*Journal of Enzyme Inhibition and Medicinal Chemistry*, June 2006; 21(3): 285–292

Taylor & Francis
Taylor & Francis Group

# Homology modelling of the serotoninergic 5-$HT_{2c}$ receptor

AMAURY FARCE, SEBASTIEN DILLY, SAID YOUS, PASCAL BERTHELOT, & PHILIPPE CHAVATTE

*Faculté des Sciences Pharmaceutiques et Biologiques, Laboratoire de Chimie Thérapeutique, EA1043, 3 rue du professeur Laguesse, BP 83, 59006 Lille Cedex, France*

*(Received 17 November 2005; in final form 11 January 2006)*

**Abstract**
Since its discovery, 5-hydroxytryptamine, more usually called serotonin, has been an elusive candidate as a major mood regulator. This capacity gives it a great importance in the treatment of depression. It is within this framework that our work takes place, as it is related more particularly to a new therapeutic class whose leader is agomelatine. This compound binds to the melatoninergic receptors and to the serotoninergic 5-$HT_{2c}$ receptor, giving rise to the MASSA concept (Melatonin Agonist and Selective Serotonin Antagonist). Like the majority of the serotoninergic receptors, the sub-type 5-$HT_{2c}$ is a G-protein coupled receptor (GPCR). The three-dimensional structure of 5-$HT_{2c}$ is not experimentally known, and we thus resorted to comparative homology modelling to build a model allowing us to study its interactions with agomelatine.

**Keywords:** *5-$HT_{2c}$ receptor, GPCRs, homology modelling, agomelatine, depression, MASSA*

## Introduction

Serotonin (5-HT) was discovered in 1935 on the basis of its contracting power on smooth muscles. Since then, the implication of the serotoninergic system in nearly all physiological situations has been uncovered. It has peripheral functions such as the aforementioned contracting capacity on smooth muscles [1], which is involved in emesis and bronchoconstriction. These properties have led to many researches for antagonists in the fields of antiemetic [2] and asthma [3]. Another well-known peripheral activity mediated by serotonin is the modulation of gastric acidity. It also has complex activities on the cardiovascular system. For example, it has proved to be a successful possibility for the treatment of headache due to the vasoconstriction of cerebral capillaries. Its central effects are also very complex. It is the precursor of melatonin and thus plays an indirect role in circadian rhythms [4], but its major effect seems to be mood regulation. It is known that a deficit in serotonin is one of the reasons for depression [5]. It also appears to be a factor in aggressiveness [6], pain perception [7], feeding behaviour [8] and body temperature regulation [9]. This diversity of effects is mediated by various 5-HT receptors. With the exception of the 5-$HT_3$ class, which are ionic channels, all of them belong to the G-Protein Coupled Receptors superfamily (GPCRs) [10]. Structurally, they display a common, characteristic fold composed of seven transmembrane domains arranged in $\alpha$–helices. Another particular characteristic of the GPCRs is the high conservation of a given pattern of residues, such as the DRY sequence of the intracellular end of transmembrane domain 3 (TM3), which is seemingly implied in the receptor activation

Correspondence. Pr. P. Chavatte, Faculté des Sciences Pharmaceutiques et Biologiques, Laboratoire de Chimie Thérapeutique, 3 rue du professeur Laguesse, 59006 Lille Cedex, France. Tel.: + 33-3-20-96-40-20. Fax: + 33-3-20-96-43-61, E-mail: philippe.chavatte@univ-lille2.fr

ISSN 1475-6366 print/ISSN 1475-6374 online © 2006 Taylor & Francis
DOI: 10.1080/14756360600700608

[11]. Serotoninergic receptors bind to the whole range of G proteins with 5-$HT_{1A}$, 5-$HT_{1B}$, and 5-$HT_{1D}$ linking to a $G_{i/o}$, 5-$HT_4$ and 5-$HT_7$ to $G_s$, and all 5-$HT_2$ to $G_q$. Despite their great interest for rational drug design, the structure of GPCRs is still largely theoretical. They are hardly amenable to standard crystallographic or NMR methods because their integral membrane protein nature renders them difficult to isolate and crystallize [12]. However, homology modelling offers an alternative pathway through these difficulties by the construction of theoretical models based on the only GPCR crystallographic structure available as yet, bovine rhodopsin [13,14]. We have devised a model of 5-$HT_{2C}$ since this particular subtype appears to be implicated in the mechanism of action of the new Melatonin Agonist and Selective Serotonin Antagonist (MASSA) antidepressant class [15,16]. Due to a longstanding interest of our laboratory in the field of melatoninergic ligands, as is demonstrated by the synthesis of agomelatine [17], it is therefore important to study at the molecular level the interactions between 5-$HT_{2c}$ and agomelatine in order to achieve a better comprehension of the Structure-Affinity-Activity of this new therapeutic class. In this paper, we report preliminary findings achieved through the construction of a 3D–structure of the receptor and the docking of agomelatine.

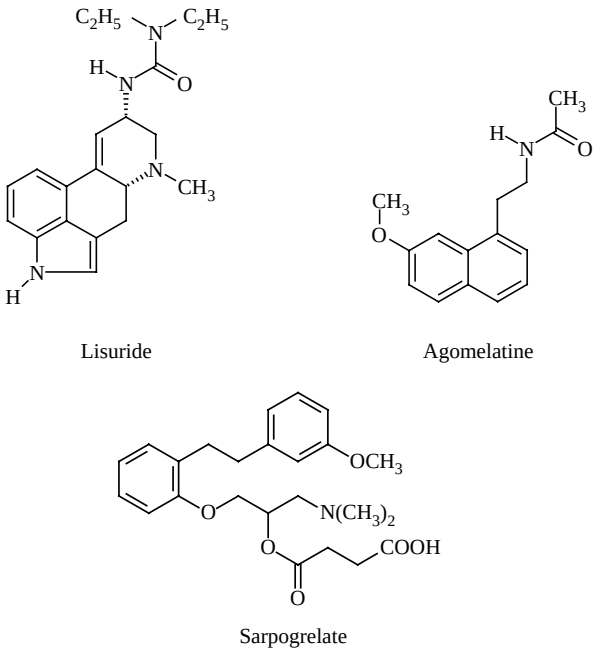


Figure 1. Structure of compounds.

## Materials and methods

Molecular modelling studies were performed using SYBYL software version 6.9.1 [18] running on a Silicon Graphics Octane 2 workstation.

### *Homology modelling*

Homology modelling is based on the idea that the structure of proteins is better preserved during the evolution than their sequence [19]. It is thus possible to build the structure of a protein belonging to a family by referring to the known structure of another protein of this family. The homology modelling process consists in several steps. The first and more critical for obtaining a correct structure is the alignment of the sequences of the target and a carefully chosen reference. In our case, this latter problem was solved by the existence of only one such possible reference, bovine rhodopsin extracted from the PIR (*http://www-nbrf.georgetown.edu/pirwww/dbinfo/pirpsd.html*) (PIR entry OOBO) [20,21]. 5-$HT_{2c}$ sequence was also extracted from the PIR (PIR entry JS0616) [22]. The alignment was performed using the sequence alignment program CLUSTAL_X [23].

The second step is the transfer of a set of constraints derived from the reference structure to the corresponding amino acids of the sequence to be modelled. These constraints allow the construction of the model skeleton. In this case, the crystallographic structure of bovine rhodopsin at a 2.2 Å resolution (PDB entry 1U19) was employed [24,25]. We used the NEST program of the JACKAL protein structure modeling package that has the advantage of placing the side chains from a library of rotamers derived from high resolution crystallographic structures, and thus taking into account the environment of a residue to orient conveniently its side chain [26].

Lastly, the energy of the model was minimized to bring it to a stable conformation. It is however necessary to preserve the geometry of the backbone to keep the tertiary structure of a GPCR. We thus carried out this minimization in three stages by using the AMBER 7.0 force field [27,28], a dielectric constant of 4 and a limit distance of non-bounded interactions of 10Å. Steric clashes of the side chains were first resolved by 200 iterations. Then the backbone of the helices was constrained while the other residues were allowed to move for 1000 iterations. A final 1000 cycles was carried out on the whole protein. The resulting structure was verified with the PROCHECK program [29] and the errors corrected before minimizing the energy of the model again.

### *Binding site identification and docking*

We docked two compounds behaving as antagonists in our model: lisuride and agomelatine (Figure 1). Both were built from a standard fragments library, and their geometry was subsequently optimized using the Tripos force field [30] including the electrostatic term calculated from Gasteiger and Hückel atomic charges. The method of Powell available in Maximin2

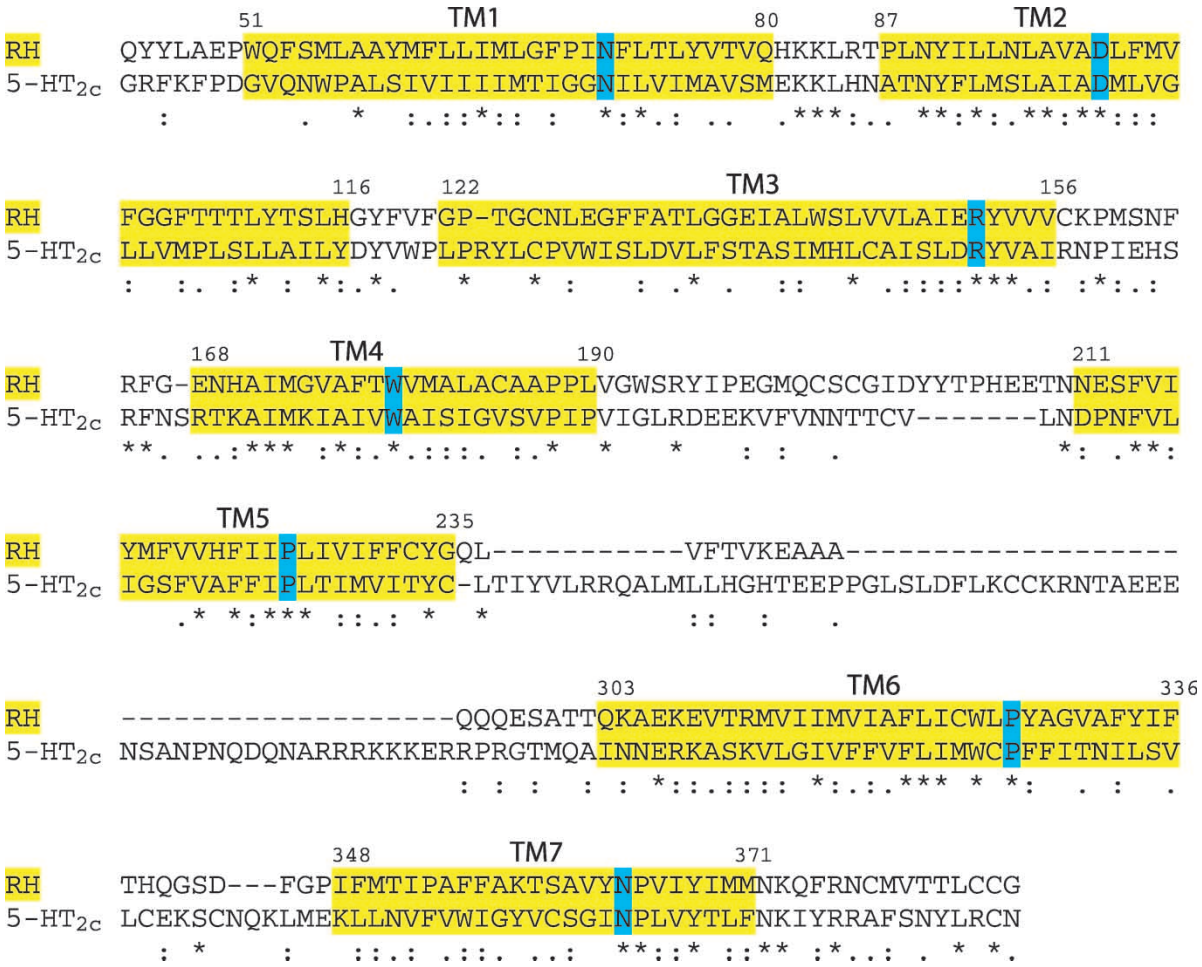


Figure 2. Alignment of the sequences of bovine rhodopsin (RH) and 5-HT$_{2c}$ (5-HT$_{2c}$). The experimental helices of bovine rhodopsin are coloured in yellow, conserved residues in the GPCRs superfamily are coloured in cyan. * means identical residues,: homolog residues,. similar properties residues. The sequence numbers are related to 5-HT$_{2c}$.

procedure was used for energy minimization until the gradient value was smaller than 0.001 kcal/mol.Å. The binding site of 5-HT$_{2c}$ is still very poorly defined, with only a theoretical study based on the antagonist sarpogrelate [31] (Figure 1). As mentioned in this paper, sarpogrelate interacts with the acidic function of Asp134 on TM3. We thus assumed that the pocket should lie in the immediate vicinity of Asp134 and defined the binding site as a 15 Å sphere centred on this residue. The ligands were positioned using the flexible docking software GOLD 2.2 (Genetic Optimization for Ligand Docking) [32,33] and the resulting conformations were classed by a consensus method involving X-Score [34] and Goldscore [32,33].

## Results and discussion

### *Alignment of the sequences*

The choice of a relevant reference for homology modelling is easily solved for GPCRs as there is only one crystallographic structure available, that of bovine rhodopsin, in a dark adaptated form complexed with 11-cis-retinal, currently resolved at a 2.2 Å resolution [25,35]. However, this does not solve the crucial need for a sound alignment of the two sequences. We therefore carried out this task with great care, in order to achieve a suitable superposition of the conserved amino acids of the GPCRs superfamily. The resulting alignment (Figure 2) displays few gaps in the experimentally identified transmembrane domains of rhodopsin as well as a good superposition of the conserved amino acids. A striking feature of this alignment is the considerably much longer third intracellular loop (located between TM5 and TM6) of the serotoninergic receptor, compared to that of rhodopsin. This result was not unexpected, but it raised a problem for modelling this loop. A first alignment displayed an inadequate superimposition on the rhodopsin Gln225-Thr243 sequence, leading to a huge gap just before TM6. This perculiarity prevented a correct conformation of the 6$^{th}$ alpha helix of the model.

Therefore, we manually adjusted the alignment in order to open the gap in the middle of the interhelical space rather than in contact with the helix. In the same vein, N and C-terminal sequences were only included to avoid a deformation of the extremities of the terminal helices during optimization. However, as these parts are highly flexible, their conformation is impossible to determine with a sufficient degree of confidence. We only constructed the fragments of the terminal sequences in direct contact with TM1 and 6 without extending our efforts beyond the gap which is the nearest to the helix.

*Building of the model*

From this alignment, we obtained a model of the 5-$HT_{2c}$ receptor. Its geometry was optimized carefully in order to avoid the loss of its tertiary structure while improving its overall quality. Therefore, steric clashes of the side chains were released first, then loops were optimized, and lastly the whole structure was permitted to relax. Intractable defaults showed up in the third intracellular loop, due to the length of the gap. We replaced it using the loop search module available in Sybyl, which seeks a suitable conformation in an internal database of fragments drawn from crystallographic structures. The resulting model was checked with PROCHECK to assess its structural validity and the detected errors were corrected before minimizing the energy again.

*Validation of the model*

The helices of the model achieved after this last optimization were submitted to a Ramachadran plot in order to verify the correctness of the core structure of the receptor (Figure 3). Loops were not taken into account as their more flexible nature renders them more prone to structural mistakes in a static single-conformation snapshot such as a Ramachadran plot. All the helices residues lie either in the favoured or allowed regions of the plot, which is a good reflexion of the care taken to minimize the energy of the protein.

*Identification of the binding site*

Few data are currently available on the binding site of the 5-$HT_{2c}$ receptor. It was assumed to be in the immediate vicinity of Asp134, a residue thought to be critical in the 5-HT receptors family. This hypothesis is also coherent with previously published modelling studies, which identified Trp130, Asp134, Phe137 on TM3, Val185 on TM4, Ala222, Phe223, Pro226 on TM5 and Phe327, Phe328, Trp324 on TM6 to form the binding site of sarpogrelate [31].

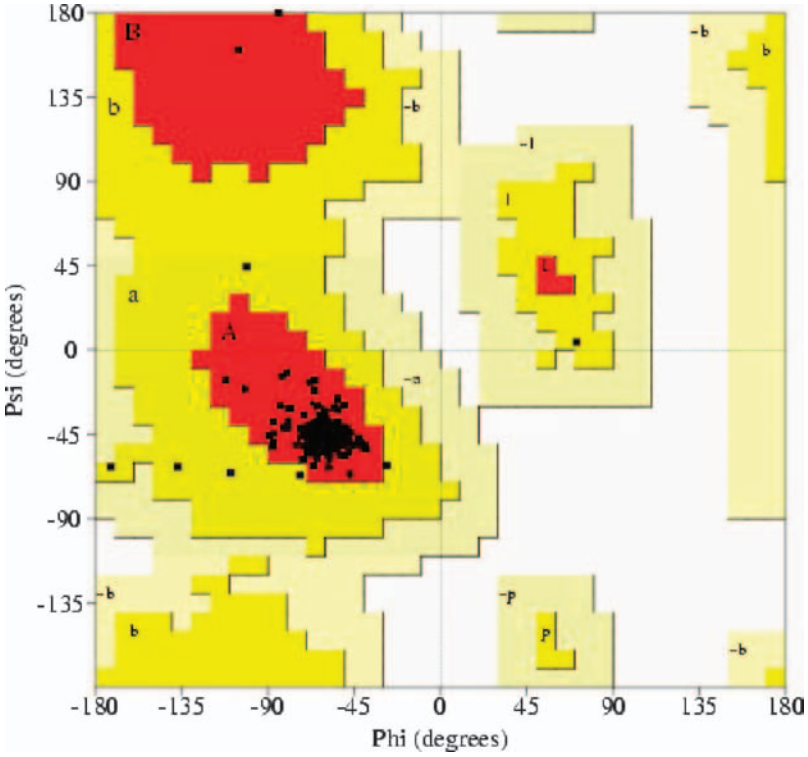


Figure 3. Ramachadran plot of the helices. Favourable regions are coloured in red, allowed regions in yellow, generously allowed regions in tan, disallowed regions in white. Each residue is represented by a point.

*Docking of lisuride*

Lisuride is a 5-HT$_{2c}$ antagonist [36] with a structure derived from lysergic acid. From a structural view, it is closer to serotonin than sarpogrelate, and thus corresponds more to our own ligands. We therefore considered that it was a better reference than sarpogrelate for our studies, even though previous work has been realized using sarpogrelate [31]. Lisuride adopted a conformation directing its urea moiety in front of Asp134 without interacting with it (Figure 4). This should contribute to a steric hindrance impeding a conformation switch of this critical region of the receptor. A closer look at the binding site revealed a two-part cavity roughly perpendicular to the helix bundle (Figure 5). The part accommodating the urea is fairly hydrophilic and seemed to readily form hydrogen bonding networks with the ligands, with Asp134, Asn204, Thr206 and Tyr358 slightly farther. The other part, where the cyclic half of lisuride lay, is hydrophobic in nature. This is largely demonstrated by the numerous electrostatic contacts strongly binding lisuride. The indolic ring of lisuride was placed in a crevice surrounded by a group of three aromatic residues (Trp324, Phe327 and Phe328) closed in front of lisuride by Met230 and Phe223. The quinolinic structure of lisuride is blocked by Val185 and Thr139, with Ser138 slightly below the ring.

*Docking of agomelatine*

Agomelatine also adopted a conformation putting its non aromatic group in front of Asp134, as expected due to the hydrophilic nature of this part of the pocket (Figure 5). However, contrary to lisuride, the hydrogen borne by the amide of agomelatine was engaged in a bond with the carboxylate of this residue (Figure 6). This difference is probably coming from the much larger volume occupied by the diethylurea of the former and the perpendicular orientation the dimethylurea adopts versus the aromatic ring, whereas the acetamide and the aromatic groups are nearly parallel for agomelatine. The acetamide chain also stabilized the position of agomelatine by another hydrogen bond involving the carbonyl group, with the Asn204 side chain. The presence of an acetyl group branched on its nitrogen could explain the antagonist character of agomelatine. Compared to serotonin, agomelatine occupies a greater volume and shows to the receptor a lower electronic density than the smaller and more heavily charged primary amine of serotonin. Assuming that the nitrogens of both compounds are lying in the receptor at the same position versus Asp134, the acetyl group of agomelatine sterically restricts possible conformational changes of the receptor. Among these, a cluster of aromatic moieties (Trp130, Tyr358) that could trap and stabilize the

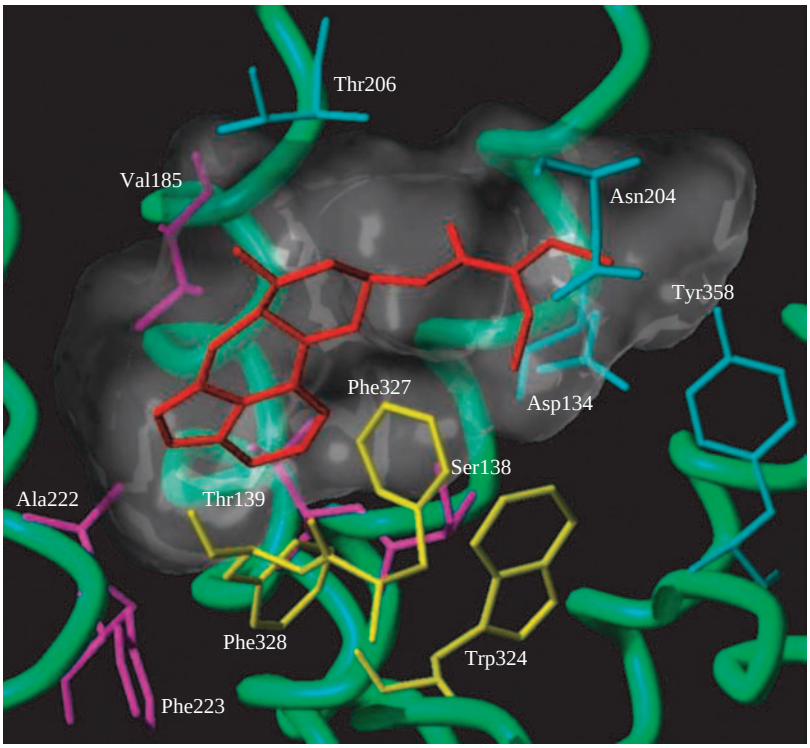


Figure 4. Docking of lisuride (coloured in red).

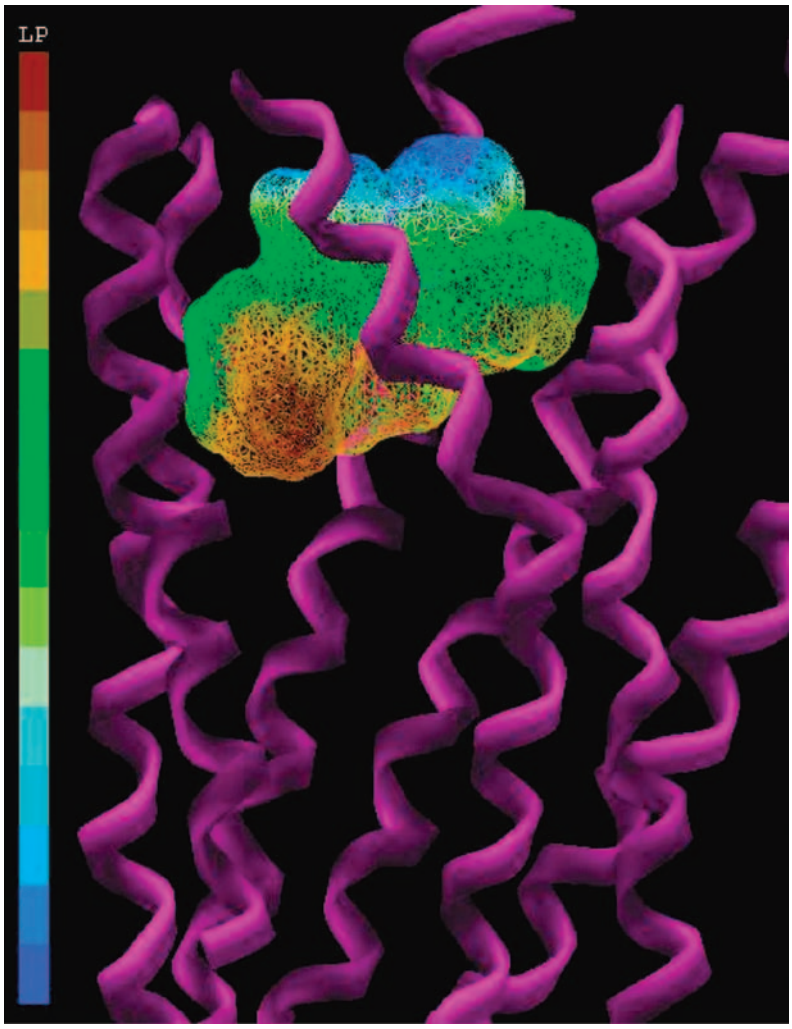


Figure 5. Lipophilic potential (LP) mapped onto the surface of the putative binding site. The color ramp for LP ranges from brown (highest lipophilic area) to blue (highest hydrophilic area).

ionic bond between Asp134 and the nitrogen is unable to fold toward the ionic pair. The other extremity of agomelatine lies in the hydrophobic part of the binding site, where it was engaged in a series of hydrophobic interactions. The naphthalene ring stacked with Trp324 and Phe327. It was also maintained in its conformation by Val135, Phe328, and a third hydrogen bond between the methoxy oxygen of agomelatine and Ser138. One interesting difference between lisuride and agomelatine is that the more straighter conformation of the latter permits a stronger hydrogen bonding pattern around the amide, which is too constrained in lisuride by a more rigid cyclic spacer. On the contrary, lisuride is able to reach more residues with its aromatic part than agomelatine, thus forming more electrostatic interactions. However, our main interest is linked to the MASSA class, to which lisuride does not belong. It contributed to the docking of agomelatine as a convenient template, but more structures of the MASSA family are needed to draw accurate conclusions concerning their structure-affinity relationships.

## Conclusion

In order to investigate the structure-affinity relationships of the new MASSA class of antidepressant drugs and their interactions with their target protein, we have constructed a 5-$HT_{2c}$ model by homology with bovine rhodopsin. The resulting model is thought to be an inactive form of the receptor, thus being able to bind antagonists in a pocket lying perpendicular to the helices. Two parts have been identified in the binding site: a hydrophilic part in front of Asp134, supposed to play a role in activation, and a more hydrophobic part accommodating the aromatic rings of the ligands. These conclusions are in good agreement with previous works on the subject and form a basis to further investigate the nature of the residues required

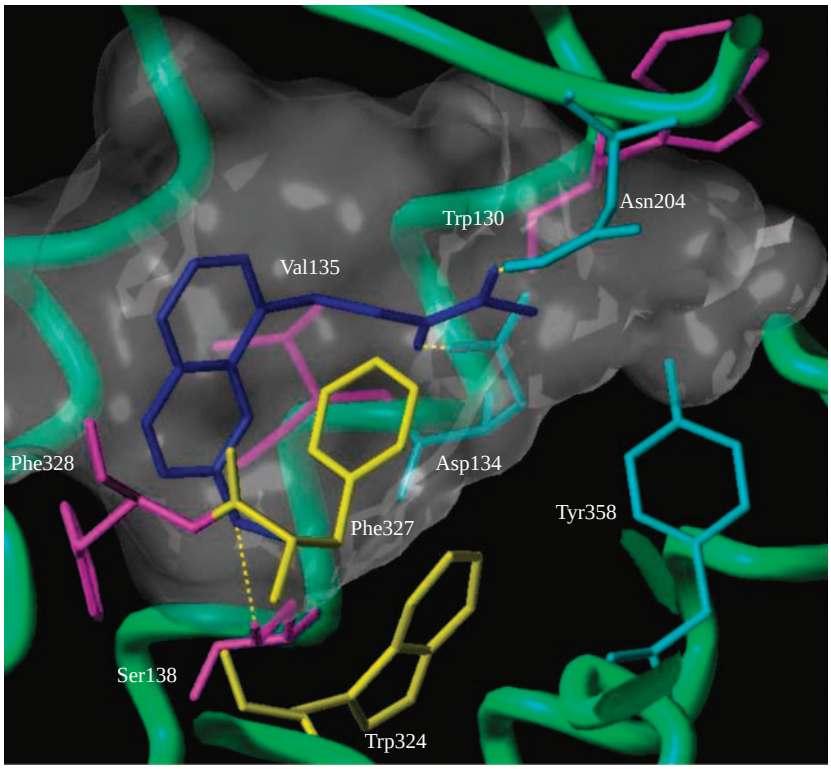


Figure 6. Docking of agomelatine (coloured in blue).

for a good affinity. The docking of agomelatine hinted that Ser138, Asp134 and Asn204 could be responsible for the binding of the ligand through a net of hydrogen bonds supplemented by a number of electrostatic interactions around the naphthalene ring. These studies of protein-ligand interactions will prove to be very valuable in the design of new MASSA compounds by improving our knowledge of their structure-affinity relationships.

## References

[1] Dalgliesh CE, Toh CC, Work TS. J Physiol 1953;120: 298–310.
[2] Yamamoto K, Nohara K, Furuya T, Yamatodani A. Pharmacol Biochem Behav 2005;82:24–29.
[3] Cazzola I, Matera MG. Trends Pharmacol Sci 2000;21:13–16.
[4] Garabette ML, Martin KF, Redfern PH. Br J Pharmacol 2000;131:1569–1576.
[5] Coppen AJ, Doogan DP. J Clin Psychiatry 1998;49:4–11.
[6] Moskowitz DS, Pinard G, Zuroff DC, Annable L, Young SN. Adv Exp Med Biol 2003;527:215–224.
[7] Holden JE, Naleway FE, Jeong Y. Neuroscience 2005;135:1255–1268.
[8] Reis LC, Almeida AC, Cedraz-Mercez PL, Olivares EL, Marinho Jr, A, Thomaz CM. Braz J Biol 2005;65:353–361.
[9] Saadat KS, O'shea E, Colado MI, Elliott JM, Green AR. Psychopharmacology (Berl) 2005;179:884–890.
[10] Hoyer D, Hannon JP, Martin GR. Pharmacol Biochem Behav 2002;71:533–554.
[11] Ballesteros JA, Jensen AD, Liapakis G, Rasmussen SG, Shi L, Gether U, Javitch JA. J Biol Chem 2001;276:29171–29177.
[12] White SH, Ladokhin AS, Jayasinghe S, Hristova K. J Biol Chem 2001;276:32395–32398.
[13] Stenkamp RE, Teller DC, Palczewski K. Arch Pharm Chem Life Sci 2005;338:209–216.
[14] Archer E, Maigret B, Escrieut C, Pradayrol L, Fourmy D. Trends Pharmacol Sci 2003;24:36–40.
[15] Millan MJ, Gobert A, Lejeune F, Dekeyne A, Newman-Tancredi A, Pasteau V, Rivet JM, Cussac D. J Pharmacol Exp Ther 2003;306:954–964.
[16] Chilman-Blair K, Castañer J, Bayès M, Silvestre JS. Drugs Future 2003;28:7–13.
[17] Depreux P, Lesieur D, DAit Mansour H, Morgan P, Howell HE, Renard P, Caignard DH, Pfeiffer B, Delagrange P, Guardiola B, Yous S, Demarque A, Adam G, Andrieux J. J J Med Chem 1994;37:3231–3239.
[18] SYBYL 6.9.1, Tripos Inc, 1699 South Hanley Road, St. Louis, MO 63144-2913, U.S.A
[19] Lesk AM, Chothia C. Philos Trans R Soc Lond B Biol Sci 1986;317:345–356.
[20] Wu CH, Yeh LS, Huang H, Arminski L, Castro-Alvear J, Chen Y, Kourtesis P, Ledley RS, Suzek BE, Vinayaka CR, Zhang J, Barker WC. Nucleic Acids Res 2003;31:345–347.
[21] Nathans J, Hogness DS. Cell 1983;34:807–814.
[22] Saltzman AG, Morse B, Whitman MM, Ivanshchenko Y, Jaye M, Felder S. Biochem Biophys Res Commun 1991;181: 1469–1478.
[23] Thompson JP, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG. Nucleic Acids Res 1997;24:4876–4882.
[24] Berman HM, Westbrook J, Feng Z, Gary G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE. Nucleic Acids Res 2000;28:235.

[25] Okada T, Sugihara M, Bondar AN, Elstner M, Entel P, Buss V. J Mol Biol 2004;10:71–83.
[26] Petrey D, Xiang Z, Tang CL, Xie L, Gimpelev M, Mitros T, Soto CS, Goldsmith-Fischman S, Kernytsky A, Schlessinger A, Koh IYY, Alexov E, Honig B. Proteins: Struct Func and Genet 2003;53:430–435.
[27] Pearlman DA, Case DA, Caldwell JW, Ross WR, Cheatham III TE, DeBoltS, Ferguson D, Seibel G, Kollman P. Comp Phys Commun 1995;91:1–41.
[28] Ponder JW, Case DA. Adv Prot Chem 2003;66:27–85.
[29] Laskowski RA, Moss DS, Thornton JM. J Mol Biol 1993;231:1049–1067.
[30] Clark M, Cramer III RD, V, an Opdenbosch N. J Comp Chem 1989;10:982–1012.
[31] Rashid M, Manivet P, Nishio H, Pratuangdejkul J, Rajab M, Ishiguro M, Launay JM, Nagatomo T. Life Sci 2003;73:193–207.
[32] Jones G, Willet P, Glen RC. J Mol Biol 1995;245:43–53.
[33] Jones G, Willet P, Glen RC, Leach AR, Taylor R. J Mol Biol 1997;267:727–748.
[34] Wang R, Lai L, Wang S. J Comp-Aided Mol Des 2002;16:11–26.
[35] Okada T. Biochem Soc Trans 2004;32:738–741.
[36] Burris KD, Breeding M, Sanders-Bush E. J Pharmacol Exp Ther 1991;258:891–896.