that spiperone is quite a promising choice for studying the conformational space of the ligand-bound dopamine receptors for several reasons. First of all, spiperone is known to bind to all three receptors with subnanomolar affinities in the range of 120−330 pM (Table 5), as determined in our lab. Therefore, it should uniformly be able to form specific interactions, responsible for substantial ligand recognition, in all receptors. Furthermore, it has been proven to be an antagonist, which is appropriate for the inactive state of the receptors we intend to model. In fact, the kink angle in TM6 and the closed ionic lock between the proximate cytoplasmatic ends of TM3 and TM6 indicate strongly that we have developed receptor models in an antagonist-like state. Moreover, because of the use of spiperone or its *N*-methyl analogue as radioligands or reference compounds, a multitude of binding results in mutant receptors is available from the literature.[45,59,62−64]

The simulations of all complexes were performed using Amber7 in combination with the FF99 force field[65] and PARM99 parameter set, which was used for the protein part, and the novel General Amber Force Field (gaff),[66] for the ligand spiperone. Parametrization of the ligand was accomplished with Antechamber after geometry optimization with Gaussian98[67] at the HF/6-31G* level of theory. Atom-centered point charges were obtained via the RESP method of charge fitting to the ab initio molecular electrostatic potential.

Because of the lack of lipid parameters in Amber, the previous modeling steps were conducted in vacuo, disregarding an explicit water/membrane environment. As a result, partial "flooding" of the binding site cleft with water and its resulting expansion is absent from the models. Thus, the ligand was carefully docked by hand into the tight binding pocket preventing obvious clashes and then deliberately minimized in five steps to avoid undesirable distortions of the receptor. First, only the ligand was allowed to move in the rigid receptor for 1000 steps of minimization, and then three times 5000 steps were performed, decreasing the harmonic force constant of the positional restraints applied to the receptor from 1000 kcal·mol$^{-1}$·Å$^{-2}$ to 10 and 0.1 kcal·mol$^{-1}$·Å$^{-2}$. Finally, 10000 steps were allowed with a DRMS convergence criterion of 0.01 kcal·mol$^{-1}$·Å$^{-1}$ and no restraints on the receptor.

We subjected the complexes thus obtained to an elaborate simulated annealing (SA) protocol for efficient sampling of the conformational space. To generate a controlled, additional dimension of diversity, three different maximal temperatures were used for the simulation of each receptor. The lowest $T_{max}$ at 400 K was employed to aim at the proximate space, while both higher temperature levels (500 and 600 K) were intended to facilitate exploration of more distant spatial regions. For each receptor at each $T_{max}$ level, 30 SA runs of 250 ps length were performed, adding up to a total simulation effort of 67.5 ns. Typical wall clock times of 6 h for the simulation of one 250 ps sequence on a dual Xeon 2.66 GHz node (Transtec IA32 Cluster) were found. A five-digit random number generator was coupled to the starting procedure of each SA run, which seeded the random setup process of initial velocities. These five-digit numbers were tested and, where necessary, rejected to prohibit duplicate runs. Each 250 ps run consisted of 3 phases: (1) The heating sequence (0−50 ps), when the temperature is gradually increased to $T_{max}$, while the heat bath coupling constant is decreased from 1.0 to 0.2 ps and the restraint multiplier is augmented from 3-fold to 5-fold. This parameter-variation combination resulted in tighter coupling of the system to the heating process and, simultaneously, enhanced stabilization of the original model topology. (2) The production phase (50−100 ps), where the simulation is constantly held at the final values of phase 1 ($T_{max}$, 0.2 ps heat bath coupling and 5-fold relative restraints). (3) The cooling sequence (100−250 ps), when the temperature is gradually decreased to 0 K. This phase is again divided into three parts: First, the cooling is started slowly (100−220 ps) by using a coupling constant of 3.0 ps, which is gradually modified

**Chart 1**




to 1.0 ps, while the restraint multiplier is decreased from 5.0 to 1.0. Next, the coupling is tightened to 0.5 ps and the restraints are further gradually reduced to 0.01 times their original value (220−240 ps). Finally, in the last 10 ps a very tight coupling of 0.05 ps ensures that the simulation adopts the character of a minimization, which is allowed to act on the completely unconstrained receptor. To ensure that comparably minimized states were achieved for each ligand−receptor complex, a final optimization was performed until the convergence criterion of 0.005 kcal·mol$^{-1}$·Å$^{-1}$ was reached.

Although high temperatures are required to prevent complexes from sticking in local minima, special care has to be taken to preserve the structure of the SCR backbone from unreasonable perturbations or even unfolding at high temperatures. Thus, we imposed square-bottom soft harmonic-wall restraints on (1) the $\varphi$- and $\psi$-torsions of the structurally conserved regions (SCR), (2) all trans-$\omega$-torsions, (3) all improper torsions at chiral centers, (4) the existing hydrogen bonds within the SCR backbone, and, because of the missing IL3, (5) the distance between the intracellular part of TM5 and TM6, as reported for the previous MD simulation. The restraint potentials consisted of a well with a square bottom (from $r_{S1}$ to $r_{S2}$) and parabolic sides (from $r_{P1}$ to $r_{S1}$ and $r_{S2}$ to $r_{P2}$) that pass into linear sides (below $r_{P1}$ and beyond $r_{P2}$) with the gradient of the harmonic function at $r_{P1}$ or $r_{P2}$, respectively. $\varphi$- and $\psi$-dihedral angles of the SCR were determined from the final models with Carnal and constrained with a force constant of 50.0 kcal·mol$^{-1}$·rad$^{-2}$ ($r_{S1/2} = \pm 5.0°$; $r_{P1/2} = \pm 20.0°$). Trans-$\omega$- and chirality-defining improper torsions were selected by the Amber tool makechir_rst and stabilized with a force constant of 100.0 kcal·mol$^{-1}$·rad$^{-2}$ ($r_{S1} = 10.0°$; $r_{P1} = 60.0°$; $r_{P2} = 80.0°$; $r_{S2} = 130.0°$) or 50.0 kcal·mol$^{-1}$·rad$^{-2}$ ($r_{S1} = 150.0°$; $r_{P1} = 170.0°$; $r_{P2} = 190.0°$; $r_{S2} = 210.0°$), respectively. SCR hydrogen bonds satisfying the standard definitions of Carnal (N···O distance ≤3.5 Å; angle cutoff: 30.0°) were maintained by a force constant of 20 kcal·mol$^{-1}$·Å$^{-2}$ ($r_{S1} = 1.3$ Å; $r_{P1} = 1.8$ Å; $r_{P2} = x_0$; $r_{S2} = x_0 + 0.5$ Å; with $x_0$ being the original distance in the final models without spiperone).

Applying the above simulated-annealing (SA) protocol, we obtained 90 receptor−ligand complexes for each subtype. The intention of this exhaustive approach was to allow the ligand to vary its position in order to optimize its binding interactions with the receptor and to allow the receptor to adapt to the ligand within the limitations of certain topology restraints. Despite the high temperatures employed in the SA runs, the results show only minor to moderate variations of the SCR backbone, indicating that the restraints have adequately stabilized the topology. Thus, the average RMSD values between the minimized SA models and their common start structure are 0.98 Å for $D_2$, 0.99 Å for $D_3$, and 1.18 Å for $D_4$. Additionally, a slight increase of the RMSD with higher $T_{max}$ values can be found in all receptors ($D_2$, 0.75 Å/0.97 Å/1.23 Å; $D_3$, 0.84 Å/0.93 Å/1.19 Å; $D_4$, 0.87 Å/1.13 Å/1.54 Å for $T_{max}$ = 400 K/500 K/600 K). Similarly, the ligand is able to relocate in a $T_{max}$-dependent manner, but the magnitude of the deviation is considerably higher, showing average RMSD values of 3.04 Å ($D_2$), 2.10 Å ($D_3$), or 3.82 Å ($D_4$) and maximal RMSDs of up to 7.31 Å ($D_2$), 5.00 Å ($D_3$), or 6.48 Å ($D_4$). Consequently, we can conclude that through this approach enough energy was provided to facilitate rearrangement of ligand binding, while the overall topology was predominantly maintained.

**8. Validation and Analysis of the Ligand−Receptor Complexes.** Quality assessment of the models produced is an important, but also challenging, task. Only a few approaches have been reported that evaluate ligand−receptor interactions as a determinant of model quality in a more sophisticated way. As cited above, Bissantz et al. have used the performance of homology models in a virtual screening task to investigate agonist and antagonist binding modes in different GPCRs.[19] Some authors report a combination of ligand docking into homology models and development of predictive 3D-QSARs based on the alignment obtained for all docked ligands. Jalaie et al. successfully used this strategy to derive a homology model from spinach photosystem II,[68] while Afzelius et al. applied it to a model of CYP2C9[69] and Tikhonova et al. studied the glycine-binding site of

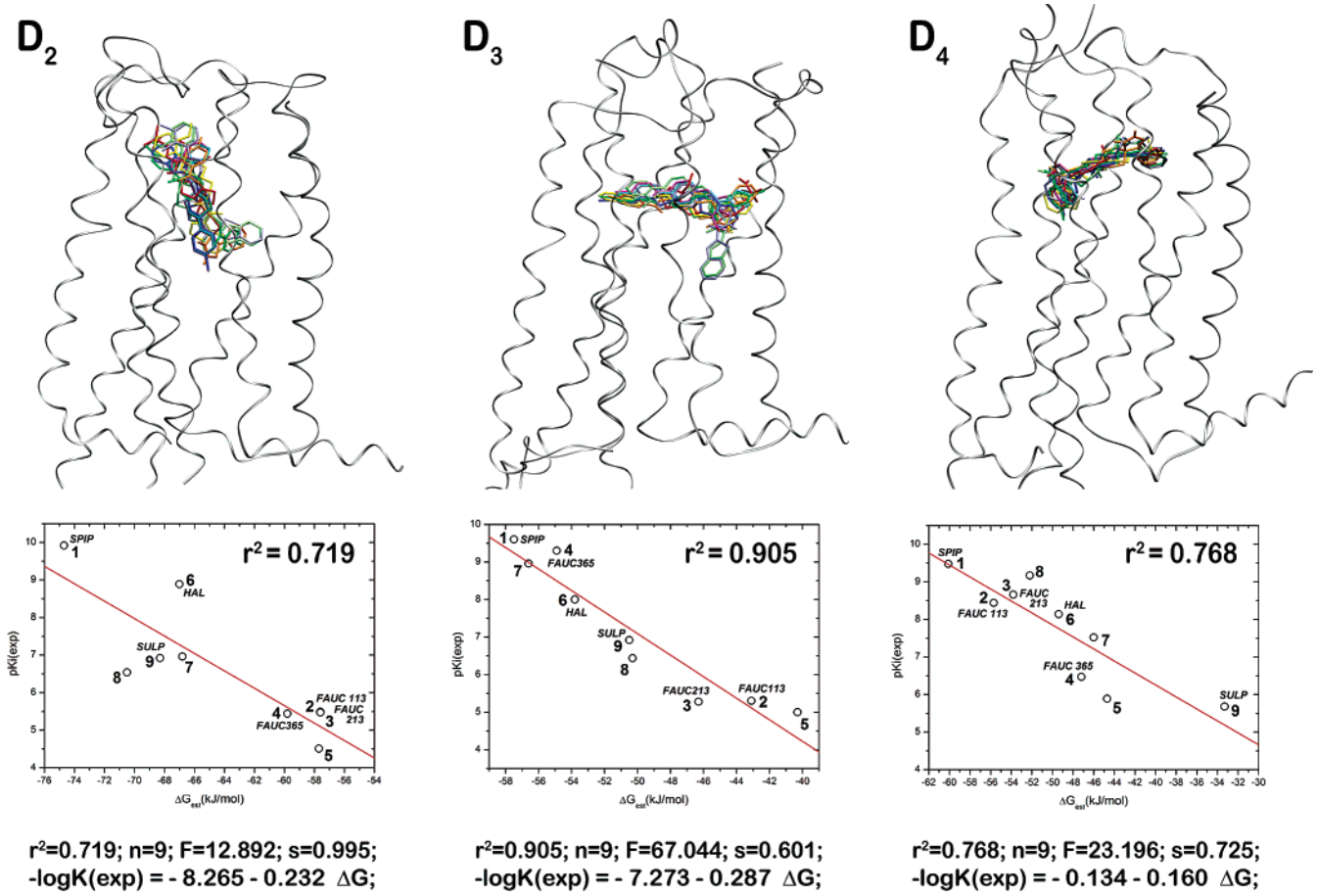



**Figure 5.** Graphical representations of the most significant complexes for each subtype and their corresponding correlations between predicted free energy of binding and experimental p$K_i$ values. In the framework of the sole ribbon of the receptor, the positioning of the docked ligands **1** to **9** is shown, indicating the variation and similarity of the respective binding modes. The regression plots were depicted and statistical data of the correlations were calculated using Origin7.[78]

an *N*-methyl-D-aspartate receptor model.[70] Schafferhans and Klebe used their novel approach DRAGHOME to dock 88 thrombin inhibitors into binding-site representations of homology models built from different serine proteases and were able to establish significant 3D-QSAR models based on the obtained ligand alignments.[71]

To evaluate the total number of 270 receptor models generated, we docked a set of diverse ligands with QXP+[72] into each binding site and ranked the models by their explanatory power, measured in terms of the correlation coefficients between the estimated free energies of binding and the experimental p$K_i$ values. Complementing spiperone, eight other ligands were selected (Chart 1) to provide substantial diversity at each subtype: $\Delta pK_i(D_2) = 5.4$; $\Delta pK_i(D_3) = 4.6$; $\Delta pK_i(D_4) = 3.8$ (Table 5). All experimental data were determined homogeneously within our lab, thus complying with the requirements to yield significant correlations suitable for ranking the models.

In order to allow for an appropriate extent of protein flexibility, docking was conducted in two steps. First, 500 cycles of Monte Carlo searching were performed with the module MCDOCK. Thereby, the ligand was flexibly docked into an almost rigid binding site, where only polar hydrogens were not fixed, but able to move within the regular limitations of the underlying AMBER force field. During the minimization steps of the MCDOCK algorithm, QXP is able to support the formation of a hydrogen bond between ASP$^{3.32}$ and the ligands, which are all considered in their protonated state. Consecutively, the five best docking solutions obtained at the first stage were subjected to an optimization with DOCKMIN, in which the ligand was again regarded as fully flexible, whereas the atoms in the binding pocket were constrained by a flat-bottom harmonic potential (<0.5 Å, no restraint/≥0.5 Å, 20 kJ/mol·Å$^2$). As documented within the Flo$^{2003}$/QXP package, a comparative docking study, evaluating the different algorithms provided in QXP, has yielded for 75 crystal structures containing ligands with up to 6 rotatable bonds an average "best RMSD from X-ray" of 1.2 Å and an average "predicted RMS energy error" of 0.5 kJ/mol, when performing 500 MC cycles with MCDOCK. Hence, these investigations underline that the method and parameters selected for our problem are appropriate. The docking results were scored using molecular mechanics energies, composed of a total energy for nonbonded interactions (van der Waals, electrostatic, and contact energy) and relative energy contributions of the ligand and binding-site conformations.

Within the scope of this validation approach, one best model with a correlation coefficient between 0.72 and 0.91 was identified for each subtype (Figure 5). In the $D_2$ receptor ($r^2 = 0.719$; $F = 17.892$; $s = 0.995$), the ligands are oriented in a more upright position, while in $D_3$ ($r^2 = 0.905$; $F = 67.044$; $s = 0.601$), the original, horizontal orientation of spiperone is maintained by the

**Table 6.** Interaction Overview about the Residues Proximal to Ligands **1** to **9** in the Most Correlating Complexes of D2, D3, and D4

| | D2 | | | | | | | | | | D3 | | | | | | | | | | D4 | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| BW[a] | aa[b] | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** | aa[b] | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** | aa[b] | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** |
| 1.39 | | | | | | | | | | | ***Y*** | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | | | | | | | | | | | |
| 1.42 | | | | | | | | | | | ***L*** | 1 | 2 | 3 | 4 | 5 | | 7 | 8 | 9 | | | | | | | | | | | |
| 1.43 | | | | | | | | | | | ***I*** | 1 | | 3 | | 5 | 6 | | 8 | 9 | | | | | | | | | | | |
| 2.50 | ***D*** | 1 | | | 4 | | 6 | 7 | 8 | | | | | | | | | | | | | | | | | | | | | | |
| 2.57 | | | | | | | | | | | | | | | | | | | | | ***V*** | 1 | 2 | | 4 | | 6 | 7 | 8 | 9 |
| 2.58 | ***M*** | | | | 4 | | | 7 | | | ***M*** | 1 | 2 | | 4 | 5 | 6 | | 8 | 9 | | | | | | | | | | | |
| 2.61 | | | | | | | | | | | ***V*** | | | | | | | | | 9 | ***F*** | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | |
| 2.64 | | | | | | | | | | | | | | | | | | | | | ***S*** | | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| 3.28 | ***F*** | 1 | 2 | | | 5 | | | 8 | | ***F*** | 1 | 2 | | 4 | 5 | 6 | 7 | 8 | 9 | ***L*** | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| 3.29 | ***V*** | 1 | | | | | | | 8 | 9 | | | | | | | | | | | ***M*** | 1 | 2 | | 4 | 5 | 6 | 7 | 8 | 9 |
| 3.32 | ***D*** | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | ***D*** | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | ***D*** | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| 3.33 | | | | | | | | | | | ***V*** | | 2 | | 4 | | | 7 | 8 | | ***V*** | 1 | | 3 | | | 6 | | 8 | 9 |
| 3.36 | ***C*** | | | | 4 | | | 7 | 8 | | ***C*** | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | | ***C*** | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | |
| 3.37 | | | | | | | | | | | | | | | | | | | | | ***T*** | 1 | | 3 | 4 | | 6 | | 8 | |
| 3.39 | ***S*** | 1 | 2 | 3 | | 5 | | | 8 | | | | | | | | | | | | | | | | | | | | | |
| 5.38 | ***F*** | | | | 4 | | 6 | 7 | 8 | 9 | ***F*** | | | | | | | 7 | | | | | | | | | | | | |
| 5.42 | ***S*** | | | 3 | | | 6 | 7 | 8 | 9 | ***S*** | | | | | | | | 8 | | ***S*** | 1 | | 3 | 4 | 5 | 6 | 7 | 8 | |
| 5.46 | | | | | | | | | | | | | | | | | | | | | ***S*** | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | |
| 6.47 | ***C*** | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | | | | | | | | | | | | | | | | | | | | |
| 6.48 | ***W*** | | | 3 | | | 6 | 7 | | 9 | ***W*** | 1 | | 3 | | | 6 | | 8 | | ***W*** | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | |
| 6.51 | ***F*** | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | ***F*** | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | | ***F*** | 1 | 2 | 3 | 4 | 5 | 6 | 7 | | 9 |
| 6.52 | | | | | | | | | | | | | | | | | | | | | ***F*** | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| 6.55 | ***H*** | | | | 4 | | 6 | 7 | 8 | 9 | ***H*** | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | | ***H*** | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| 6.56 | | | | | | | | | | | | | | | | | | | | | ***I*** | 1 | 2 | | | 5 | 6 | 7 | | 9 |
| 7.35 | ***Y*** | 1 | 2 | | 4 | 5 | | 7 | | | | | | | | | | | | | ***V*** | 1 | | | | | | | | |
| 7.42 | ***G*** | 1 | 2 | 3 | | 5 | 6 | | 8 | | ***G*** | | | | 4 | | | | | | | | | | | | | | | |
| 7.43 | ***Y*** | 1 | | | 4 | | 6 | 7 | 8 | | ***Y*** | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | ***Y*** | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| 7.45 | ***N*** | 1 | 2 | 3 | | 5 | 6 | | | | ***N*** | 1 | | 3 | 4 | | 6 | 7 | 8 | | | | | | | | | | | |
| 7.46 | ***S*** | 1 | 2 | 3 | | 5 | 6 | 7 | 8 | | ***S*** | 1 | | 3 | 4 | | 6 | | | 9 | | | | | | | | | | |

[a] Ballesteros−Weinstein[25] notation of the residue position. [b] One-letter amino acid code, showing the residue type for each receptor. In each row, appearance of a ligand number indicates that the residue at the given position is closer than 3.5 Å to the denoted ligand in the most correlating complex of this subtype. For clarity reasons, only residues that are adjacent to more than half the ligands (>4) in any receptor are listed.

ligands. Similarly, the ligand orientation in the $D_4$ receptor ($r^2 = 0.768$; $F = 23.196$; $s = 0.725$) is horizontal, but also slightly shifted in the area TM2, TM3, and TM7 toward the extracellular side. For the docking/scoring process, we found that the constrained minimization is essential for obtaining these significant correlations. Interestingly, the complexes are all uniformly altered during this second step, shown by consistent RMSDs (Cα atoms of the binding site) between complexes before and after minimization (0.35 to 0.38) and small RMSDs between the minimized complexes and their average structure (0.04 to 0.09).

On the basis of spatial proximity to the ligands (minimal distance < 3.5 Å), we evaluated and compared the patterns of key interactions for receptor−ligand recognition (Table 6 and Figure 6). Some residues were found to interact with most of the ligands in all three subtypes. ASP$^{3.32}$ typically forms ionic interactions with the protonated amine functions of the ligands. PHE/LEU$^{3.28}$, CYS$^{3.36}$, TRP$^{6.48}$, PHE$^{6.51}$, HIS$^{6.55}$, and TYR$^{7.43}$

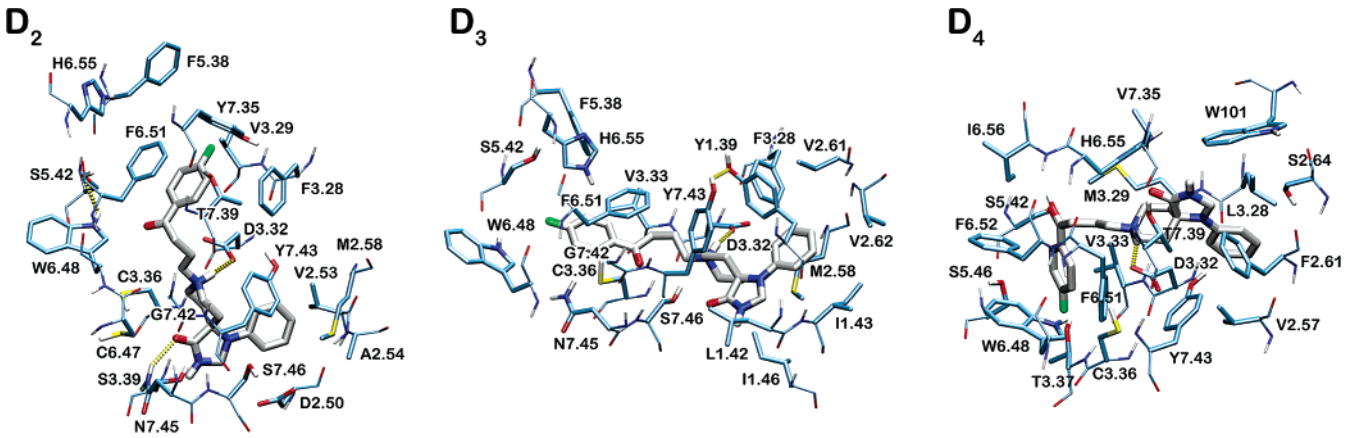



**Figure 6.** Binding-site residues shown in complex with spiperone for the $D_2$, $D_3$, and $D_4$ receptors. Only residues are displayed which are found to be typically proximal (≤3.5 Å) to most of the ligands (>4) in the respective subtype model or which are relevant for the spiperone complex only. The figure was prepared using VMD.[75]

represent a prominent part of the binding pocket, featuring various unspecific and specific interactions with the different ligands. Interactions with some other residues, such as MET$^{2.58}$, ASN$^{7.45}$ and SER$^{7.46}$ ($D_2$/$D_3$), VAL/MET$^{3.29}$ and SER$^{5.42}$ ($D_2$/$D_4$) or VAL$^{3.33}$ ($D_3$/$D_4$), are commonly shared by just two subtypes. Of course, several residues interact more exclusively with only one receptor subtype. ASP$^{2.50}$, SER$^{3.39}$, PHE$^{5.38}$, CYS$^{6.47}$, TYR$^{7.35}$, and GLY$^{7.42}$ are primarily relevant for the $D_2$ receptor, while close contact to residues in TM1 (TYR$^{1.39}$, LEU$^{1.42}$, and ILE$^{1.43}$) can only be found in the $D_3$ receptor. Several interactions with TM2, TM3, and TM6 (VAL$^{2.57}$, PHE$^{2.61}$, SER$^{2.64}$, THR$^{3.37}$, SER$^{5.46}$, PHE$^{6.52}$, and ILE$^{6.56}$) seem to be most prevalent in the $D_4$ receptor.

Comparing our modeling results to the site-directed mutagenesis data found for spiperone or its *N*-methyl analogue, the experimental effects of several mutants are found to be in good agreement with the validated complex structures. For instance, the 2.9-, 11.5-, and 3.9-fold decrease of ligand binding reported for the F3.28L, V3.29M, and Y7.35V mutants[62] can be explained by direct interactions of these residues with the *p*-fluorobenzoyl moiety of the ligand. The moderate change shown by F3.28L compared to V3.29M may indicate that the ligand−PHE$^{3.28}$ interaction is merely of unspecific, hydrophobic nature, corresponding to the absence of specific $\pi-\pi$ interactions with this residue in the model. The huge effect of V3.29M, in contrast, is probably based on hardly resolvable steric clashes of the more voluminous MET$^{3.29}$ with the ligand in the mutant. An even stronger impairment of ligand binding is found for the mutation F6.51A (33.7-fold for spiperone)[63] or F6.51C (83.5-fold for *N*-methylspiperone).[59] This can be explained by the intensive $\pi-\pi$-stacking interaction with the *p*-fluorobenzoyl system, which is reflected in the model by the coplanar orientation of the aromatic rings and their fairly exact superimposed centroids. When mutated to ALA, PHE$^{5.47}$ and LEU$^{6.49}$ have shown almost no effects on spiperone binding[63] and, in fact, having both a minimal distance of 6.7 Å to the ligand, no direct or indirect interaction is visible in the complex. Likewise, single and multiple ALA mutants of the three conserved SER residues in TM5 do not affect spiperone affinity at $D_2$ receptors.[73] Correspondingly, the minimal distances are 4.5 Å (SER$^{5.42}$), 8.4 Å (SER$^{5.43}$), and 7.8 Å (SER$^{5.46}$), implying no direct interactions with spiperone. In the model complex, TRP$^{6.48}$ lies adjacent to spiperone (minimum distance: 3.6 Å), although it does not interact strongly. Therefore, the vast effect of W6.48C on binding of *N*-methylspiperone (>1350-fold decrease) is difficult to explain by direct interactions. Nonetheless, a substantial destabilization of the aromatic cluster in TM6 (sequence pattern: F$^{6.44}$-x-x-x-*W*-x-x-F-F-x-x-H$^{6.55}$) may be responsible for such an enormous reduction of ligand binding. Both residues, ASP$^{2.50}$ and TYR$^{7.43}$, are found close to the phenyl moiety, which is attached to the triazaspiro[4.5]decanone substructure. However, while the D2.50C mutant produces an improved hydrophobic microenvironment around the phenyl moiety and, thus, increases the affinity 2.8-fold,[45] in the mutant Y7.43C, favorable hydrophobic interactions with this phenyl moiety are eliminated, leading to a 3.0-fold[64] decrease. Very recently, Shi and Javitch reported that the EL2 lines the binding site crevice in the $D_2$ receptor.[51] With a 3.6-fold reduction, the mutant I184C was identified to have the strongest impact on *N*-methylspiperone binding. Although not found to interact directly with spiperone in our model, ILE184 is the nearest residue of the EL2 with a minimal distance of 5.4 Å, except for CYS182, which is covalently bound to CYS$^{3.25}$ in TM3.

In the $D_4$ receptor, the double-mutant LM3.28-3.29FV gave a 2.0-fold increase, while the V7.35Y mutant gave a 1.8-fold decrease of binding affinity toward *N*-methylspiperone. All three residues lie adjacent to the ligand in our final model, and the mutated PHE$^{3.28}$ is able to form additional stabilizing hydrophobic or $\pi-\pi$ interactions with the phenyl moiety of spiperone, while the mutated TYR$^{7.35}$ is likely to cause clashes with the ligand directly or indirectly via other proximal residues.

## Conclusion

In this paper, we present a computational approach to modeling the similarity and diversity of dopamine $D_2$-like receptors. On the basis of the currently available high-resolution X-ray structures of bovine rhodopsin, we generated homology models of all three subtypes. The structures were refined using extensive MM and MD protocols with a total simulation time of 18 ns, in order to allow structural characteristics to develop that depend on the individual sequences. We extracted structures from homogeneous parts of the trajectory,

checked them by statistics-based procedures, and validated them using the available experimental data on SCAM and site-directed mutagenesis. As a result of these analyses, we selected the most reliable model for further detailed investigations. Focusing on these models and the associated trajectory parts, we examined various aspects of the differential receptor topology, such as inter- and intrahelical angles of the transmembrane domains, as well as a broad range of typical structural features. Our investigations revealed subtype-specific positioning of TM1, 2, 6, and 7 as well as key similarities, such as the rotamer state of $ASP^{3.32}$, the bending of TM6, and the presence of the closed DRY-lock between TM3 and TM6. On the basis of a two-level strategy, we gained insights into the requirements of ligand−receptor recognition. Starting from selected models, we performed 90 simulated annealing runs of each receptor with the $D_2$, $D_3$, $D_4$ high-affinity antagonist spiperone docked into the putative binding areas of the receptor subtypes. Thus, a total effort of 22.5 ns simulation time per receptor was used to explore the conformational space of the three complexes, allowing the ligand to vary its position in the binding pocket and the receptor to adapt to the ligand within the limitations of a tight set of topology constraints imposed on torsions and hydrogen bonds in the backbone. At the subsequent level, the receptor conformations obtained were validated for their capability to discriminate ligands according to their subtype affinities. Thus, spiperone and eight other ligands, which were selected to span a range of $pK_i$ values from 3.8 to 5.4 for each subtype, were flexibly docked into all 270 rigid binding sites using QXP/FLO[2003]. Afterward, the resulting complexes were minimized, allowing for a constrained movement of the receptor to incorporate some protein flexibility. Using this strategy, we were able to identify one high-quality model for each subtype, for which the estimated free energy of binding correlates to the experimental ligand affinity with $r^2 = 0.72$ ($D_2$), 0.91 ($D_3$), or 0.77 ($D_4$). These complexes were further analyzed to reveal key interactions of ligand binding, which may define novel targets for site-directed mutagenesis studies. We have found both key interactions common to all receptors, such as with $ASP^{3.32}$ or aromatic residues in TM6 and 7, as well as subtype-specific interactions, for instance with TM1 residues in the $D_3$ or TM2 residues in the $D_4$ receptor model. The comparably high degree of correlation suggests that the models we have developed in this study may prove to be valuable for guiding subsequent drug design efforts and for predicting and explaining the binding properties of novel dopamine-receptor antagonists. The successful strategy used to derive these models will be helpful for modeling further GPCR−ligand complexes. A reliable prediction of receptor−agonist complexes, however, will require further insights into GPCR activation.

**Acknowledgment.** The authors wish to thank Dr. W. Utz for helpful discussions and continuous technical support and S. Haerterich for helpful technical assistance. We also gratefully acknowledge Dr. G. Wellein and G. Hager (High Performance Computing division of the "Regionales Rechenzentrum Erlangen") for providing the opportunity to perform calculations on a 28 processor SGI Origin 3400 and a 170 Intel Xeon processor IA32 cluster. This work was supported by the BMBF and Fonds der Chemischen Industrie. We thank Prof. T. Clark for helpful discussions.

**Supporting Information Available:** Detailed discussions of an additional validation approach for the refined receptor models (5b), general structural and dynamic features of the models and model trajectories (6b), and specific structural features of aminergic GPCRs found in the dopamine receptor models (6c). Tables S1 to S5 and Figures S1 to S8 complementing the data given in the article and discussed in the Supporting Information dealing with the following topics: trajectory clustering and minimization of averaged structures, model selection and validation (Ramachandran plots, minimal distances of characteristic interactions), general structural and dynamic characteristics (relative SASA, atomic fluctuations), and specific structural features (rotamer state of $ASP^{3.32}$, aromatic cluster, and hydrogen bond network). This Material is available free of charge via the Internet at http://pubs.acs.org. A 3D-VRML file of the $D_2$ model displaying SCAM data can be obtained on request from the authors.

## References

(1) Sibley, D. R.; Monsma, J. F., Jr. Molecular biology of dopamine receptors. *Trends Pharmacol. Sci.* **1992**, *13*, 61−69.
(2) Hall, D. A.; Strange, P. G. Comparison of the ability of dopamine receptor agonists to inhibit forskolin-stimulated adenosine 3′5′-cyclic monophosphate (cAMP) accumulation via D2L (long isoform) and D3 receptors expressed in Chinese hamster ovary (CHO) cells. *Biochem. Pharmacol.* **1999**, *58*, 285−289.
(3) Palczewski, K.; Kumasaka, T.; Hori, T.; Behnke, C. A.; Motoshima, H.; Fox, B. A.; Trong, I. L.; Teller, D. C.; Okada, T.; Stenkamp, R. E.; Yamamoto, M.; Miyano, M. Crystal Structure of Rhodopsin: A G Protein-Coupled Receptor. *Science* **2000**, *289*, 739−745.
(4) Richtand, N. M.; Woods, S. C.; Berger, S. P.; Strakowski, S. M. D3 dopamine receptor, behavioral sensitization, and psychosis. *Neurosci. Biobehav. Rev.* **2001**, *25*, 427−443.
(5) Joyce, J. N.; Ryoo, H. L.; Beach, T. B.; Caviness, J. N.; Stacy, M.; Gurevich, E. V.; Reiser, M.; Adler, C. H. Loss of response to levodopa in Parkinson's disease and co-occurrence with dementia: role of D3 and not D2 receptors. *Brain Res.* **2002**, *955*, 138−152.
(6) Bezard, E.; Ferry, S.; Mach, U.; Stark, H.; Leriche, L.; Boraud, T.; Gross, C.; Sokoloff, P. Attenuation of levodopa-induced dyskinesia by normalizing dopamine D3 receptor function. *Nat. Med.* **2003**, *9*, 762−767.
(7) Pilla, M.; Perachon, S.; Sautel, F.; Garrido, F.; Mann, A.; Wermuth, C. G.; Schwartz, J. C.; Everitt, B. J.; Sokoloff, P. Selective inhibition of cocaine-seeking behaviour by a partial dopamine D3 receptor agonist. *Nature* **1999**, *400*, 371−375.
(8) Hrib, N. J. The dopamine D4 receptor: A controversial therapeutic target. *Drugs of the Future* **2000**, *25*, 587−611 and references cited therein.
(9) Brioni, J. D.; Moreland, R. B.; Cowart, M.; Hsieh, G. C.; Stewart, A. O.; Hedlund, P.; Donnelly-Roberts, D. L.; Nakane, M.; Lynch, J. J., 3rd; Kolasa, T.; Polakowski, J. S.; Osinski, M. A.; Marsh, K.; Andersson, K. E.; Sullivan, J. P. Activation of dopamine D4 receptors by ABT-724 induces penile erection in rats. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 6758−6763.
(10) Hackling, A. E.; Stark, H. Dopamine D3 receptor ligands with antagonist properties. *ChemBioChem* **2002**, *3*, 946−961.
(11) Loeber, S.; Huebner, H.; Utz, W.; Gmeiner, P. Rationally Based Efficacy Tuning of Selective Dopamine D4 Receptor Ligands Leading to the Complete Antagonist 2-[4-(4-Chlorophenyl)-piperazin-1-ylmethyl]pyrazolo[1,5-*a*]pyridine (FAUC 213). *J. Med. Chem.* **2001**, *44*, 2691−2694.
(12) Boeckler, F.; Russig, H.; Zhang, W.; Löber, S.; Schetz, J.; Hübner, H.; Ferger, B.; Gmeiner, P.; Feldon, J. FAUC 213, a highly selective dopamine D4 receptor full antagonist, exhibits atypical antipsychotic properties in behavioural and neurochemical models of schizophrenia. *Psychopharmacology* **2004**, *175*, 7−17.
(13) Bettinetti, L.; Schlotter, K.; Huebner, H.; Gmeiner, P. Interactive SAR Studies: Rational Discovery of Super-Potent and Highly Selective Dopamine D3 Receptor Antagonists and Partial Agonists. *J. Med. Chem.* **2002**, *45*, 4594−4597.
(14) Boeckler, F.; Leng, A.; Mura, A.; Bettinetti, L.; Feldon, J.; Gmeiner, P.; Ferger, B. Attenuation of 1-methyl-4-phenyl-1,2,3,6-tetrahydropyridine (MPTP) neurotoxicity by the novel selective dopamine D3-receptor partial agonist FAUC 329 predominantly in the nucleus accumbens of mice. *Biochem. Pharmacol.* **2003**, *66*, 1025−1032.

(15) Hibert, M. F.; Trumpp-Kallmeyer, S.; Bruinvels, A.; Hoflack, J. Three-dimensional models of neurotransmitter G-binding protein-coupled receptors. *Mol. Pharmacol.* **1991**, *40*, 8−15.
(16) Hulme, E. C.; Birdsall, N. J.; Buckley, N. J. Muscarinic receptor subtypes. *Annu. Rev. Pharmacol. Toxicol.* **1990**, *30*, 633−673.
(17) Livingstone, C. D.; Strange, P. G.; Naylor, L. H. Molecular modelling of D2-like dopamine receptors. *Biochem. J.* **1992**, *287* (Part 1), 277−282.
(18) Ballesteros, J. A.; Shi, L.; Javitch, J. A. Structural mimicry in G protein-coupled receptors: implications of the high-resolution structure of rhodopsin for structure-function analysis of rhodopsin-like receptors. *Mol. Pharmacol.* **2001**, *60*, 1−19.
(19) Bissantz, C.; Bernard, P.; Hibert, M.; Rognan, D. Protein-based virtual screening of chemical databases. II. Are homology models of G-Protein Coupled Receptors suitable targets? *Proteins* **2003**, *50*, 5−25.
(20) Kapp, O. H.; Siemion, J.; Kuo, J.; Johnson, B. A.; Shankaran, V.; Reba, R. C.; Mukherjee, J. Comparison of the interaction of dopamine and high affinity positron emission tomography radiotracer fallypride with the dopamine D-2 receptor: a molecular modeling study. *J. Mol. Model.* **2001**, *7*, 6−18.
(21) Hackling, A.; Ghosh, R.; Perachon, S.; Mann, A.; Holtje, H. D.; Wermuth, C. G.; Schwartz, J. C.; Sippl, W.; Sokoloff, P.; Stark, H. *N*-(ω-(4-(2-methoxyphenyl)piperazin-1-yl)alkyl)carboxamides as dopamine $D_2$ and $D_3$ receptor ligands. *J. Med. Chem.* **2003**, *46*, 3883−3899.
(22) Varady, J.; Wu, X.; Fang, X.; Min, J.; Hu, Z.; Levant, B.; Wang, S. Molecular modeling of the three-dimensional structure of dopamine 3 ($D_3$) subtype receptor: discovery of novel and potent $D_3$ ligands through a hybrid pharmacophore- and structure-based database searching approach. *J. Med. Chem.* **2003**, *46*, 4377−4392.
(23) Klabunde, T.; Hessler, G. Drug design strategies for targeting G-protein-coupled receptors. *ChemBioChem* **2002**, *3*, 928−944.
(24) Baker, D.; Sali, A. Protein structure prediction and structural genomics. *Science* **2001**, *294*, 93−96.
(25) Ballesteros, J.; Weinstein, H. Integrated methods for the construction of three-dimensional models of structure-function relations in G protein-coupled receptors. In *Methods in Neurosciences: Receptor Molecular Biology*; Sealfon, S. C., Conn, M. P., Eds.; Academic Press: San Diego, 1995; pp 366−428.
(26) Visiers, I.; Ballesteros, J. A.; Weinstein, H. Three-dimensional representations of G protein-coupled receptor structures and mechanisms. *Methods Enzymol.* **2002**, *343*, 329−371.
(27) The term SCR is typically used in the context of homology modeling software, such as COMPOSER. In our appoach, the SCR comprises the transmembrane regions complemented by some adjacent, conserved residues. The exact definition of the SCRs can be found in Figure 1 (blue frames).
(28) Gonnet, G. H.; Cohen, M. A.; Benner, S. A. Exhaustive matching of the entire protein sequence database. *Science* **1992**, *256*, 1443−1445.
(29) Lu, Z. L.; Saldanha, J. W.; Hulme, E. C. Seven-transmembrane receptors: crystals clarify. *Trends Pharmacol. Sci.* **2002**, *23*, 140−146.
(30) Shin, N.; Coates, E.; Murgolo, N. J.; Morse, K. L.; Bayne, M.; Strader, C. D.; Monsma, F. J., Jr. Molecular modeling and site-specific mutagenesis of the histamine-binding site of the histamine H4 receptor. *Mol. Pharmacol.* **2002**, *62*, 38−47.
(31) Archer, E.; Maigret, B.; Escrieut, C.; Pradayrol, L.; Fourmy, D. Rhodopsin crystal: new template yielding realistic models of G-protein-coupled receptors? *Trends Pharmacol. Sci.* **2003**, *24*, 36−40.
(32) Thompson, J. D.; Gibson, T. J.; Plewniak, F.; Jeanmougin, F.; Higgins, D. G. The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res.* **1997**, *25*, 4876−4882.
(33) Nicholas, K. B.; Nicholas, H. B. J.; Deerfield, D. W. I. *GeneDoc*; http://www.psc.edu/biomed/genedoc.
(34) Hall, T. A. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp. Ser.* **1999**, *41*, 95−98; http://www.mbio.ncsu.edu/BioEdit/bioedit.html.
(35) Blundell, T.; Carney, D.; Gardner, S.; Hayes, F.; Howlin, B.; Hubbard, T.; Overington, J.; Singh, D. A.; Sibanda, B. L.; Sutcliffe, M. Knowledge-based protein modelling and design. *Eur. J. Biochem.* **1988**, *172*, 513−520.
(36) *SYBYL 6.9*; Tripos Inc., 1699 South Hanley Road, St. Louis, Missouri, 63144.
(37) Filteau, F.; Veilleux, F.; Levesque, D. Effects of reciprocal chimeras between the C-terminal portion of third intracellular loops of the human dopamine D2 and D3 receptors. *FEBS Lett.* **1999**, *447*, 251−256.
(38) Van Leeuwen, D. H.; Eisenstein, J.; O'Malley, K.; MacKenzie, R. G. Characterization of a chimeric human dopamine D3/D2 receptor functionally coupled to adenylyl cyclase in Chinese hamster ovary cells. *Mol. Pharmacol.* **1995**, *48*, 344−351.
(39) Case, D. A.; Pearlman, D. A.; Caldwell, J. W.; Cheatham, T. E., III; Wang, J.; Ross, W. S.; Simmerling, C. L.; Darden, T. A.; Merz, K. M.; Stanton, R. V.; Cheng, A. L.; Vincent, J. J.; Crowley, M.; Tsui, V.; Gohlke, H.; Radmer, R. J.; Duan, Y.; Pitera, J.; Massova, I.; Seibel, G. L.; Singh, U. C.; Weiner, P. K.; Kollman, P. A. *AMBER 5*/*7*; University of California: San Francisco, 1997/2002.
(40) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M., Jr.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules. *J. Am. Chem. Soc.* **1995**, *117*, 5179−5197.
(41) Strahs, D.; Weinstein, H. Comparative modeling and molecular dynamics studies of the delta, kappa and mu opioid receptors. *Protein Eng.* **1997**, *10*, 1019−1038.
(42) Vriend, G. WHAT IF: A molecular modeling and drug design program. *J. Mol. Graphics* **1990**, *8*, 52−56.
(43) Hooft, R. W. W.; Vriend, G.; Sander, C.; Abola, E. E. Errors in protein structures. *Nature* **1996**, *381*, 272.
(44) Lovell, S. C.; Davis, I. W.; Arendall, W. B., 3rd; de Bakker, P. I.; Word, J. M.; Prisant, M. G.; Richardson, J. S.; Richardson, D. C. Structure validation by Calpha geometry: phi,psi and Cbeta deviation. *Proteins* **2003**, *50*, 437−450.
(45) Javitch, J. A.; Ballesteros, J. A.; Chen, J.; Chiappa, V.; Simpson, M. M. Electrostatic and aromatic microdomains within the binding-site crevice of the D2 receptor: contributions of the second membrane-spanning segment. *Biochemistry* **1999**, *38*, 7961−7968.
(46) Javitch, J. A.; Shi, L.; Simpson, M. M.; Chen, J.; Chiappa, V.; Visiers, I.; Weinstein, H.; Ballesteros, J. A. The fourth transmembrane segment of the dopamine D2 receptor: accessibility in the binding-site crevice and position in the transmembrane bundle. *Biochemistry* **2000**, *39*, 12190−12199.
(47) Shi, L.; Javitch, J. A. THE BINDING SITE OF AMINERGIC G PROTEIN-COUPLED RECEPTORS: The Transmembrane Segments and Second Extracellular Loop. *Annu. Rev. Pharmacol. Toxicol.* **2002**, *42*, 437−467 and references cited therein.
(48) Guo, W.; Shi, L.; Javitch, J. A. The Fourth Transmembrane Segment Forms the Interface of the Dopamine D2 Receptor Homodimer. *J. Biol. Chem.* **2003**, *278*, 4385−4388.
(49) Davies, A.; Gowen, B. E.; Krebs, A. M.; Schertler, G. F. X.; Saibil, H. R. Three-dimensional structure of an invertebrate rhodopsin and basis for ordered alignment in the photoreceptor membrane1. *J. Mol. Biol.* **2001**, *314*, 455−463.
(50) Javitch, J. A.; Fu, D.; Chen, J. Residues in the fifth membrane-spanning segment of the dopamine D2 receptor exposed in the binding-site crevice. *Biochemistry* **1995**, *34*, 16433−16439.
(51) Shi, L.; Javitch, J. A. The second extracellular loop of the dopamine D2 receptor lines the binding-site crevice. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 440−445.
(52) Kahn, P. C. Defining the axis of a helix. *Comput. Chem.* **1989**, *13*, 185−189.
(53) Sansom, M. S.; Weinstein, H. Hinges, swivels and switches: the role of prolines in signaling via transmembrane alpha-helices. *Trends Pharmacol. Sci.* **2000**, *21*, 445−451.
(54) Sankararamakrishnan, R.; Vishveshwara, S. Geometry of proline-containing alpha-helices in proteins. *Int. J. Pept. Protein Res.* **1992**, *39*, 356−363.
(55) Mezei, M. *SIMULAID*: Department of Physiology and Biophysics, Mount Sinai School of Medicine, New York, http://fulcrum.physbio.mssm.edu/~mezei/simulaid/simulaid.html.
(56) Ballesteros, J. A.; Jensen, A. D.; Liapakis, G.; Rasmussen, S. G.; Shi, L.; Gether, U.; Javitch, J. A. Activation of the beta 2-adrenergic receptor involves disruption of an ionic lock between the cytoplasmic ends of transmembrane segments 3 and 6. *J. Biol. Chem.* **2001**, *276*, 29171−29177.
(57) Jensen, A. D.; Guarnieri, F.; Rasmussen, S. G.; Asmar, F.; Ballesteros, J. A.; Gether, U. Agonist-induced conformational changes at the cytoplasmic side of transmembrane segment 6 in the beta 2 adrenergic receptor mapped by site-selective fluorescent labeling. *J. Biol. Chem.* **2001**, *276*, 9279−9290.
(58) Shi, L.; Liapakis, G.; Xu, R.; Guarnieri, F.; Ballesteros, J. A.; Javitch, J. A. Beta2 adrenergic receptor activation. Modulation of the proline kink in transmembrane 6 by a rotamer toggle switch. *J. Biol. Chem.* **2002**, *277*, 40989−40996.
(59) Javitch, J. A.; Ballesteros, J. A.; Weinstein, H.; Chen, J. A cluster of aromatic residues in the sixth membrane-spanning segment of the dopamine D2 receptor is accessible in the binding-site crevice. *Biochemistry* **1998**, *37*, 998−1006.
(60) Evers, A.; Gohlke, H.; Klebe, G. Ligand-supported homology modelling of protein binding-sites using knowledge-based potentials. *J. Mol. Biol.* **2003**, *334*, 327−345.
(61) Evers, A.; Klebe, G. Ligand-supported homology modeling of g-protein-coupled receptor sites: models sufficient for successful virtual screening. *Angew. Chem., Int. Ed. Engl.* **2004**, *43*, 248−251.

(62) Simpson, M. M.; Ballesteros, J. A.; Chiappa, V.; Chen, J.; Suehiro, M.; Hartman, D. S.; Godel, T.; Snyder, L. A.; Sakmar, T. P.; Javitch, J. A. Dopamine D4/D2 receptor selectivity is determined by A divergent aromatic microdomain contained within the second, third, and seventh membrane-spanning segments. *Mol. Pharmacol.* **1999**, *56*, 1116−1126.
(63) Cho, W.; Taylor, L. P.; Mansour, A.; Akil, H. Hydrophobic residues of the D2 dopamine receptor are important for binding and signal transduction. *J. Neurochem.* **1995**, *65*, 2105−2115.
(64) Fu, D.; Ballesteros, J. A.; Weinstein, H.; Chen, J.; Javitch, J. A. Residues in the seventh membrane-spanning segment of the dopamine D2 receptor accessible in the binding-site crevice. *Biochemistry* **1996**, *35*, 11278−11285.
(65) Wang, J.; Cieplak, P.; Kollman, P. A. How well does a restrained electrostatic potential (RESP) model perform in calculating conformational energies of organic and biological molecules? *J. Comput. Chem.* **2000**, *21*, 1049−1074.
(66) Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. Development and testing of a general amber force field. *J. Comput. Chem.* **2004**, *25*, 1157−1174.
(67) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Zakrzewski, V. G.; Montgomery, J. A. J.; Stratmann, R. E.; Burant, J. C.; Dapprich, S.; Millam, J. M.; Daniels, A. D.; Kudin, K. N.; Strain, M. C.; Farkas, O.; Tomasi, J.; Barone, V.; Cossi, M.; Cammi, R.; Mennucci, B.; Pomelli, C.; Adamo, C.; Clifford, S.; Ochterski, J.; Petersson, G. A.; Ayala, P. Y.; Cui, Q.; Morokuma, K.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Cioslowski, J.; Ortiz, J. V.; Baboul, A. G.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Gonzalez, C.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Andres, J. L.; Gonzalez, C.; Head-Gordon, M.; Replogle, E. S.; Pople, J. A. *Gaussian 98*, revision A.7; Gaussian, Inc.: Pittsburgh, PA, 1998.
(68) Jalaie, M.; Erickson, J. A. Homology model directed alignment selection for comparative molecular field analysis: Application to photosystem II inhibitors. *J. Comput.-Aided Mol. Des.* **2000**, *14*, 181−197.
(69) Afzelius, L.; Zamora, I.; Ridderstrom, M.; Andersson, T. B.; Karlen, A.; Masimirembwa, C. M. Competitive CYP2C9 Inhibitors: Enzyme Inhibition Studies, Protein Homology Modeling, and Three-Dimensional Quantitative Structure-Activity Relationship Analysis. *Mol. Pharmacol.* **2001**, *59*, 909−919.
(70) Tikhonova, I. G.; Baskin, I. I.; Palyulin, V. A.; Zefirov, N. S. CoMFA and homology-based models of the glycine binding site of *N*-methyl-D-aspartate receptor. *J. Med. Chem.* **2003**, *46*, 1609−1616.
(71) Schafferhans, A.; Klebe, G. Docking ligands onto binding site representations derived from proteins built by homology modelling. *J. Mol. Biol.* **2001**, *307*, 407−427.
(72) McMartin, C.; Bohacek, R. S. QXP: powerful, rapid computer algorithms for structure-based drug design. *J. Comput.-Aided. Mol. Des.* **1997**, *11*, 333−344.
(73) Coley, C.; Woodward, R.; Johansson, A. M.; Strange, P. G.; Naylor, L. H. Effect of multiple serine/alanine mutations in the transmembrane spanning region V of the D2 dopamine receptor on ligand binding. *J. Neurochem.* **2000**, *74*, 358−366.
(74) Corpet, F. Multiple sequence alignment with hierarchical clustering. *Nucleic Acids Res.* **1988**, *16*, 10881−10890.
(75) Humphrey, W.; Dalke, A.; Schulten, K. VMD—Visual Molecular Dynamics. *J. Mol. Graphics* **1996**, *14*, 33−38.
(76) Visiers, I.; Braunheim, B. B.; Weinstein, H. Prokink: a protocol for numerical evaluation of helix distortions by proline. *Protein Eng.* **2000**, *13*, 603−606.
(77) Koradi, R.; Billeter, M.; Wuthrich, K. MOLMOL: a program for display and analysis of macromolecular structures. *J. Mol. Graphics* **1996**, *14*, 51−55, 29−32.
(78) *Origin 7*; OriginLab Corporation, One Roundhouse Plaza, Northampton, MA 01060, www.OriginLab.com.

JM049612A