Multistructure 3D-QSAR Studies on a Series of Conformationally Constrained Butyrophenones Docked into a New Homology Model of the 5-HT$_{2A}$ Receptor

Cristina Dezi,[†] José Brea,[†,‡] Mario Alvarado,[§] Enrique Raviña,[§] Christian F. Masaguer,[§] María Isabel Loza,[‡] Ferran Sanz,[†] and Manuel Pastor*[,†]

Research Unit on Biomedical Informatics (GRIB), IMIM, Universitat Pompeu Fabra, Dr. Aiguader 88, E-08003 Barcelona, Spain, and Department of Pharmacology, Department of Organic Chemistry, Faculty of Pharmacy, Universidade de Santiago de Compostela, E-15782 Santiago de Compostela, Spain

Received March 11, 2007

The present study is part of a long-term research project aiming to gain insight into the mechanism of action of atypical antipsychotics. Here we describe a 3D-QSAR study carried out on a series of butyrophenones with affinity for the serotonin-2A receptor, aligned by docking into the binding site of a receptor model. The series studied has two peculiarities: (i) all the compounds have a chiral center and can be represented by two enantiomeric structures, and (ii) many of the structures can bind the receptor in two alternative orientations, posing the problem of how to select a single representative structure for every compound. We have used an original solution consisting of the simultaneous use of multiple structures, representing different configurations, binding conformations, and positions. The final model showed good statistical quality (n = 426, r^2 = 0.84, q^2_{LOO} = 0.81) and its interpretation provided useful information, not obtainable from the simple inspection of the ligand−receptor complexes.

Introduction

Drugs active at 5-HT$_2$ receptors are used in the treatment of depression, mania, anxiety, and schizophrenia. The blockade of these serotonin receptors has been postulated to play a critical role in the action of a new generation of antipsychotic drugs, characterized by producing less extrapyramidal side effects (EPS[a]), usually named atypical antipsychotic drugs (AAPD). The involvement of 5-HT$_2$ receptors in the pharmacological profile of atypical antipsychotics is supported by a large number of biological, pharmacological, and clinical studies.[1−3] Unlike classical antipsychotics like haloperidol, which mainly block type 2 dopamine receptors (D_2), clozapine and other AAPD are relatively more potent at blocking serotonin-2A (5-HT$_{2A}$) than dopamine D_2 receptors. This finding gave rise to the serotonin-dopamine hypothesis, suggesting that for AAPD the blockade of presynaptic 5-HT$_{2A}$ receptors is the predominant mechanism in the nigrostriatal, mesocortical, and tuberoinfundibular dopaminergic pathways, where they increase dopamine release. This effect will counteract the drugs action at D_2 receptors, thus decreasing the incidence of adverse EPS, cognitive deficits, hyperprolactemia, and/or negative symptoms. In contrast, D_2 blockade would prevail over 5-HT$_{2A}$ antagonism in the dopaminergic mesolimbic pathway, resulting in the mitigation of the positive symptoms of psychosis.

Meltzer and co-workers[4,5] suggested that the ratio between the pK_i of antipsychotic agents at the 5-HT$_{2A}$ and D_2 receptors reflects the atypical profile (pK_i 5-HT$_{2A}$/pK_i D_2 >1.12).[4] This ratio led to the definition of the Meltzer's index (MI) as a screening criteria for candidate AAPD. Although the exceptional therapeutic performance of clozapine, not completely achieved by other AAPD that fulfill the MI criteria, has been recently proposed to be extended to its interaction with a large number of receptors in addition to 5-HT$_{2A}$ and D_2 receptors.[6]

For a long time, our group has been involved in research aiming to improve the understanding of the pharmacological mechanisms of AAPD action and to discover novel compounds with interest in psychosis therapy. Many of these studies have been carried out on a series of conformationally restricted butyrophenone analogues, still in active development, containing compounds with high affinity for the 5-HT$_{2A}$ receptor,[7−9] and with particularly interesting pharmacological profiles.[10] Current computational studies carried out by our group on this series aim to unveil the structural determinants of their receptor binding affinity and selectivity and to identify the receptor binding profile, which correlates with their therapeutic usefulness. These studies, of which this work is a first report, involve building homology models for receptors potentially involved in psychosis and using them in a computational protocol that helps to unveil the ligand−receptor interaction mechanisms.

Computational methods can be used in this context for obtaining structural models of the ligand−receptor complexes, but for a large series of compounds (as in this case), the visual inspection of the complex structures is not helpful. Instead, 3D-QSAR methodologies can use these receptor-docked ligand structures to obtain models that extract from them the most relevant information. This use of ligand−receptor complexes as a starting point for 3D-QSAR has been previously reported.[11−18] It is well-known that one of the main drawbacks of 3D-QSAR methods is the difficulty of obtaining suitable ligand alignments. In this study in which a receptor structure has been built ad hoc, the ligands can be docked into the receptor binding site, and the resulting conformations can be used as aligned structures for the 3D-QSAR analysis. The use of the receptor-docked structures has the advantage of providing a more realistic

* To whom correspondence should be addressed. Tel.: (+34) 933 160 512. Fax: (+34) 933 160 550. E-mail: manuel.pastor@upf.edu.
[†] Universitat Pompeu Fabra.
[‡] Department of Pharmacology, Universidade de Santiago de Compostela.
[§] Department of Organic Chemistry, Universidade de Santiago de Compostela.

[a] Abbreviations: 5-HT, serotonin; AAPD, atypical antipsychotic drugs; EPD, extrapyramidal side effects; D2, dopamine receptors of type 2; MI, Meltzer's index; 3D-QSAR, three-dimensional quantitative structure−activity relationships; GPCR, G-protein coupling receptors; Rh, bovine rhodopsin; MD, molecular dynamics; SRD/FFD, smart region definition/fractional factorial design; PLS, partial least squares; LV, latent variables; LOO, leave-one-out; CoMFA, comparative field analysis; BR, bacteriorhodopsin; EM, energy minimization; MIF, molecular interaction fields.

10.1021/jm070277a CCC: $37.00
Published on Web 06/19/2007

representation of the ligands bioactive conformations and normally does not introduce any difficulty in the methodology. Unfortunately, in this particular series, these efforts led to obtaining not one, but a set of candidate structures for every compound, from which it was not obvious how to pick a single representative structure. Most of the criteria that are used in these circumstances to make a choice seem to us arbitrary or dangerous, thus, we decided to build the 3D-QSAR models using a balanced set of structures representing the different species that bind the receptor. The procedure is unusual but not entirely new. A similar approach has been reported in the past in a different context[19] with good results. Moreover, the peculiar characteristics of the series studied offers an excellent opportunity to validate the usefulness of the proposed approach by interpreting the results using the structure of the receptor and by comparing the results with those obtained using more conservative approaches.[7]

Results and Discussion

Synthesis. We will start by reporting the synthesis of seven novel compounds added to the series subject of this study (Table 1; Chart 1). Pyridazinone derivatives **3**, **5**, and **6** were synthesized as shown in Scheme 1: compounds **1** and **2** were prepared by the reaction of 3-benzoylpropionic acid with formaldehyde and heterocyclic amines under Mannich conditions. The synthesis of the aminomethylpyridazinones **3** and **4** was readily accomplished in high yields by refluxing the required 3-benzoyl-4-aminobutyric acid with hydrazine hydrate in ethanol. Compounds **5** and **6** were obtained by alkylation of the piperazine derivative **4** with 5-bromo-1-(*p*-fluorophenyl)-1-pentanone,[20] for compound **6**, or 4-chloro-1,1-ethylenedioxy-1-(*p*-fluorophenyl)butane[21] followed by acid hydrolysis of the ketal, for compound **5**.

Compounds **9** and **10** were prepared (Scheme 2) from 3-(*p*-toluenesulfonyloxymethyl)-1,2,3,4-tetrahydronaphthalen-1-one **7**[22] or 6,7-dimethoxy-3-(*p*-toluenesulfonyloxymethyl)-1,2,3,4-tetrahydronaphthalen-1-one **8**,[23] respectively, via nucleophilic displacement of the tosylate with 1,3,8-triazaspiro[4.5]-decan-4-one[24] (for **9**) or 1,3,8-triazaspiro[4.5]decane-2,4-dione[25] (for **10**) in *N*,*N*-dimethylformamide (DMF) and provided, after bulb to bulb distillation of DMF and chromatography, the target compounds **9** and **10** in moderate yields.

For the preparation of compound **16** (Scheme 3) 1-(2-hydroxyethyl)piperazine was used as starting material. BOC protection of the secondary amine followed by Swern oxidation of the hydroxy group gave the unstable aldehyde **12** in 45% yield. Aldol condensation with 4,5,6,7-tetrahydrobenzofuran-4-one in the presence of LDA-HMPA and subsequent Pd−C catalyzed hydrogenation afforded the piperazine derivative **14** in 46% yield. BOC-deprotection with TFA and finally direct coupling with 4-fluorobenzoic acid by means of carboxylate activation by 1,3-dicyclohexylcarbodiimide (DCC) in the presence of 1-hydroxybenzotriazole (HOBt) led to the target amine **16** in 70% yield (two steps).

The preparation of compound **24** (Scheme 4) was carried out following a similar synthetic strategy reported for compound **69**;[7] Friedel−Crafts acylation of 2-pentylthiophene with succinic anhydride in the presence of $AlCl_3$, subsequent Clemmensen reduction of the ketonic group of the ketoacid **17**, and cyclization with a 20:1 mixture of trifluoroacetic acid (TFA) and trifluoroacetic anhydride (TFAA) afforded the pentylbenzothiophene **19** in 40% yield (three steps). Alkylation of **19** with lithium diisopropylamide at −78 °C, followed by quenching with ethyl bromoacetate, gave ethyl ester **20** that was then hydrolyzed to the tetrahydrobenzothiophenylacetic acid **21** in 50% overall yield. This was condensed with 4-(6-fluorobenzisoxazol-3-yl)-piperidine in the presence of DCC and HOBt to give the corresponding amide **22**. This compound was protected by ketalization with ethylene glycol and *p*-toluenesulfonic acid (*p*-TsOH), the amide group was reduced to amine with $LiAlH_4$, and the ethylene ketal was cleaved in acidic medium to afford the amine **24** in 54% yield from the acid **21**. Further details about the experimental procedures and data about the novel compounds are available as Supporting Information.

Building of 5-HT_{2A} Model. At present, the best starting material for the homology building of G-protein coupling receptor (GPCR) models is the crystal structure of bovine rhodopsin (Rh),[26] even if the low-sequence homology (∼20%) between the 5-HT_2 receptors and the Rh makes difficult the application of standard homology modeling methods. In addition, the availability of only one experimental template might introduce an undesirable bias in the modeling, making the modeled structure "too similar" to the template. In this work, we attempted to reach a certain compromise in-between making use of the Rh template and the introduction of other sources of structural information. The model was built introducing in the MODELLER suite of programs[27] a set of structural templates, including the structure of Rh as well as other relevant structures. The program yielded a consensus solution consisting of a set of 25 candidate structures from which a single structure was selected (Figure 1). A detailed description of the procedure used for building and validating this model can be found in the Experimental Section.

In the receptor model, the location of the binding pocket is hinted by the position of the Asp3.32, which is known to establish a charge-reinforced hydrogen bond with the protonated nitrogen present in aminergic ligands. Before using this structure for docking simulations, the binding pocket must be widened and the side chains of the residues located there must be reoriented. This operation was done by carrying out a preliminary docking of the 5-HT_{2A} antagonist/inverse agonist ketanserin[28,29] (see Chart 3) and extensive molecular dynamic (MD) simulations using a stepwise strategy (as described in the Experimental Section) until the observed ligand−receptor interactions were consistent with experimental data. The reasons for choosing ketanserin were twofold: on the one hand, it is a well-known ligand of 5-HT_{2A} for which many results of mutagenesis experiments are available, thus allowing a validation of the proposed binding site using experimental data; on the other hand, ketanserin exhibits a rather large chemical similarity with the compounds of our series in terms of shape and in terms of presence of similar functional groups at equivalent positions.

The best 5-HT_{2A}−ketanserin complex obtained is shown in Figures 1 and 2. The binding site is located approximately parallel to the helices, slightly tilted with respect to the axis of the helices bundle and nearly parallel to the helix 3. The ligand makes polar contacts mainly with residues in helices 3 and 7, as well as with some residues of the extracellular loop 2. Hydrophobic contacts are observed with residues of helices 5 and 6. The amino acids making the most relevant interactions with ketanserin are Asp3.32, Ser3.36, Trp3.28, Phe6.51, Tyr7.43, and Ser161. Table 2 lists the percentages of hydrogen bonding between these residues and ketanserin, obtained by monitoring a 1ns MD simulation (see the Experimental Section for details).

The comparison of the ketanserin complex with the results of site-directed mutagenesis experiments shows a good agree-

Table 1. Chemical Structure and Binding Affinity Measured for the 5-HT_{2A} Receptor

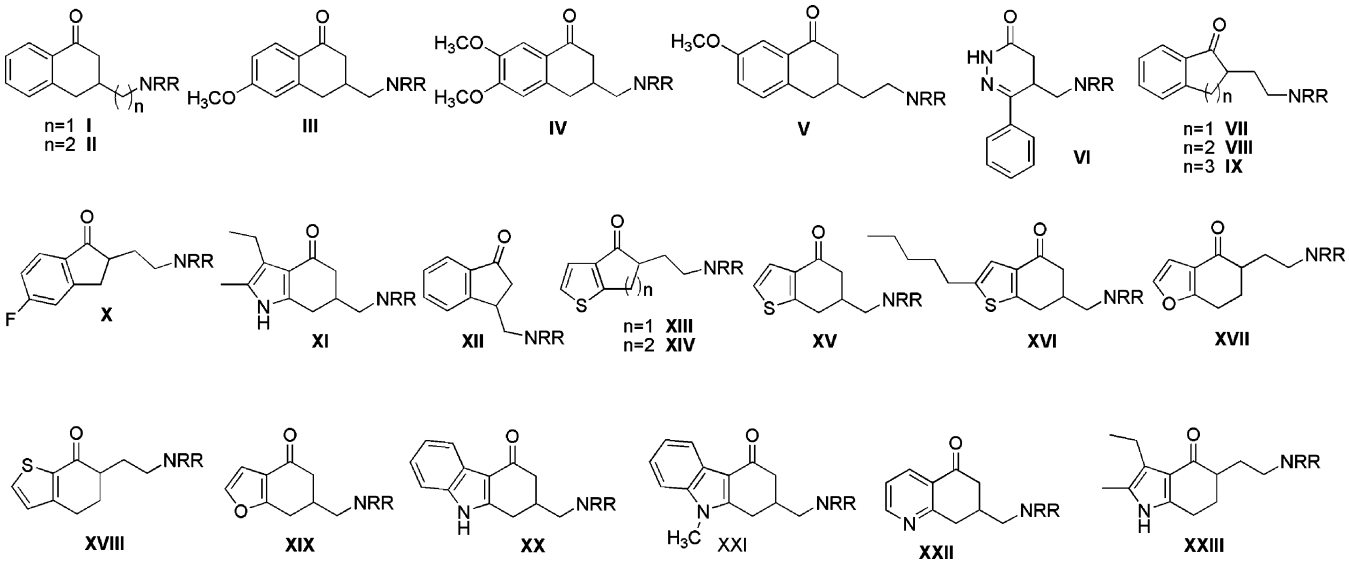
cmpd	code	alkanone fragment[a]	amine fragment[b]	pK_i 5-HT_{2A} rat	pK_i 5-HT_{2A} human	ref
3	QF0225B	VI	b	7.24 ± 0.02		
5	QF0226B	VI	a	6.56 ± 0.02		
6	QF0227B	VI	h	6.00 ± 0.01		
9	QF0109B	I	c	<5		
10	QF0140B	IV	g	<5		
16	QF0712B	XVII	i	5.70 ± 0.20		
24	QF0620B	XVI	b	7.68 ± 0.13		
25	QF0102B	I	a	7.75 ± 0.60	6.94 ± 0.31	38
26	QF0107B	I	f	6.01 ± 0.20		23
(*R*)-**27**	(*R*)-QF0108B	I	b	9.33 ± 0.12	8.20 ± 0.37	23
(*S*)-**27**	(*S*)-QF0108B	I	b		7.53 ± 0.32	23
28	QF0120V	II	d	6.59 ± 0.29		39
(*R*)-**29**	(*R*)-QF0124B	III	e		7.79 ± 0.46	23
(*S*)-**29**	(*S*)-QF0124B	III	e		7.95 ± 0.41	23
30	QF0127B	III	f	6.16 ± 0.21		23
(*R*)-**31**	(*R*)-QF0128B	III	b		7.58 ± 0.19	23
(*S*)-**31**	(*S*)-QF0128B	III	b		8.83 ± 0.49	23
32	QF0128V	II	b	8.29 ± 0.25		39
33	QF0129V	II	c	<5		39
34	QF0130V	II	g	6.15 ± 0.13		39
35	QF0140V	V	d	6.65 ± 0.07		39
(*R*)-**36**	(*R*)-QF0144B	IV	e		8.25 ± 0.23	23
(*S*)-**36**	(*S*)-QF0144B	IV	e		7.36 ± 0.30	23
37	QF0147V	V	f	6.33 ± 0.11		39
38	QF0148V	V	b	8.23 ± 0.14		39
39	QF0149V	V	c	<5		39
40	QF0150V	V	g	5.98 ± 0.15		39
41	QF0301B	VIII	d	7.39 ± 0.70		40
42	QF0303B	VIII	e	8.11 ± 0.80		41
43	QF0304B	VIII	a	7.14 ± 0.70		42
44	QF0307B	VII	e	8.60 ± 0.80		41
45	QF0308B	VII	d	6.71 ± 0.14		8
46	QF0311B	IX	e	7.88 ± 0.70		41
47	QF0313B	X	e	8.42 ± 0.30		41
48	QF0315B	X	a	7.27 ± 0.10		42
49	QF0402B	XXIII	d	7.04 ± 0.70		43
50	QF0407B	XI	d	6.55 ± 0.80		43
51	QF0408B	XI	e	8.06 ± 0.10		43
52	QF0409B	XI	b	8.37 ± 0.80		43
53	QF0410B	XI	a	6.04 ± 0.60		43
54	QF0501B	XII	e	7.58 ± 0.14		7
55	QF0502B	XII	a	7.34 ± 0.10		7
56	QF0503B	XII	d	6.05 ± 0.21		7
57	QF0504B	XII	f	<5		8
58	QF0505B	XIII	e	8.15 ± 0.13		7
59	QF0506B	XIII	a	7.37 ± 0.15		7
60	QF0510B	XIII	b	8.76 ± 0.20		7
61	QF0601B	XIV	e	8.84 ± 0.17		7
62	QF0602B	XIV	a	6.54 ± 0.13		7
63	QF0603B	XIV	d	6.84 ± 0.12		7
64	QF0604B	XIV	f	6.76 ± 0.11		8
65	QF0605B	XV	e	7.24 ± 0.12		7
66	QF0606B	XV	a	6.97 ± 0.13		7
67	QF0607B	XV	d	5.82 ± 0.06		7
68	QF0608B	XV	f	5.75 ± 0.10		8
68	QF0609B	XV	b	7.93 ± 0.05		7
69	QF0610B	XIV	b	8.56 ± 0.20		7
70	QF0701B	XVII	d	6.72 ± 0.17		8
71	QF0702B	XVII	e	7.71 ± 0.90		8
72	QF0703B	XVII	b	8.97 ± 0.09	9.14 ± 0.11	8
73	QF0704B	XVII	f	6.45 ± 0.10		8
74	QF0901B	XVIII	e	7.95 ± 0.09		7
75	QF0902B	XVIII	b	8.17 ± 0.18		7
76	QF1003B	XIX	e	7.29 ± 0.20		46
77	QF1004B	XIX	b	7.97 ± 0.03	8.26 ± 0.12	46
78	QF1005B	XIX	a	6.02 ± 0.03		46
79	QF1006B	XIX	d	<5		46
80	QF1007B	XIX	j	7.66 ± 0.05		46
81	QF1008B	XIX	f	5.08 ± 0.10		46
82	QF1010B	XIX	k	<5		46
83	QF1011B	XIX	h	5.90 ± 0.04		46

Table 1 (Continued)

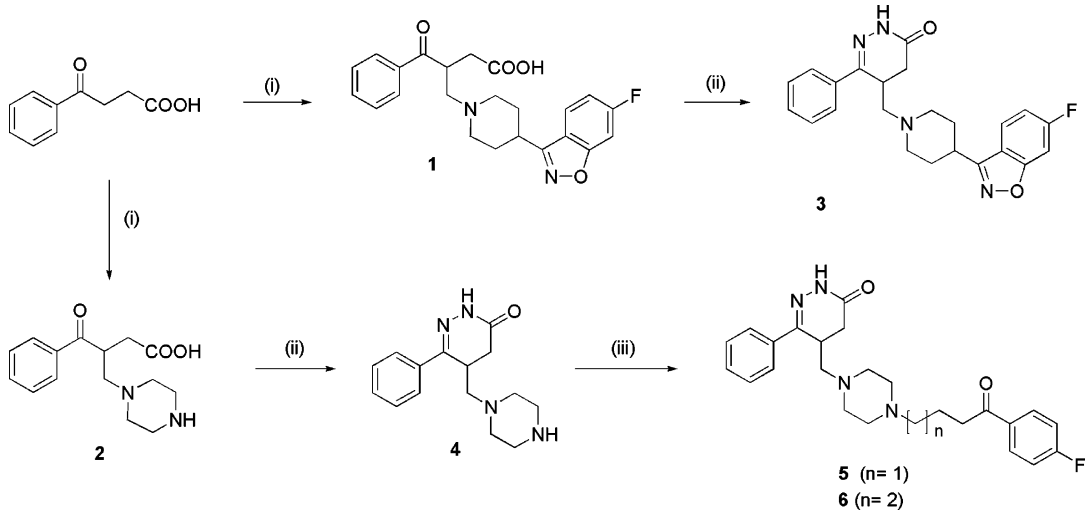
cmpd	code	alkanone fragment[a]	amine fragment[b]	pK_i 5-HT$_{2A}$ rat	pK_i 5-HT$_{2A}$ human	ref
84	QF2001B	XX	l	<5		44
85	QF2002B	XX	m	<5		44
86	QF2003B	XX	e	8.04 ± 0.88		44
87	QF2004B	XX	b	8.8 ± 0.88	8.93 ± 0.15	44
88	QF2006B	XX	d	6.20 ± 0.01		44
89	QF2014B	XXI	m	<5		44
90	QF3000B	XXII	d	<5		45
91	QF3004B	XXII	e	7.58 ± 0.16		45
92	QF3008B	XXII	b	8.86 ± 0.07		45

[a] See Chart 1. [b] See Chart 2.

Chart 1. Alkanone Fragments Present in Compounds of the Series



Scheme 1[a]

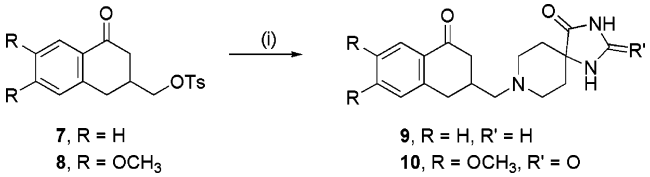


[a] Reagents and conditions: (i) 37% CH_2O, HNRR, rt; (ii) $N_2H_4 \cdot H_2O$, EtOH, reflux; (iii) for **5**, (a) 4-chloro-1,1-ethylenedioxy-1-(*p*-fluorophenyl)butane, Na_2CO_3, KI, DMF, reflux; (b) 2 N HCl, reflux, and then rt; for **6**, 5-bromo-1-(*p*-fluorophenyl)-1-pentanone, Na_2CO_3, KI, DMF, reflux.

ment, as the most important results[30−33] (Table 3) can be easily justified by the proximity of the ligand to the residues.

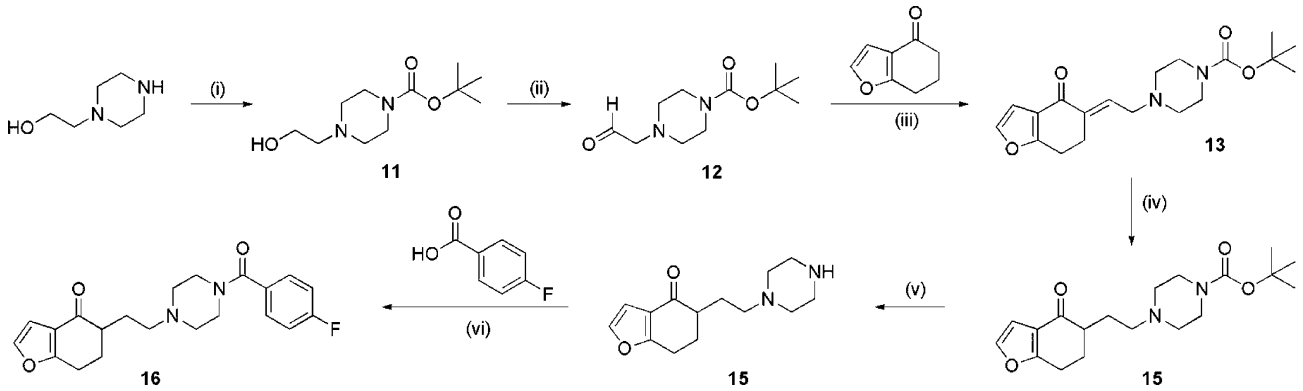
The observation of the complex (Figure 2) shows that the side chain of Asp3.32 and the charged N of ketanserin interact making a charge-reinforced hydrogen bond. The phenyl ring of the Phe6.51 side chain adopts a conformation parallel to the piperidine ring of ketanserin, where the aliphatic carbons directly linked to the charged ketanserin nitrogen (holding a positive partial charge) can make favorable electrostatic interactions with

Scheme 2[a]



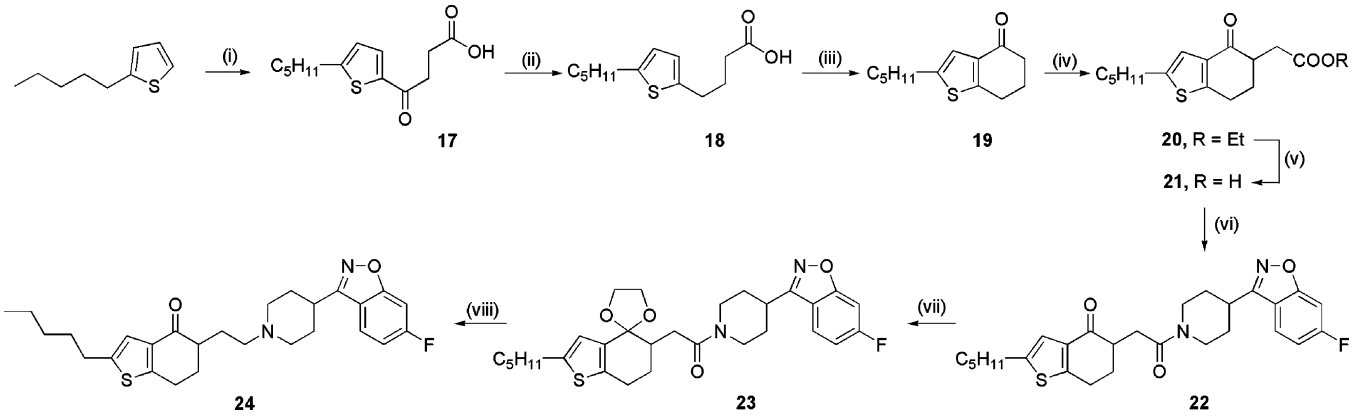
[a] Reagents and conditions: (i) 1,3,8-triazaspiro[4.5]decane-4-one (for **9**) or 1,3,8-triazaspiro[4.5]decane-2,4-dione (for **10**), Et_3N, IK, DMF, 85 °C.

Scheme 3[a]



[a] Reagents and conditions: (i) $[(CH_3)_3COCO]_2O$, $CHCl_3$, rt; (ii) DMSO, oxalyl chloride, Et_3N, CH_2Cl_2, −60 °C; (iii) *n*-BuLi, diisopropylamine, HMPA, THF, −78 °C; (iv) H_2, Pd−C, THF, rt; (v) TFA, CH_2Cl_2, rt; (vi) DCC, HOBt, CH_2Cl_2, rt.

Scheme 4[a]



[a] Reagents and conditions: (i) succinic anhydride, $AlCl_3$, CH_2Cl_2, rt; (ii) Zn, $HgCl_2$, HCl, H_2O, toluene, AcOH, reflux; (iii) TFA, $(CF_3CO)_2O$, rt; (iv) *n*-BuLi, diisopropylamine, $BrCH_2CO_2Et$, THF, −70 °C; (v) LiOH, THF, H_2O, rt; (vi) 4-(6-fluorobenzisoxazol-3-yl)piperidine, DCC, HOBt, CH_2Cl_2, rt; (vii) ethylene glycol, *p*-TsOH, toluene, reflux; (viii) (a) $LiAlH_4$, Et_2O, rt, (b) HCl, Et_2O, reflux.

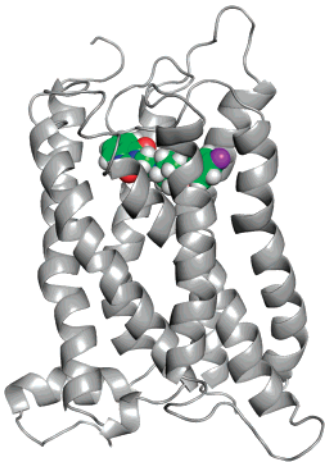


Figure 1. Best candidate structure of the 5-HT_{2A} receptor model produced by MODELLER and used for subsequent analysis. To illustrate the location of the binding site, the graphic shows a molecule of ketanserin docked according to the method explained in the text.

the phenyl ring of the Phe6.51 side chain. The result is a kind of electrostatic "sandwich", in which the ketanserin makes favorable interactions with Asp3.32 on one side and with Phe6.51 on the opposite side. The other oxygen atom of the Asp3.32 side chain is at hydrogen-bonding distance of one of the carbonyl oxygen atoms of the ketanserin benzouracil moiety, which seems to be also involved in an additional hydrogen bond with the indolic nitrogen of Trp3.28. The benzouracil moiety seem to participate in two additional polar interactions; one between the remaining carbonyl oxygen and the hydroxyl of Ser161, the other between one nitrogen and the Tyr7.43 phenolic hydroxyl group. This last interaction seems to stabilize the position of Tyr7.43 in an orientation in which it can establish an additional hydrogen bond with Ser2.61.

With respect to the *p*-fluorobenzoyl moiety, the carbonyl oxygen makes a hydrogen bond with the side chain of Ser3.36. Deeper inside the binding pocket, the *p*-fluorobenzyl is immersed in a hydrophobic pocket involving residues Val5.39 and Leu163. The size of this hydrophobic pocket is delimited by Phe6.52, a residue that is part of the so-called hydrophobic toggle switch, a set of three residues (Phe6.51, Phe6.52, and Trp6.48), which were known to participate in a coordinated conformational change, potentially important for the activation process.[34−37] Apart from this role, the conformational changes of these three residues are very important for determining the size of the aforementioned hydrophobic pocket, which in turn conditions the orientation of the *p*-fluorobenzoyl moiety. To conclude, the fluorine substituent of the aromatic ring is located close to the Ser5.43 and Ser5.46, which are residues that have been postulated to be relevant for the serotonin binding.[38,39]

Docking. The series subject of this study (Table 1) contains previously published conformationally constrained butyrophenones,[7,8,23,40−48] as well as the seven new, unpublished compounds (**3**, **5**, **6**, **9**, **10**, **16**, and **24**), the preparation of which was described above. The affinities of these 76 compounds for the 5-HT_{2A} receptor have been determined experimentally and

Chart 2. NRR of Structures in Chart 1

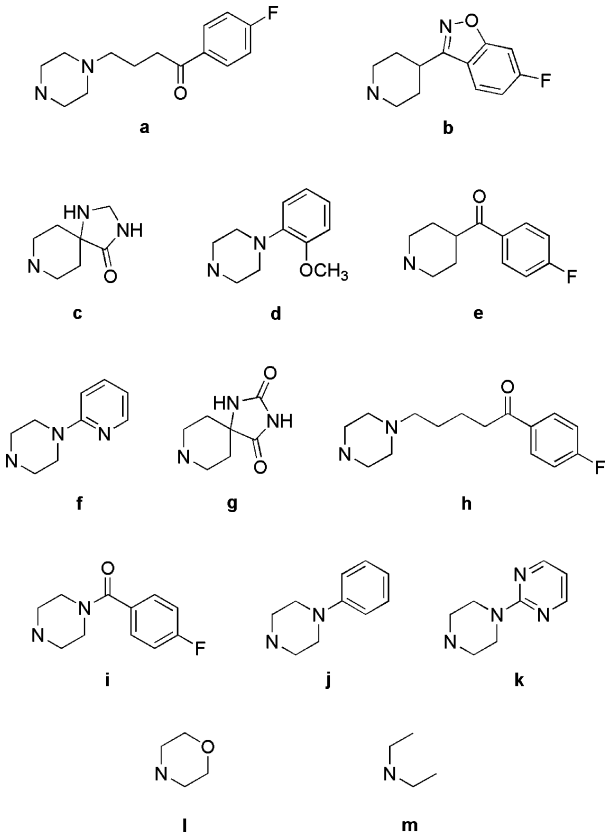
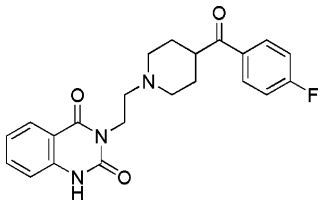


Chart 3. Structure of Ketanserin



were listed in Table 1, expressed as pK_i values. For QSAR analysis purposes, inactive compounds were assigned an arbitrary pK_i value of 5. In this table, the values range from 5 to 9.14, covering more than four logarithmic units, and the experimental errors (expressed as standard error of the mean) fall typically under 0.5 logarithmic units.

Most of the compounds of this series share relevant structural features with ketanserin, like the presence of a piperidine ring, the linear shape with aromatic rings at both ends, and hydrogen-bond acceptor and donor groups at an intermediate position. This similarity justifies the use of the above-described ketanserin-conditioned binding pocket to carry out docking simulations in this series. The objective was to obtain complexes that highlight the more relevant features that determine the binding affinity, but without carrying out extensive MD simulations. With this aim, every compound of the series was docked into the modeled receptor binding site using program GOLD 2.2.[49] Unlike ketanserin, all the compounds in this series contain a chiral center and, therefore, either their *R* or the *S* enantiomers can be used for the docking. Both enantiomers were built in 2D and converted to 3D structures using CORINA 2.4.[50] The structures were all modeled with a formal positive charge centered on the piperidine ring nitrogen. For the compounds with a piperazine ring, the nitrogen linked to an aliphatic substituent was preferred for protonation, and when both nitrogen atoms were linked to aliphatic carbons, the one linked to the alkanone fragment was chosen for consistency with other members of the series, even if a tautomeric equilibrium between species protonated in both positions can be expected. At the end, 152 different structures were obtained and submitted to the docking protocol (see Experimental Section for details).

The method was able to produce docked structures with a fitting score (GOLDscore)[51] above a reasonable prefixed cutoff (35) for most of the compounds in the series (142 out of 152, 93%). Figure 3 shows, superimposed, the docked structures for the *R* and for the *S* series. In the figure, the observation of both series clearly shows both enantiomers bind in approximately the same position, but some orientations are more frequent in one series than in the other. For example, in the *R* series, many compounds place the carbonyl group on the lower right side, while in the *S* series, the fluorine substituent of the *p*-fluorobenzoyl moiety is much more frequent on the upper right side.

The analysis of the docking results revealed that most structures can bind in two alternative orientations, one with the alkanone group oriented toward the inner part of the binding pocket and another orienting this group in the opposite direction. In both orientations, the charged nitrogen can bind the Asp3.32 and the polar groups can make approximately equivalent contacts. The presence of these two alternative binding modes was also identified in a previous work[7] in which a much earlier version of the present series, containing only 25 compounds, was docked into the binding site of a 5-HT_{2A} receptor model, even if the receptor modeling and the ligand docking followed a completely different methodology. The authors of this work justified this duplicity due to the presence of the same pharmacophoric groups at both sides of the protonated amino group and suggested that this finding "hints the possibility of multiple binding modes".

The preference of the structures for binding in a certain orientation follows no obvious rules. In some structures (like *R*-**38**, *R*-**61**, or *S*-**62**), both orientations were present among the three best solutions provided by GOLD, thus showing no strong preference for any orientation. Other compounds seem to bind only in one orientation, but the introduction of a small structural change induces a preference for the other alternative. This is the case for *R*-**72** and *R*-**75**, which were inserted in opposite directions even if their structures differ only in the substitution of the furan ring by a thiophene (Figure 3). For some compounds, like **75**, the *R*- and *S*-enantiomers show a preference for binding in opposite orientations (Figure 3). A particularly interesting case are compounds **69** and **24**. In the docking simulation, both **69** enantiomers are found to bind with the alkanone oriented toward the inner part of the binding pocket. Compound **24**, a derivative of **69** prepared ad hoc to prevent the original binding mode (by adding a bulky pentane substituent to the alkanone), was observed to dock in the opposite orientation (Figure 3). The experimentally measured binding affinity of **24** was only slightly lower than the binding affinity of its parent (8.56 for **69** and 7.68 for **24**), thus suggesting that for compound **69** both binding orientations would also be accessible and contribute to the experimental binding affinity, something that probably applies to most of the compounds in this series.

In summary, these results seem to confirm the coexistence of ligand molecules bound in two alternative orientations, the proportion of which might be different for different structures, considering also as such both enantiomeric forms of the same compound. With all probability, none of the experimentally

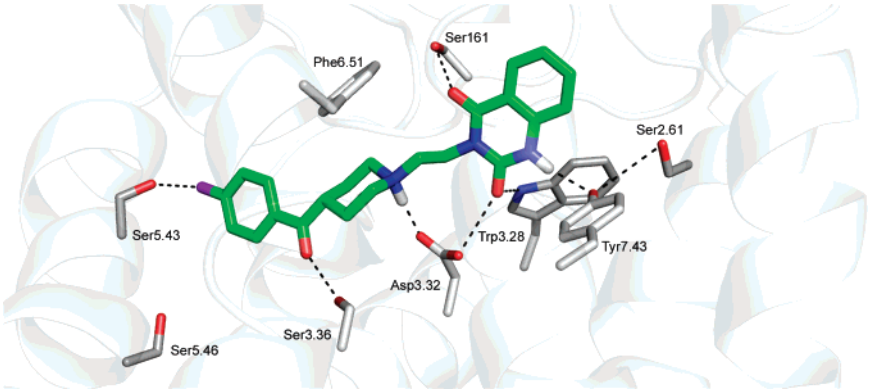


Figure 2. Binding site of the 5-HT_{2A} receptor complexed with a molecule of ketanserin. Only residues establishing relevant interactions are depicted. The dashed lines represent hydrogen bonds and polar interactions described in the text.

Table 2. Percentage of Structures with a Geometry Compatible with a Hydrogen Bond between the Ketanserin and the Listed Residues[a]

residue	% hydrogen bond
Asp3.32	100
Ser3.36	21
Ser161	27
Trp3.28	76
Tyr7.43	18

[a] As observed during the 1 ns MD simulation described in the text.

Table 3. Residues Involved in the Binding of Ketanserin, According to Mutagenesis Experiments, and the Variation of K_i Associated to Each Mutation

mutation	K_i variation > (fold)	interactions observed in complex
Trp6.48Ala	100	part of the aromatic cluster, delimiting binding site
Asp3.32Asn	75	charge-reinforced hydrogen bond with charged piperidine nitrogen atom
Phe6.52Tyr	73	near the *p*-flurobenzyl moiety
Tyr7.43Ala	20	hydrogen bonded to the benzouracil moiety
Phe6.51Ala	12	Van der Waals contact with piperidine ring
Asp2.50Asn	10	
Trp1.34Ala	10	

determined binding affinity constants listed in Table 1 represent the affinity of a single binding mode, and in all instances, they are a weighted average of the affinity values of the different species present in the solution (corresponding to the two alternative binding modes in different proportions for the two enantiomers).

Apart from these mainly descriptive observations, our main interest in the study was to identify in the complex structures interaction patterns that can be correlated with the activity (e.g., polar interactions present in the most active compounds and absent in the less active), but the preliminary inspection of the complexes did not reveal any evident relationship. The direct analysis of all these complexes would be cumbersome and prone to introduce subjectivity in the observations. Another problem associated with the use of a "rigid-receptor" docking method that does not take into account the flexibility of the binding site (GOLD), is that the complex structures often contain small errors and inaccuracies, the observation of which can produce misleading interpretations.

For all these reasons, we decided to use 3D-QSAR methods. The 3D-QSAR models can highlight the structural differences that correlate with binding affinity differences, thus summarizing the information contained in many complexes into a few interpretable graphic plots. Moreover, the procedure would be less influenced by the aforementioned small "random" docking errors than the individual complex inspection, because 3D-QSAR models highlight only correlations that are present in several compounds.

3D-QSAR. The output structures obtained from the docking are an excellent material for building a 3D-QSAR model, because they provide a more realistic representation of the ligand bioactive conformation than any alignment obtained by structural superimposition. When proceeding this way it must be remembered that the spatial framework for the interpretation of the resulting models is not the cluster of superimposed compounds, but the structure of the receptor binding site, because this structure was used to place the ligands within a consistent reference frame.

In the present study, the biological activity values were obtained from binding affinity measurements carried out (in the vast majority of cases) using the compounds racemic mixture, thus making available only one affinity value for the *R* and *S* enantiomers present in the mixture. From a biological point of view, both enantiomers are not equivalent and their binding affinity might also be different. Indeed, in the few examples in which the activity has been measured separately for the *R* and *S* enantiomers, the values obtained were slightly different (see Table 1).

Obviously, the lack of separate affinity values for both enantiomers is a problem for the QSAR modeling. A potential solution used in a previous work[7] is to select one of the enantiomers and carry out the whole 3D-QSAR procedure using only this structure. The alternative solution proposed here is to include both enantiomers in the model as separate structures, but associate them to the single binding affinity value measured for the racemic mixture. By proceeding so, the binding affinity used corresponds to neither of the individual enantiomers and can be considered an average of both enantiomers true values. The observation of the differences in binding affinity between *R* and *S* enantiomers in the few examples reported in Table 1 shows that these differences are not large (0.67, 0.17, 1.25, 0.89) but certainly not negligible. From a statistical point of view, it is important to notice that the use of both values will introduce "noise" into the model but not bias, because the true binding affinity values should deviate from the estimate in opposite directions. Such a model could not be expected to have a perfect fitting, because part of the model lack-of-fit will be produced by this "noise" and, therefore, the model r^2 value will be lower than the r^2 obtained for a regular model. From the point of view of the interpretation, a model obtained in this way could not be

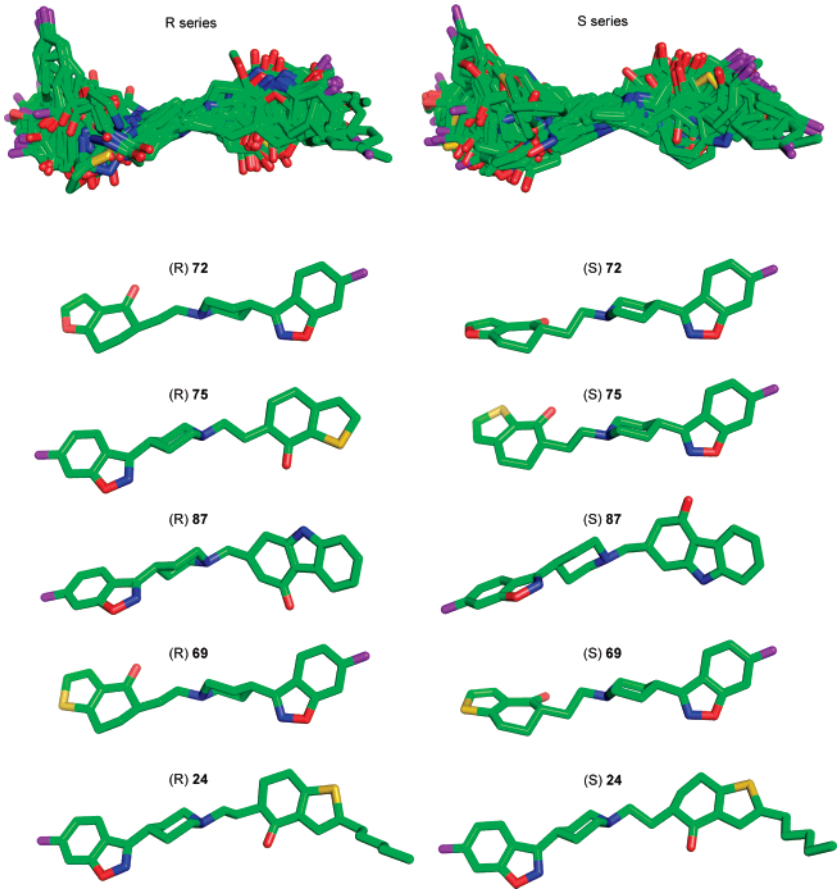


Figure 3. From top to bottom: superimposition of all the docking solutions for the *R* enantiomers (left) and the *S* enantiomers (right); side by side view of the docking solution for the *R* (left) and *S* (right) enantiomers obtained for compounds **72**, **75**, **87**, **69**, and **24**. In the case of compound **72**, both enantiomers are inserted in the same orientation, while for **75**, the *R* and *S* enantiomers are inserted in opposite directions. Notice that the only difference between **72** and **75** is the substitution of the oxygen by a sulfur in the heteroaromatic ring of the alkanone moiety. In compound **87**, both enantiomers orient the acceptor and donor groups in opposite directions. Both enantiomers of compound **69** are inserted in the same direction, but the incorporation of a pentane substituent to produce compound **24** induces a change to the alternative orientation, with no relevant change in the binding affinity.

expected to explain "fine effects", because no structural difference between the *R* and *S* complexes could be correlated with differences in their binding affinities (the same binding affinity is used for representing both).

Accordingly, the first 3D-QSAR model (M1) was built including both the *R* and *S* enantiomers for every compound. The enantiomer structures were represented by the docking solution with the highest fitting scores. Their binding affinities were represented by the experimental measurements listed in Table 1 corresponding, for the vast majority of the cases, to the affinity of the racemic mixture, but for a few exceptions in which separate values for either enantiomers were available. The GRID/GOLPE methodology[14] was used. Every structure was first analyzed using GRID[52] to obtain molecular interaction fields (MIF) with three probes, DRY, O, and N1, representing, respectively, a hydrophobic, a hydrogen bond acceptor, and a hydrogen bond donor group. These MIF were imported into the program GOLPE[53] together with the affinity values. The variables were pretreated and scaled before using the partial least-squares (PLS) regression analysis and the SRD/FFD[54] variable selection method. At the end, the PLS model obtained with three latent variables (LV) showed good statistical quality ($r^2 = 0.90$, $q^2_{LOO} = 0.78$, see Table 4 and Figure 4). To further validate the model and to estimate its true predictive ability, we have split the series into a training and a test set. The objects assigned to the test set (15 structures, approximately 10% of the whole set) were selected applying the Most Descriptive Compounds algorithm,[55] as it was implemented in the GOLPE software,[53] on the space of the first three principal components extracted from the Principal Component Analysis of the whole series. Then the whole 3D-QSAR modeling procedure was carried out as described above (PLS modeling and SRD/FFD variables selection) using only the compounds of the training set, while the compounds in the test set were used only for the predictions. In this case, the Standard Deviation of Error of Prediction (SDEP) obtained for the test set was of 0.53, and the external r^2 between experimental and predicted values was of 0.80. These values are perfectly comparable with the values obtained by standard cross-validation methods and further confirm the good predictive ability of M1.

Table 4. Statistical Parameters of the 3D-QSAR Models Obtained

model	LV	objects	var.	r^2	q^2_{LOO}	$SDEP_{LOO}$	q^2_{5RG}	q^2_{2RG}
M1	3	142	1441	0.90	0.78	0.55	0.77	0.73
M2	3	426	1579	0.84	0.81	0.52	0.80	0.79

Even if the model described above seems suitable for our purposes we decided to go one step further and apply to the problem of the multiple binding modes the same solution applied to address the presence of two enantiomers. In the same way that two enantiomers were used to characterize each compound, it is possible to use more than one docking solution to characterize each enantiomer. For those compounds in which the three docking solutions are rather similar the use of multiple structures will have very little effect (the median of the RMSD

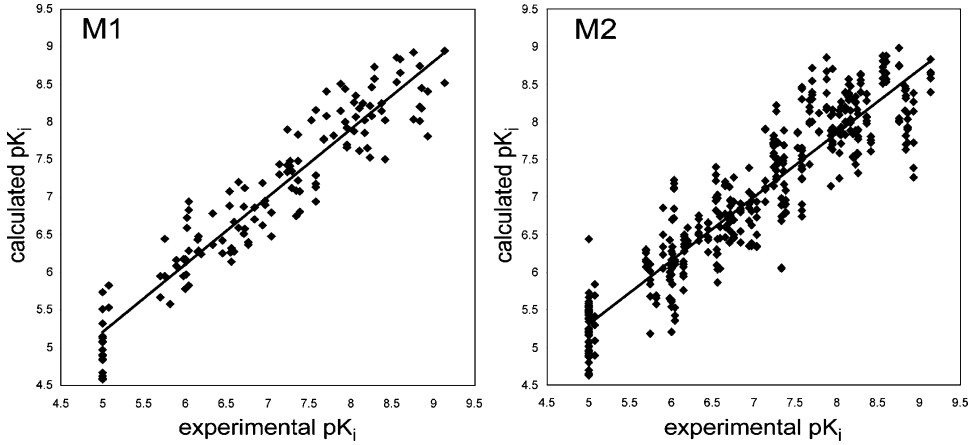


Figure 4. Scatterplots representing the experimental versus calculated pK_i values for the models M1 (left) and M2 (right).

was 0.62 for the solutions 1 and 2, 0.87 for the solutions 1−3, and 0.9 for the solutions 2 and 3), but in some cases, these solutions represent alternative binding modes. In our opinion, the choice of the first solution to represent the docking structure was arbitrary in these cases, much in the same sense that the choice of a certain enantiomer would be arbitrary, because the limited accuracy of the scoring functions question the reliability of the solution ranking provided by the docking programs. Moreover, most compounds probably interact with the receptors in both orientations and in proportions that are difficult to estimate. Therefore, the contemporary use of all these solutions will provide a much more realistic picture of their true binding ability. As in the previous case, we must be ready to accept a small decrease in the model fitting, lower r^2 values, and a less detailed description of fine effects, but hopefully, the models will be more robust and have better predictive ability, as it has been reported in other similar situations.[19]

A second 3D-QSAR model (M2) was built using a similar methodology, but in this case, including the first three solutions provided by GOLD to characterize every enantiomer (six structures per compound) amounted to a total of 426 structures. At the end of the modeling procedure (see Experimental Section for details), we obtained a PLS model with three LV of remarkable quality ($r^2 = 0.84$ and a $q^2_{LOO} = 0.81$, Table 4, Figure 4).

From a statistical point of view, the r^2 of M2 is lower than M1, as can be expected for the aforementioned reasons, but only slightly (from 0.90 to 0.84), and the difference is small if we consider that the number of objects included in the model has been increased from 142 to 426. With respect to the values of q^2, they should be interpreted with care. In this particular series, the application of leave-one-out (LOO) method can produce overoptimistic results, because the removal of a single structure will not completely remove the presence of the compound from the reduced models. Therefore, it is not surprising to obtain a higher q^2_{LOO} of 0.81 for M2 than the 0.78 value obtained for M1. A better estimation of the true predictive ability of M1 and M2 can be obtained using more strict cross-validation utilizing either two or five randomly assigned groups (see Experimental Section). The results listed in Table 4 agree with the previous results, indicating that both models have a rather high predictive ability and that the incorporation of multiple structures does not decrease the predictive ability of the original model. To further confirm the predictive ability of M2 and to compare it with M1, an external validation protocol identical to the one described for M1 was carried out. In this case, the test set included 45 compounds, corresponding to the three best docking solutions obtained for the 15 structures of the M1 test set. As in the previous case, the structures in the test set were left out of the analysis, which was repeated from the beginning using only the compounds in the training set. To obtain comparable results with M1, the predictions obtained for the three docking solutions used to characterize each compound in the test set were averaged to provide a single estimate. The results of such predictions were good (SDEP = 0.52, $r^2 = 0.81$) and slightly better than those obtained with M1, thus confirming the improved predictive ability of M2 over M1, as suspected.

The comparison of the experimental versus calculated plots for M1 and M2 represented in Figure 4 shows that in M2 the objects exhibit a larger spread, and as it was expected, not all structures fit well into the model. The coefficient plots obtained for M1 and M2 look rather similar and contain essentially the same information. The main difference observed is that the plots for M2 are "cleaner", devoid of the small regions produced by the effect of single compounds that usually have no general relevance. All these results suggest that the model M2, obtained with multiple structures, can be a better alternative than M1, but the similarity observed between their coefficient plots indicate that the results of the M1 and M2 interpretation would be qualitatively similar.

Before beginning the interpretation of the M2 coefficient plots, we must remind the reader that the MIF assign negative energy values to favorable probe−ligand interactions and positive values to unfavorable (repulsive) probe−ligand interactions. Positive field values represent mainly the molecular shape, while negative values represent regions where the ligand can make energetically favorable interactions with the binding site. The PLS coefficient plots shown in Figure 5 contain both positive and negative values for the three probes used. Positive (yellow) coefficients might correspond to binding site regions where the presence of negative fields have an inverse correlation with activity or to regions where the presence of the molecular bulk (positive field) correlates directly with the activity. Conversely, negative (cyan) coefficients might represent regions where the presence of negative fields correlate directly with the activity or regions where the ligand positive fields correlate inversely with the activity. These considerations add complexity to the interpretation and require a careful analysis of the coefficients, the field values produced for different compounds, and the activity contribution plots. It should also be stressed that the coefficient positions will make reference to the binding site residues (even if the binding site has not been used directly for obtaining the model), because the ligand alignment was based on the docking solutions. Therefore, the reference framework of the whole analysis is the structure of the binding site instead of the superimposed ligand structures.

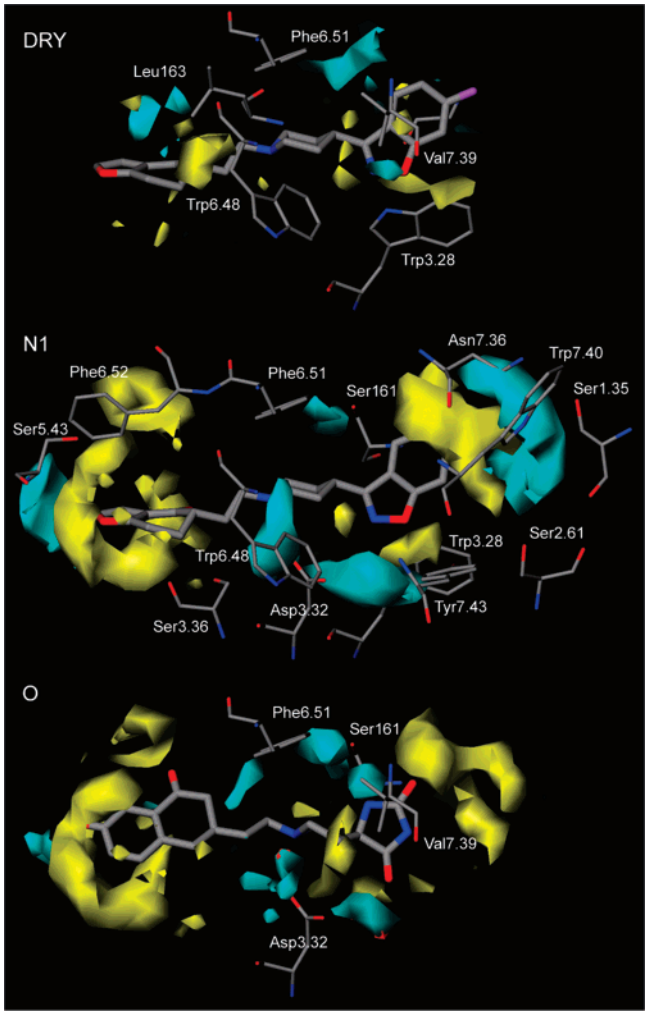


Figure 5. Coefficient plots obtained in model M2 for different blocks of variables representing (from top to bottom) the probes DRY, N1, and O. Positive coefficient values are represented in yellow and negative coefficient values are represented in cyan. Every coefficient plot included the structure of a representative ligand (**72** for DRY and N1 and **40** for O), as well as the structures of the more relevant residues, as mentioned in the text. Even if the structure of the binding site was not used for obtaining the 3D-QSAR model, the coefficients represented should reflect ligand−receptor interactions, which are relevant for the ligand binding affinity.

DRY Probe (Figure 5, Top). The highest negative coefficients were located in two areas, one at the bottom of the pocket near Leu163 and the other in the center near the side chains of Val7.39 and Phe6.51. The values of the coefficients indicate that the presence of hydrophobic groups in these positions increases the binding affinity, which is consistent with the hydrophobic nature of the neighbor residues. The main positive region obtained for this probe is located near the main chain of Trp6.48, indicating that the presence in this region of hydrophobic groups is detrimental for the binding. A similar interpretation can be given to a small positive region located near to the indolic nitrogen of Trp3.28 or to the polar atoms of the main chain near to the Gly160.

N1 Probe (Figure 5, Middle). Not surprisingly, most negative regions overlap some of the residues able to act as hydrogen bond donors. This is the case of a large region in the center that extends between the oxygen atoms of Asp3.32 and the indolic nitrogen of Trp3.28. Also in the center, there is a smaller region near the side chain of Ser161. Apart from these, there are two large negative regions in both extremes of the pocket. One is located at the bottom near the Ser5.43 and presumably indicates that ligands able to make a favorable interaction with this residue have higher binding constants. On the other side of the receptor there is a much wider area enclosing the side chain of Asn7.36 and close to other polar side chains (Ser1.35, Ser2.61, Trp7.40). Even if the regions do not overlap these residues, we must remember that the structure represented here is only a static snapshot of the receptor, and the above-mentioned residues could move their side chains to establish interactions at these positions. This can also be the case for an elongated region in the center of the pocket near to the Ser3.36 and Trp6.48. In the conformation shown in Figure 5, the polar part of the Trp6.48 side chain could not overlap this region, but this residue forms part of the toggle switch that has been suggested to participate in a concerted conformational change together with Phe6.51 and Phe6.52. During this conformational change the Trp6.48 rotates, orienting the polar nitrogen toward the ligand and locating this atom precisely in the region highlighted by the negative coefficients. With respect to the positive coefficients, most of them describe binding site regions that are overlapped by compounds with high binding affinity. The two main positive regions are located on the extremes of the binding site, indicating that the compounds with the highest affinity are long and fill the whole binding pocket, while shorter compounds usually show low affinity.

O Probe (Figure 5, Bottom). The negative regions observed for this probe are smaller and less defined than those observed for N1, probably because the series contains much fewer hydrogen bond donor groups. The most important regions are located near the oxygen atoms of Asp3.32 and have a straightforward interpretation. The negative regions located on the central part of the binding pocket, near the hydrophobic side chains of Ser161, Phe6.51, and Val7.39, seem to represent the interactions of some compounds that contain a spiranic hydantoin ring (**g** in Chart 2), like **40** shown in Figure 5. In these compounds, one of the hydrogen bond donor atoms of **g** generates two regions for potential hydrogen bond interactions, one near the side chain of Ser161 and the other oriented toward the Val7.39, both of which are shown in the coefficient plots even if only one of them actually represents a plausible interaction. The positive regions are rather similar to the positive regions observed for the N1 probe and can be interpreted in a similar way.

To summarize, the model stresses the importance of some of the polar interaction already recognized in the description of the ketanserin complex, like those with Asp3.32, Trp3.28, Ser161, and Ser 3.36, but adding some original information: (i) highlights some biologically relevant interactions with hydrogen bond donor residues in both extremes of the pocket (Ser5.43 on one side, Ser1.35 and/or Asn7.36 on the other side) and a putative interaction with the indolic nitrogen of Trp6.48 when its side chain adopts an alternative conformation; (ii) shows the importance of the hydrophobic interactions with residues located at the bottom of the pocket (Leu163) as well as in the center (Phe6.51, Val7.39); (iii) stresses the importance of the ligand size, indicating that ligands with high affinity tend to completely fill the cavity; and (iv) indicates that the few compounds in the series with hydrogen bond donor groups located in the alkanone moiety can make interactions that lead to an increase in the binding affinity, probably with Ser161.

These results are in agreement with others obtained in our group in a much earlier work[7] in which a CoMFA model[56] was obtained for a subset of 24 compounds of the present series. In this work, the model was obtained using only one enantiomer,

one binding orientation, and carrying out a classical ligand superimposition based on a few pharmacophoric points. The quality of the best model obtained was not better than that of M1 or M2 ($r^2 = 0.88$ and $q^2 = 0.64$), and the coefficient plots are far less detailed than the ones shown in Figure 5, confirming that the methodology presented here leads to results more useful than those obtained with classic methods. Remarkably, the interpretation of the previous model shows some points of coincidence with the interpretation of M2, as it also points out the importance of the hydrophobic regions in the center and in the extremes of the ligand, the importance of the hydrophilic regions at both ends, and hints the importance of the right ligand size.

Conclusions

The work presented here is a good example of the practical limitations of well-established computational methods. The analysis of the docked structures was not successful in uncovering relevant relationships between the ligand structures and their binding affinities, in part due to the large number of complexes and in part due to the duality of the binding orientations obtained. Then the application of 3D-QSAR methodologies was envisaged as a potential solution, but its application was also complicated by the presence of two enantiomers for each ligand as well as by their binding diversity. One of the goals of this article is to show how we have confronted these practical problems, with the hope that other researchers can take advantage of the proposed solutions.

The 3D-QSAR is a powerful methodology, and techniques like the one presented here can widen its range of application. It is noteworthy that our efforts were directed to building a representation of the phenomena under study as realistic as possible, even if we were aware that such models will have a poorer fitting and lower statistical indexes than those obtained with the standard methodologies. In our opinion, these "coarse grain" QSAR models are more valuable, because they should be able to identify more consistent relationships and provide better predictions than standard "fine grain" methods. Indeed, one of the reasons why the QSAR methodologies are being so heavily criticized lately[57] is probably the tendency to obtain "too good" models with very high fitting in situations in which experimental errors and methodological limitations make such expectations unrealistic.

In our series, the simultaneous use of several receptor-docked structures representing the same compound produces an interesting model that compares favorably, from a statistical point of view, with models obtained using classical methods, while being presumably more robust and having more general validity. The model is interpretable in terms similar to those used in structure-based drug design studies, allowing to extract information that can be easily translated as guidelines for the design of novel compounds. Moreover, estimation of the models predictive ability based on strict cross-validation methods and in external datasets suggest that the predictions for structurally related compounds will fall within an interval of ± 1.0. Even if this cannot be considered a very accurate estimation, it is more than enough for identifying potentially active or inactive compounds. Indeed, preliminary results obtained for new, unpublished compounds confirm the robustness and the reliability of our approach.

Experimental Section

Numbering of Residues. For residues belonging to helix regions, the generalized numbering scheme proposed by Ballesteros and Weinstein[58] was used through the text. According to this scheme, the most conserved residue in each helix was given the number 50, so for every helix, the most conserved residue was indicated with the number of the helix followed by 50. For residues belonging to loop regions, the sequence number was used instead.

GPCR Modeling. (i) Sequence Alignment and Prediction of Secondary Structure and Transmembrane Helices. The sequence of 58 serotonin receptors cloned so far from different species (excluding the subtype 3, because they are not GPCR) were retrieved from the Swiss-Prot database[59] and aligned with the ClustalX software,[60,61] using the PAM250 matrix. "Gap open" and "gap elongation" penalties of 10 and 0.05, respectively, were used in the alignment. The resulting multiple sequence alignment was realigned with the sequence of Rh (OPSD_BOVIN), introducing secondary structure information derived from the crystal structure (residues belonging to the transmembrane alpha helices) to avoid gaps within the seven helical segments.

The alignment was then manually refined to ensure a perfect alignment of the highly conserved residues of GPCR superfamily, according to Baldwin et al.,[62] and looking for a consensus between the predictions of transmembrane segments provided by TMAP,[63] Swiss-Prot (http://www.expasy.ch), and TMPRED (http://www.ch.embnet.org), as well as by considering that Arg and Lys residues are often present at the membrane boundaries.[64] Alternative lengths for the helices were considered by taking into account the experimental length of the Rh helices, the secondary structure prediction by JPRED,[65] and the sequence conservation.

(ii) Structural Alignment, Disulfide Bond Assignment, and Model Building. To enrich the structural information given by the Rh template with additional information, other membrane protein structures stored in the "Membrane Proteins of Known 3D-Structure" database[66] were also considered, but with the condition of being proteins sharing the GPCR folding (seven transmembrane helices) and displaying a minimum homology with subtype 2 serotonin receptors of 20%. Only bacteriorhodopsin (BR) satisfied both requirements, even if it is not a GPCR itself and presents major structural differences between GPCR and BR.[34] For this reason, the use of the BR structure was limited to individual helices and helices pairs to avoid introducing biases due to the different folding of BR. Therefore, BR helices were separated into pairs and superimposed to Rh using program STAMP.[67] In addition and for the same reasons, ideal alpha helices built de novo using the sequence of human 5-HT_{2A} were also superimposed to the Rh structure and added to the templates library.

Regarding the modeling of the loops, suitable templates were searched in the ArchDB loops database[68] to enrich the structural information provided by Rho. Template candidates must link a couple of alpha helices, should have a similar number of residues forming the loop (± 2 residues), and the distance (between alpha helices extremes) that braces the loops should be similar to those found in rhodopsin (± 5 Å). The modeling of the EL2 loop was largely simplified by the presence of two important structural constraints. First, the Cys163 and the Cys3.25 form a disulfide bond, a feature highly common among GPCR receptors, and the secondary structure prediction analysis suggested that a large part of the loop adopts a β-sheet conformation. For the rest of the loops, suitable templates were found and their structures were superimposed with STAMP to the crystal structure of Rh to reach a consensus between all the structures used.

3D models were then built using the MODELLER suite of programs,[27] which yielded 25 candidate models for the 5-HT_{2A} receptor final structure. From these candidates, the best structures according to the MODELLER objective function and to visual inspection were selected.

(iii) Assessing the Quality of the Structures. Models with interruptions or gaps in the transmembrane regions, as identified by visual inspection, were discarded. PROCHECK software[69] was used to assess the quality of the backbone dihedrals, resulting in good quality parameters, with an excellent distribution of φ and ψ angles in the Ramachandran plot. Also, the resulting models must reproduce the correct orientation of the side chains for the set of highly conserved amino acids in the GPCR superfamily.[34−37] The

conformation of these residues was set to the "inactive state", which probably is more appropriate for modeling the docking of antagonists and more consistent with the inactive state of the main template structure (Rh).

Docking Simulation of Ketanserin. The docking of ketanserin was carried out using the program GOLD.[49] The structure of ketanserin was docked in the active site of the 5-HT_{2A} receptor, forcing the interaction between ketanserin and the Asp3.32 side chain with a harmonic force restraint. The rest of the parameters of the program were set to their default values. The best docking solution, according to the scoring function of GOLD, was then refined using energy minimization (EM) and MD simulations, as explained in the following paragraphs.

Energy Minimization (EM). The best docking solution was submitted to EM with the SANDER classic module of AMBER6 suite of programs[70] by applying 500 cycles of steepest descent minimization, followed by conjugate gradient minimization until the RMS gradient was lower than 0.001 kcal/mol·Å. At the beginning, only the positions of the hydrogen atoms were optimized and then the whole complex (ligand and protein) was allowed to relax. In all the minimizations, a harmonic constraint of 5 kcal/mol·$Å^2$ was imposed to the Cα trace, in order to maintain the overall seven helices architecture. All the calculations were carried out using the all atom force field of Cornell et al.,[71] as implemented in AMBER6, with a distance-dependent dielectric constant of 4*r* and a cutoff for nonbonded interactions of 10 Å. The charges for the ligand atoms were calculated using Gaussian 98,[72] with a 6-31G** basis set, and then the RESP procedure was applied as described in literature.[73] A formal charge of +1 was attributed to the ligand, because the piperidine nitrogen of ketanserin is assumed to be protonated at physiological pH, and the positively charged nitrogen of the ligand is assumed to bind the negatively charged side chain of Asp3.32.

Molecular Dynamics (MD). The minimized structure obtained from the previous step was submitted to MD simulations using the SANDER classic module of AMBER6 suite of programs during 1 ns at 300 °K. The time step was set to 2 fs, a value that is justified by the application of the SHAKE algorithm.[74] The nonbonded pair interaction list was updated every 25 fs. For the aforementioned reasons, we kept the harmonic force constant of 5 kcal/mol·$Å^2$ to the Cα trace, but the EL2 loop was left without restrictions. The rest of the parameters were set to their default values. The MD trajectories were analyzed with the CARNAL module of AMBER, monitoring the formation of H-bonds, the profile for the total energy, the temperature, as well as the RMS deviations of side chains. The frames with the lowest total energy were selected, minimized, and analyzed to detect the interaction between the ligand and the protein. The best complex structure (with lower energy) was minimized and used as the starting point for the ligand docking studies.

Ligand Structures. The structure of all the compounds in Table 1 were modeled in 2D in their protonated state (charge +1) and then converted to 3D using program CORINA v2.4.[50] Two enantiomers (*R* and *S*) were obtained for every compound.

Ligand Docking. The ligands were docked within the modeled binding site using the program GOLD.[49] For every compound, the program made 10 docking trials, and only the three best solutions were retained. The fitness of the solutions were assessed using the GOLDscore function.[51] Solutions with a score under a prefixed value of 35 were discarded. To obtain consistent orientations, a distance constraint between the charged nitrogen of the ligand and the Asp3.32 was defined. To compensate the binding site rigidity, up to three bump contacts were allowed. As a test, this protocol was used to dock the structure of the ketanserin, obtaining a solution that showed a very low RMSD (0.58 Å) with respect to the MD solution.

Pharmacology. For most of the compounds, the binding was measured using rat 5-HT_{2A} receptors, except in a few for which the affinity was measured in human receptors (see Table 1). Radioligand binding assays in both human and rat were performed by labeling 5-HT_{2A} receptors with [^{3}H]ketanserin, as previously described.[8] In the compounds for which both values were available, the human values were used, but in general, both values do not show large differences (under 5% in all instances).

GRID/GOLPE Analysis. The structures corresponding to the best docking solution (M1) or to the three best solutions (M2) were imported into GRID[52] where MIF were computed for all of them using three probes (DRY, O, N1). All the MIF were computed using a grid spacing of 1 Å. The analysis was made in a box of 22 × 36 × 25 nodes, containing 19 800 energy measures per structure. The resulting MIF were then imported into program GOLPE 4.6.0,[53] and the **X** matrix of variables obtained was pretreated, first applying a zeroing of very small values (under 0.01), removing variables with small standard deviation (under 0.05), and ill-conditioned variables that take only two or three different values, one of which is assigned to a single compound. Then a Block Unscaled Weight scaling was applied to the whole matrix to equalize the importance of the different blocks. To apply regression methods, the pK_i values listed in Table 1 were imported as the **Y** block. A first PLS model was obtained and then the SRD/FFD variable selection method[54] was applied, using as parameters for the SRD method 1000 seeds, filtering with the PLS weights obtained for a 2 LV model. Two consecutive FFD variable selection iterations were consecutively applied using the groups obtained in the SRD procedure. For the selection, LOO cross-validation and 2 LV were used. The statistical parameters for the M1 and M2 models are shown in Table 4. The q^2 in this table were obtained using the classical LOO method and two more strict methods (called 5RG and 2RG) consisting of splitting the series randomly into either two (2RG) or five groups (5RG), one of which is removed and predicted in turn until every group has been removed once. This whole procedure is then repeated either 20 (for the 5RG method) or 200 times (for the 2RG method). As can be seen in Table 4, the q^2 values obtained with the most strict methods are rather similar to the LOO values.

Hardware. All the computations were carried out in Linux workstations with Intel Pentium4 and Xeon CPUs. The visualization of structures and complexes were made with VMD[75] and our own in-house developed software.

Acknowledgment. The authors acknowledge Nuria B. Centeno, Hugo Gutiérrez-de-Terán, and Fabien Fontaine for their valuable scientific comments. The present work has been supported in part by CICYT (SAF2002-04195-C03 and SAF2005-08025-C03) and "Fundació La Marató de TV3" grants. A Ph.D. fellowship (FI) has been granted to C.D. by the Generalitat de Catalunya.

Supporting Information Available: Experimental procedures, spectroscopic data, and elemental analyses for compounds **1**−**6** and **9**−**24**. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Horacek, J. Novel antipsychotics and extrapyramidal side effects. Theory and reality. *Pharmacopsychiatry* **2000**, *33* (Suppl 1), 34−42.
(2) Aghajanian, G. K.; Marek, G. J. Serotonin model of schizophrenia: Emerging role of glutamate mechanisms. *Brain Res. Rev.* **2000**, *31*, 302−312.
(3) Reynolds, G. P. Receptor mechanisms in the treatment of schizophrenia. *J. Psychopharmacol.* **2004**, *18*, 340−345.
(4) Meltzer, H. Y.; Matsubara, S.; Lee, J. C. Classification of typical and atypical antipsychotic drugs on the basis of dopamine D-1, D-2, and serotonin2 pK_i values. *J. Pharmacol. Exp. Ther.* **1989**, *251*, 238−246.
(5) Meltzer, H. Y.; Li, Z.; Kaneda, Y.; Ichikawa, J. Serotonin receptors: Their key role in drugs to treat schizophrenia. *Prog. Neuropsychopharmacol. Biol. Psychiatry* **2003**, *27*, 1159−1172.
(6) Roth, B. L.; Sheffler, D. J.; Kroeze, W. K. Magic shotguns versus magic bullets: Selectively nonselective drugs for mood disorders and schizophrenia. *Nat. Rev. Drug Discovery* **2004**, *3*, 353−359.
(7) Ravina, E.; Negreira, J.; Cid, J.; Masaguer, C. F.; Rosa, E.; Rivas, M. E.; Fontenla, J. A.; Loza, M. I.; Tristan, H.; Cadavid, M. I.; Sanz, F.; Lozoya, E.; Carotti, A.; Carrieri, A. Conformationally constrained butyrophenones with mixed dopaminergic (D(2)) and serotoninergic

(5-HT(2A), 5-HT(2C)) affinities: Synthesis, pharmacology, 3D-QSAR, and molecular modeling of (aminoalkyl)benzo- and -thienocycloalkanones as putative atypical antipsychotics. *J. Med. Chem.* **1999**, *42*, 2774−2797.

(8) Brea, J.; Rodrigo, J.; Carrieri, A.; Sanz, F.; Cadavid, M. I.; Enguix, M. J.; Villazon, M.; Mengod, G.; Caro, Y.; Masaguer, C. F.; Ravina, E.; Centeno, N. B.; Carotti, A.; Loza, M. I. New serotonin 5-HT-(2A), 5-HT(2B), and 5-HT(2C) receptor antagonists: Synthesis, pharmacology, 3D-QSAR, and molecular modeling of (aminoalkyl)-benzo and heterocycloalkanones. *J. Med. Chem.* **2002**, *45*, 54−71.

(9) Brea, J.; Masaguer, C. F.; Villazon, M.; Cadavid, M. I.; Ravina, E.; Fontaine, F.; Dezi, C.; Pastor, M.; Sanz, F.; Loza, M. I. Conformationally constrained butyrophenones as new pharmacological tools to study 5-HT 2A and 5-HT 2C receptor behaviours. *Eur. J. Med. Chem.* **2003**, *38*, 433−440.

(10) Brea, J.; Castro, M.; Loza, M. I.; Masaguer, C. F.; Ravina, E.; Dezi, C.; Pastor, M.; Sanz, F.; Cabrero-Castel, A.; Galan-Rodriguez, B.; Fernandez-Espejo, E.; Maldonado, R.; Robledo, P. QF2004B, a potential antipsychotic butyrophenone derivative with similar pharmacological properties to clozapine. *Neuropharmacology* **2006**, *51*, 251−262.

(11) Waller, C. L.; Oprea, T. I.; Giolitti, A.; Marshall, G. R. Three-dimensional QSAR of human immunodeficiency virus (I) protease inhibitors. 1. A CoMFA study employing experimentally determined alignment rules. *J. Med. Chem.* **1993**, *36*, 4152−4160.

(12) Waller, C. L.; Marshall, G. R. Three-dimensional quantitative structure−activity relationship of angiotesin-converting enzyme and thermolysin inhibitors. II. A comparison of CoMFA models incorporating molecular orbital fields and desolvation free energies based on active-analog and complementary-receptor-field alignment rules. *J. Med. Chem.* **1993**, *36*, 2390−2403.

(13) Cho, S. J.; Garsia, M. L.; Bier, J.; Tropsha, A. Structure-based alignment and comparative molecular field analysis of acetylcholinesterase inhibitors. *J. Med. Chem.* **1996**, *39*, 5064−5071.

(14) Pastor, M.; Cruciani, G.; Watson, K. A. A strategy for the incorporation of water molecules present in a ligand binding site into a three-dimensional quantitative structure−activity relationship analysis. *J. Med. Chem.* **1997**, *40*, 4089−4102.

(15) Lozano, J. J.; Pastor, M.; Cruciani, G.; Gaedt, K.; Centeno, N. B.; Gago, F.; Sanz, F. 3D-QSAR methods on the basis of ligand-receptor complexes. Application of COMBINE and GRID/GOLPE methodologies to a series of CYP1A2 ligands. *J. Comput.-Aided Mol. Des.* **2000**, *14*, 341−353.

(16) Sippl, W. Receptor-based 3D QSAR analysis of estrogen receptor ligands−merging the accuracy of receptor-based alignments with the computational efficiency of ligand-based methods. *J. Comput.-Aided Mol. Des.* **2000**, *14*, 559−572.

(17) Sippl, W.; Contreras, J. M.; Parrot, I.; Rival, Y. M.; Wermuth, C. G. Structure-based 3D QSAR and design of novel acetylcholinesterase inhibitors. *J. Comput.-Aided Mol. Des.* **2001**, *15*, 395−410.

(18) Sippl, W. Binding affinity prediction of novel estrogen receptor ligands using receptor-based 3-D QSAR methods. *Bioorg. Med. Chem.* **2002**, *10*, 3741−3755.

(19) Pastor, M.; Perez, C.; Gago, F. Simulation of alternative binding modes in a structure-based QSAR study of HIV-1 protease inhibitors. *J. Mol. Graphics Modell.* **1997**, *15*, 364−71, 389.

(20) Mewshaw, R. E.; Silverman, L. S.; Mathew, R. M.; Kaiser, C.; Sherrill, R. G.; Cheng, M.; Tiffany, C. W.; Karbon, E. W.; Bailey, M. A.; Borosky, S. A. Bridged gamma-carbolines and derivatives possessing selective and combined affinity for 5-HT2 and D2 receptors. *J. Med. Chem.* **1993**, *36*, 1488−1495.

(21) Repke, D. B.; Clark, R. D.; Kluge, A. F.; Muchowski, J. M.; Strosberg, A. M.; Lee, C. H.; Whiting, R. L. Syntheses and hypotensive properties of substituted 2-aminotetralins. *J. Pharm. Sci.* **1985**, *74*, 37−39.

(22) Julia, S.; Bonnet, Y. Etude des cétones avec noyau cyclopropane I. Les benzo- et naphtho- bicyclo- (0,1,4) hepténones. *Bull. Soc. Chim. Fr.* **1957**, 1340−1347.

(23) Caro, Y.; Torrado, M.; Masaguer, C. F.; Ravina, E.; Padin, F.; Brea, J.; Loza, M. I. Chemoenzymatic synthesis and binding affinity of novel (*R*)- and (*S*)-3-aminomethyl-1-tetralones, potential atypical antipsychotics. *Bioorg. Med. Chem. Lett.* **2004**, *14*, 585−589.

(24) Metwally, K. A.; Dukat, M.; Egan, C. T.; Smith, C.; DuPre, A.; Gauthier, C. B.; Herrick-Davis, K.; Teitler, M.; Glennon, R. A. Spiperone: Influence of spiro ring substituents on 5-HT2A serotonin receptor binding. *J. Med. Chem.* **1998**, *41*, 5084−5093.

(25) Wysong, C. L.; Yokum, T. S.; Morales, G. A.; Gundry, R. L.; McLaughlin, M. L.; Hammer, R. P. 4-Aminopiperidine-4-carboxylic acid: A cyclic alpha,alpha-disubstituted amino acid for preparation of water-soluble highly helical peptides. *J. Org. Chem.* **1996**, *61*, 7650−7651.

(26) Palczewski, K.; Kumasaka, T.; Hori, T.; Behnke, C. A.; Motoshima, H.; Fox, B. A.; Le Trong, I.; 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.

(27) Sali, A.; Blundell, T. L. Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol.* **1993**, *234*, 779−815.

(28) Egan, C.; Herrick-Davis, K.; Teitler, M. Creation of a constitutively activated state of the 5-HT2A receptor by site-directed mutagenesis: Revelation of inverse agonist activity of antagonists. *Ann. N. Y. Acad. Sci.* **1998**, *861*, 136−139.

(29) Berg, K. A.; Harvey, J. A.; Spampinato, U.; Clarke, W. P. Physiological relevance of constitutive activity of 5-HT2A and 5-HT2C receptors. *Trends Pharmacol. Sci.* **2005**, *26*, 625−630.

(30) Roth, B. L.; McLean, S.; Zhu, X. Z.; Chuang, D. M. Characterization of two [3*H*]ketanserin recognition sites in rat striatum. *J. Neurochem.* **1987**, *49*, 1833−1838.

(31) Choudhary, M. S.; Craigo, S.; Roth, B. L. A single point mutation (Phe340−>Leu340) of a conserved phenylalanine abolishes 4-[^{125}I]-iodo-(2,5-dimethoxy) phenylisopropylamine and [3*H*]mesulergine but not [3*H*]ketanserin binding to 5-hydroxytryptamine2 receptors. *Mol. Pharmacol.* **1993**, *43*, 755−761.

(32) Wurch, T.; Colpaert, F. C.; Pauwels, P. J. Chimeric receptor analysis of the ketanserin binding site in the human 5-hydroxytryptamine1D receptor: Importance of the second extracellular loop and fifth transmembrane domain in antagonist binding. *Mol. Pharmacol.* **1998**, *54*, 1088−1096.

(33) Wurch, T.; Palmier, C.; Pauwels, P. J. Induction of a high-affinity ketanserin binding site at the 5-hydroxytryptamine(1B) receptor by modification of its carboxy-terminal intracellular portion. *Biochem. Pharmacol.* **2000**, *59*, 1117−1121.

(34) 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.

(35) Bissantz C Conformational changes of G protein-coupled receptors during their activation by agonist binding. *J. Recept. Signal Transduct. Res.* **2003**, *23*, 123−153.

(36) Shapiro, D. A.; Kristiansen, K.; Weiner, D. M.; Kroeze, W. K.; Roth, B. L. Evidence for a model of agonist-induced activation of 5-hydroxytryptamine 2A serotonin receptors that involves the disruption of a strong ionic interaction between helices 3 and 6. *J. Biol. Chem.* **2002**, *277*, 11441−11449.

(37) Sylte, I. B. A.; Dahl, S. G. Ligand induced conformational states of the 5-HT(1A) receptor. *Eur. J. Pharmacol.* **2001**, *416*, 33−41.

(38) Almaula, N.; Ebersole, B. J.; Ballesteros, J. A.; Weinstein, H.; Sealfon, S. C. Contribution of a helix 5 locus to selectivity of hallucinogenic and nonhallucinogenic ligands for the human 5-hydroxytryptamine2A and 5-hydroxytryptamine2C receptors: Direct and indirect effects on ligand affinity mediated by the same locus. *Mol. Pharmacol.* **1996**, *50*, 34−42.

(39) Shapiro, D. A.; Kristiansen, K.; Kroeze, W. K.; Roth, B. L. Differential modes of agonist binding to 5-hydroxytryptamine(2A) serotonin receptors revealed by mutation and molecular modeling of conserved residues in transmembrane region 5. *Mol. Pharmacol.* **2000**, *58*, 877−886.

(40) Cortizo, L.; Santana, L.; Ravina, E.; Orallo, F.; Fontenla, J. A.; Castro, E.; Calleja, J. M.; De Ceballos, M. L. Synthesis and antidopaminergic activity of some 3-(aminomethyl)tetralones as analogs of butyrophenone. *J. Med. Chem.* **1991**, *34*, 2242−2247.

(41) Alvarado, M.; Coelho, A.; Masaguer, C. F.; Raviña, E.; Brea, J.; Padín, J. F.; Loza, M. I. Synthesis and binding affinity of novel 3-aminoethyl-1-tetralones, potential atypical antipsychotics. *Bioorg. Med. Chem. Lett.* **2005**, *15*, 3063−3066.

(42) Loza, M. I.; Ferreiro, T. G.; Sanz, F.; Lozoya, E.; Rodriguez, J.; Manaut, F.; Verde, I.; Castro, E.; Fontenla, J. A.; Cadavid, I. Antiserotoninergic activity of 2-aminoethylbenzocyclanones in rat aorta: Structure−activity relationships. *J. Pharm. Sci.* **1993**, *82*, 513−517.

(43) Fontenla, J. A.; Osuna, J.; Rosa, E.; Castro, M. E.; G-Ferreiro, T.; Loza-Garcia, I.; Calleja, J. M.; Sanz, F.; Rodriguez, J.; et al. Synthesis and atypical antipsychotic profile of some 2-(2-piperidinoethyl)-benzocycloalkanones as analogs of butyrophenone. *J. Med. Chem.* **1994**, *37*, 2564−2573.

(44) Ravina, E.; Fueyo, J.; Masaguer, C. F.; Negreira, J.; Cid, J.; Loza, I.; Honrubia, A.; Tristan, H.; G-Ferreiro, T.; Fontenla, J. A.; Rosa, E.; Calleja, J. M.; De Ceballos, M. L. Synthesis and affinities for dopamine (D2) and 5-hydroxytryptamine (5-HT2A) receptors of 1-(benzoylpropyl)-4-(1-oxocycloalkyl-2-ethyl)-piperazines as cyclic butyrophenone derivatives. *Chem. Pharm. Bull. (Tokyo)* **1996**, *44*, 534−541.

(45) Masaguer, C. F.; Casariego, I.; Ravina, E. Conformationally restricted butyrophenones with mixed dopaminergic (D[2]) and serotoninergic (5-HT[2A]) affinities. Synthesis of 5-aminoethyl-and 6-aminomethyl-4-oxotetrahydroindoles as potential atypical antipsychotics. *Chem. Pharm. Bull.* **1999**, *47*, 621−632.
(46) Masaguer, C. F.; Formoso, E.; Raviña, E.; Tristán, H.; Loza, M. I.; Rivas, E.; Fontenla, J. Butyrophenone analogues in the carbazole series: Synthesis and determination of affinities at D2 and 5-HT2A receptors. *Bioorg. Med. Chem. Lett.* **1998**, *8*, 3571−3576.
(47) Pita, B.; Masaguer, C. F.; Raviña, E. A simple, efficient method for regioselective synthesis of 7-aminomethyl-7,8-dihydro-6*H*-quinolin-5-ones, new potential CNS agents. *Tetrahedron Lett.* **2000**, *41*, 9829−9833.
(48) Ravina, E.; Casariego, I.; Masaguer, C. F.; Fontenla, J. A.; Montenegro, G. Y.; Rivas, M. E.; Loza, M. I.; Enguix, M. J.; Villazon, M.; Cadavid, M. I.; Demontis, G. C. Conformationally constrained butyrophenones with affinity for dopamine (D1, D2, D4) and serotonin (5-HT2A, 5-HT2B, 5-HT2C) receptors: Synthesis of aminomethylbenzo[*b*]furanones and their evaluation as antipsychotics. *J. Med. Chem.* **2000**, *43*, 4678−4693.
(49) Verdonk, M. L.; Cole, J. C.; Hartshorn, M. J.; Murray, C. W.; Taylor, R. D. Improved protein−ligand docking using GOLD. *Proteins* **2003**, *52*, 609−623.
(50) Gasteiger, J.; Rudolph, C.; Sadowski, J. Automatic generation of 3D-atomic coordinates for organic molecules. *Tetrahedron Comput. Methodol.* **1990**, *3*, 537−547.
(51) Jones, G.; Willett, P.; Glen, R. C.; Leach, A. R.; Taylor, R. Development and validation of a genetic algorithm for flexible docking. *J. Mol. Biol.* **1997**, *267*, 727−748.
(52) *GRID 22*; Molecular Discovery, Ltd.: London, U.K., 2006.
(53) *GOLPE 4.6.0.*; Multivariate Infometric Analysis: Srl. Perugia, Italy, 2003.
(54) Pastor, M.; Cruciani, G.; Clementi, S. Smart region definition: A new way to improve the predictive ability and interpretability of three-dimensional quantitative structure-activity relationships. *J. Med. Chem.* **1997**, *40*, 1455−1464.
(55) Hudson, B. D.; Hyde, R. M.; Rah, E. r.; Wood, J.; Osman, J. Parameter based methods for compound selection from chemical databases. *Quant. Struct.-Act. Relat.* **2006**, *15*, 285−289.
(56) Cramer, R. D., III; Patterson, D. E.; Bunce, J. D. Comparative molecular field analysis (CoMFA). 1. Effect of shape on binding of steroids to carrier proteins. *J. Am. Chem. Soc.* **1988**, *110*, 5959−5967.
(57) Doweyko, A. M. 3D-QSAR illusions. *J. Comput.-Aided Mol. Des.* **2004**, *18*, 587−596.
(58) Ballesteros, J.; Weinstein, H. Integrated methods for the construction of three dimensional models and computational probing of structure−function relations in G-protein coupled receptors. *Methods Neurosci.* **1995**, *25*, 366−428.
(59) Boeckmann, B.; Bairoch, A.; Apweiler, R.; Blatter, M. C.; Estreicher, A.; Gasteiger, E.; Martin, M. J.; Michoud, K.; O'Donovan, C.; Phan, I.; Pilbout, S.; Schneider, M. The SWISS-PROT protein knowledgebase and its supplement TrEMBL in 2003. *Nucleic Acids Res.* **2003**, *31*, 365−370.
(60) 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.
(61) Thompson, J. D.; Higgins, D. G.; Gibson, T. J. CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties, and weight matrix choice. *Nucleic Acids Res.* **1994**, *22*, 4673−4680.
(62) Baldwin, J. M.; Schertler, G. F.; Unger, V. M. An alpha-carbon template for the transmembrane helices in the rhodopsin family of G-protein-coupled receptors. *J. Mol. Biol.* **1997**, *272*, 144−164.
(63) Persson, B.; Argos, P. Prediction of membrane protein topology utilizing multiple sequence alignments. *J. Protein Chem.* **1997**, *16*, 453−457.
(64) Zhang, D.; Weinstein, H. Polarity conserved positions in transmembrane domains of G-protein coupled receptors and bacteriorhodopsin. *FEBS Lett.* **1994**, *337*, 207−212.
(65) Cuff, J. A.; Clamp, M. E.; Siddiqui, A. S.; Finlay, M.; Barton, G. J. JPred: A consensus secondary structure prediction server. *Bioinformatics* **1998**, *14*, 892−893.
(66) White, S. H.; Wimley, W. C. Membrane protein folding and stability: Physical principles. *Annu. Rev. Biophys. Biomol. Struct.* **1999**, *28*, 319−365.
(67) Russell, R. B.; Barton, G. J. Multiple protein sequence alignment from tertiary structure comparison: assignment of global and residue confidence levels. *Proteins* **1992**, *14*, 309−323.
(68) Espadaler, J.; Fernandez-Fuentes, N.; Hermoso, A.; Querol, E.; Aviles, F. X.; Sternberg, M. J.; Oliva, B. ArchDB: Automated protein loop classification as a tool for structural genomics. *Nucleic Acids Res.* **2004**, *32*, D185−D188.
(69) Laskowski, R. A.; MacArthur, M. W.; Moss, D. S.; Thornton, J. M. PROCHECK: A program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.* **1993**, *26*, 283−291.
(70) Case, D. A.; Pearlman, D. A.; Caldwell, J. M.; Cheatman, T. E., III; 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.; Radmer, R. J.; Duan, Y.; Pitera, J.; Massova, I.; Seibel, G. L.; Singh, U. C. *AMBER 6.0.*; University of California, San Francisco, San Francisco, CA, 1999.
(71) Cornell, W. D.; Cieplak, P.; Bayly, I. R.; Gould, I. R.; Merz, K. M., Jr.; Ferguson, D. M.; Spellmeir, 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.
(72) Frisch, M. J.; et al. *Gaussian 98*, revision A.7; Gaussian, Inc.; Pittsburgh, PA, 1998.
(73) Bayly, C. I.; Cieplak, P.; Cornell, W.; Kollman, P. A. A well-behaved electrostatic potential based method using charge restraints for determining atom-centered charges: The resp model. *J. Phys. Chem.* **1993**, *97*, 10269.
(74) Miyamoto, S.; Kollman, P. A. An analytical version of the SHAKE and RATTLE algorithm for rigid water molecules. *J. Comput. Chem.* **1992**, *13*, 952−962.
(75) Humphrey, W.; Dalke, A.; Schulten, K. VMD: Visual molecular dynamics. *J. Mol. Graph.* **1996**, *14*, 33−8, 27−8.

JM070277A