Special issue article

Journal of Physical Organic Chemistry

Received: 12 January 2016, Revised: 12 April 2016, Accepted: 26 April 2016, Published online in Wiley Online Library

(wileyonlinelibrary.com) DOI: 10.1002/poc.3586

Halogen bonding in biological context: a computational study of D2 dopamine receptor

Adriano M. Luchi[a], Emilio L. Angelina[a]*, Sebastián A. Andujar[b], Ricardo D. Enriz[b] and Nélida M. Peruchena[a]

In this work, Halogen Bond (X-bond) interactions formed by halogenated ligands (LX) at the Dopamine Receptor D2 (DRD2) binding pocket were studied by Molecular Dynamics (MD) and charge density analysis. The X-bonds were contrasted with the Hydrogen Bond (H-bond) interactions established by hydroxylated analogs (LOH, where X was replaced by OH). The ligands for this study were extracted from a dataset of compounds deposited in ZINC database that were active in binding assays to DRD2. This dataset was subjected to the filtering rules by employing cheminformatics tools to find the LX/LOH pairs that were then submitted to MD simulations. A homology model of DRD2 was employed for the simulations because no crystal structure is yet available for the receptor. To mimic the positive cap (σ-hole) on the halogen atom, a massless, positive charged extra-point was introduced in the force field. An analysis of the charge density (QTAIM) was performed on reduced models of simulated complexes to explain their binding differences. Results show that the halogen atom tends to form X-bond with protein backbone oxygen atom. Two out of the four halogenated ligands studied form a specific X-bond with the carbonyl oxygen of Ser193. This specific X-bond decreases the inherent propensity of transmembrane 5 to unfolding. These results suggest a possible role of the X-bond as a protein secondary structure modulator because of the ability of the halogen to interact with the protein backbone. Copyright © 2016 John Wiley & Sons, Ltd.

Supporting information may be found in the online version of this paper.

Keywords: GPCR; molecular dynamics; QTAIM; σ-hole

INTRODUCTION

Molecular interactions play a key role in biological systems. As noted by Boeckler et al.[1] "molecular interactions embody the dialog between chemistry and biology". However, such relationship is not so clear in many cases, because of the inherent complexity of the biological systems. In the field of drug discovery, this is evidenced by the numerous examples of serendipitous discoveries in drug history, from Sir Alexander Fleming's accidental finding up to the more recent examples.[2] Even screening, especially the automated high-throughput screening (HTS) can be considered as a systematic approach to benefit from mere chance. Fortunately nowadays, with the advance in structural biology and the computational techniques these fortuitous findings give the Medicinal Chemists the opportunity to learn retrospectively from them, by means of structure-based approaches.

An example of a more empirical rather than rational approach was the use of halogen atoms to optimize the properties of lead series. Halogen chemistry has been exploited by medicinal chemists for nearly 70 years. They were regarded as useful for optimization of ADMET properties because they improve oral absorption and facilitate crossing biological barriers; they are useful for filling small hydrophobic cavities present in many protein targets and they prolong lifetime of the drug.[3] Much in contrast to that perception of halogens as being just hydrophobic moieties, now we know that chlorine (Cl), bromine (Br) and iodine (I) can also form directed close contacts of the type R—X···Y—R′, where the halogen X acts as a Lewis acid and Y can be any electron donor moiety.[4] This interaction, referred to as "Halogen bonding" has its origin in the anisotropy of charge distribution around the halogen atom, when it is bound to an electron-withdrawing atom. Unexpectedly, despite of being highly electronegative, halogens have a region on the hind side of X along the R—X bond axis which remains positively charged. This region, called σ-hole, is responsible for the directional and stabilizing character of halogen bonding with other electronegative atoms, such as oxygen or nitrogen.[5] This region of electron density deficiency is compensated by an electron-rich belt around the halogen that allows it to engage in "side-on" contacts with electrophiles, at the same time that form contacts with nucleophiles in a "head-on" fashion, through the σ-hole.[4]

Because of the ADMET related properties of halogens, compound databases contain a lot of biological activity records for halogenated lead-like compounds. However in most cases the

* *Correspondence to: E. M. Angelina, Lab. Estructura Molecular y Propiedades, IQUIBA-NEA, Universidad Nacional del Nordeste, CONICET, FACENA, Av. Libertad 5470, Corrientes 3400, Argentina.*
E-mail: emilioluisangelina@hotmail.com

a *A. M. Luchi, E. L. Angelina, N. M. Peruchena*
Lab. Estructura Molecular y Propiedades, IQUIBA-NEA, Universidad Nacional del Nordeste, CONICET, FACENA, Av. Libertad 5470, Corrientes 3400, Argentina

b *S. A. Andujar, R. D. Enriz*
Instituto Multidisciplinario de Investigaciones Biológicas San Luis, Universidad Nacional de San Luis, CONICET, FQBF, Chacabuco 917, San Luis 5700, Argentina

Copyright © 2016 John Wiley & Sons, Ltd.

relationship between such activities and halogen bonding occurrence has not been addressed just because the importance of this kind of interactions was unknown at the time of developing the compounds. In view of the current recognition of halogens as atoms being able to form specific interactions, it would be interesting to retrospectively study such interactions to test, with the aid of affinity annotations, the hypothesis that halogen bond interactions might play a key role in biomolecular systems.

In this work, Halogen Bond (X-bond) interactions formed by known halogenated ligands (LX, with X = Cl, Br, I) of the Dopamine Receptor D2 (DRD2) were studied at the receptor binding pocket. For this study we compiled a dataset of compounds that were active in direct binding assays to DRD2 from literature and from ZINC database.[6] In the compiled dataset there are about 3.9K compounds with affinity annotations by DRD2 and ~1.2K contain Cl, Br and/or I. For each halogenated ligand (LX) in the dataset we searched for the existence of the hydroxylated analog (LOH, where X was replaced by OH) by applying filtering rules based in SMARTS patterns (daylight.com). If both LX and LOH were in the database, they were appended to the dataset of LX/LOH pairs. To verify (or otherwise refute) our hypothesis that halogen bond interactions might play an important role in biomolecular systems we compared interactions of the halogenated ligand with the corresponding interactions in the hydroxylated analog. For this purpose the topological analysis of the charge density based on the Quantum Theory of Atoms in Molecules (QTAIM) was performed on reduced models of the complexes.

Because crystallization of G-Protein Coupled Receptors (GPCRs) is very challenging, there is scarce structural evidence on how these ligands actually bind at the receptor binding pocket. Even the structure of DRD2 receptor alone has never been resolved. Therefore, a structure-based analysis of the interactions of the LX/LOH pairs at the DRD2 binding pocket requires a homology model of the receptor tridimensional structure and molecular modeling tools to model the interactions. Accordingly, we performed Molecular Dynamics (DM) simulations of selected LX/LOH pairs in the binding pocket of a homology model of DRD2 previously reported.[7]

Because the opening of the σ-hole on the halogen atom is clearly a quantum effect, it poses a serious challenge to the current modeling approaches employing conventional force fields, which treat halogen atoms as having all-negative electrostatic potential, thus failing to correctly describe the halogen-bonded systems.[5] Overlooking of the σ-hole might lead to errors in predictions of structure and energetic of drug-protein complexes and thus to failure in drug development. Therefore, to mimic the positive cap (σ-hole) on the halogen atom, a massless, positive charged extra-point (EP) was introduced in the force field during the MD simulations following the procedure of Ibrahim.[8]

Hopefully, this retrospective study might help to understand the intricate relationships between the chemistry of halogen bonding and ligand–receptor recognition which, in turn might lead to a more rational use of halogen substitution in drug development.

COMPUTATIONAL DETAILS

Compilation of the LX/LOH pairs

DRD2 halogenated ligands (LX) were compiled from literature[9,10] and from ZINC database.[6] These compounds were active in direct binding assays to DRD2. For each LX, its hydroxylated analog (LOH, where halogen X was replaced by OH) was then searched in the database, by applying filtering rules based in SMARTS (Smiles ARbitrary Target Specification) patterns. If both LX and LOH were in the database, they were appended to the dataset of halogenated/hydroxylated pairs (Fig. 1). This resulted in a total number of 27 LX/LOH pairs.

Scaffolds selected for further modeling studies

From the 27 LX/LOH pairs, a small sample of them were selected for Molecular Dynamics simulations, based in the following selection criteria: (i) there must be a significant binding affinity difference between LX and LOH; (ii) if there are isomers, the one that is active should be known; and (iii) there must be some evidence of their binding mode in the receptor binding site. Figure 2 shows the scaffolds selected according to these criteria: a pair of haloperidol-like ligands (1-LOH and 1-LX), two dopamine-like pairs (2-LOH and two halogenated isomers, 2-

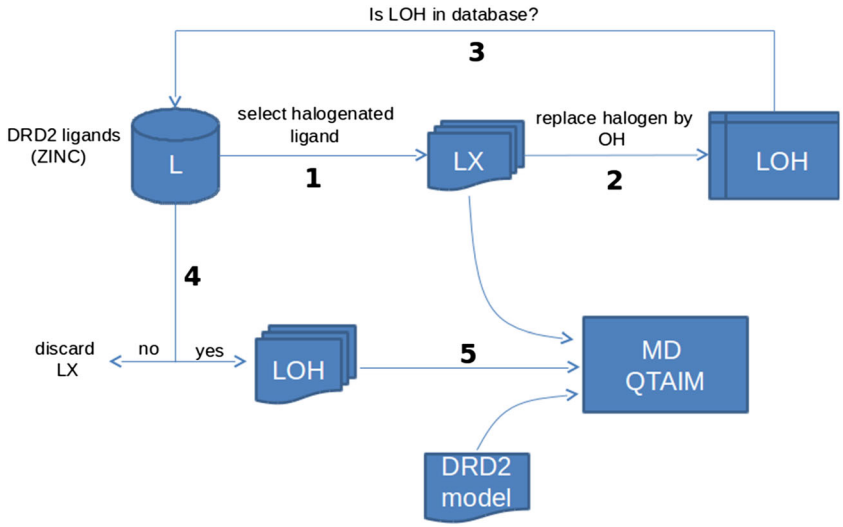

Figure 1. Workflow employed for the compilation of the LX/LOH pairs

Copyright © 2016 John Wiley & Sons, Ltd.

Journal of Physical Organic Chemistry

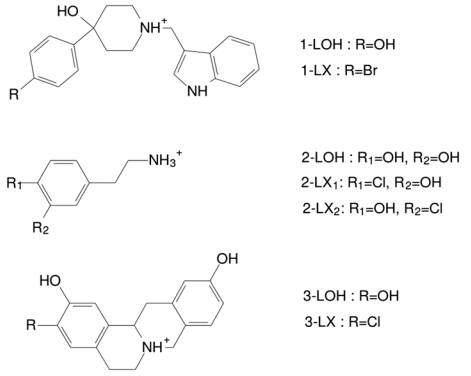

Figure 2. Scaffolds selected from the dataset of LX/LOH pairs

LX1 and 2-LX2) and one pair of tetrahydroprotoberberines (3-LOH and 3-LX) were selected for further modeling studies.

Figure S1 in Supporting Information depicts a spatial view of dopamine bound to DRD2. Relevant residues from the binding pocket are highlighted. The starting poses of the ligands inside the DRD2 orthosteric site were constructed based in previous experimental evidence of the dopamine binding mode into the binding pocket. Mansour et al.[11] have proven the importance of Aspartate 114 and Serines 194 and 197 in ligand binding by mutating these residues at the human dopamine D2 receptor by site-directed mutagenesis. The distance between D114 and the relevant serines of TM5 is about the distance between the amine and catecholic hydroxyl groups of dopamine. Salt bridge between D114 and the protonated amine of dopamine has been recognized as the guideline interaction for anchoring of aminergic ligands into the DRD2 binding site.[12] Therefore, the catecholic OH groups must be close to the relevant serines. There is still a degree of freedom that remains unset in order to define the dopamine location into the binding site, i.e. whether the meta OH group (m-OH) in the dopamine catecholic ring points outward (up) or inward (down) into the receptor seven trans-membrane channel. Therefore, for each one of the dopaminergic ligands, i.e. dopamine and the chlorinated dopamine analogues we performed two molecular dynamics simulations, with m-OH/m-Cl up and down and reported the one with the lowest binding free energy.

As to the case of the tetrahidroprotoberberine derivatives, because the ammonium group is in a ring asymmetric structure, the ligand pose is defined by choosing the right enantiomer. Sun et al.[13] have shown for a large set of tetrahidroprotoberberine derivatives that the S form is the more active, so we performed the MD simulations using the S form.

For the haloperidol-like derivatives the starting binding pose was defined according to the results reported by Wang et al.[14]

Molecular dynamics with explicit treatment of halogen σ-hole

The MD simulations of the unbound state of DRD2 and its complexes with the selected LX/LOH pairs were conducted using the AMBER software package[15] at 300K target temperature and extended to 10-ns overall simulation time. To neutralize the complex positive net charge, 26 negative counter-ions (i.e. chlorine ions) were added. For further details of the MD protocol applied see references.[9,10] A homology model of the long form of the human DRD2 was employed because no crystal structure is yet available for the receptor. This model was built by using as template the structure of the recently crystallized human D3 dopamine receptor (DRD3).[7] Positional restraints were applied to all the backbone α carbons from α-helix transmembrane segments except those of the transmembrane 5 (TM5). It has been stated that receptor models incorporating a flexible TM5 backbone allow reliable prediction of binding affinities for a set of diverse ligands.[16]

To mimic the positive cap (σ-hole) on the halogen atom of the halogenated ligands, a massless, positive charged extra-point (EP) was introduced in the force field during the MD simulations. The parametrization of the EP was performed following the procedure in Ref.[8] In brief, electrostatic potential was generated for the studied halogenated molecules at the HF/6-31G* level. A massless atom was included in the ligand force field. The C(ar)–X–EP angle was set to 180°, and the X–EP distance was set to the halogen atomic radius, i.e. 2.22 Å for bromine and 1.95 Å for chlorine. Atomic partial charges (including the EP coordinates) were assigned according to the restrained electrostatic potential (RESP) approach.

Finally, the MM-GBSA protocol included in AMBER was applied to calculate the relative binding energies of the complexes by taking snapshots at 200-ps time intervals from the entire MD trajectory.

Construction of reduced models and QTAIM (atoms in molecules theory)

Reduced 3D model systems representing the DRD2 binding pocket of the selected LX/LOH pairs were constructed from the MD simulation. Forty-nine residues were included in the model including V91 and L94 from TM2, E99 from extracellular loop 1 (ECL1), D108, F110, V111, D114, V115, C118 and T119 from TM3, G173-N186 segment from ECL2, P187-I203 segment from TM5, W386, F389, F390 and H393 from TM6 and Y408, T412 and Y416 from TM7.

Charge density topological analysis based in the QTAIM[17] was performed on these reduced models to evaluate the ligand–receptor interactions. These calculations were performed with the help of Multiwfn[18] and AIMAll[19] software. The wavefunction used as input for these calculations were computed with the Gaussian 09 package[20] by employing the B3LYP functional,[21–24] with dispersion correction (B3LYP-D)[25] and 6-31G(d)[26] as basis set. The empirical dispersion correction for the B3LYP functional was applied by invoking the IOp 3/124 = 3 keyword in Gaussian 09.

RESULTS AND DISCUSSION

The root-mean-square deviation (RMSD) of each snapshot relative to the initial structure was calculated to monitor the stability of each trajectory. The RMSD obtained for the backbone atoms of each complex does not change appreciably during the 10 ns of the production phase of the simulation (see Fig. S2 in Supporting Information).

Table 1 shows the experimental affinity data for the selected LX/LOH ligands together with the relative binding free energies (ΔΔG) calculated from the MD trajectories with the MM-GBSA protocol. As can be seen, our MD protocol qualitatively reproduces the trend of the binding data within each scaffold, i.e. $\Delta\Delta G(1\text{-LX}) < \Delta\Delta G(1\text{-LOH})$, $\Delta\Delta G(2\text{-LX2}) < \Delta\Delta G(2\text{-LOH}) < \Delta G(2\text{-LX1})$ and $\Delta\Delta G(3\text{-LX}) < \Delta\Delta G(3\text{-LOH})$.

Copyright © 2016 John Wiley & Sons, Ltd.

Table 1. Relative binding free energies and experimental affinities for complexes listed in Figure 2

Complexes	Affinity (nm)	ΔΔG (kcal/mol)	Reference
1-LOH	5360.0	2.15	[27]
1-LX	4.8	0.00	[28]
2-LOH	520.0	4.40	[29]
2-LX1	26 310.0	5.64	[30]
2-LX2	150.0	0.00	[9]
3-LOH	396.0	0.84	[10]
3-LX	188.0	0.00	[10]

Having validated the MD protocol employed in the simulations, in the next sections we focus on the analysis of the molecular interactions.

Evaluating the molecular interactions

To inquire about the origin of the binding affinity differences between LX/LOH pairs, the molecular interactions at the receptor binding pocket were analyzed with the aim of QTAIM theory.

Figures 3–5 show the backbone superposition of DRD2 bound to LOH and LX analog. Also the charge density molecular graphs for the interactions involving the halogen atom and the hydroxyl groups are depicted.

By comparing each LX/LOH pair, one can see from the backbone superpositions that the substitution of the hydroxyl group by an halogen atom in general does not change drastically the overall binding mode of the ligand. At first glance, this resemblance in the binding modes might be explained by the fact that the halogen atom having a explicit treatment of the σ-hole might behave as a bioisoster of the hydroxyl group: the lateral region of the halogen can emulate the oxygen and the σ-hole the proton of the OH group. However, a more careful inspection of the intermolecular interactions at the substitution point reveals important differences in the interaction patterns. The charge density molecular graphs from Figures 3–5 as well as previous results[9,31] show that the ligand phenolic OH groups tend to act either as H-bond donor against the hydroxyl oxygen of the serine residues in the binding pocket or to form oxygen-oxygen interactions with them.[32] On the other hand, the halogen rarely interacts with the hydroxyl oxygen of the serine residues. In fact, in most cases when the halogen atom interacts with the side chain of any of the relevant serines of the binding site, this occurs through the serines non polar hydrogen atoms (i.e. with halogen acting as H-bond acceptor). Instead, the halogen atom tends to act as X-bond donor (i.e. through the σ-hole) with the backbone carbonyl oxygen of Ser193 (O@S193). These findings are in line with the evidence that the backbone carbonyl oxygen function is the most prominent Lewis base involved in halogen bonds in protein binding sites, as found from an analysis of the Protein Data Bank (PDB).[4]

As can be seen in the molecular graphs, in three out of the four halogenated ligands (i.e. 1-LX, 2-LX1 and 3-LX) the halogen atom points to the O@S193 and in two of these complexes a critical point with the corresponding bond paths connecting both atoms can be observed (i.e. in 1-LX and 3-LX). In the case of 2-LX1, even when no topological elements are connecting the chlorine atom with O@Ser193 atom in the depicted molecular graph, most of the simulation time the distance between both atoms is closer and therefore it is very likely that the halogen bond is formed at times during the MD simulation. On the other hand, in 2-LX2 the halogen atom does not point to O@Ser193, but instead it is anchored in a hydrophobic sub-pocket comprising Val190, Phe186, Ile184 and His393.

Figures S3–S9 in the Supporting Information depict expanded charge density molecular graphs for the complexes together

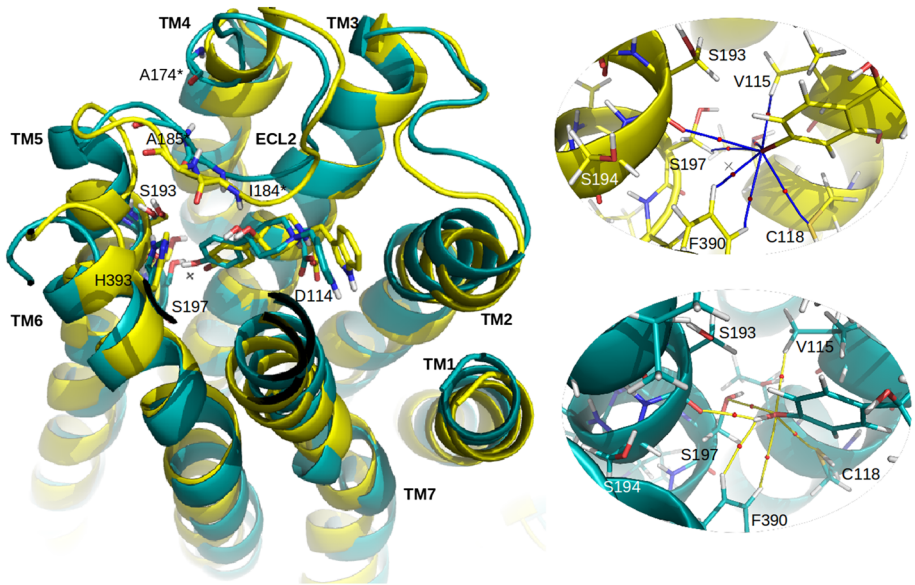

Figure 3. Backbone superposition of DRD2 bound to 1-LX in yellow and 1-LOH in cyan (left side) and charge density molecular graphs showing the interactions of the halogen and hydroxyl groups, respectively (right side). The molecular graphs were constructed from the potential energy minimum of the MD trajectory. Blue and yellow lines connecting the nuclei are the bond paths, and the small red spheres on them are the bond critical points (BCPs). The extra-point representing the σ-hole is depicted with a small gray cross. In residues labeled with an asterisk only the backbone is shown for clarity

Copyright © 2016 John Wiley & Sons, Ltd.

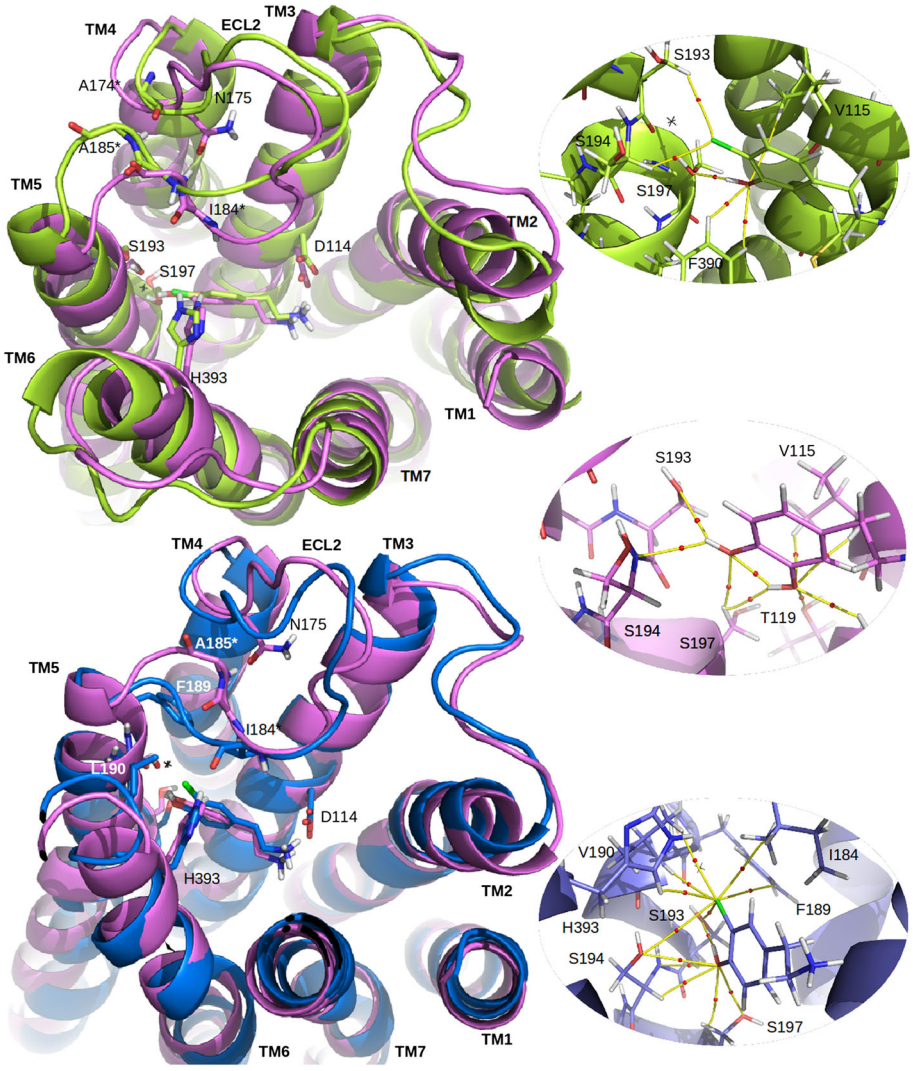

Figure 4. Backbone superposition of complex 2-LOH/DRD2 in purple with its halogenated analogs 2-LX1/DRD2 in green (top left panel) and 2-LX2/DRD2 in blue (bottom left panel). The molecular graphs showing the ligands halogen and hydroxyl interactions are shown in the right side. These graphs were constructed from the potential energy minimum of the MD trajectory. Yellow lines connecting the nuclei are the bond paths, and the small red spheres on them are the bond critical points (BCPs). The extra-point representing the σ-hole is depicted with a small gray cross

with geometrical parameters and local charge density values for the most relevant intermolecular interactions.

Wilcken et al.[4] have analyzed X-bond occurrence and interaction geometries in protein–ligand complexes from a survey of the Protein Data Bank. For all halobenzene systems, a significant decrease of the quality of the interaction was observed when increasing the deviation from 180° in the σ-hole angle (i.e. α(C(ar)—X···O)). Because of the anisotropic electron distribution around the halogen, deviations from the optimal orientation of the σ-hole onto the carbonyl oxygen are not well tolerated. They found that beyond a deviation of 40° from linearity, no significant attractive interaction is found and claimed that below an angle of α(C(ar)—X···O) = 140° the term halogen bonding should not be used, as there is no favorable overlap of the Lewis base with the σ-hole on the halogen.

Accordingly, complexes of 2-LX1 and 3-LX show a σ-hole angle that is compatible with an X-bond while in complex of 1-LX that angle is far from linearity α(C(ar)—X···O) = 64.8° at the potential energy minimum structure of the MD trajectory. Furthermore, in this last complex the distance between the extra-point (EP) and O@Ser93 is about 4 Å most of the simulation time. Because the EP was placed at the atomic radius of the halogen atoms, this distance should be about 2 Å regardless of the size of the halogen if the X-bond were formed.

Laplacian of the charge density

The Laplacian of the charge density ($\nabla^2\rho(r)$) determines spatial regions where the charge density is locally concentrated or depleted. The topology of the Laplacian distribution reflects the shell structure of the isolated atoms in terms of alternating shells of charge concentration followed by a shell of charge depletion. The outer shell of charge concentration is called the valence shell charge concentration (VSCC).[17] When an atom is involved in bonding the spherical symmetry of the VSCC is broken and points of maximum and minimum charge concentration are formed on this shell. It has been shown that when the halogen atom is bound to an electron withdrawing group a point of minimum charge concentration on the VSCC is observed at the same location of the σ-hole in electrostatic potentials.[33]

Copyright © 2016 John Wiley & Sons, Ltd.

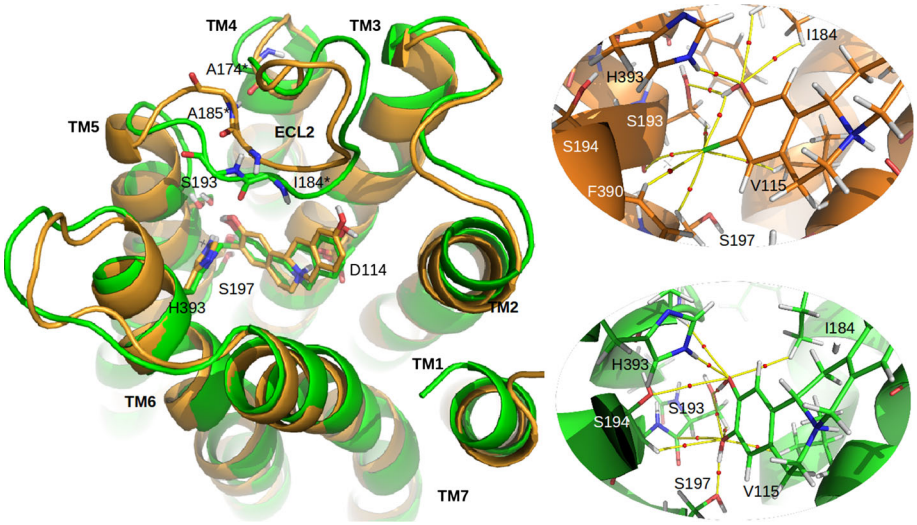

Figure 5. Backbone superposition of DRD2 bound to 3-LX in orange and 3-LOH in green (left side) and charge density molecular graphs showing the interactions of the halogen and hydroxyl groups (right side). The molecular graphs were constructed from the potential energy minimum of the MD trajectory. Yellow lines connecting the nuclei are the bond paths, and the small red spheres on them are the bond critical points (BCPs). The extra-point representing the σ-hole is depicted with a small gray cross

Contour plots in Figure 6 show the Laplacian distribution for the X-bonded complexes 2-LX1/DRD2 and 3-LX/DRD2.

The figure shows the presence of a bond critical point and the bond paths connecting the chlorine and the protein backbone oxygen atom, in both complexes. The structures shown in this figure represent different points in the MD trajectories to that shown in Figures 4 and 5. In these structures the chlorine atom is closer to O@Ser193 so that the X-bond formed competes with the main chain H-bond O@Ser193•H@Ser197 that defines protein secondary structure. The small charge density value at the BCP of this last interaction (in atomic units) indicates that it is almost completely broken because of X-bond formation.

The Laplacian distribution shows a point of minimum charge concentration on the VSCC of the chlorine atom that is pointing toward the VSCC of O@Ser193. Also note that the VSCC of chlorine is thinner in the region of the minimum charge concentration. Thus, chlorine nucleus is less shielded in that region of the VSCC allowing the electrostatic attraction with O@Ser193. It should be noted that the points of maximum charge concentration on the VSCC of O@Ser193, which are associated with the lone pairs of the Lewis base, are located in the plane perpendicular to the one that is shown in the figure (not shown). While these points of maximum charge concentration (lump) should be aligned with the minimum charge concentration on the halogen (hole) for a proper hole-lump interaction to take place, it has been shown that backbone carbonyl oxygen atoms tolerate a broad range of X-bonding geometries with high binding energy. Thus, unlike the σ-hole angle preferences, the angular dependence around the acceptor is very week, which has sense because the backbone is often fixed within more rigid secondary structure elements and cannot easily adapt to the ligand.[4]

Charge density evaluation of complexes stability

Besides the display of the molecular graphs and Laplacian contour plots, the charge density values at the intermolecular critical points (ρ_b) are a good estimate of the strength of the intermolecular interactions. Moreover, it has been proven that the sum of the ρ_b values ($\Sigma\rho$) for all the interactions established by the ligand molecule is a good measure of its anchoring strength to the binding pocket. Furthermore, the $\Sigma\rho$ value corresponding to the interactions established by a particular group of atoms of the ligand molecule is a measure of the anchoring strength of that group to the binding pocket.[9,31,34] The ability to decompose the interaction energy in contributions by atom or group of atoms makes the QTAIM analysis particularly useful in analysis, design and optimization of ligand molecules.

Bar plot in Figure 7 shows the $\Sigma\rho$ values for the interactions of the substituted group (i.e. X/OH) as well as the sum of all the intermolecular interactions for each LX/LOH pair.

As can be seen in the figure, the substitution of the hydroxyl group by an halogen follows the expected trend in 1-LX/1-LOH and 2-LX1/2-LOH pairs, in the first case the halogen improves the binding at the substitution point as well as the total ligand binding, while in the second case the opposite occurs. In both cases the trend is in agreement with the experimental data in Table 1.

On the other hand the replacement of OH by X in the remaining pairs does not improve the ligand binding as indicated by the binding assays. This discrepancy might be reflecting an oversimplification of the interaction model employed, in which only the intermolecular interactions are considered. Sometimes biological receptors experience strong conformational changes after ligand binding (i.e. induced fit effects). In those situations not only the intermolecular but also the intramolecular interactions should be taken into account to judge the overall stability of the ligand–receptor complex.

Accordingly, Figure 8A shows in stacked bars the sum of the electronic charge density values ($\Sigma\rho$) at the intermolecular (DRD2-L) and intramolecular (intra-DRD2 and intra-L) bond critical points computed on the reduced models of the complexes.

The value of the charge density corresponding to the total height of the stacked bars can be considered as a measure of the overall stability of the ligand–receptor complexes. These values classify correctly the compounds according to the experimental affinity data in Table 1. Thus, $\Sigma\rho$(1-LX) $>$ $\Sigma\rho$(1-LOH) within the

Copyright © 2016 John Wiley & Sons, Ltd.

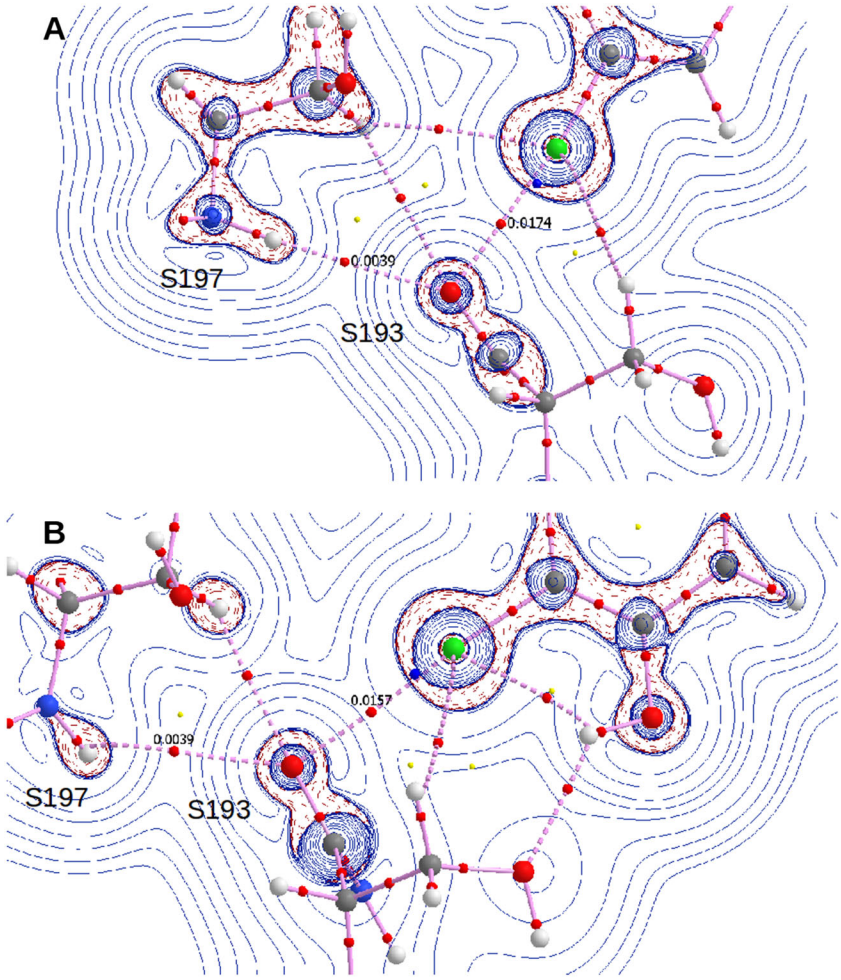

Figure 6. Laplacian distribution contour plots superimposed to the charge density molecular graphs of complexes 2-LX1/DRD2 (A) and 3-LX/DRD2 (B). Molecular graph elements are represented with critical points (bond and ring critical points in red and yellow spheres, respectively) and bond paths that extend from the BCPs to the bonded nuclei. The Laplacian distribution is shown with blue solid curves and red dashed curves for regions of charge depletion (i.e. $\nabla 2\rho(r) > 0$) and charge concentration ($\nabla 2\rho(r) < 0$), respectively. A point of minimum charge concentration (blue sphere) can be observed on the VSCC of the chlorine atom. The structures shown in this figure represent different points in the MD trajectories to that shown in Figures 4 and 5

haloperidol-like compounds, $\Sigma\rho(2\text{-LX2}) > \Sigma\rho(2\text{-LOH}) > \Sigma\rho(2\text{-LX1})$ within the dopamine-like analogues and $\Sigma\rho(3\text{-LX}) > \Sigma\rho(3\text{-LOH})$ within the THP derivatives.

An interesting trend that is observed among the three series of compounds is the inverse relationship between the ligand–receptor intermolecular interactions (i.e. DRD2-L, blue bars) and the receptor–receptor interactions (i.e. intra-DRD2, red bars). In general, as the charge density value for the intermolecular interactions rises within each series of halogenated/hydroxylated analogues, the intra-receptor interactions get weaker. A more detailed analysis of the intra-receptor interactions reveals that the segment of the receptor that is more sensible to ligand binding is TM5.

The plot in Figure 8B shows the inverse relationship between the intermolecular and intra-TM5 interactions. As depicted in that plot, the stronger the binding of the ligand to the receptor binding pocket, the greater the weakening of the intra-TM5 interactions, within each series. The superposed complexes in Figures 3–5 reveal how these changes are manifested structurally. As can be seen in these figures, the ligand showing the strongest binding to the receptor within each series (i.e. 1-LX/DRD2, 2-LX2/DRD2 and 2-LOH/DRD2 and 3-LOH/DRD2) are also the ones that present the closest contacts with ECL2. The movement of ECL2 toward the ligand pocket in these complexes allows the formation of an H-bond between backbone of Ile184 from ECL2 and His393 from TM6 (see Figs 3–5). This inter-domain interaction places ECL2 loop over the top of the binding pocket, closing the cavity and ensuring the correct ligand binding. This is evidenced also by the stronger interaction with Asp114 (see Fig. 8C) in these complexes. The salt bridge involving Asp114 is considered as the guideline interaction for anchoring of aminergic ligands into the DRD2 binding site.[12]

On the other hand, the ligand interaction with ECL2 in these complexes also causes the partial unfolding (and therefore the weakening) of TM5 in its outer end (see Figs 3–5). These findings might explain in part the inverse relationship observed between the intermolecular and the intra-TM5 interactions. However, the ligand interaction with ECL2 could not be the driven force for TM5 unfolding because there are no structural differences between the ligand LX/LOH pairs in the interaction interface with ECL2. The only structural difference among the pairs is the halogen/hydroxyl group, which is close to TM5, so that there must be some direct interactions between the halogen atom

Copyright © 2016 John Wiley & Sons, Ltd.

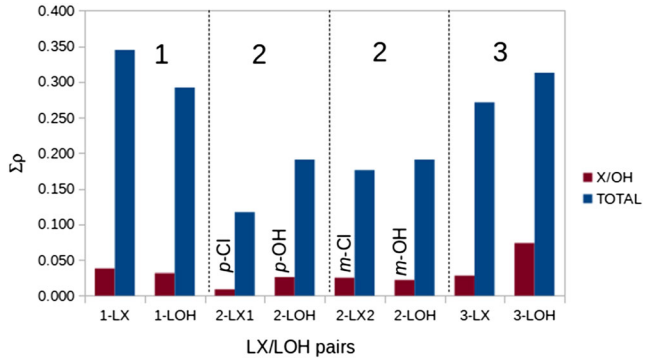

Figure 7. Σρ values for the interactions of the substituted group (i.e. X/OH) and for the total number of intermolecular interactions in each LX/LOH pair. The values were computed from the structure of the potential energy minimum of the MD simulations. Note that 2-LOH appears twice because the substituted OH group is different in each pair (i.e. *meta*-OH and *para*-OH of the dopamine catecholic ring)

(or otherwise the hydroxyl group) and TM5 that influence the folding of that transmembrane segment at its extracellular end. In turn, the folding/unfolding of TM5 then might place the ECL2 segment far away/close to the ligand.

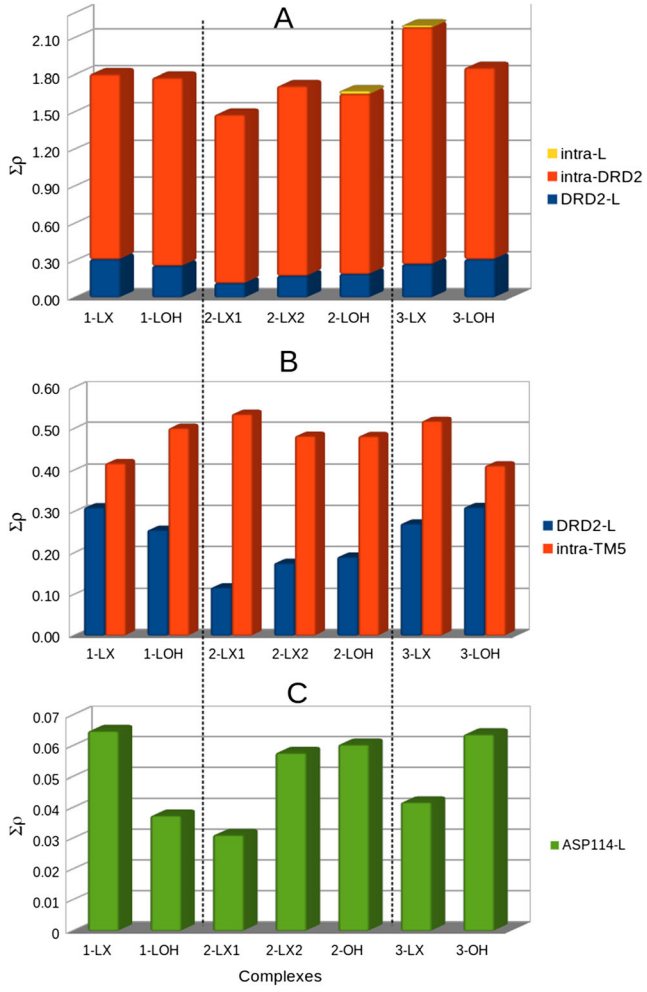

Figure 8. Σρ values corresponding to all intra and intermolecular BCPs (A), all TM5 intramolecular BCPs (B) and the intermolecular BCPs involving Asp114 (C). The values were computed from the structure of the potential energy minimum of the MD simulations

Structure of the outer (extracellular) segment of TM5

Before discussing the ligand effect on TM5 folding we consider convenient to analyze the overall architecture of the outer part of this transmembrane segment. By analyzing the MD trajectories of the studied complexes we noted that in general the secondary structure of the outer segment of TM5 is quite distorted and it shows a tendency to unfolding. The distorted appearance of TM5 in its extracellular end is, at least in part, a consequence of a proline residue P201 which is halfway between the inner and outer end of TM5.

Figure 9(A) show the location of P201 as well as relevant main chain interactions of the outer segment of TM5.

Proline residues in transmembrane segments are believed to be involved in GPCR activation. They act as pivot points for helix movements and bending that occur during the receptor activation process.[35] DRD2 receptor has proline residues in TM2, TM4, TM5, TM6 and TM7. In particular, P201 at TM5 as having no N-H group available, it cannot donate a proton to the C=O group of S197 four residues earlier to conform a proper i→i+4 alpha helix turn. Instead, the following residue in the helix Val200 donates the proton to Ile195 to close an i→i+5 H-bond loop of 16-atoms. As a consequence a broader helix turn that

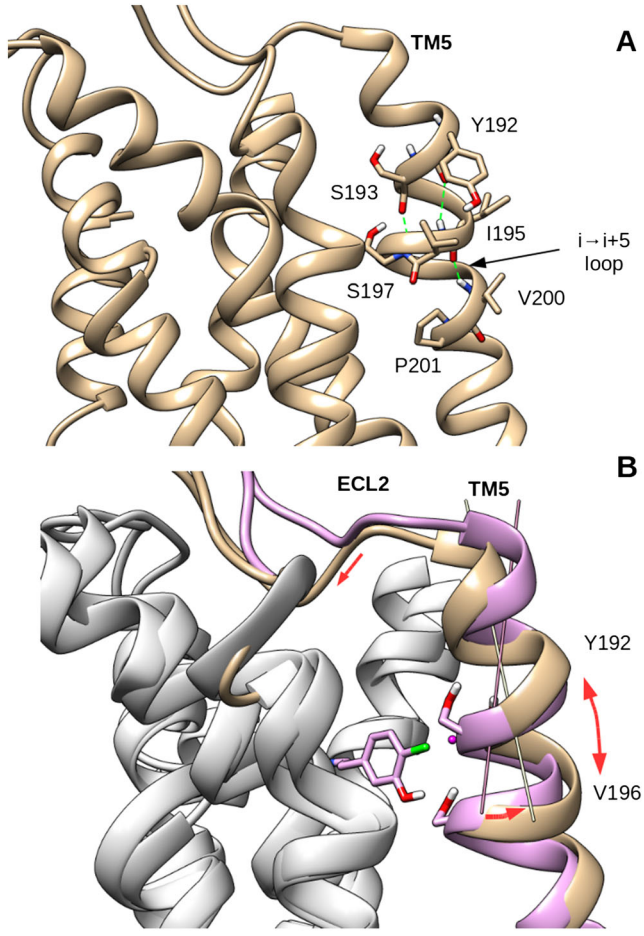

Figure 9. (A) Structure of the outer segment of TM5 in the unbound form of the dopamine receptor and (B) backbone superposition of the unbound state (gray) and bound state of DRD2 in complex with 2-LX1 (magenta). Principal axis of the outer segment of TM5 helix is shown for both states. The extra-point of 2-LX1 is indicated with a small purple sphere

Copyright © 2016 John Wiley & Sons, Ltd.

protrudes into the ligand binding site (see Fig. 9) is formed. Moreover, this local defect in helix folding is transmitted beyond the involved turn impacting on the overall folding of the outer part of TM5.

Effects of halogen bonding in TM5 folding

An interesting trend that was observed in two out of the three series of complexes is that the halogen atom, when interacting with backbone of TM5, it tends to attenuate the unfolding propensity of the outer (extracellular) side of TM5. Figure 10 is intended to show the ligand effect on architecture of the outer part of TM5 through the MD simulations.

Left column of the plot shows either the X-bond distance in halogenated ligands or the H-bond distance in the hydroxylated analogs, along simulation time.

As can be seen, 2-LX1 and 3-LX are X-bonded to O@S193 most of the simulation time while their hydroxylated counterparts 2-LOH and 3-LOH are not H-bonded to that TM5 residue. Moreover, 3-LX2 does not form any X-bond with the main chain of TM5 either because chlorine atom is anchored in a different sub-pocket formed by hydrophobic residues from TM5 and ECL2 (see Fig. 4).

Because the X-bonded O@Ser193 is engaged in a intra-TM5 H-bond with backbone of Ser197 the effect of X-bond formation is the destabilization of the intra-TM5 backbone interaction as evidenced by the larger distance and fluctuation of the H-bond O@S193•H@S197 in complexes of 2-LX1 and 3-LX as compared with the unbound state of DRD2 and its complexes with 2-LOH, 2-LX2 and 3-LOH that do not interact with TM5 backbone (Fig. 10, second panel, from left to right).

Unexpectedly, the destabilization of the O@S193•H@S197 H-bond because of X-bond formation is not followed by a distortion in the architecture of TM5 at its outer side. Instead, the overall backbone interactions are maintained or even reinforced in some cases, as evidenced by the shorter average main chain H-bonds in complexes of 2-LX1 and 3-LX with respect to the analogs 2-LOH, 2-LX2 and 3-LOH respectively (Fig. 10, third panel, from left to right). The average H-bond distance was computed considering the O@res$_i$…H@res$_{i+4}$ distance of the last six backbone interactions of TM5 at its outer end, where i runs from P187 to Y192.

A more careful inspection of TM5 backbone interactions along simulation time reveals that the main chain H-bond O@Y192•H@V196 is highly sensitive to the changes in the O@S193•H@S197 H-bond distance. The weaker the last interaction (i.e. longer H-bond distance), the stronger the first one (i.e. shorter H-bond distance) and vice versa. This trend is observed in all the complexes. As can be observed in Figure 10, the H-bond O@Y192•H@V196 is partially disrupted in the unbound form of the receptor and in complexes of 2-LOH, 2-LX2 and 3-LOH but it is stabilized in complexes of 2-LX1 and 3-LX.

A possible explanation for this paradoxical behavior is that the H-bond O@S193•H@S197 which is disrupted by X-bond formation, lies on the defective interaction interface between the broader i → i + 5 turn produced by P201 and the outer segment

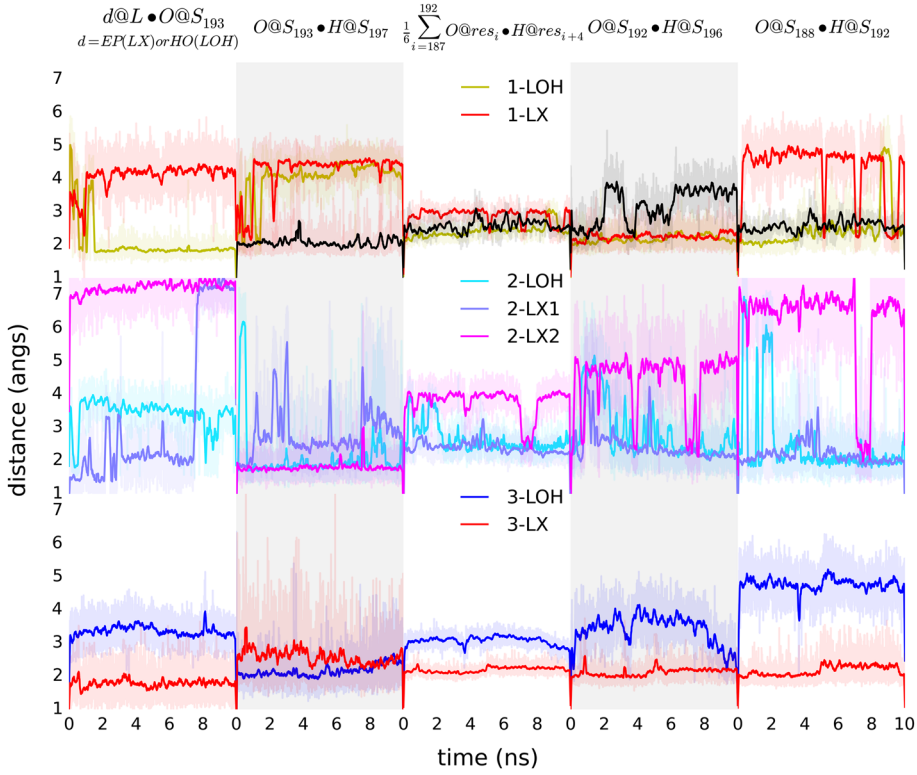

Figure 10. Ligand effect on folding of the outer side of TM5 along simulation time. From left to right: first panel shows the distance from carbonyl oxygen of S193 (O@S193) to the extra-point or to hydroxyl hydrogen atom in halogenated and hydroxylated analogs, respectively; second panel shows the distance from O@S193 to amide hydrogen of S197 (H@S197); third panel shows the average O@resi•H@res$_{i+4}$ distance for the last six main chain interactions in the outer end of TM5 with i running from P187 to Y192 and fourth and fifth panel show the distances of main chain interactions involving residue Y192. Interaction distances for the unbound form of DRD2 are depicted with black lines in the top panels

Copyright © 2016 John Wiley & Sons, Ltd.

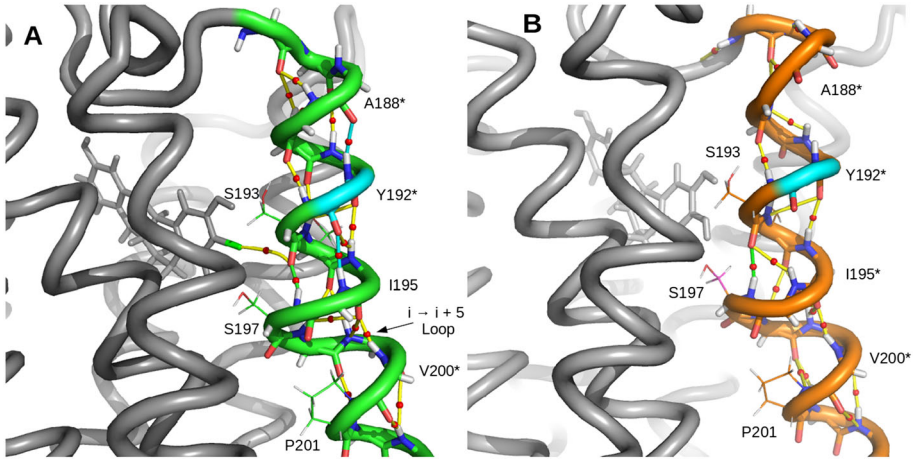

Figure 11. Structure of the outer segment of TM5 in complexes 3-LX/DRD2 (A) and 3-LOH/DRD2 (B). Charge density topological elements for non covalent interactions of TM5 backbone atoms are depicted as well as the intermolecular halogen bond. BCPs are depicted with small red circles over the bond paths which are shown in yellow except for those connecting S193 to S197 which are highlighted in green and the bond paths linking residue Y192 to V196 and A188 which are highlighted in cyan. Note that these last bond paths are absent in complex 3-LOH/DRD2. The lack of the two Y192 backbone interactions results in a distortion of TM5 secondary structure in its outer end. For residues labeled with an asterisk only the backbone atoms are shown for clarity

of TM5 having normal i → i + 4 folding, so that the disruption or weakening of this interaction might allow the readjustment of the remaining backbone interactions at this defective interface, especially the H-bond O@Y192•H@V196.

Because the H-bond O@Y192•H@V196 is located at the back of TM5, namely in the opposite side to the binding site, when it is disrupted it causes the bending of the outer part of TM5 toward the binding site as illustrated in Figure 9(B). As indicated in that figure, the bending of this segment could bring the ECL2 in contact with the ligand which might enhance the unfolding of TM5 at its extracellular end. Conversely, the strengthening of the H-bond O@Y192•H@V196, such as observed in complexes of 2-LX1 and 3-LX, causes the straightening of the outer part of TM5 and the opposite changes are observed, i.e. ECL2 moves away from the ligand binding site and the outer part of TM5 is refolded.

Another interaction that is involved in the folding of the outer end of TM5 is the main chain H-bond O@A188•H@Y192. This interaction is formed in the X-bonded complexes 2-LX1/DRD2 and 3-LX/DRD2 but it is disrupted in complexes of 3-LOH and 2-LX2 (Fig. 10, last to panels, from left to right). Because this interaction is preserved in the unbound state of DRD2 it is likely that the disruption of this interaction is secondary to the ligand pulling from the ECL2.

Figure 11 clearly shows how the presence of the two Y192 backbone interactions (i.e. O@Y192•H@V196 and O@A188•H@Y192) restores the folding of TM5 at its outer end in complex 3-LOX/DRD2 as compared with complex 3-LOH/DRD2 where these interactions are absent.

On the other hand, the MD trajectories for complexes of the haloperidol-like analogs 1-LX and 1-LOH show outliers results for this scaffold. Left panel in Figure 10 shows that 1-LX is not X-bonded to the backbone of TM5 according to distance measurement from the halogen extra-point to O@Ser193 while 1-LOH is H-bonded to that TM5 backbone atom. However, in both complexes the main chain H-bond O@S193•H@S197 is completely broken most of the simulation time regardless of whether Ser193 carbonyl oxygen is bonded or not to ligand atoms. Therefore, the differences in TM5 folding in complexes of 1-LX and 1-LOH cannot be explained in the same way as in complexes of 2-OH/2-LX1/2-LX2 and 3-LOH/3-LX. Steric effects caused by the bulky bromine atom on TM5 seem to be related with the distortion of the secondary structure of this transmembrane segment in complex 1-LX/DRD2.

CONCLUSIONS

In this work, Halogen Bond (X-bond) interactions formed by known halogenated ligands (LX) of the Dopamine Receptor D2 (DRD2) were studied at the receptor binding pocket. To prove our hypothesis that halogen bond interactions might play an important role in biomolecular systems, we compared interactions of the halogenated ligand with the corresponding interactions in the hydroxylated analog. The complexes were subjected to MD simulations with explicit treatment of the σ-hole on the halogen atom. Then, the interactions were analyzed along the MD trajectory with the help of the QTAIM theory.

The results show that the replacement of an hydroxyl by halogen does not change drastically the overall binding mode of the ligand into the receptor. However a closer look at the molecular interactions reveals differences in the interaction patterns. The halogen atom tends to act as X-bond donor (i.e. through the σ-hole) with protein backbone carbonyl oxygen atoms. The geometrical as well as the QTAIM analysis revealed that two out of the four halogenated ligands studied (i.e. 2-LX1 and 3-LX) form a specific X-bond with the carbonyl oxygen of Ser193.

On the other hand, when analyzing the effect of substitution of hydroxyl by halogen on the anchoring strength of the ligand to the receptor binding pocket, we found that neither the intermolecular interactions of the substituted group itself (i.e. halogen or hydroxyl group) nor the totality of the ligand intermolecular interactions can explain the trend of the experimental affinity data. However, when the receptor intramolecular interactions are also considered, the theoretical predictions correlate well with the experimental results. This indicates that some conformational changes are taking place in the receptor structure because of ligand binding. Thus, the stability of the complexes results from an interplay between inter and intramolecular interactions. In particular, it was found that TM5 is very

Copyright © 2016 John Wiley & Sons, Ltd.

sensible to ligand binding: the stronger the binding of the ligand to the receptor binding pocket, the greater the weakening of the intra-TM5 interactions.

By analyzing the MD trajectories we noted that the secondary structure of the outer segment of TM5 is quite distorted, and it shows a tendency to unfolding. It was found that in complexes engaged by an X-bond to O@Ser193 (i.e. 2-LX1 and 3-LX), this interaction decreases the unfolding propensity of the outer side of TM5. In the case of 3-LX/DRD2 the stabilization of TM5 because of X-bond formation might explain why this complex is more stable than 3-LOH/DRD2 even when the intermolecular interactions are stronger in the last complex.

Thus, while X-bonds involving carbonyl oxygen are generally only comparable in strength to weaker C—H···O H-bonds, the ability of the halogen atom to interact with the protein backbone might trigger conformational changes that in turn could modify the overall stability of the ligand–receptor complex.

Acknowledgements

We acknowledge SECyT-UNNE and CONICET for financial support. The authors also acknowledge the use of CPUs from the High Performance Computing Center of the Northeastern of Argentina (CECONEA). The Geforce GTX Titan X GPU used for this research was donated by the NVIDIA Corporation. This work was supported by the Grants PIP-00678 CONICET and PI F017-2014 SECyT-UNNE.

REFERENCES

[1] M. O. Zimmermann, A. Lange, R. Wilcken, M. B. Cieslik, T. E. Exner, A. C. Joerger, P. Koch, F. M. Boeckler, *Future Med. Chem.* **2014**, *6*, 617–639.
[2] H. Kubinyi, *J. Recept. Sig. Transd.* **1999**, *19*, 15–39.
[3] Science Daily Website: www.sciencedaily.com/releases/2013/01/130118064729.htm.
[4] R. Wilcken, M. O. Zimmermann, A. Lange, A. C. Joerger, F. M. Boeckler, *J. Med. Chem.* **2013**, *56*, 1363–1388.
[5] M. Kolář, P. Hobza, A. K. Bronowska, *Chem. Commun. (Camb.)*. **2013**, *49*, 981–983.
[6] J. J. Irwin, T. Sterling, M. M. Mysinger, E. S. Bolstad, R. G. Coleman, *J. Chem. Inf. Model.* **2012**, *52*, 1757–1768.
[7] M. A. Soriano-Ursúa, J. O. Ocampo-López, K. Ocampo-Mendoza, J. G. Trujillo-Ferrara, J. Correa-Basurto, *Comput. Biol. Med.* **2011**, *41*, 537–545.
[8] M. A. A. Ibrahim, *J. Comput. Chem.* **2011**, *32*, 2564–2574.
[9] E. L. Angelina, S. A. Andujar, L. Moreno, F. Garibotto, J. Párraga, N. M. Peruchena, N. Cabedo, M. Villecco, D. Cortés, R. D. Enriz, *Mol. Inform.* **2015**, *34*, 28–43.
[10] J. Párraga, N. Cabedo, S. A. Andujar, L. Piqueras, L. Moreno, A. Galán, E. Angelina, R. D. Enriz, M. D. Ivorra, M. J. Sanz, D. Cortés, *Eur. J. Med. Chem.* **2013**, *68*, 150–166.
[11] A. Mansour, F. Meng, J. H. Meador-Woodruff, L. P. Taylor, O. Civelli, H. Akil, *Eur. J. Pharmacol.—Mol. Pharmacol.* **1992**, *227*, 205–214.
[12] S. A. Andujar, R. D. Tosso, F. D. Suvire, E. Angelina, N. M. Peruchena, N. Cabedo, D. Cortes, R. D. Enriz, *J. Chem. Inf. Model.* **2012**, *52*, 99–112.
[13] H. Sun, L. Zhu, H. Yang, W. Qian, L. Guo, S. Zhou, B. Gao, Z. Li, Y. Zhou, H. Jiang, K. Chen, X. Zhen, H. Liu, *Bioorg. Med. Chem.* **2013**, *21*, 856–868.
[14] Q. Wang, R. H. MacH, R. R. Luedtke, D. E. Reichert, *J. Chem. Inf. Model.* **2010**, *50*, 1970–1985.
[15] D. A. Case, T. E. Cheatham, T. Darden, H. Gohlke, R. Luo, K. M. Merz, A. Onufriev, C. Simmerling, B. Wang, R. Woods, *J. Comput. Chem.* **2005**, *26*, 1668–1688.
[16] V. Katritch, K. A. Reynolds, V. Cherezov, M. A. Hanson, C. B. Roth, M. Yeager, R. Abagyan, *J. Mol. Recognit.* **2009**, *22*, 307–318.
[17] R. F. W. Bader, *Atoms in Molecules: A Quantum Theory*, Oxford University Press, New York, **1990**.
[18] T. Lu, F. Chen, *J. Comput. Chem.* **2012**, *33*, 580–592.
[19] T. A. Keith, *AIMAll (Version 15.05.18)*, TK Gristmill Software, Overland Park KS, USA, **2015** (aim.tkgristmill.com).
[20] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery Jr. , J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, O. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski, D. J. Fox, *Gaussian 09, Revision A.01* Gaussian, Inc., Wallingford CT, **2009**.
[21] A. D. Becke, *J. Chem. Phys.* **1993**, *98*, 5648–5652.
[22] C. Lee, W. Yang, R. G. Parr, *Phys. Rev. B.* **1988**, *37*, 785–789.
[23] S. H. Vosko, L. Wilk, M. Nusair, *Can. J. Phys.* **1980**, *58*, 1200–1211.
[24] F. J. Devlin, J. W. Finley, P. J. Stephens, M. J. Frisch, *J. Phys. Chem.* **1995**, *99*, 16883–16902.
[25] S. Grimme, *J. Comput. Chem.* **2006**, *27*, 1787–1799.
[26] R. Ditchfield, W. J. Hehre, J. A. Pople, *J. Chem. Phys.* **1971**, *54*, 724–728.
[27] P. Grundt, S. L. Jane Husband, R. R. Luedtke, M. Taylor, A. H. Newman, *Bioorg. Med. Chem. Lett.* **2007**, *17*, 745–749.
[28] S. Vangveravong, M. Taylor, J. Xu, J. Cui, W. Calvin, S. Babic, R. R. Luedtke, R. H. Mach, *Bioorg. Med. Chem.* **2010**, *18*, 5291–5300.
[29] L. B. Kozell, K. A. Neve, *Mol. Pharmacol.* **1997**, *52*, 1137–1149.
[30] F. Claudi, G. Giorgioni, A. Di Stefano, M. P. Abbracchio, A. M. Paoletti, W. Balduini, *J. Med. Chem.* **1992**, *35*, 4408–4414.
[31] E. L. Angelina, S. A. Andujar, R. D. Tosso, R. D. Enriz, N. M. Peruchena, *J. Phys. Org. Chem.* **2014**, *27*, 128–134.
[32] E. A. Zhurova, V. G. Tsirelson, A. I. Stash, A. A. Pinkerton, *J. Am. Chem. Soc.* **2002**, *124*, 4574–4575.
[33] D. J. R. Duarte, E. L. Angelina, N. M. Peruchena, *Comput. Theor. Chem.* **2012**, *998*, 164–172.
[34] R. D. Tosso, S. A. Andujar, L. Gutierrez, E. Angelina, R. Rodríguez, M. Nogueras, H. Baldoni, F. D. Suvire, J. Cobo, R. D. Enriz, *J. Chem. Inf. Model.* **2013**, *53*, 2018–2032.
[35] B. Trzaskowski, D. Latek, S. Yuan, U. Ghoshdastider, A. Debinski, S. Filipek, *Curr. Med. Chem.* **2012**, *19*, 1090–1109.

SUPPORTING INFORMATION

Additional supporting information can be found in the online version of this article at the publisher's website.

Copyright © 2016 John Wiley & Sons, Ltd.