Generalization of a Targeted Library Design Protocol: Application to 5-HT_7 Receptor Ligands

Erik Nordling* and Evert Homan

Computational Chemistry, Department of Structural Chemistry, Biovitrum AB, S-112 76 Stockholm, Sweden

Received May 28, 2004

Herein a general concept for the design of targeted libraries for proteins with binding sites that are divided into subsites is laid out, including several practical aspects and their solutions. The design is based on a chemogenomic classification of the subsites followed by collection of bioactive molecular fragments and virtual library generation. The general process is outlined and applied to the assembly of a library of 500 molecules targeting the serotonin type 7 (5-HT_7) receptor, a class A G-Protein Coupled Receptor (GPCR). Utilizing commercially available building blocks of similar size and composition, a reference library was created. Control sets of known ligands for the 5-HT_7 receptor, other GPCRs, and nuclear receptors were collected from literature sources. Principal component analysis of molecular descriptors for the two libraries and the literature sets, displayed a focusing of the targeted library to the region in the chemical space defined by the literature actives, suggesting a denser coverage of the bioactive region than for the more diverse reference library. Additional computational validations, including PCA class predictions, 3D pharmacophore modeling, and docking calculations all indicated an enrichment factor of 5-HT_7 ligand-like molecules in the range of 2−4 for the targeted library compared to the reference library.

INTRODUCTION

G-Protein Coupled Receptors (GPCRs) constitute the largest family of cell-surface receptors with more than 700 occurrences reported in Uniprot[1] for human full-length forms. The GPCRs are divided into at least six classes with low sequence similarity between them. The common theme is the overall retained fold of 7 trans-membrane (TM) helices that form an anticlockwise helix bundle and the retained coupling to G-proteins. One member of the family has an available crystal structure, that of rhodopsin,[2] the light sensitive protein in the eye, a member of the class A GPCRs. Previously the only related protein with experimentally determined 3D structure was bacteriorhodopsin,[3] which shares 7 TM helices with GPCRs, but does not couple to G-proteins and shares low sequence similarity with GPCRs. This class of receptors recognizes and transduces messages ranging from photons, through organic molecules to peptides and proteins. The signal is transmitted through the rearrangement of the TM bundle so that the G-protein binding site is exposed at the interior of the membrane. The rearrangement of the TM bundle may be spontaneous, as in the case of constitutive active receptors, or mediated by the binding of a small molecule or peptide in a binding site in the interior of the TM bundle or to an extracellular domain of the receptors. The abundance of receptors and stimuli make the receptors an important class of targets for the pharmaceutical industry, as illustrated by the fact that half of all currently prescribed drugs are targeting only 20 GPCRs.[4]

These vast numbers of unexploited GPCRs and the previous successes makes the GPCR drug discovery area attractive for further studies. A cost-efficient approach to lead discovery is to incorporate structural knowledge of both the receptor and the ligand in a targeted library. This allows for a better coverage of the active region of the chemical space of a receptor or of a receptor class.

In a series of articles by Jacoby and co-workers a targeted library concept tracing physicochemical relationships of the binding site in monoaminergic GPCRs were introduced.[5−7] Mutagenesis studies have mapped three distinct binding regions in the type 1A 5-hydroxytryptamine (serotonin) receptor (5-HT_{1A} receptor), which correspond to the binding sites of the ligands 5-hydroxytryptamine (5-HT), 8-hydroxy-*N*,*N*-dipropylaminotetralin (8-OH-DPAT), and propranolol, where mutations in spatially distinct regions affect the binding of the three ligands independently. The three regions are located within the TM region and overlap at a strictly conserved aspartic acid in TM3 in the monoaminergic family, where it acts as a key anchor point for the biogenic basic amine-containing ligands. This is further supported by the architecture of known high-affinity ligands that cross-link two or three spatially distinct groups around an ionizable nitrogen containing group. By defining each site by the binding environment of the ligands, represented by the amino acids that line the binding site, each subsite can be compared to other receptors in a chemogenomic approach.

An extension of this methodology to the entire class A GPCRs should be possible since the possibility to affect the conformation of the receptor through binding to the central cavity probably is a common feature of the family A GPCRs. Although the conserved aspartic acid in TM3 is not present throughout the class A GPCRs, the central TM3 aspartic acid is present in other GPCR subfamilies, and those can be activated by nitrogen-containing compounds, similar to those active at the aminergic receptors. The rat somatostatin receptor subtype SSTR[8] is one example. In the chemokine

* Corresponding author phone: +46 8 697 3401; fax: +46 8 697 23 19; e-mail: Erik.Nordling@biovitrum.com.

10.1021/ci049822w CCC: $27.50
Published on Web 10/07/2004

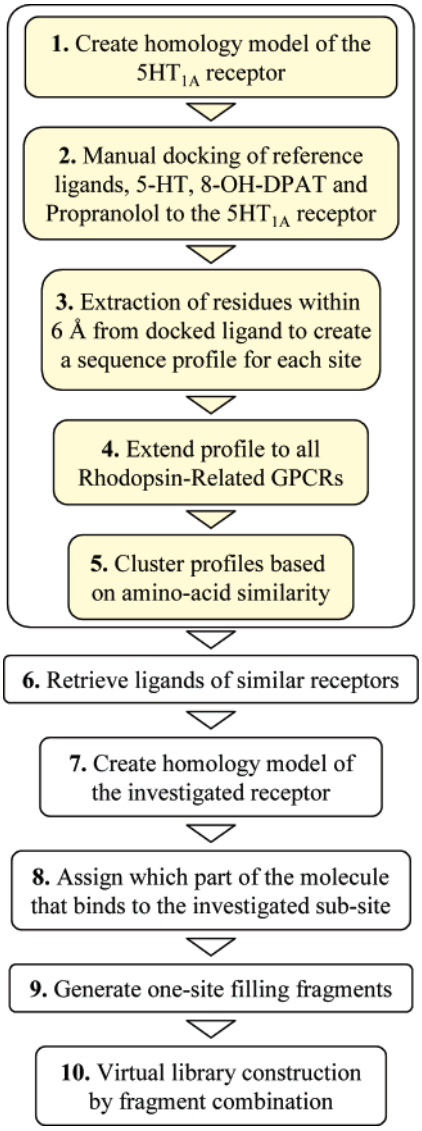

Figure 1. Scheme depicting the design and the construction of a targeted library. Steps 1−5 are identical for a receptor class, while the last five steps are individual for each targeted library.

family, many receptors have an aspartic acid in TM7 that is arranged to create similar interactions as in the aminergic receptors with the result of activation by similar compounds,[9] even though their natural ligands are proteins that bind to the extracellular domain. Other families such as the purinergic receptors share the central binding site, although it is primarily basic residues that interact with the negatively charged natural ligands.[10] These examples highlight the possibility to induce the required conformational change for activation of GPCRs by binding of a small-molecule compound in the central cavity in the TM region, regardless of the nature of the natural ligand, and its native binding site. Taken together the methodology should be extendable to target non-aminergic receptors as well, with the added potential problem of clear subsite assignment for the ligands, due to the lack of a centrally positioned conserved residue.

The design process follows the scheme depicted in Figure 1 and can be divided into one part that assigns the pharmacogenomic relationship among the receptor class (steps 1−5 in Figure 1). This analysis does not have to be repeated until new members of the class are added. The last five steps are individual for each investigated receptor and constitute the retrieval of ligands for similar receptors, generation of one-site targeting fragments, and the creation of the targeted library.

The current work shows that the design strategy proposed by Jacoby and co-workers is a feasible route in the process of designing targeted libraries. Here we address practical issues in the creation of a library and design a library against the 5-HT_7 receptor. This receptor has been implicated in psychosis,[11] circadian rhythms,[12,13] and migraines.[14]

MATERIALS AND METHODS

Homology Model Construction. A homology model of the human 5-HT_{1A} receptor was created by homology modeling in the software package ICM[15−17] (step 1 in Figure 1). The model was based on an in-house alignment[18] of the GPCRs' TM regions. The alignment was extended to the loop regions by using the profile/structural alignment feature of ClustalX,[19] using the aligned TM regions of bovine rhodopsin and the 5-HT_{1A} receptor[20] as a seed for the alignment of the full-length sequences. The template for the homology modeling was the crystal structure of bovine rhodopsin (pdb code 1f88).[2] The third intracellular loop (residue 175−266) was omitted from the model due to lack of structural information from the template for this region. The loop is situated far from the active site and does not directly interact with the ligands investigated in this study.

A homology model of the 5-HT_7 receptor[21] was built by the same procedure, with the added criteria that the N-terminus (first 40 residues), C-terminus (residue 431−479), and intracellular loop 3 (residue 277−309) were truncated due to large insertions in these regions compared to the template structure of rhodopsin (step 7 in Figure 1).

Classification of the Receptors. Manual docking of 5-HT, propranolol, and 8-OH-DPAT to fulfill experimentally verified interactions in the 5-HT_{1A} receptor were used to define the three binding sites (step 2 in Figure 1). A profile for each site was constructed by extracting amino acids from within 6 Å of each docked reference ligand (step 3 in Figure 1). The corresponding amino acids in the human class A GPCRs were collected from an in-house alignment of the TM regions (step 4 in Figure 1). Loop residues, primarily from extracellular loop 3, which form part of the lid of the binding site, were excluded from the profile, due to the large uncertainties in assigning the corresponding amino acids in other receptors.

Quicktree[22] was used to calculate evolutionary trees of the individual subsites sequence motifs by the neighbor-joining method (step 5 in Figure 1). Evolutionary trees were developed to trace the evolution of biological sequences. They are based on amino acid substitution matrices, which penalize substitutions of amino acids with similar properties less. Thus, a measure of the similarity of the physicochemical characteristics of the binding sites is incorporated in the trees.

Construction of the Virtual Targeted Library. Based on the physicochemical clustering of the binding site profiles, the receptor with the most similar physicochemical characteristics can be determined from the trees. Ligands for that receptor were extracted from the Beilstein database using the MDL CrossFire Commander interface. This procedure was repeated for each subsite to yield 3 independent sets of molecules (step 6 in Figure 1).

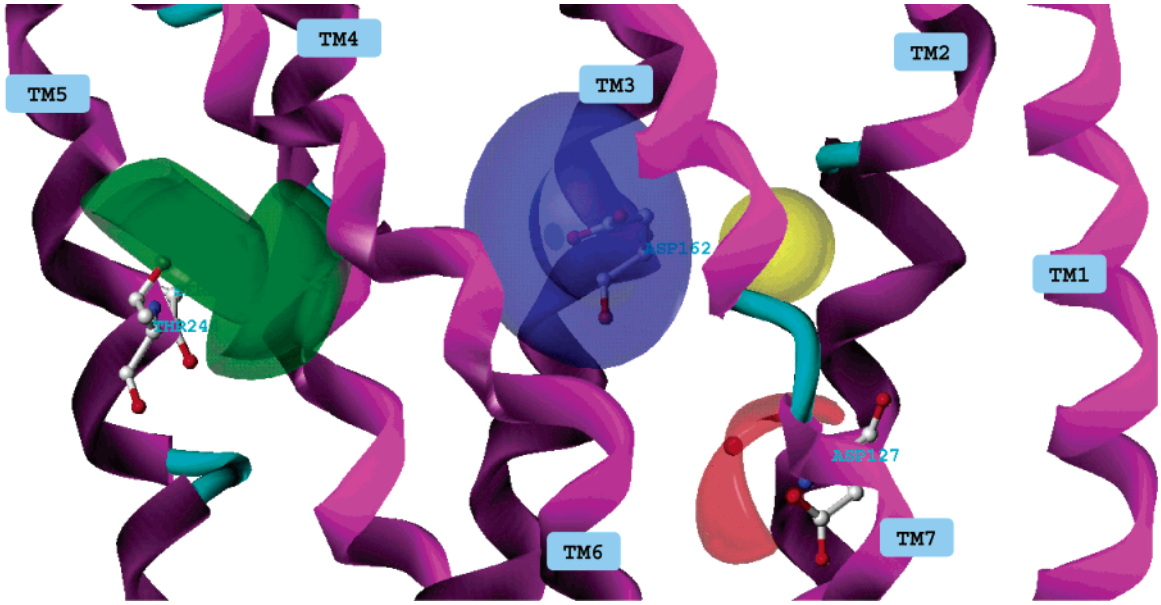

Figure 2. Overview of the TM bundle of the 5-HT$_7$ receptor with the UNITY query features displayed. In blue is the central interaction for the positively charged amine colored blue. The green features to the left are the 5-HT site interactions with Thr244 and/or Ser243 in TM5. The hydrophobic interaction for the propranolol site is colored yellow, and the red fields are the interactions with Asp127 in TM2 for the 8-OH-DPAT site.

Since each set is chosen to target one subsite and each ligand may bind to any number of subsites, it is necessary to assign which part of the molecules bind to the investigated subsite. UNITY[23] queries were constructed for each subsite based on the 5-HT$_7$ model and the three reference one-site filling ligands (step 8 in Figure 1). The queries were constructed to incorporate the mutationally verified interactions between the receptor and the ligands inferred from the 5-HT$_{1A}$ receptor. In Figure 2 the TM region of the 5-HT$_7$ receptor is schematically represented. The conserved aspartic acids interaction field that is shared between the three sites is colored blue. The hydrogen bonding possibilities of the 5-HT site at TM5 are colored green. They were modeled as partial match constraints in UNITY, requiring only one of them to be matched. The 8-OH-DPAT site is located between TM3 and TM2, stretching into the core of the TM region, with the key interaction site colored red at aspartic acid 102. The propranolol site is devoid of hydrogen bonding partners, as most residues in the binding site are hydrophobic, leaving only a hydrophobic interaction colored yellow. The matching part of the molecule, up to and including the central nitrogen, was cut out to produce a one-site filling fragment. As many reports of biological activity present series of close structural analogues this procedure commonly introduces an amount of redundancy among the fragments, which have to be removed (step 9 in Figure 1).

A virtual combinatorial library was constructed with LEGION[24] by joining the one-site filling fragments of each set at the common nitrogen, to produce 2-site filling molecules (step 10 in Figure 1). All standard molecular descriptors and Molconn-Z indices were calculated in ChemEnlighten.[25] A Principal Component Analysis (PCA) of the descriptors was performed in SIMCA-P+,[26] to project the high dimensional space to fewer dimensions. The resulting principal components were subjected to hierarchical clustering to divide the library into 500 discrete clusters. Selecting one compound from each cluster created the final library, a diverse subset with 500 molecules.

Construction and Compilation of Control Sets. A reference library was constructed by extracting primary and secondary amines from commercially available sources, removing molecules with a molecular weight below 100 and above 400 to mimic the selection process for the targeted library. As previously described, a 2-site filling library was generated and resulted in a library of more than 22 000 molecules, from which a set of 500 representative members were selected using the previously described clustering and selection technique.

Known actives against the 5-HT$_7$ receptor were collected from several publications[27−37] and from CrossFire Beilstein. GPCR ligands targeting the 5-HT$_{2A}$, the dopamine D$_3$, and the adrenergic α_1 from other subbranches of the main aminergic cluster were collected from CrossFire Beilstein (set name: GPCR actives). A set of nuclear receptor (NR) ligands, targeting the liver X receptor, retinoid X receptor, retinoic acid receptor, peroxisome proliferator-activated receptor, and triiodothyronine receptor was collected from the same source (set name: NR actives). These sets were assembled for the evaluation of the virtual targeted library.

Evaluations of the Targeted Library. The purpose of the validation is to establish if libraries created by the design scheme will contain molecules that would be selected for a HTS of the targeted receptor using common computational techniques. Three computational methods have been used to assess the degree of focusing of the targeted library: PCA on physicochemical descriptors, docking, and 3D pharmacophore modeling. All standard molecular descriptors and Molconn-Z indices for the libraries and the control sets were calculated in ChemEnlighten.[25] PCA of the descriptors of the investigated sets of molecules was performed in SIMCA-P+.[26] To further analyze the similarities of the investigated sets, PCA class analysis was used. This is based on PCA of the individual sets, and it indicates whether an observation is similar to the observations in the training set based on the distance to the class model.

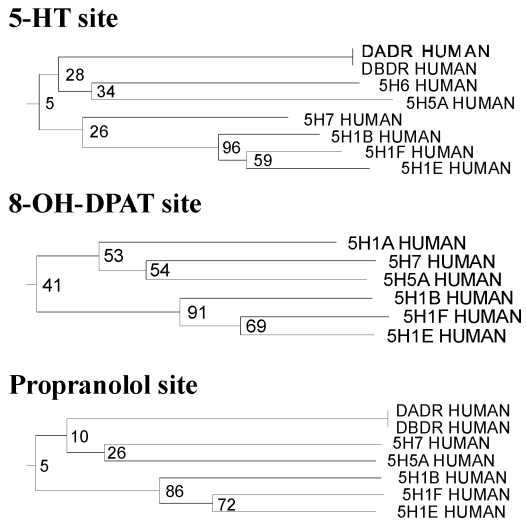

Figure 3. The branch of the evolutionary tree for each subsite containing the 5-HT_7 receptor. For the sake of clarity the complete trees have not been included as the entire tree contain several hundred entries.

Docking of libraries and controls were performed in GLIDE[38] using the OPLS-AA force field[39] with the binding site defined by the conserved aspartic acid (Asp162) in TM 3 and a box size of 14 Å, covering the three subsites. The maximum ligand length was set to 16 Å, restricting the ligands to the TM domain of the receptor. To enhance the ranking of the docked compounds the scores were recalculated using the multiple active site correction (MASC) method,[40] where the ligands are docked to multiple proteins and the score is retrieved for each target. The MASC score is calculated as the original score minus the average score for the different targets divided by the standard deviation of the scores of the different dockings. The new scores are then used to re-rank the docked poses. The method should penalize ligands favored by the scoring function, that is ligands that will score high against several targets. Proteins used in the analysis were thrombin, carbanhydrase, stromelysin, and the estrogen receptor with pdb codes 1etr, 1bn1, 1b3d, and 1ere, respectively.

3D pharmacophore models were calculated in GASP[41] with a population size of 150 and max operations set to 500000. The pharmacophore calculations ended before the max operations were reached by fulfillment of the convergence criteria. The 5-HT_7 antagonist pharmacophore was based on work presented by Lopez-Rodriguez and co-workers[37,42] and is based on their set of selective antagonists. The 5-HT_7 agonist pharmacophore was recreated after a model presented by Vermeulen and co-workers.[43] The pharmacophores were then used as UNITY queries to find matching compounds in the investigated sets.

RESULTS AND DISCUSSION

Classification of the Receptors and Collection of Bioactive Fragments. The biogenic amine receptors clustered together in the evolutionary analysis of all three subsites, indicating the related function and the conserved ability to bind compounds containing a central basic nitrogen. In Figure 3 close-ups of the branches containing the 5-HT_7 receptor for all subsites are displayed. As the 5-HT_{5A} receptor lacked known selective ligands at the time of this study, the receptors with the most similar binding sites are the 5-HT_{1A}, the 5-HT_{1B}, and the dopamine $D_{1/5}$ receptors. Thus, an obvious limiting factor to the usefulness of the methodology is that only receptors with known ligands can drive the design-process.

The number of ligands retrieved for the 5-HT_{1A} receptor, the 5-HT_{1B} receptor, and the dopamine $D_{1/5}$ receptor were 751, 134, and 128, respectively. After the subsite matching process, fragment generation, and removal of redundant fragments there were 137, 38, and 61 unique fragments left for the three subsites. The size of the fragments is similar for the 5-HT and the 8-OH-DPAT-site, with an average molecular weight of 245 and 265, respectively, while the fragments targeting the propranolol-site are lighter with an average weight of 208. The average ring count of the fragments are 2.5, and they are generally composed of a terminal ring system that either directly contains a secondary nitrogen that can be ionized (Figure 4A,F) or has a linker that connects it to a primary (Figure 4D,E) or secondary amine (Figure 4B,C). In a few cases for the 8-OH-DPAT site, the fragment lack a cyclic structure and is constituted by a substituted alkyl chain. The use of the UNITY-queries is a way to bridge the sequence-to-structure gap and aims at incorporating structural information such as the effect of mutations in the active site in the library design process, without relying too heavily on the homology model.

Based on these fragments a virtual library of 9171 molecules was assembled. To obtain a reasonable ADME profile for the library all compounds with molecular weight over 550 and ClogP larger than 5.5 were removed, leaving a library of 5206 molecules.[44] Finally a diverse subset of 500 compounds was chosen to represent the entire library. The choice to build a two-site filling library is made because

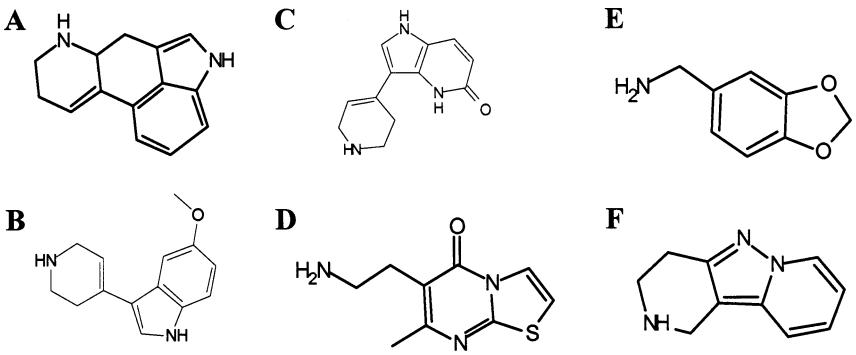

Figure 4. Example of fragments generated in step 9 of the design scheme. A, B: fragments targeting the 5-HT site; C, D: fragments targeting the 8-OH-DPAT site; E, F: fragments targeting the propranolol site.

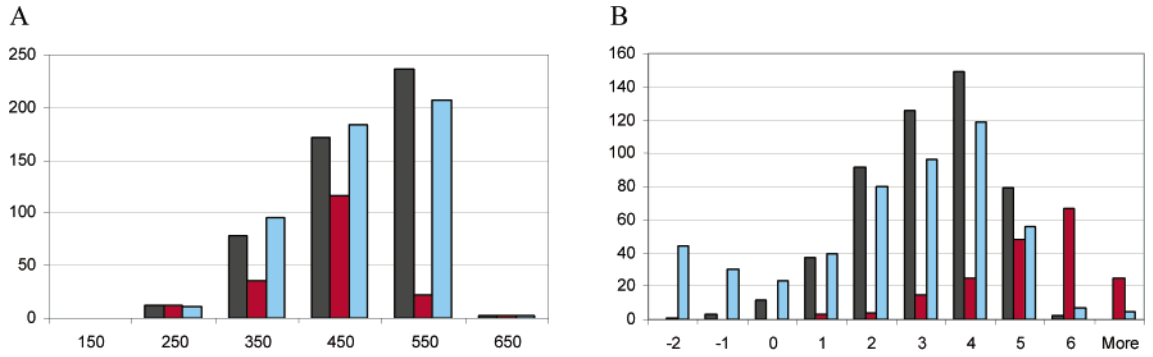

Figure 5. Distribution of molecular weight (A) and ClogP (B) for the targeted library (black bars), the 5-HT$_7$ actives (red bars), and the reference library (blue bars).

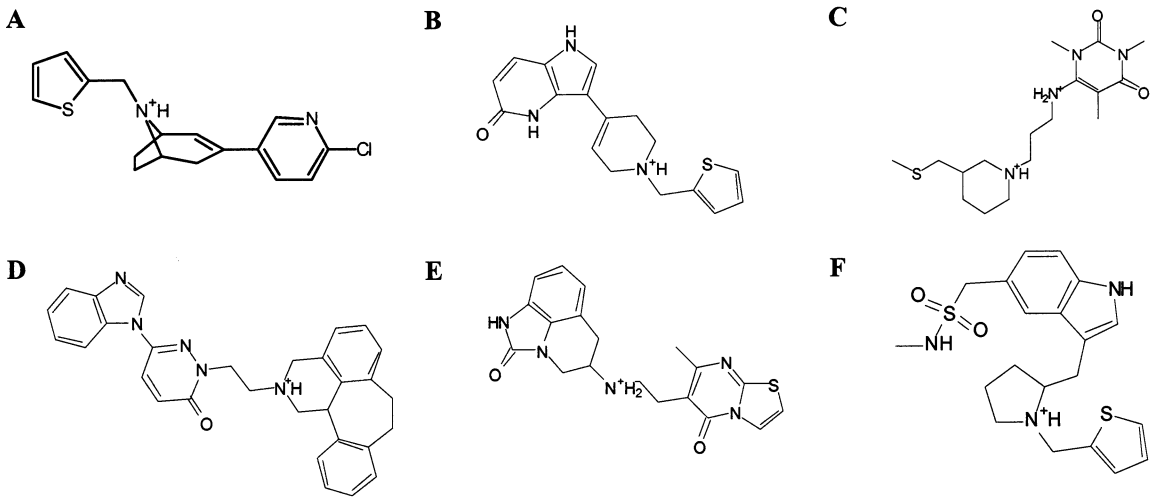

Figure 6. A−F shows final library members, with molecular weights ranging from 312 to 475 coming from the mid-segment of the final library.

a library combining three bioactive fragments to one molecule would have unfavorable physicochemical characteristics from an ADME perspective. The likelihood to achieve a successful hit-to-lead process based on hits from a library based on three fragments would be slim as the hits would be already large and leave little room for potency optimization by additional substitution. The patent situation for a library created by this scheme is perhaps not the strongest since known active fragments are used in the creation of the library. However a more favorable IP situation could be easily obtained by using the generated bioactive fragments to decorate novel scaffolds.

Overview of the Library. The molecular properties of the targeted library have a reasonable profile regarding ClogP and molecular weight for a library of lead-like structures.[44] In Figure 5A the distribution of molecular weight of the targeted library, the reference library, and the literature 5-HT$_7$ actives are plotted. The distributions of the two virtual libraries do not deviate considerably from each other, which speak in favor for the use of bioactive fragments in the library construction process, as they should meet the requirements for biological activity. However in comparison with the known actives the size distribution is shifted slightly in favor of molecules with larger molecular weight. The hydrophobicity profile (Figure 5B) of the directed library is slightly more in phase with the known actives as very hydrophilic molecules are minimized when using fragments from molecules with documented biological activity. Both libraries do have a lower ClogP average than the known actives due to the ClogP filter used in the design process.

The targeted library contains more complex structures compared to the reference library, as the average ring count for it is 4.4 compared to 3 for the reference library, although they share the same extremes with a minimum of 1 and a maximum of 7 rings. In Figure 6 a few examples of the members of the targeted library are displayed.

PCA-Analysis of the Library. The first 6 principal components explained 59% of the variability of the original descriptors. In Figure 5 the first two components of the PCA, which explained 38% of the variability, are plotted against each other. The targeted library (black markers in Figure 7) displays focusing on the region occupied by the 5-HT$_7$ actives (red markers), providing a thorough coverage of the relevant chemical space. The reference library covers the relevant area as well, but, due to its diversity, less depth in the coverage is obtained. Another sign of the diversity is that the reference library also encompasses the chemical space occupied by the nuclear receptor actives (green markers).

Figure 8 shows a Cooman plot for the investigated sets, where the distances to the PCA class model for the target library and the 5-HT$_7$ actives are plotted against each other. Solid lines corresponds to cutoffs for class membership at a 95% tolerance level, leaving the lower left quadrant with molecules classified as members of both sets, indicating similar characteristics and possibly similar biological activities. 12% of the targeted library falls into this category

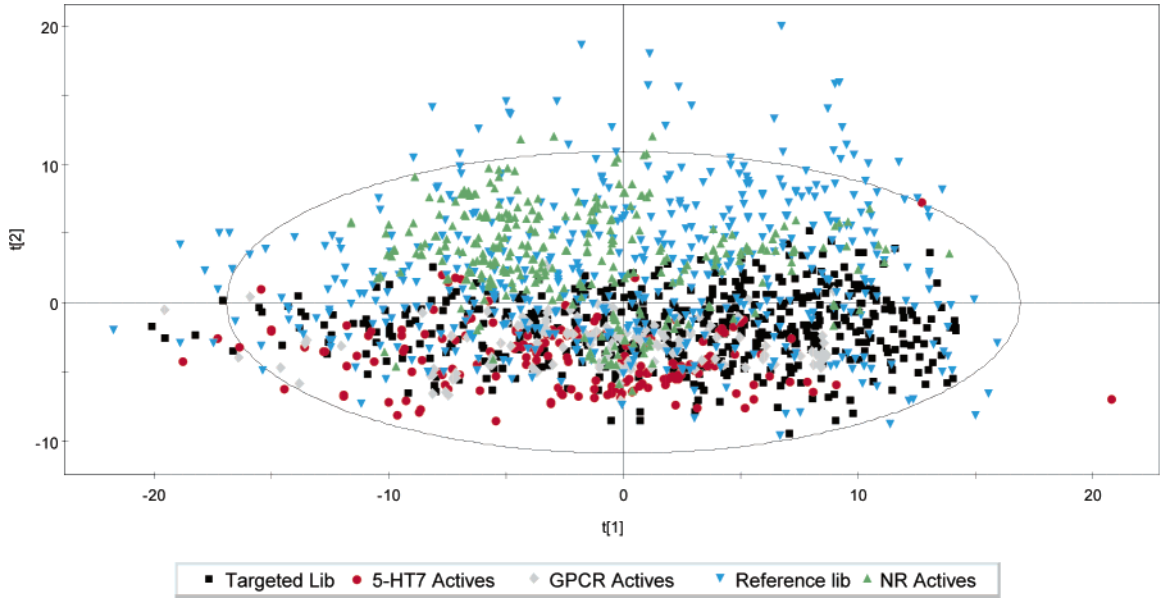

Figure 7. PCA overview of the different sets, with the first two principal components on the axis, the oval corresponds to a confidence level of 95%.

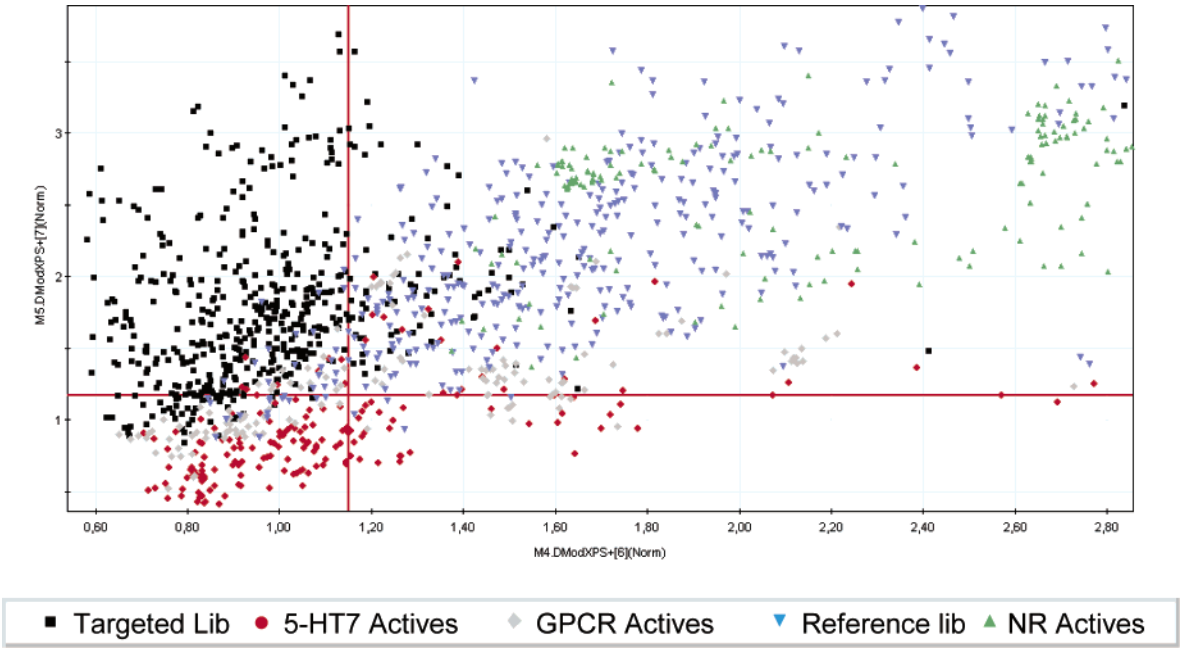

Figure 8. Cooman-plot of the sets in the investigation, where the distance to the PCA class models for the target library (x) and the 5-HT$_7$ actives (y) are plotted against each other. The red lines correspond to a membership probability of 95%. Entries to the left of the vertical line match the targeted library model, and those below the horizontal line match the 5-HT$_7$ actives model.

compared with 3% for the control library. Of the compounds with biological activity, 63% of the 5-HT$_7$ actives are found here and 29% of the GPCR actives. Worth mentioning is that none of the NR actives are found in this space.

From Figure 7 is it evident that the biogenic aminergic receptors share similar ligand characteristics, which is further supported by the concept of GPCR privileged substructures.[45] As active fragments from ligands for other receptors are combined in the library, the cross-receptor specificity should be even more pronounced. But in the Cooman plot in Figure 8, a 2-fold higher membership probability rate for the 5-HT$_7$ actives than the other GPCR actives is evident. This speaks in favor of that the combination of bioactive fragments from ligands targeting physicochemical similar receptors infers some of the desired binding requirements of the targeted receptor.

3D Pharmacophore Models for 5-HT$_7$ Agonism and Antagonism. The 5-HT$_7$ agonist pharmacophore (Figure 9A) consists of two hydrophobic features corresponding to the indole ring of 5-HT. For two hydrogen bond donor features, the topmost one corresponds to the positive nitrogen and the other is located to the indole ring. It also contains a hydrogen bond acceptor that can be positioned as either of the red features in Figure 9A but coordinates to the same partner in the receptor.[43] This was modeled as a partial match constraint in UNITY. The pharmacophore matched 104 compounds of the targeted library and 11 of the reference library. The 5-HT$_7$ antagonist pharmacophore (Figure 9B) has a hydrogen bond

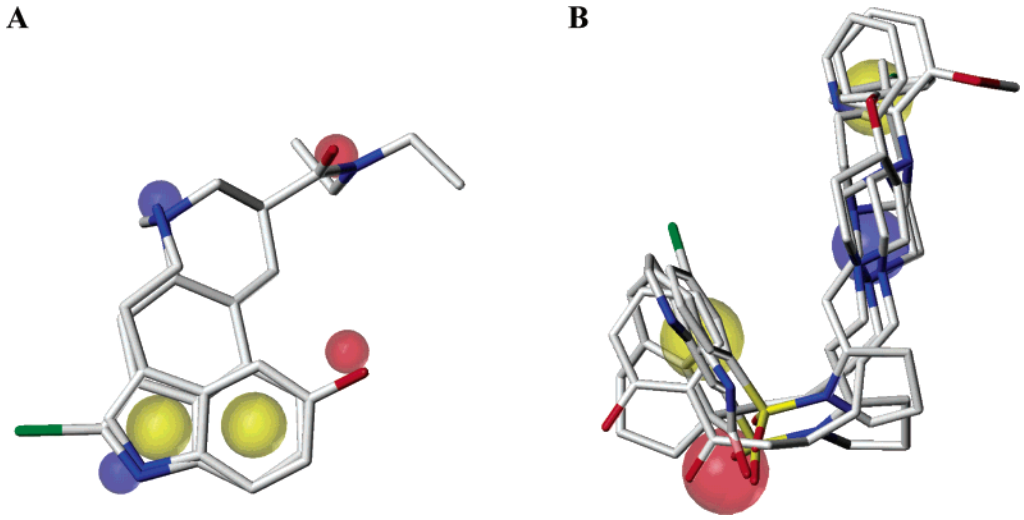

Figure 9. 5-HT$_7$ pharmacophore models: yellow color indicates a hydrophobic feature, blue a hydrogen bond donor, and red a hydrogen bond acceptor. A. Agonist pharmacophore model, the red hydrogen bonding acceptor sites are modeled as a partial match constraint, at least one is required for agonist activity. 2-Br-LSD and 5-HT is aligned to the model. B. Antagonist pharmacophore, representative selective antagonists are aligned to the model.

Table 1. Hit Rates of the Pharmacophore Models[a]

		antagonist pharmacophore		agonist pharmacophore		total
	set size	hits	%	hits	%	hits,%
5-HT$_7$ actives	188	119	63	9	5	66
targeted library	500	293	59	104	21	61
reference library	500	154	31	11	2	32
GPCR actives	176	71	40	0	0	40
NR actives	332	37	11	0	0	11

[a] The column of total hits are composed of unique hits against either pharmacophore.

donor corresponding to the positively charged nitrogen, a hydrophobic feature, and a hydrogen bond acceptor in the lower part of the figure. It also carries an additional hydrophobic feature on the far side of the central hydrogen bond acceptor compared to the agonist model. It matched 293 molecules from the targeted library and 154 from the control library. The pharmacophores suggest that the main influences for agonism or antagonism action of the ligand are the existence of an extra hydrophobic feature in the antagonists and that the distance to the hydrogen bond acceptor from the central nitrogen is 7 Å in the antagonist model compared to 6 Å for the agonist model, while the other comparable features are at equivalent positions. These differences may be enough to induce different conformations in the receptor. The pharmacophore modeling appears to incorporate important features of the known 5-HT$_7$ ligands, as the hit rate is higher for known actives than for other GPCR actives (see Table 1). The members of the targeted library nearly matched the pharmacophore at the same levels as the literature actives and far better than the control sets and the reference library. The uneven distribution of the 5-HT$_7$ actives hits between the two pharmacophores reflects most reported compounds being antagonists.

Docking to the 5-HT$_7$ Receptor. The top 100 poses based on their MASC score show an enrichment of 5-HT$_7$ actives, target library members, and GPCR actives compared to the control library and the NR actives. The docking against the 5-HT$_7$ receptor homology model shows an enrichment of the focused library compared with the control library of 2.4 in the top 500 best scoring hits.

The design is in part based on the assumption that all class A GPCRs share a common fold, at least for the TM region. The existence of easily recognizable conserved sequence stretches, such as the NPXXY stretch in TM7, where X denotes any amino acid, involved in G-protein coupling combined with the hydrophobicity of the TM helices allows for a reasonable accurate alignment of the sequences. According to several studies the overall fold of the TM region should be fairly similar among the class A receptors.[45,46] A recent article highlights the potential possibilities to accurately use docking calculations in GPCR drug discovery by correctly predicting the binding mode of retinal in rhodopsin and bacteriorhodopsin after randomization of the side-chain angles, followed by a global minimization scheme.[47] This should simulate the case of a homology model with correctly predicted backbone conformation but with low reliability in the side-chain placements.

CONCLUSION

Taken together we have demonstrated that it is possible to apply this methodology in the design of a targeted library against GPCRs. One of the advantages of this design strategy is that the initial classification of the receptor class remains unchanged, which shortens the development time for new libraries designed against other receptors of the same class. With the completion of the human genome a complete analysis of all human receptors can be performed. This could serve both as a platform for designing targeted libraries as well as a crude basis for the initial exploration of functionality based on binding site similarities. The computational methods used to validate the targeted library design protocol combine to show that some of the molecular characteristics required for 5-HT$_7$ activity can be inferred from the structure of the ligands. The validation methods also imply that the targeted library has incorporated characteristics required for 5-HT$_7$ activity, as its members match the models at a higher rate than the other sets included in the analyses.

Table 2. Docking Results to the 5-HT_7 Receptor, Ranked after the MASC Score

set name	set size	top 100 poses	% of set	top 500 poses	% of set
5-HT_7 actives	188	23	12	108	57
targeted library	500	45	9	220	44
GPCR actives	177	21	11	73	41
reference library	500	9	2	91	18
NR actives	332	2	1	8	2
total hits	1703	100	6	500	29

Supporting Information Available: Information of the literature 5-HT_7 actives and their activities. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES AND NOTES

(1) Apweiler, R.; Bairoch, A.; Wu, C. H.; Barker, W. C.; Boeckmann, B. et al. UniProt: the Universal Protein knowledgebase. *Nucleic Acids Res.* **2004**, *32 Database issue*, D115−119.

(2) Palczewski, K.; Kumasaka, T.; Hori, T.; Behnke, C. A.; Motoshima, H. et al. Crystal structure of rhodopsin: A G protein-coupled receptor. *Science* **2000**, *289*, 739−745.

(3) Luecke, H.; Schobert, B.; Lanyi, J. K.; Spudich, E. N.; Spudich, J. L. Crystal structure of sensory rhodopsin II at 2.4 angstroms: insights into color tuning and transducer interaction. *Science* **2001**, *293*, 1499−1503.

(4) Christopoulos, A. Allosteric binding sites on cell-surface receptors: novel targets for drug discovery. *Nat. Rev. Drug Discov.* **2002**, *1*, 198−210.

(5) Jacoby, E.; Fauchère, J. L.; Raimbaud, E.; Ollivier, E.; Michel, A. et al. A Three Binding Site Hypothesis for the Interaction of Ligands with Monoamine G Protein-coupled Receptors: Implications for Combinatorial Ligand Design. *Quant. Struct.−Act. Relat.* **1999**, *18*, 561−572.

(6) Jacoby, E. A Novel Chemogenomics Knowledge-Based Ligand Design Strategy − Application to G Protein-Coupled Receptors. *Quant. Struct.−Act. Relat.* **2001**, *20*, 115−123.

(7) Milligan, G.; Kellett, E.; Dacquet, C.; Dubreuil, V.; Jacoby, E. et al. S 14506: novel receptor coupling at 5-HT(1A) receptors. *Neuropharmacology* **2001**, *40*, 334−344.

(8) Strnad, J.; Hadcock, J. R. Identification of a critical aspartate residue in transmembrane domain three necessary for the binding of somatostatin to the somatostatin receptor SSTR2. *Biochem. Biophys. Res. Commun.* **1995**, *216*, 913−921.

(9) Mirzadegan, T.; Diehl, F.; Ebi, B.; Bhakta, S.; Polsky, I. et al. Identification of the binding site for a novel class of CCR2b chemokine receptor antagonists: binding to a common chemokine receptor motif within the helical bundle. *J. Biol. Chem.* **2000**, *275*, 25562−25571.

(10) Moro, S.; Guo, D.; Camaioni, E.; Boyer, J. L.; Harden, T. K. et al. Human P2Y1 receptor: molecular modeling and site-directed mutagenesis as tools to identify agonist and antagonist recognition sites. *J. Med. Chem.* **1998**, *41*, 1456−1466.

(11) Roth, B. L.; Craigo, S. C.; Choudhary, M. S.; Uluer, A.; Monsma, F. J., Jr. et al. Binding of typical and atypical antipsychotic agents to 5-hydroxytryptamine-6 and 5-hydroxytryptamine-7 receptors. *J. Pharmacol. Exp. Ther.* **1994**, *268*, 1403−1410.

(12) Vanhoenacker, P.; Haegeman, G.; Leysen, J. E. 5-HT7 receptors: current knowledge and future prospects. *Trends Pharmacol. Sci.* **2000**, *21*, 70−77.

(13) Jovanovska, A.; Prosser, R. A. Translational and transcriptional inhibitors block serotonergic phase advances of the suprachiasmatic nucleus circadian pacemaker in vitro. *J. Biol. Rhythms* **2002**, *17*, 137−146.

(14) Villalon, C. M.; Centurion, D.; Valdivia, L. F.; De Vries, P.; Saxena, P. R. An introduction to migraine: from ancient treatment to functional pharmacology and antimigraine therapy. *Proc. West Pharmacol. Soc.* **2002**, *45*, 199−210.

(15) ICM version 3; Molsoft LLC, San Diego, CA.

(16) Abagyan, R.; Batalov, S.; Cardozo, T.; Totrov, M.; Webber, J. et al. Homology modeling with internal coordinate mechanics: deformation zone mapping and improvements of models via conformational search. *Proteins* **1997**, *Suppl*, 29−37.

(17) Cardozo, T.; Totrov, M.; Abagyan, R. Homology modeling by the ICM method. *Proteins* **1995**, *23*, 403−414.

(18) Personal communication, Torsten Sejlitz.

(19) 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.

(20) Fargin, A.; Raymond, J. R.; Lohse, M. J.; Kobilka, B. K.; Caron, M. G. et al. The genomic clone G-21 which resembles a beta-adrenergic receptor sequence encodes the 5-HT1A receptor. *Nature* **1988**, *335*, 358−360.

(21) Bard, J. A.; Zgombick, J.; Adham, N.; Vaysse, P.; Branchek, T. A. et al. Cloning of a novel human serotonin receptor (5-HT7) positively linked to adenylate cyclase. *J. Biol. Chem.* **1993**, *268*, 23422−23426.

(22) Howe, K.; Bateman, A.; Durbin, R. QuickTree: building huge Neighbour-Joining trees of protein sequences. *Bioinformatics* **2002**, *18*, 1546−1547.

(23) UNITY, version 4.3; Tripos Inc., 1699 South Hanley Rd, St. Louis, MO, 63144.

(24) LEGION; Tripos Inc., 1699 South Hanley Rd, St. Louis, MO, 63144.

(25) ChemEnlighten; Tripos Inc., 1699 South Hanley Rd, St. Louis, MO.

(26) SIMCA-P+; Umetrics AB, Sweden.

(27) Perrone, R.; Berardi, F.; Colabufo, N. A.; Lacivita, E.; Leopoldo, M. et al. Synthesis and structure-affinity relationships of 1-[omega-(4-aryl-1-piperazinyl)alkyl]-1-aryl ketones as 5-HT(7) receptor ligands. *J. Med. Chem.* **2003**, *46*, 646−649.

(28) Thomas, D. R.; Melotto, S.; Massagrande, M.; Gribble, A. D.; Jeffrey, P. et al. SB-656104-A, a novel selective 5-HT7 receptor antagonist, modulates REM sleep in rats. *Br. J. Pharmacol.* **2003**, *139*, 705−714.

(29) Hagan, J. J.; Price, G. W.; Jeffrey, P.; Deeks, N. J.; Stean, T. et al. Characterization of SB-269970-A, a selective 5-HT(7) receptor antagonist. *Br. J. Pharmacol.* **2000**, *130*, 539−548.

(30) Lovell, P. J.; Bromidge, S. M.; Dabbs, S.; Duckworth, D. M.; Forbes, I. T. et al. A novel, potent, and selective 5-HT(7) antagonist: (R)-3-(2-(2-(4-methylpiperidin-1-yl)ethyl)pyrrolidine-1-sulfonyl)phenol (SB-269970). *J. Med. Chem.* **2000**, *43*, 342−345.

(31) Parikh, V.; Welch, W. M.; Schmidt, A. W. Discovery of a series of (4,5-dihydroimidazol-2-yl)-biphenylamine 5-HT7 agonists. *Bioorg. Med. Chem. Lett.* **2003**, *13*, 269−271.

(32) Linnanen, T.; Brisander, M.; Unelius, L.; Rosqvist, S.; Nordvall, G. et al. Atropisomeric derivatives of 2′,6′-disubstituted (R)-11-phenylaporphine: selective serotonin 5-HT(7) receptor antagonists. *J. Med. Chem.* **2001**, *44*, 1337−1340.

(33) Linnanen, T.; Brisander, M.; Mohell, N.; Johansson, A. M. Serotonergic and dopaminergic activities of rigidified (R)-aporphine derivatives. *Bioorg. Med. Chem. Lett.* **2001**, *11*, 367−370.

(34) Kikuchi, C.; Ando, T.; Watanabe, T.; Nagaso, H.; Okuno, M. et al. 2a-[4-(Tetrahydropyridoindol-2-yl)butyl]tetrahydrobenzindole derivatives: new selective antagonists of the 5-hydroxytryptamine7 receptor. *J. Med. Chem.* **2002**, *45*, 2197−2206.

(35) Kikuchi, C.; Suzuki, H.; Hiranuma, T.; Koyama, M. New tetrahydrobenzindoles as potent and selective 5-HT(7) antagonists with increased In vitro metabolic stability. *Bioorg. Med. Chem. Lett.* **2003**, *13*, 61−64.

(36) Forbes, I. T.; Douglas, S.; Gribble, A. D.; Ife, R. J.; Lightfoot, A. P. et al. SB-656104-A: a novel 5-HT(7) receptor antagonist with improved in vivo properties. *Bioorg. Med. Chem. Lett.* **2002**, *12*, 3341−3344.

(37) Lopez-Rodriguez, M. L.; Porras, E.; Benhamu, B.; Ramos, J. A.; Morcillo, M. J. et al. First pharmacophoric hypothesis for 5-HT7 antagonism. *Bioorg. Med. Chem. Lett.* **2000**, *10*, 1097−1100.

(38) Glide; Schrodinger, 1500 SW First Avenue, Suite 1180, Portland, OR 97201.

(39) Jorgensen, W. L.; Maxwell, D. S.; Tirado-Rives, J. Development and Testing of the OPLS All-Atom Force Field on Conformational Energetics and Properties of Organic Liquids. *J. Am. Chem. Soc.* **1996**, *118*, 11225−11236.

(40) Vigers, G. P.; Rizzi, J. P. Multiple active site corrections for docking and virtual screening. *J. Med. Chem.* **2004**, *47*, 80−89.

(41) Jones, G.; Willett, P.; Glen, R. C. A genetic algorithm for flexible molecular overlay and pharmacophore elucidation. *J. Comput.-Aided Mol. Des.* **1995**, *9*, 532−549.

(42) Lopez-Rodriguez, M. L.; Porras, E.; Morcillo, M. J.; Benhamu, B.; Soto, L. J. et al. Optimization of the pharmacophore model for 5-HT7R antagonism. Design and synthesis of new naphtholactam and naphthosultam derivatives. *J. Med. Chem.* **2003**, *46*, 5638−5650.

(43) Vermeulen, E. S.; Schmidt, A. W.; Sprouse, J. S.; Wikstrom, H. V.; Grol, C. J. Characterization of the 5-HT(7) receptor. Determination

of the pharmacophore for 5-HT(7) receptor agonism and CoMFA-based modeling of the agonist binding site. *J. Med. Chem.* **2003**, *46*, 5365−5374.

(44) Lipinski, C. A.; Lombardo, F.; Dominy, B. W.; Feeney, P. J. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv. Drug Deliv. Rev.* **2001**, *46*, 3−26.

(45) Bondensgaard, K.; Ankersen, M.; Thogersen, H.; Hansen, B. S.; Wulff, B. S. et al. Recognition of privileged structures by G-protein coupled receptors. *J. Med. Chem.* **2004**, *47*, 888−899.

(46) 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.

(47) Cavasotto, C. N.; Orry, A. J.; Abagyan, R. A. Structure-based identification of binding sites, native ligands and potential inhibitors for G-protein coupled receptors. *Proteins* **2003**, *51*, 423−433.

CI049822W