TITLE

Multidimensional analysis of extended molecular dynamics simulations shows the complexity of signal transduction by the histamine H3 membrane receptor

Short title: Molecular dynamics of signal transduction

AUTHORS

Herrera-Zuniga LD[1,4], Moreno-Vargas LM[2,4], Correa-Basurto J[1,3], Prada D[2], Curmi P[1], Arrang, JM[4], Maroun, RC[1,4*]

[1] SABNP, UMR-S U1204, INSERM/Université d'Evry-Val d'Essonne/Université Paris-Saclay, 91025 Evry, FRANCE

[2] Computational Biology and Drug Design Research Unit. Federico Gomez Children's Hospital of Mexico City, MEXICO

[3] Laboratory for the Design and Development of New Drugs and Biotechnological Innovation) SEPI-ESM, MEXICO

[4] INSERM U894, Paris, FRANCE

[*]Corresponding author

ABSTRACT

In this work, we study the mechanisms of activation and inactivation of signal transduction by the histamine H3 receptor (H3R), a 7TM GPCR through extended molecular dynamics (MD) simulations of the receptor embedded in a hydrated double layer of dipalmitoyl phosphatidyl choline (DPPC), a zwitterionic poly-saturated ordered lipid. Three systems were prepared: the apo H3R, representing the constitutively active receptor; and the holo-systems: the H3R coupled to an antagonist/inverse agonist (ciproxifan) and representing the inactive state of the receptor; and the H3R coupled to the endogenous agonist histamine and representing the active state of the receptor.

An extensive structural and dynamical analysis of the MD simulation trajectories shows that the three states of H3R present important structural and dynamic differences in several geometric and energy properties and that the behavior of this system is complex given that the measured properties interact in multiple and inter-dependent ways. For instance, rotamer toggle switches involved in the mechanism are multiple and not just single nor double, as reported before. In addition, the MD simulations describe an unexpected escape of histamine from the binding site, in agreement with the experimental rapid off-rates of agonists.

KEYWORDS: 7TM receptors, GPCR, signal transduction, histamine H3 receptor, histamine, structural bioinformatics, molecular dynamics, molecular complexity

INTRODUCTION

The G-protein coupled receptors (GPCR) is the largest family of integral-membrane signaling receptors containing seven transmembrane helices and belonging to the superfamily of 7TM receptors[1–3]. The histamine receptors (H1, H2, H3, and H4) are members of the biogenic amine receptor subfamily of GPCRs. Their interaction with histamine, a neurotransmitter, elicits a variety of physiological effects, including allergic reactions (H1)[4], gastric acid secretion (H2)[5], mediation of neurotransmitter release and the inhibition of cAMP production (H3)[6,7], and immunological response (H4)[8]. Histamine, the endogenous agonist of this family of receptors is a neurotransmitter synthetized and released by histaminergic neurons (Fig. 1a). It plays a major role in cognition and in other physiological functions such as vigilance, attention, impulsivity and feeding/weight regulation[9]. It is stocked in vesicles and released after an electrical stimulus. Histamine will bind pre- or post-synaptic receptors. Ciproxifan[10–12] (CPX or FUB-359; CAS No. 184025-18-1; GRAC database, guidetopharmacology.org; Fig. 1b) is a highly potent and selective competitive H3-receptor antagonist/inverse agonist with pro-cognitive properties and a nanomolar affinity[13]. For a given signaling assay used, inverse agonism refers to the ability of a compound to inhibit constitutive GPCR signaling, presenting thus negative efficacy.

The H3 receptor was originally discovered in the brain on histaminergic neurons as a presynaptic auto-receptor inhibiting the synthesis and the depolarization-induced release of histamine[7]. H3R is predominantly expressed in the CNS and to a lesser extent in the peripheral nervous system[14].

The cloning of the histamine H3 receptors cDNA in 1999[6] allowed detailed studies of its molecular aspects and indicated that the H3R can activate several signal transduction pathways. The H3R is regarded as a potential therapeutic target because of its location in the CNS and for the modulation of a variety of functions such as cognitive processes, epilepsy, food intake and sleep-wakefulness[15,16]. The transmembrane region of H3R is often the site of ligand and drug interaction. Several H3R antagonists/inverse agonists appear to be promising drug candidates[17–19]. Three-dimensional (3D) atomistic models of antagonist-receptor complexes have been used to investigate the details of ligand and drug interactions with H3R and have been successful in providing important insights regarding their binding; additionally, several groups have reported the features of the general H3R pharmacophore. This approach has been particularly successful for investigating GPCR/ligand binding modes and is complementary to 3D receptor/ligand modeling. The features of this antagonist

pharmacophore are a primary basic group, either a piperidine or pyrrolidine, which is connected by an alkyl linkage to a second group; other groups observed that the addition of a second basic group increases the binding affinity[20]. Across the superfamily of GPCRs, there exist many residues that have been conserved throughout evolution and are thus thought to play key roles in receptor structure and/or function. Site-directed mutagenesis has demonstrated the importance of many of these residues in several different biogenic amine receptors, including some of the histamine receptors. For example, mutation to Ala of the conserved Asp 2.50 in a multitude of GPCRs had pronounced effects on binding the positively charged biogenic amine ligands and on receptor signaling. In this case, Asp 2.50, involved in sodium ion coordination, is responsible for an indirect allosteric regulation of ligand binding by directly affecting the mobility of sodium ions in the sodium pocket[21] (and references therein). Indeed, human H3R is sensitive to monovalent cations such as sodium[22]. The interaction of the ligand with Asp 2.50 facilitates binding to residues in other TM domains, particularly TMs 4 through 7. The critical role of TM5 has been demonstrated in many receptors (β2-adrenergic, H1, H2, etc.) including the H3R[23]. Sequence alignments show that TM5 of the histamine receptors is poorly conserved, suggesting a potential difference in the mechanism in which histamine binds to the H3 receptor. To investigate the interaction of histaminergic ligands with TM5 of the H3 receptor, several groups have performed in the past an alanine scan of the first 14 amino acids of TM5 and investigated their effect using radioligand binding, cAMP assays, and the fluorometric imaging plate reader (FLIPR), in which receptor activation is coupled to calcium mobilization. Modeling of this complex suggested that Glu 5.46 (206) in TM5 plays key roles in agonist-induced activation of the H3 receptor through a reduction in potency and affinity[23].

On another hand, the existence of distinct active and inactive conformations of the H3R has been established *in vitro* and *in* vivo via the pharmacological concept of protean agonism[24,25].

To the best of our knowledge, the crystal structure of H3R is still unavailable. Therefore, several computational studies have been carried out for determination of the binding of several antagonists to the H3R using a homology model for the receptor and employing the continuum dielectric approximation for the surrounding bilayer environment. Thus, based on the crystals of rhodopsin[26], 3D *in silico* models for H3R have been obtained in the past[20,23,27–37]. Nevertheless, these models were built using incorrect or absent alignments, or are based on other models or on the coordinates for bovine rhodopsin (a far homolog) or include manual

adjustments and manipulations or the cartesian coordinates are not available or the refinement is done *in vacuo* or the molecular dynamics (MD) trajectories are very short (1-10 ns) or do not take into consideration the state of the template structure (active/inactive). Others do not contain any ligands[38]. In other instances, modeling of GPCRs (class A, B and C) was performed on a high-throughput basis, but the models only pretended to represent receptor structures aimed at studying receptor-antagonist interactions and not their activated states[39]. In addition, only TM domains were considered. In the MemProtMD database of Sansom and colleagues[40], the structures of membrane proteins in the PDB are inserted in an explicit lipid bilayer and MD simulations in a coarse-grained (CG) representation are undertaken. The structures are taken from the PDB as they are, with missing loops and/or segments, so that the reported analyses (contacts, displacements) might not be accurate. Nevertheless, the H3R not having an experimentally-derived structure, is absent from the database. As expected, significant divergences among the models exist.

The need of dynamical models of the structure of 7TMs prompted many computational studies in recent years[41]. High performance computing for numerical simulation has become an essential tool as simulation can replace experiments. Although long MD simulations with explicit representation of the lipid bilayer and its aqueous environment are very costly computationally, several research groups report studies of this type[42–53], and are more frequent nowadays[54] given the increase in computer power and data storage, the optimization of algorithms and the development of special-purpose machines[55], allowing multi-scale simulations going from the nano- to the meso-scale[56].

In this work, we generate our own curated *in silico* 3D models of H3R in the active, inactive and constitutive states, this latter being represented by a ligand-devoid receptor and proceed to embed the receptor in a hydrated, ionized and electrically neutral DPPC phospholipid bilayer for studying its behavior more realistically. Our approach consists in docking and simulating the binding of known H3 ligand compounds to the receptor model and then in studying and analyzing the static and dynamic behavior of the resulting H3R-ligand complexes through MD simulations, which total more than 3 μs. An extensive analysis of the results of the trajectories indicates that the states of the receptor cannot be described by a few properties. Given that these properties are interdependent, the behavior of the system is complex *stricto sensu* showing an intricate network of short- and long-distance crosstalk between distinct regions of the receptor. In addition, the MD

simulations show a spontaneous escape of histamine from the binding site, preceded by a short binding step in the extracellular vestibule during the unbinding pathway. Nevertheless, histamine can return and bind from time to time to this vestibule.

In establishing the mechanism of activation of H3R, identifying the large-scale movements taking place, the structural changes conveying from the ligand binding pocket to the G protein binding site, the interaction networks taking place and how those changes propagate, is of utmost importance.

RESULTS

The template 3D structures that we used were closely related and highly homologous to H3R, justifying our knowledge-based approach for obtaining reliable models for H3R. For instance, with human β2-AR, the sequence identities between TM segments go from ~19% for TM2 to ~36% for TM1, and of 25% for the whole sequences, excluding the ICL3s. We obtained the initial structures for the two holo-systems, i.e., a histamine-H3R complex and a ciproxifan-H3R complex through ligand docking to the orthosteric site of the receptor.

We found that there were multiple bound states of the **histamine** ligand corresponding to the activated state of the receptor due to the small size of the ligand and to its ability to establish stabilizing interactions in different orientations with surrounding residues. Indeed, during the docking simulation histamine showed several binding modes of which two major poses in the orthosteric binding pocket. We chose the one compatible with experimental results with H2R in which the quaternary N_ζ of histamine interacts with Asp 3.32 (Asp98 of H2R; Asp114 of H3R), highly conserved through class A GPCRs and essential for histamine binding and action, serving as a counter-anion to the cationic amine moiety of histamine[57,58] and agonist and antagonist binding. The found histamine poses are compatible with the experimental findings of Gantz et al.[57] in which removal of the negatively charged amino acid abolished histamine stimulated increases in cellular cAMP, but not histamine binding to the receptor. For **ciproxifan (CPX)**, we chose a pose analogous to that of histamine in which the imidazole moiety is in contact with Asp 3.32. The binding mode of the two ligands is partially shared. It needs to be mentioned that there may be a secondary low-affinity binding site for agonists as seen by binding kinetics[59,60], but in this study we focus on the orthosteric site only.

Interestingly, during the MD simulation the agonist-H3R complex showed an unusual behavior in which the histamine ligand, even though at given times deep in the binding cleft, eventually escaped spontaneously from the binding site. We thus defined four intervals to better describe the unbinding process. The first interval (section 1) during which the ligand is bound to the binding cavity, comprises from nanosecond 112 of the production phase, when the root mean-square deviation of Cα atoms (RMSD) attains a plateau, to nanosecond 590. In this interval, the histamine displays, nevertheless, a very high mobility despite the solvent molecules filling the binding cleft. The N-terminus (N-ter) folds over ECL2, and ECL2 folds over the cleft. The second interval (section 2) includes the beginning of the motion of the ligand out of the H3R binding site but still bound

to the transmembrane part of the receptor -it lasts 1.15 ns and represents a transition period. The N-ter begins to unfold and ECL2 moves aside. The third interval (section 3) comprises the ligand trapped between the extra-membrane N-ter and ECL2 regions, and the anionic heads of the upper leaflet of the phospholipids; it lasts almost 1 ns. Thus, it involves the extracellular vestibule during the unbinding pathway, but it does not seem to represent a metastable binding site[61]. The binding cleft is open to the extra-cellular space. The fourth and last interval (section 4) describes the entrance of the ligand into the bulk of the solvent, starting at nanosecond 592 until the end of the trajectory at 1020 ns. The N-ter now forms the lid of the cleft, with ECL2 remaining "aside". In the bulk, histamine shows a hydration shell of about 30 water molecules and interacts eventually with a chloride ion with its alkyl amine. NOTE: From now on, the starting frame for all our analyses will be the frame at which the RMSD reaches a first plateau. For the antagonist complex, this corresponds to nanosecond 150, and for the agonist and apo receptors to nanosecond 112. Given the behavior of histamine, only the first segment of the agonist complex trajectory, when the ligand is well bound to the receptor, is considered for the analyzes (112-590 ns).

Energy landscapes and Multiplicity of conformations

Fig. 2a shows the 2D plot of the cartesian PCA obtained for the **antagonist-H3R** complex. In the presence of CPX, the PC1-PC2 plot shows a V-shaped free energy landscape with the vortex to the left and with four populations: clusters C1, C2, C3 and C4, located in quadrants I to IV, respectively (quadrant I is in the upper right-hand side; the turn is clockwise). Cluster 1 in the upper-right hand side is the one with the deepest relative energy minimum (not shown). The V-shape shows a large span of about 65 Å over PC1 and a smaller one of about 43 Å over PC2. The PC1-PC3 map (Fig. 3a) has a laid S-shape and shows no energy barrier between the lowest-energy cluster (lower center) and the closest cluster (at the center of the map). There are also four clusters. The PC2-PC3 map (Fig. 4a) shows an energy landscape in which the four clusters are centered in the map and fusion into essentially two wide clusters –a quadrant III-IV cluster with clusters C2 and C3 with not so deep wells, and C1 and C4 in quadrant I, II that fusion into a cluster containing in the right-hand region the lowest energy cluster. Fig. 2b shows the PC1-PC2 plot for the **agonist-H3R** complex when the agonist is still bound to H3R. It is composed of four clusters, with a major cluster in the center of the map and spanning a

wide range on PC1. Fig. 3b shows the corresponding PC analysis for components 1 and 3, with a deep cluster on the right-hand side of the map. The corresponding PC2-PC3 map is shown in Fig. 4b showing a reduced area and one cluster in the middle of the map. The PC1-PC2 plot for the **apo** receptor is in Fig. 2c. The occupied region in the map is rather large and diffuse, presenting several not so deep energy clusters. The map of components 1 and 3 shows a rather deep cluster in the middle of plot (Fig. 3c). The PC2-PC3 map (Fig. 4c) is rather round, with two deep clusters again, around the center of the plot.

Even though the principal component maps cannot be interpreted directly in physical terms, it is noteworthy to observe that, in general, the maps for the different systems show differences in the morphologies and the number and depths of the conformation populations due solely to the presence or absence of ligand and to its nature –agonist or antagonist- corresponding to various ensembles of a given state of the receptor. This reflects the conformational isomerism taking place for a given state of the receptor.

Characterization of the internal cavity of the receptor

Table 1 shows that there are 38 residues that form the cavity of the **CPX-bound** receptor. These residues are contributed by TM2, TM3, ECL2, TM5, TM6 and TM7. The environment of the ligand is rather hydrophobic with the three aromatic amino acid side chains being represented. Twenty-six residues compose the internal cavity of the **agonist-H3R** complex, with contributions coming this time from TM1, TM2, TM3, ECL2, TM5, TM6, TM7. For the **apo** receptor, the cavity is composed of only 24 residues. As expected, this cavity is the smallest one, given the absence of ligand and the subsequent contraction of the binding cavity. Secondary structures contributing to it are TM2, TM3, ECL2, TM5, TM6 and TM7. In all three systems, TM4 being eccentric contributes with no residues to the receptor's cavity, just like all the loops (except ECL2) and, of course, H8. Notice the contribution of ECL2 to the morphology of the binding cavity.

The residues that compose the internal cavity of the receptor in its different states are of diverse types, with Tyr and Leu being the most abundant, followed by Phe and Ser. Residues Pro, Gly, Gln, Lys, Thr and His are absent (Table 1). Interestingly, the antagonist and agonist pockets are negatively charged at 4-, whereas the

pocket of the apo receptor is neutral (Asp 3.32 + Arg 6.58). The cationic ligands histamine and CPX thus induce changes in the pockets that bring into play acidic residues.

The role of water molecules in the internal cavity

A continuous dynamic network of water molecules settles in the interior of the receptor, with waters penetrating and exiting the receptor from the extracellular region. These waters establish several H-bonded networks; nevertheless, the receptor is not a water channel since it presents a molecular plug formed by the lower-leaflet internal walls of TM2, 3, 6 and 7 and the TM7-H8 loop that keeps the internal water molecules from exiting to the cytoplasm or entering from it. About 61 water molecules are constantly present in the internal cavity of the **antagonist**-bound receptor during the trajectory, 35 for the **agonist** one, and 49 for the **apo** (Table 1), so that water penetration is largest upon antagonist binding, and smallest for the agonist-H3R complex, just like in the high-resolution structure of α2A-AR that reveals about 60 internal waters[62]. In contrast, Yuan et al.[63] mention an increased penetration of water into the receptor cavity of μ- and κ-opioid receptors, which has been linked to the activation mechanism upon agonist binding. In any event, the presence of water in the binding site clearly demonstrates its influence in the dynamics and conformation of the receptor.

Rotamer toggle switches are multiple and concerted

The goal of this section is to determine those amino acid side chains that undergo concomitant side-chain conformational changes. We focus on the aromatic side chains of Tyr, Trp and Phe, since several rotamer toggle switches dealing with those residues have been reported in the literature[64].

In Fig. 5a and Fig. 5b are shown the time evolution of side chain dihedral angle χ_1 (N-Cα-Cβ-Cγ) torsion angles of Trp 6.48 of the CWXP conserved motif, and Phe 5.47, respectively, for the **antagonist-H3R** complex. We can see that χ_1 of Trp 6.48 adopts values of -60° (corresponding to a g- configuration) most of the time, until 665 ns, when it undergoes a transition to a trans conformation (+/-180°). Concomitantly, χ_1 of Phe 5.47 undergoes a transition from trans to g- (Fig. 5c). These transitions are thus correlated. As compared to the MD simulations

of Nygaard et al.[65] (their Fig. 5a), Trp 6.48 (TrpVI:13) in our extended simulations adopts an intermediate conformation between the active and inactive states. It is interesting to notice that Phe 5.47 (PheV:13) is rather distanced from the Trp 6.48 side chain. Starting at about half of the trajectory, the average distance between the center of masses (COMs) of the two side chains is ~13 Å. Trp 6.48 is highly solvated and establishes h-bonds with water molecules. The Arg 3.50 switch of the DRF motif ($^{3.49}$AspArgPhe$^{3.51}$) is found in a bifurcated h-bond with two other residues: Asp 6.30 and Asn 5.64, coming from TM6 and TM5, respectively (Fig. 6). This interaction is equivalent to that reported for bovine rhodopsin in the inactive state[65,66] (Fig. 5c of Nygaard et al.2009), since H3R contains no Glu at position 6.40, but rather an Ile. Trp 7.43 makes incursion into the -90 and 180° states of χ_1. Tyr 7.53 (NPXXY motif) is in the active conformer state throughout the trajectory (not shown), showing no switch; it engages in a hydrophobic interaction with Ile 6.40 (Fig. 5b of Nygaard et al 2009) (Fig. 7). In a remarkable fashion, other aromatic side chains undergo simultaneous transitions in the same time interval: Trp 1.31, 23.50 (ECL1), Phe 3.51, 45.36 (ECL2) and Trp 7.43 (not shown). Thus, a multiple toggle rotamer switch of certain aromatic chains, along with bi-modal switches appears to be important for receptor H3R <u>inactivation</u> (1.31, 23.50, 3.51, 45.36, 5.47, 6.48, 7.43). For the **agonist-H3R** complex (section 1), the configuration of χ_1 of Phe 5.47 remains constant at g- whereas that of Trp 6.48 shifts from 90° at the beginning of the trajectory to -25° for the rest of the time. Concomitant transitions in χ_1 of several aromatic side chains – Phe29 of the N-ter, Trp 3.28, Tyr 45.51 (TM4-TM5 loop), Tyr 45.56 (TM4-TM5 loop), Trp 5.36, Phe 5.38 and Tyr 7.33- take place starting at about 0.5 μs. Phe 3.41 undergoes a g- to trans transition in the last fourth of the trajectory. Again, a multiple toggle rotamer switch of aromatic side chains takes place during <u>activation</u> but involves several different amino acid residues. Only Trp 6.48 is common to both mechanisms -activation, inactivation. The number of aromatic side chains that undergo conformational transitions in the **apo** receptor is much reduced and includes only Phe 45.54, Tyr 45.56, Tyr 7.33 and Phe 8.54 side chains, with the first and the third toggle switches interconverting between g+ (+60°), g- and trans states, and the second and third toggle switches common to the activation mechanism. None of the residues of the DRF motif are thus involved in H3R activation or constitutive activity implicating that the ionic lock involving Arg3.50-(Asp 6.30 and Asn 5.64) is not conserved, consistent with the lack of direct evidence for a role of that motif in those two situations[67].

In résumé, there are multiple concerted long-range rotamer switches for each system as several residues are several tens of angstroms away from each other. On the other hand, the number of aromatic residues

undergoing simultaneous transitions upon activation is larger than in the other two systems (antagonist-H3R complex and apo receptor) and involves residues from the N-ter, TM3, TM4-TM5 loop, TM5, TM6 and TM7. For the apo receptor, the aromatic side chain rotamer transitions involve just ECL2, TM7, plus the intracellular C-terminus (C-ter) of helix H8. Finally, it is interesting to observe that aromatic side chains switches from the ECL2 participate in the different mechanisms.

Inter-residue distances and novel ionic locks

We have measured many selected inter-residue distances using the COMs of the side chains to detect when proximities or interactions take place.

For the **antagonist-H3R** complex, the conserved arginine of the ionic lock (motif DRF) among rhodopsin-like G protein-coupled receptors, Arg 3.50, interacts with Asp 3.49 and Asp 6.30 in a constant but dynamic fashion throughout the trajectory, in agreement with the behavior of the ionic lock in the MD simulations of the β2-AR receptor[68]. In addition, there is formation of a hydrophobic interaction between Met 1.39 and Trp 7.40/Trp 7.43, followed by disappearance of an initial Met 1.39-Tyr 2.61 interaction; the interaction between Met 6.55 and Tyr 3.33/Phe 7.39 is sporadic. Lastly, the distribution of the Met 6.55-Tyr 6.51 distance with a major peak at 5 Å and a minor one at 9 Å is analogous to the one described in the literature[69] (Nygaard et al 2013, their Fig. 4c). For the **agonist-H3R** complex the following interactions emerge: Met 1.39-Trp 7.40, Met 6.55-Phe 5.48, while the Phe 5.47-Trp 6.48 becomes distended; Met 6.55-Tyr 6.51/Tyr 3.33, and Tyr 7.53-Phe 8.50 distances are constant throughout. For the **apo** receptor, are stable throughout the trajectory: Asn 1.50-Pro 7.50, Asn 2.39-Asp 3.49/Arg 3.50, Asp 3.32-Trp 7.43, Asp 2.50-Pro 7.50, Met 6.55-Tyr 6.51, Met 1.39-Tyr 2.61/Trp 7.40/Trp 7.43, Met 1.54-Phe 2.51, and Trp 6.48-Ser 7.46. A Met 6.55-Tyr 3.33 interaction goes gradually away. Those distances that undergo a notable change are: formation of interactions Phe 5.47-Trp 6.48, and Tyr 7.53-Phe 8.50/Phe 8.54; sporadic interactions Arg 3.50-Asn 2.40/Asp 3.49/Asp 6.30, and the Met 4.46-Trp 4.50 couple (Table 2).

In conclusion, only the apo receptor presents interactions between TM1 and TM2; and between TM2/TM3 and TM7. Interactions between TM3 and TM6 are unique to the agonist and antagonist-H3R complexes. Asp 3.32 in

the apo receptor not being involved in interaction with a ligand interacts now with a Trp residue from TM7. Notice that the Arg 3.50-Asp 3.49/Asp 6.30 h-bond and electrostatic interaction is unique to the antagonist-H3R complex. The Asn 2.39-Asp 3.49/Arg 3.50 interaction is absent in the agonist-H3R complex, just like interactions between TM7 and TM8 are absent in the antagonist-H3R complex (not shown).

Ligand-receptor interactions

For **CPX**, the carbonyl oxygen between the cyclopropane and the phenoxy ring is in a long-lasting (>70% of the trajectory) bifurcated bond with the N atoms of the main chain of Glu 45.53 and Phe 45.54 of the TM4-TM5 loop (ECL2) (Fig. 8a). Whereas, due to the mobility that the single bonds give to the imidazole ring, this one interacts at one time or another with either the internal solvent molecules or different neighboring residues, none of these interactions being significantly populated. As shown in the LigPlot+ diagram[70] of Fig. 8a, the imidazole moiety is associated to three water molecules. In Fig. 8b, this moiety is H-bonded to Asp 3.32 and a water molecule, whereas in Fig. 8c, d, e, f, the imidazole ring shows no h-bonds to water molecules. In Fig. 8e, the imidazole is H-bonded to the N3 of the indole of Trp 7.43. Other contacts along the trajectory are listed in Table 3 and include all residues within 4 Å of the CPX ligand with large residence times. The residues with 70% or more in contact with CPX include three Cys, two Phe and one of each Leu, Ala, Val, Tyr, Trp, Asp, Glu, and Gly. Notice the presence of two acidic residues -Asp and Glu. The LigPlot+ plots give an indication of the diverse conformations adopted by the imidazole-containing moiety of CPX. These conformations include an extended or all-trans chain (Fig. 8a); a "zig-zag" chain (Fig. 8b); a "paddle" chain (Fig. 8c); a "cyclic" conformation that approaches Cδ2 of the imidazole with C11 (the one bound to the ether oxygen) forming a virtual 5-membered ring (Fig. 8d); and two "closed" or cis conformations (Fig. 8e and Fig. 8f). Ciproxifan then adopts essentially six bioactive conformations. In the last frame of the trajectory, CPX has already adopted an all-trans conformation for the chain connecting the imidazole ring to the phenoxy ring (ϕ_1- ϕ_4), and for the dihedral connecting the carbonyl oxygen and the cyclopropane; ϕ_5 adopts two states: -90 to -110°, and +/-180°. The flexibility of CPX is reflected in the wide range of values the different torsion angles corresponding to single bonds can adopt during the MD trajectory: t, g+, g-, -120° (not shown).

For the endogenous ligand **histamine**, the number of residues most in contact are Asp 3.32, Trp 3.28 and 7.43, and Phe 45.54 (ECL2) (Table 4; Fig. 9 SEC1). As for the effects of the Glu 5.46 to Ala mutation[23], instead of a long-lasting direct interaction between the imidazole ring of either ligand and Glu 5.46, we suggest rather an indirect/allosteric interaction involving the sodium allosteric binding site. Indeed, thanks to the long simulation time, a potassium cation binds in the sodium allosteric binding pocket of the antagonist-H3R complex, near Asp2.50. In the last frame of the trajectory, the cation is coordinated by Asp 2.50, Asp 3.32 and by three structured water molecules (r<3.0 Å), just like in the α2A-AR[62]. The presence of the potassium cation only in the antagonist-H3R complex indicates a role in the stabilization of the inactive state, in agreement with experimental findings[71].

Finally, Arrang and co-workers[60] observed that the presence of Ca^{2+} down-regulated the H3R receptor. We attribute this effect to an allosteric binding site for Ca^{2+} at the level of the extra-cytoplasmic region and/or at the sodium binding site.

Each state of the receptor presents different lipid binding sites

To detect those lipids in direct contact with the membrane protein, we computed the residence times of the lipids about the receptor. Of a total of 9 highest-occupancy lipid molecules associated to the receptor, four upper-leaflet and five lower-leaflet lipids are bound by the receptor and are distinctly different from bulk lipids, as seen from their long residence times. As an illustration of the interaction of DPPC lipids with the inactivated H3R receptor, Fig. 10 shows a 2D LigPlot+ diagram of lipid 149, showing h-bonds between Arg 8.51 and one oxygen atom from the phosphate head, the hydrophobic interactions with surrounding residues, and the proximity of lipid 125. Both lipids share Leu 7.55 as an interacting residue. In Fig. 11, extra- and intra-cytoplasmic views of the inactivated receptor for the 490 ns frame, we can see that DPPC molecules essentially bind to opposite sides of the exposed surfaces of the receptor. On one side of the receptor TM1, 6 and 7 lodge three lower-leaflet lipids and one upper-leaflet lipid. The other side, with TM3, 4 and 5 offers another binding site for two lower- and two upper-leaflet lipids. On that side, an upper-leaflet lipid finds a binding site provided by TM3 and 4. Both leaflets contribute thus with lipids for receptor binding. For the **histamine** complex, all in all, there are five lipids in the lower part of the membrane, one in the upper part, and one in the middle. A

LigPlot+ diagram illustrates the surroundings of lipid 3: lipids 31, 48 and 66, Lys 6.32, Lys 6.35, Ile 6.39, Leu 6.46, Gly 6.45, Val 7.48, Leu 7.55 and Cys 7.56. The Lys residues interact with the phosphate moiety of the DPPC molecule, whereas the hydrophobic residues with its fatty acyl tail chains (Fig. 12a). For the **apo** receptor, there are five lipids binding to sites in the upper leaflet of the double bilayer. The LigPlot+ representation of Fig. 12b shows lipid 184 and its neighboring residues, Leu 1.42, Tyr 7.33, Glu 7.36, Thr 7.37, Trp 7.40, Ala 7.44 and Ala 7.47 most of them from TM7. Several water molecules surround the phosphate and choline moieties. Ne1 of Trp 7.40 interacts with O22 of the DPPC lipid; Leu 1.42 and Ala 7.47 interact with one of the acyl tails of the lipid.

Table 5 for the **antagonist-H3R** complex indicates that the percentage of five classes of amino acids (ILVFM, KRH, ED, STYCNQW and GAP, respectively) in contact with DPPC molecules are of 41, 19, 1.4, 24 and 14%, respectively. The hydrophobic set of residues represent the major part and will be in contact with the larger-area acyl tails of the lipids; the polar class of residues follows. The third largest population corresponds to the basic residues that can interact with the lipid head groups. For the antagonist-H3R complex, 13 high-residence-time DPPC lipids find binding sites in all TM helices except TM2 (Fig. 13). For the **agonist-H3R** complex, the following secondary structures participate to the binding of 7 lipids: ICL1s, TM2, TM5, TM6 and the TM7-TM8 loop. Interestingly, for the **apo** receptor, 5 high-residence-time lipids associated with it belong to the upper leaflet of the membrane only and to all helices, except TM5 -there is no contribution from the lower leaflet of the bilayer to lipid binding.

Additional results

In the Supplemental Data section, we describe the outcome of a multidimensional analysis of many microscopic properties of each of the three systems dealt with in this work:

- RMSD of Cα atoms and RMSD matrices
- Cα-Cα distance maps
- Rg and RMSF

- Principal Component Analysis
    - Eigenvectors
- Characterization of the internal cavity of the receptor
    - Charged residue, H-bonded and hydrophobic residue networks
- H-bonded water networks
- Hydrophobic clusters
- Mechanical properties of the helices
    - Global tilt
    - Global and local helix rotations
    - Displacement, tilt, global toggle switch,
    - Bend (Kink), Wobble, Face-shift (winding-unwinding) angles and Turn angle per residue
    - Rotation, compression-expansion, winding-unwinding and bending
    - Compression-extension movements
    - Total helix displacement
- Lipid binding sites

DISCUSSION and CONCLUSIONS

In this work, we have explored with extensive MD simulations, many structural properties and phenomena of the activation and inactivation pathways of the second step of signal transduction by the histamine H3R that are not easily accessible by experiment, including new undocumented interactions.

In our simulations, we have several indications that the sampling of the phase space was complete since the average values of the properties of interest remain roughly constant with increased sampling. Although many functional mechanisms appear to be intrinsically defined by the 3D structure of the protein itself, the effects of the interactions with the environment, such as the lipids and water molecules, are also crucial for membrane proteins. An apparent paradox is that we can observe significant changes during the MD simulations that span 1 μs, whereas the time constant for the activation switch of a GPCR can be in the tens to hundreds of ms. For example, for α2A-AR in living cells the rate constant was < 40 ms[72]; these authors also refer to other results that indicate that the entire GPCR-signaling chain can be activated within 200-500 ms. But these numbers depend on the intrinsic properties of the receptor at hand, the type of ligand and the mode of ligand binding, additional interactions of receptors with other receptors and presence of adapter proteins[73], so that different types of ligands cause changes in FRET[74] (Forster resonance energy transfer) with very different speeds[75]. The time-scale difference between theory and experiment in this case is due to the fact that we have studied only a fragment of the entire GPCR signaling cascade whose individual elements comprise at least (1) ligand binding, (2) conformational change of the receptor(activation or inactivation), (3) interaction between ligand-complex receptor and G protein, (4) G protein conformational changes including GDP release and GTP binding, (5) G protein-effector interaction, (6) change in effector activity, and (7) the resulting ion conductance or second messenger concentration changes (Lohse et al British journal of Pharmacology 2008). Steps (1) to (3) are consistent with the ternary complex model for GPCRs [76,77] and correspond to the Ri ⇔ ARi ⇔ ARa schemes of the extended or cubic complex models (Ri, initial state unbound receptor; ARi, ligand-complexed receptor in the initial state; ARa, ligand-complexed receptor in the active-inactive state). We can add receptor oligomerization between steps (2) and (3). But our simulations deal with the second step only -the initial structures represent already the docked complexes- and can be considered to span it well, even though the H3R in any of the two ligand-bound described states is not in a full-active or inactive state since it is not bound to its G-protein (Gi/o).

An illustration of this step is the β2-AR, whose agonist-induced conformational changes include a large outward movement at the cytoplasmic end of TM6 and an α-helical extension of the cytoplasmic end of TM5[78,79]. This is valid also for H3R as seen in the superposition of the three states of the receptor in their representative structures, and the respective collective motions of some of the helices, such as TM5, 6 and 7 (**Fig.** 14), supporting the view that several receptor collective movements are common throughout class A GPCRs.

Ghamari et al.[80] built a 3D model of human and rat H3R (with N-ter absent and with ICL3 modeled as the lysozyme fusion protein in the crystal structure) by the homology comparative approach using the M3 muscarinic acetylcholine receptor in complex with tiotropium, a potent muscarinic inverse agonist (PDB Code 4DAJ). Thus, the model, just like the template receptor, is in the inactivated state. For the structurally conserved segments of the molecule, i.e. the TM helices, the RMSD (as estimated by PyMol's Align function) with our ciproxifan-bound H3R is of 3.2 Å, the main deviations located at TM4 and 5. Over the years, Stark, Sippl and coworkers have published several *in silico* 3D models of H3R for different purposes[34,81]. The 2001 publication, a study of receptor–antagonist interactions, presented the first models of rH3R and hH3R and was based on the rhodopsin structure. In the 2005 work[81], the rhodopsin-based homology model of hH3R focused only in the hH3R binding pocket suitable for virtual drug design. The 2007 homology model of hH3R was also based on the crystal structure of bovine rhodopsin and was simulated by small MD trajectories in a DPPC/water membrane mimic, in view of obtaining binding sites not overfitted to a given ligand. Different binding modes with different ligands were obtained by allowing side-chain rotamerism, but ciproxifan was not among the ligands probed. Yet, we find several features in common with our rH3R binding site, including the presence of important residues in the binding pocket, such as Asp 3.32, Tyr 3.33, and Glu 5.46. Levoin et al.[29] generated several hH3R models based on the crystal structure of hH1R (PDB Code 3RZE) and described four distinct binding modes out of the published models of H3R. A binding mode of our histamine agonist corresponds to the β-binding mode (perpendicular to the membrane plane and bridging Asp 3.32 (114) and Asp 2.50 (80)). Another of our histamine binding modes is represented by mode γ, implying an interaction between Asp 3.32 and the amine of the ligand. Nevertheless, the ligand can be found in head-to-tail or tail-to-head orientation. The binding mode of the CPX inverse agonist in our complex is oblique to the membrane plane and has some features with mode δ.

Finally, Kong, Y. and Karplus, M.[51] studied the signal transduction mechanism of rhodopsin by MD simulations of the high resolution, inactive structure in an explicit membrane environment. MD simulations with restraints on each transmembrane helix indicated that (1) the two highly conserved motifs, D(E)RY and NPXXY, were found to have strong interaction correlation with retinal, (2) that the major signal-transduction pathway involves the interdigitating side chains of helices VI and VII, (3) that Glu 3.49 of the (D/E)R(Y/F) motif may act as a "signal amplifier", and (4) that Asp 2.50 may introduce a threshold to prevent background noise from activating rhodopsin. Even though we are dealing with H3R and not rhodopsin, and that the ligands are not the same, several of our results are coherent with those results, indicating common features in the general mechanism of receptor activation. Thus, we observe correlated movements of TM6 and TM7 around Pro 6.50, and Pro 7.50 of the NPXXY motif.

Different indicators in this work, such as the time evolution of the RMSD and the PCA[82], lead us to believe that the 1 μs simulations for each of the three systems fully sample their dynamics. Several new interactions are observed to contribute to the mechanism of signal propagation from the binding pocket to the G protein-binding sites in the CP domain. Our results show that the biological state of the receptor is closely linked to its conformational dynamics and that the binding of a given ligand results in shifts of pre-existing equilibria states, in coherence with experimental data[25,83], and thus in a conformational variability. We thus expect functionally different ligands to induce different conformations in the receptor, given that the agonist and the antagonist used in this work and in cited experimental results induce them.

The present work confirms several findings reported in the literature and describes a number of new phenomena that contribute to the mechanism of signal transduction: (i) events that lead to signal propagation from the binding pocket to the G protein binding sites in the CP domain upon activation by the endogenous agonist histamine; (ii) the propagation of a different signal upon binding of the ciproxifan inverse agonist and subsequent inactivation of the receptor; (iii) the state of the signal in the absence of ligand binding, corresponding, in the case of H3R, to constitutive activity; and (iv) the elucidation of specific lipid-binding sites on the receptor. Our results show that the allosteric propagation of H3R's conformational changes coupling the extracellular, ligand-binding region to the transmembrane and intracellular regions, and associated with the molecular activation or inactivation mechanisms is much more complex than expected, involving several

networks of residues and water molecules interacting among them through electrostatic, hydrophobic and h-bonds; multiple rotamer switches; rigid-body motions of helices, loops and N-terminus, with many of these properties being interdependent. The internal water molecules are very mobile, even if their number remains fixed. Notwithstanding, this number depends on the state of the receptor -activated, inactivated or constitutively activated. Several of these water molecules mediate h-bonds between the receptor and the ligands. Water molecules play then important roles not only in the activation processes, but also in inactivation and spontaneous activation by influencing the dynamics and conformation of H3R.

We made other compelling observations. The high mobility in the binding site of the endogenous agonist ligand histamine leads to several poses and eventually to it leaving the binding site after a certain time lag. Before escaping into the solvent, it interacts with the N-ter and ECL2, consistent with the reports that ECL2 seems to be involved in ligand binding, selectivity and activation of GPCRs[84–86]. The pathway of ligand unbinding involves unfolding of the N-ter of the receptor. This very short-termed transition between the agonist-bound state and the constitutively activated state implies that the histamine-H3R complex is thus not a stable one and suggests that this orthosteric ligand can also act as an allosteric modulator. Experimentally, the agonist efficacy and the dissociation rate constant are highly correlated[87], explaining the high efficacy of the histamine endogenous agonist. The time interval of the unbinding of histamine, ~2 ns, is very fast and is consistent with the experimental rapid off-rates of agonists[59,78]. Histamine binding is thus transient, and a short MD trajectory would not have witnessed this phenomenon. In contrast, the binding of the antagonist was stable all along the trajectory. This is consistent with the experimental affinities measured for histamine and ciproxifan, which indicate that the affinity of this latter for the receptor is at least an order of magnitude larger at the human H3R[13]. The escape pathway observed for histamine in this simulation suggests what the converse binding pathway could be. Moreover, histamine bound to the extracellular vestibule of the receptor several times during its perambulation in the solvent bulk. One binding mode was equivalent to section 3 of the getaway.

Arrang and co-workers[13] compared potencies of H3-receptor ligands on inhibition of [^{125}I]-iodoproxyfan binding to rat H3R and human H3R. CPX displayed significantly higher potency at the rH3R when compared to the hH3R. Production of a partially humanized chimeric rH3R allowed them to identify two residues responsible for this heterogeneity: Ala 3.37 (119) and Val 3.40 (122), both from TM3. In our structures of the CPX-rH3R

complex, those residues are not in a direct contact with CPX, suggesting short-range allosteric effects behind the experimental observations and an increased, favorable hydrophobic environment for rH3R, as compared to hH3R with Val 3.37 and Thr 3.40.

The differences in many properties between the three states of the system are significant and help us disentangle how the signaling complex transmits the information from the binding of a given ligand to the intra-cellular space to trigger a desired cellular response. Nevertheless, the complexity of the atomic phenomena involved in the structure and dynamics of the receptor makes the design of ligands for the selective activation or inactivation of the receptor to produce the desired molecular and conformational effects difficult. This study confirms that a GPCR can explore a large conformational space, allowing a given receptor to start many signaling processes despite a limited number of signaling partners, displaying thus a pleiotropic behavior as far as signaling pathways is concerned. The macromolecular surfaces involved are rather flat, with no large concavities, as compared to globular proteins. The RMSD matrices and the PCA illustrate how subfamilies of conformations interconvert for a given state of the receptor indicating the existence of heterogeneous receptor conformations.

We have also seen that the sites in the receptor for lipid binding will depend on the state of the receptor. In other words, the location, number and structure of the lipid binding site will distinguish the different activity states of the receptor. Tellingly, the lipids will bind to different TM helices and will be localized in different regions -upper and, lower leaflets- of the bilayer. Therefore, just like cholesterol binding to the β2-AR, there are specific binding sites for DPPC in the H3R receptor, specific to the state of the receptor. Like palmitic acid and cholesterol in the β2-AR-carazolol crystal structure (PDB code 2RH1), the contacts between the lipids' hydrocarbon chains and hydrophobic receptor residues follow the hydrophobic matching principle. But, as we found, the specific lipid-binding motifs on the membrane protein surface consist also of positively charged residues that specifically interact with the negatively charged phosphodiester groups. Our findings are thus consistent with the experimentally reported lipid binding sites in membrane proteins[88,89] and with the protein-head and protein-tail contacts for the H1R (PDB ID 3RZE), as reported in the MemProtMD database[40]. Our analyses highlight also different lipid allosteric regulatory sites according to the state of the receptor[90] and

point to the mechanism of modulation of lipid on receptor function[91–93] -each specific protein-lipid interaction stabilizes the corresponding state of the receptor.

The totality of our results and the corresponding analyses lead us to believe that unraveling the mechanisms of signal transduction -activation, inactivation, constitutive activity- cannot be based on changes in a single microscopic feature or a small number of them. These mechanisms are complex and need multiple descriptors for a better understanding. The descriptors involve rigid-body motions of helices, along with different changes in their mechanical properties (kink, displacement, tilt, rotation, compression-expansion, winding-unwinding), multiple simultaneous rotamer toggle switches, ionic locks and inter-residue distances. We have observed differences as well in other physicochemical characteristics, such as the formation of clusters of charged residues, networks of h-bonds between amino acid residues and water molecules, the composition in amino acid residues of the internal cavity of the receptor, the presence of mobile networks of internal water molecules, and the presence of hydrophobic and aromatic clusters.

The H3R binding pocket experiences a significant plasticity since it can accommodate different types of ligands, just as what is observed experimentally for the CB1 cannabinoid receptor, a class A, monoamine GPCR -the ligand induces the corresponding conformational changes. Thus, the helices that are in contact with CPX are TM2, TM3 and TM7, as well as ECL2 (Table 3); for histamine, the corresponding helices are TM3 and TM7 (Table 4).

More realistic simulations should now include cholesterol, an abundant membrane component, in addition to the major DPPC lipid component and study its diffusion to the receptor surface, as well as the competitive effects between cholesterol and phospholipid[94]. On another hand, even though we observed distortions of the lipid bilayer for each system, we do not describe them since this was not the scope of this work.

Altogether, we hope that our extensive study of the histamine H3 receptor will contribute to the elucidation of the molecular events leading to signal transduction by first messengers on 7TM signal transducers.

METHODS and MATERIALS

We use throughout this paper the Ballesteros-Weinstein generic numbering scheme for the amino acid residues of class A GPCRs[95,96]. The numbers attributed correspond to frame 0 of the trajectory. There is no dynamical reattribution of numbers along the trajectory.

Homology modeling

We built the 3D structure of the transmembrane regions of the rat histamine H3 receptor (Hrh3 gene, UniProt accession number Q9QYN8, isoform 1 or rH3R(445) or H3L) with MODELLER's sequence homology method[97] by taking advantage of the different experimental X-ray structures of several homologous class A 7TM receptors that have become available in the last years, such as the β2-AR (PDB codes 2RH1, 2R4R, 2R4S, etc.). Despite the lower-than-the-threshold sequence identity between H3R and other receptors needed to generate reliable 3D models, the pattern of highly conserved residues on each or the TM helices allows us to "anchor" the H3R sequence to the template sequences and thus use homology modeling. The *R. norvegicus* sequence is 93.7% identical to *H. sapiens*', but we took the former because its pharmacological characterization is more complete. We paid attention by choosing active state structures as templates for the H3R-histamine and H3R apo homology models, and inactive state structure templates for the H3R-ciproxifan homology model. We have modeled also the long N-ter of the H3R, not reported in experimental structures of homologous receptors. We removed the native ligands and other foreign molecules, such as T4 lysozyme. We excluded the several 3D structures of rhodopsin in its different states as templates. As far as the loops connecting the TM helices is concerned, we employed a truncated form of the receptor by reducing the size of the IC3 loop (~140 residues), given its conformational heterogeneity and the fact that its long sequence shows no homology to any experimental 3D structure. In addition, ICL3 has been often replaced through protein engineering to produce stabilized versions of the receptor for structure determination. Nevertheless, some teams have tried to study the effects of ICL3 on the β2-AR[98]. Our use of a short isoform H3R is justified by the finding that the histamine auto-receptor is reported to be a short isoform of the H3R, H3R(413)[99]. The lengths of all other loops were those of the native sequence. Since they are small-sized, we were in no need to generate their conformations. The starting conformation for ECL2, which contains about 22 residues was unstructured, but it folded unto

specific conformations as the simulations evolved. In all models obtained, the disulfide bridge between Cys 3.25 of TM3 and Cys 45.50 of ECL2 was satisfied. In addition, helix 8 in the C-ter was stable throughout the trajectory and remained bound to the cytoplasmic side of the membrane, despite the lack of palmitoylation at Cys 8.59. All these features are an indication that our 3D homology models present a correct folding and are thermodynamically stable.

Docking simulations

We generated three systems: two holoreceptors - an agonist-receptor complex and an antagonist-receptor complex; and an apo receptor, i.e. in the absence of any ligand. For the first complex, we used as the prototype ligand the endogenous agonist histamine in the (major) tele-tautomeric form; the quaternary amine at nitrogen Nζ carries a positive charge. For the second complex, we took ciproxifan as the prototype antagonist/inverse agonist. The structures of the agonist-receptor and antagonist-receptor complexes were obtained after steered molecular docking calculations in the putative binding pocket with the AutoDock Vina program[100]. For that purpose, we centered the grid box of about 30 Å of side length in the ligand pocket of the receptor, amid the TMs, with values of 12, 10 and 10 for the exhaustiveness, num_modes and energy_range variables, respectively. All single bonds in the ligands (histamine and ciproxifan) were free to rotate. After the docking of the ligand to the receptor, the most representative complexes were further optimized by means of energy minimization. The stereochemical quality of the 3D models was verified with the ProCheck[101] and WhatIf[102] programs. Given the significative sequence homologies of H3R to other GPCRs used as templates as far as the transmembrane regions is concerned, we estimate our 3D models to be of high accuracy and corresponding to a crystal structure of ~3 Å resolution.

Molecular dynamics simulations

We used the Membrane Builder tool in the CHARMM-GUI server[103,104] for construction of the membrane and for immersing the receptor in a pre-equilibrated symmetric lipid bilayer with the Insertion method. For the protein-membrane-water system, a 1,2-Dipalmitoylphosphatidylcholine (DPPC)-based bilayer (PubChem CID 6138) was generated in a rectangular water box in which the ionic strength was kept at 0.15 M by KCl. In

addition, the relative number of K^+ and Cl^- counterions allowed to obtain an electrically-neutral system. The cubic box lipid bilayer consisted of about 188 phospholipids solvated with a shell of water molecules and ions placed around the bilayer with the program Solvate (https://www.mpibpc.mpg.de/grubmueller/solvate). We also included buried waters in the internal cavity of the receptor with the program Dowser (http://danger.med.unc.edu/hermans/dowser/dowser.htm). The space surrounding the bilayer was cropped to a rectangular, periodic simulation box with periodic boundary conditions. This configuration allows the receptor, placed in the center of the layer, and its periodic image, to be separated by a significant distance, avoiding thus unwanted receptor-receptor interactions. The dimensions of the box were of about 80 nm by 80 nm by 100 nm and depended on the system at hand. In addition to the receptor, the explicit water solvent, the phospholipids and the K^+, Cl^- counterions, the box contained, if appropriate, the ligand. The resulting systems contained 70 000-100 000 atoms. Among the several output files generated by the Charmm-gui.org server, we used the topology files for the subsequent MD simulations. The dielectric constant of the interior of the receptor is high and like that of the aqueous solvent (~78) due to the presence of internal waters.

The Propka[105] program was utilized to assign the protonation states of the titratable groups of H3R at the value pH of 7.4. We payed attention to the fact that the side chains exposed to the central part of the membrane are embedded in a low- dielectric, hydrophobic environment. The ionization states of ionizable side chains in the interior of the receptor (since it is saturated with water), in the phosphatidylcholine head-group zone and in the extra- and intra-cellular zones were assigned as those corresponding to exposure to a high-dielectric surrounding. DPPC is a phospholipid that is well characterized physicochemically and whose experimental average surface area per headgroup in the gel phase is known[106] (63+/-1 A^2 at 323K). We performed MD experiments on a test membrane and noticed that it was not necessary to apply an external surface tension in the calculations as we reproduced the experimental DPPC area per phospholipid in its absence. In addition, certain GPCRs have been shown to function cholesterol-free membranes[90].

We employed the Nanoscale Molecular Dynamics software package (NAMD), version 2.7b1, using the CHARMM-27 force field for proteins for the simulations. Histamine and CPX were parametrized with the CHARMM General Force Field (CGenFF) program of the ParamChem initiative (https://www.paramchem.org) or with the CHARMM "patch" (https://www.charmm.org/charmm/documentation/by-

version/c40b1/params/doc/struct/#Patch) command that allowed us to obtain the topology and parameter files.

We used the following parameters for the molecular simulations: Leapfrog Verlet algorithm for Newton's equation integration; integration step of 2 fs allowing the SHAKE algorithm for keeping fixed all bonds involving hydrogen atoms; update of the lists of non-covalent pairs of atoms every 20 fs; long range electrostatic interactions with Particle Mesh Ewald algorithm, spacing 1 Å, 4 fs update; non-covalent cut-off radius of 10 Å and the non-bonded pair-list distance of 15 Å; microcanonical/NPT ensemble during equilibration and production runs; TIP3P model of water; temperature and pressure-coupled Langevin baths to ensure an isotherm and isobaric ensemble, with T=323.15 K and P = 1.013 bar. The water thickness on either side of the membrane could be up to 35 Å, depending on the system.

The simulation of the membrane-water system was performed in three stages. During the first stage, restraints on all non-solvent heavy atoms were applied; the system was energy-minimized for several hundred thousand steps using steepest-descent first and then conjugated gradient. Afterwards, several hundred thousand steps of conjugate gradient were used to minimize the side chains of the receptor and the aliphatic chains of DPPC, keeping the heads of the phospholipids and the backbone of the receptor fixed. Lastly, other several hundred thousand steps were used to minimize the free system with harmonic constraints on protein atoms, then by freeing all atoms. This procedure was followed by a slow warming-up of the system for 50 ps. In the second stage, a full equilibration of the system was achieved by using the NPT ensemble for 100 ns, smoothly removing the applied constraints. The third stage of the simulation consisted of 800-900 ns production run under the NPT ensemble. We performed the equilibration of the system in the absence of external pressure on the system, allowing energy dissipation without restriction.

SUPPLEMENTAL DATA

TITLE

The complexity of the signal transduction mechanism by the histamine H3 receptor as assessed by extended molecular dynamics simulations

AUTHORS

Herrera-Zuniga LD[1], Moreno-Vargas LM[4,2], Correa-Basurto J[1, 3], Prada D[2], Curmi P[1], Arrang, JM[4], Maroun, RC[1,*]

- The RMSD of the Cα atoms measures the structural drift of a molecule. In the absence of determining free energy landscapes for the membrane-protein system, the time dependence of the RMSD is a good indicator of the convergence of the 3D structure of the protein along the simulation to a stable state.
- The RMSD matrices contain the RMS deviations using Cα atoms only between all possible pairs of structures from the trajectory. For generating the matrices, we used every 500th structure for the calculation. The RMSD color gradient goes from dark blue for small values, through yellow for medium values, to dark red for large values, with the origin at the upper left-hand corner. The different blue squares around the diagonal show time periods within which the structures resemble each other more than to frames outside these intervals. The blue squares may be taken to represent subfamilies or ensembles of conformations. The passage from one square to the other indicates different conformers. The local conformational transitions involved in these changes in RMSD are reflected in the high number of blue regions around the diagonal of the corresponding RMSD matrix.

- The radius of gyration (Rg) is the RMS distance of a chain segment from the center of mass of the molecule. It is used to describe the dimensions of a polymer chain, i.e. to estimate protein expansion or compactness along the simulation[107].

- The RMSF as a function of the residue number is a measure of the thermal mobility or structural heterogeneity.

- Principal Component Analysis (PCA) represents a classic dimension reduction approach by constructing orthogonal linear combinations of many properties (in this case conformations), called principal components (PC). The greatest variance of the data lies on the first component, the second greatest variance on the second component, and so on. The PCA allows the identification of conformational communities or clusters, dividing the conformations into several populations and covering in this way a well-defined region in diversity space given by the principal components. These essential degrees of freedom describe major collective modes of fluctuation that are relevant for the function of the protein[108]. PCs may be referred here to as 'meta- or super-conformations' and are very useful in investigating the molecular motions of proteins[82,109]. Through the PCA, MD allows the sampling and the grouping of similar conformational states.

- Characterization of the structural changes of the α-helices, such as displacement, tilt, rotation, compression-expansion, winding-unwinding and bending is important for clarifying the contribution of helices to the mechanical properties of proteins, contribution that can be important for understanding the mechanics of signal transduction in the case of 7TMs.

- Global tilt defines the orientation of each helix in a protein. It is the angle that the helix axis forms with the laboratory frame's coordinate axes. The evolution of this variable is shown in dial plots in which the concentric circles represent the passage of time.

- As defined by Mezei and Filizola[110], the global helix rotation is a measure of the rotation of the helix as a rigid body around its own axis.

- The local tilt, which defines the orientation of the helical axis with respect to the reference helix configuration, shows the extent of change in helix direction.

- The Bend (Kink), Wobble, Face-shift (winding-unwinding) angles and the Turn angle per residue (TPR) are defined by the ProKink methodology of Visiers et al, 2000[111]. The bend angle measures the angle between the axis of the helix before and after a proline residue. The closer to 0° the bend angle is, the smaller the kink. The wobble angle is close to zero when the post-proline helix is bent, so that its axis is moved towards the proline Cα, and close to positive or negative values when this axis is moved away from the Cα. The face-shift angle is a measure of the winding-unwinding of the helix. When the face-shift angle adopts positive values, it means that the helices are under-wound.

- During MD simulations, the membrane-spanning helices can undergo compression-extension movements beyond just random fluctuations. The measurement of the length of the helix serves to estimate its compression-extension.

- The center of mass (COM) of the helix is used to characterize its position and the program TRAJHELIX from the SIMULAID package[110] allows monitoring of the total helix displacement as a function of time. The total helix displacement gives the contribution from the x- and y-components, i.e. along the membrane plane. The displacement of each helix in the z-direction, perpendicular to the plane of the membrane, can also be calculated.

- We did not deem necessary to energy minimize the complexes at the frames we looked at during the analysis. Instead, we selected ten frames at regular intervals of the MD trajectory to obtain the interactions between the ligand and the receptor, since we compiled the long-lasting neighbors of the former all along the trajectory.

- When dealing with the lipid components of the membrane, we wished to find out whether there were any DPPC-specific binding sites on the receptor. For this purpose, we found out the amino acids that are in contact at least 80% of the time with at least one atom of a lipid molecule defining high-affinity, specific binding sites with the membrane protein.

For the analysis of the data generated by the MD production trajectories, we used the following programs:

- EUCB[112] for the computation of the following properties, with 10 ns jumps and a switch distance of 5 Å.

- Root mean-square fluctuations (RMSF) of each amino acid residue
- H-bonds
- Ionic pair matrices or charged-residue clusters
- Conformational switches
- Side chain rotamer angles
- Ionic bridges

- SIMULAID[110,113]
  - Mechanical properties of the helices (TRAJHELIX)
  - Analysis of proline residues with bends (PROKINK)
  - H-bonds between protein and solvent
- CARMA[114]
  - Root mean-square deviation (RMSD) of the cartesian coordinates of Cα atoms with respect to the initial frame
  - Radius of gyration (Rg)
  - RMSD matrices between structures
  - Distance maps
  - The principal component analysis (PCA)
- ProKink[111]
  - Bend angles
  - Wobble angles

    - Winding-unwinding angles

- VMD[115] and Tcl/TK home-brewed scripts run in VMD
    - Distances between amino acid residues
    - Detection of the residues of the internal cavity of the receptor
    - Non-covalent contacts between cavity residues and ligand
    - Distribution and residence times of lipids around the protein
    - Lipid binding sites on the protein
    - Positions and residence times of water molecules in the interior of the protein
    - H-bonds between residues
    - Visualization of the trajectories
    - Production of molecular figures

- PyMol (PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC.)
    - Molecular graphics and visualization

The residues composing the internal cavity of the representative structures of the receptor produced by CARMA were defined by finding the "wet" residues, i.e. those residues that were in contact (3 Å or less) with an internal water for more than 80% of the trajectory. The "wet" atoms of the ligands were determined in the same fashion. This approach determined the residues that had an interaction with a water molecule that had itself an interaction with the ligand. On another hand, pocket detection, ligand binding site and analysis and visualization of tunnels and channels were performed with the Fpocket[116], SiteHound[117] and Caver software tools[118]. Those residues found in at least three out of four searches were considered cavity residues.

We performed the computations in a Linux cluster with one master node of 8 CPUs, 10 To stocking; and 192 CPUs (16 nodes of 12 Intel XEON E5630@2.53GHz, 24 GB RAM, CPUs each). The NAMD parallel jobs were executed through the mpiexec application.

RMSD

Fig. S1a shows for the **antagonist-H3R** complex that the RMSD attains a plateau at about 150 ns, remains constant at 4.2 Å until ≅550 ns and then rises slightly to 4.8 Å. The **agonist-H3R** complex shows an RMSD that attains values ≅5.0 Å at 600 ns until the end of section 1 (Fig. S1b). This observation is in consonant with the fact that agonist ligands tend to be less efficient in the stabilization of the structure of the protein[119]. The **apo** receptor shows considerable changes in the RMSD (Fig. S1c), especially in the middle of the trajectory, just before 600 ns, when it reaches almost 5 Å of RMSD, to fall afterwards to around 3.5 Å at the end of the 930 ns trajectory.

For the structurally-conserved TM helices of the receptor, the RMSDs of each of the three systems with respect to the crystal structure of the H1R-doxepin complex at 3.1 Å resolution (PDB code 3RZE) are as follows: 3.1 Å for the last frame of the antagonist-H3R complex, 2.5 Å for the last frame of the agonist-H3R complex, and 2.0 Å for the apo receptor. These data indicate that the MD trajectories lead to physically stable systems.

RMSD matrices

Fig. S2a for the **antagonist-H3R** complex shows a tendency for the blue areas to get darker as the trajectory unfolds, indicating that subsets of similar conformations persist. For the first period (section 1) of the **agonist-H3R** complex (Fig. S2b), i.e. when histamine is bound to the receptor (up until nanosecond 590), the structures up to 328 ns resemble each other a lot (small RMSD values) at the beginning of the first phase of the period. Then about five subsets of conformations appear, of which the last one is maintained sometime after histamine unbinding. Then another subset appears centered around 726 ns. For the **apo** receptor (Fig. S2c), the RMSD matrix is rather sparse and weak in blue regions, i.e. in small values of RMSD, indicating higher structural variations.

Cα-Cα distance maps

The Cα-Cα distance map of the average structure of the **antagonist-H3R** complex shows that CPX (represented in the map after all the amino acid residues) is in contact with TM2, the N-ter of TM3, ECL2, the C-ter of TM6 and the N-ter of TM7 (Fig. S3a). For the first section of the trajectory of the **agonist-H3R** complex, the proximity relationships are in Fig. S3b and for the **apo** receptor in Fig. S3c. All relative positions of the helices are shown in the form of a matrix for each of the three systems in Fig. S3d, in which an x represents the mentioned proximities between TM helices. The matrices are symmetrical. We can see that the distance maps of the average structure of the agonist-H3R complex and apo systems are equal, with a clear off-diagonal line showing a sequential array of close distances: TM1 with 2, TM2 with 3, TM3 with 4, TM4 with 5, TM5 with 6, and TM6 with 7. Instead, for the antagonist-H3R complex the inter-helix distribution of spatial distances shows no proximity between TM2 and 4, TM3 and 7, and TM4 and 5. In place, this complex shows a proximity between TM2 and 6. As mentioned in the Introduction, H3R possesses a high constitutive activity in the absence of agonist. Our results show indeed that the overall relative positions of TM helices in the apo receptor and the agonist-H3R complex resemble each other. In addition, ECL2 is in contact with the antagonist, in conformity with the importance of this loop in ligand binding, selectivity and (in)activation[85], as well as stabilization of the inactive state of the receptor[84].

Rg

The Rg for the **antagonist-H3R** complex (Fig. S4a) shows initial values around 21.1 Å until about 360 ns, when it drops to values around 20.8 Å until the end of the trajectory. For the complex with **histamine**, the Rg starts at around 22.8 Å, and after minor fluctuations ends up in 22.6 Å at the end of section 1 (Fig. S4b). There is, nevertheless, a jump to 23.7 Å just before 600 ns. As we shall see below, this is due to the particular behavior of the histamine agonist. The **apo** system, Fig. S4c, begins at about 21.75 Å and ends up in 21.25 Å. Even though the changes are small -less than 1 Å- the Rg of all systems decreases as compared to the initial values. The reason may be because the receptor is embedded in a phospholipid bilayer in our MD simulations, whereas the different homologous receptor structures that were used as templates for the H3R model come

from crystals in the presence of detergents, fatty acids, cholesterol and other substances such as N-acetyl glycosamine, and not a membrane. These experimental conditions present membrane proteins with a physicochemical environment different from that of native phospholipids, altering thus native protein structure. The MD scheme is thus more reflective of a biological membrane environment. Thus, the starting crystal receptor structure undergoes a compaction in the calculations due to the surface tension exerted by the lipid bilayer. Comparing the values of Rg for each of the three systems, antagonist, agonist and apo at the end of their trajectories -20.8, 22.6 and 21.3 Å, respectively- indicates that the agonist-H3R complex is the most voluminous one.

RMSF

Fig. S5a for the **antagonist-H3R** complex shows that the highest values come from TM5 (17 Å) and TM3 (16 Å). All trans-membrane helices, except TM3 and TM5, show a decrease in RMSF when going from the N- to the C-ter. Instead, TM3 and 5 show large increases towards the C-ter. This amounts to larger structural fluctuations for the fragments of the TM helices 2-5 in the lower leaflet of the membrane and for the segments of the TM helices 1 and 7 in the upper leaflet of the membrane. As far as the first section of the **agonist-H3R** complex is concerned (Fig. S5b), the fluctuations are much more homogeneous between the ends of the helices but compared to those of the antagonist-H3R complex, they are larger, in the 21-22 Å range. Astonishingly, the values of the fluctuations of the **apo** receptor are much lower, in the 7 to 9 Å range (Fig. S5c). Like the antagonist-H3R complex, the structural heterogeneity is predominantly localized to the lower leaf of the lipid bilayer, although of about half the magnitude.

In conclusion, the overall values of the RMSFs are in the order agonist > antagonist > apo receptor. This is consistent with the fact that the crystallization of agonist-H3R complexes of GPCRs is much more unwieldy due to their increased structural heterogeneity.

With regards to the ligands, the RMSF average values of histamine and CPX are 13.5 and 40.2 Å, respectively, indicating the increased flexibility of CPX with respect to histamine.

Eigenvectors

The eigenvectors represent the displacement of each residue in the system and are used to generate new conformations. The image sequences in Fig. S6 show the eigenvectors for each of the four clusters generated by the PCA for the **antagonist-H3R** complex. The sequence of conformations in Fig. S6a corresponds to cluster 1 (C1) and shows for the frame at 140 ns that this mode involves essentially large movements of the N-ter extracellular head of the receptor and its ECL2, which act as a flexible lid that opens and closes on top of the ligand-binding cavity. Loops ICL2 and ECL2 exert a push-n-pull motion on TM4 as a rigid body. For TM5, we observe a compression-extension movement of the fragment above Pro 5.50 and a kink movement around this residue. We also observe kink movements around Pro 6.50 of the conserved $^{6.47}$CWXP$^{6.50}$ motif, and Pro 7.50 of the $^{7.49}$NPXXY$^{7.53}$ motif for each, TM6 and TM7, respectively. All these movements are correlated. No other important movements are detected in the other helices. Cluster 2 (C2) in Fig. S6b shows large correlated movements of ECL2, ICL2 and to a lesser extent ICL3, with slight kink motions of TM1, 5, 6 and 7 around their middle-helix proline residues. In addition, TM5-7 move collectively in a breathing motion along the plane of the membrane. The mode corresponding to cluster 3 (C3) shows large amplitude motions for ECL2, and a pendulum-like rigid-body motion of the lower zone of TM5-7 with their IC loops (Fig. S6c). Fig. S6d shows the collective modes of movement corresponding to the last cluster, C4. First is an angular N-ter to C-ter motion of H8, whose C-ter gets close and far from the membrane. This is accompanied by a concerted motion of the N-ter, ECL3 and ICL2, indicating long range communication between the extra- and intra-cellular regions of the receptor. With regards to the **agonist-H3R** complex, Fig. S7 shows the eigenvectors of cluster 1. We can observe big correlated motions of the N-terminal and ECL2 extra-cellular segments of the receptor with a large displacement of the N-ter end of TM5 and TM6 on the plane of the membrane. Along with large conformational changes in ICL2, this latter motion leads to an opening and closing of the far-away intra-cellular domain of the receptor, in agreement with experimental 3D structures of active state GPCRs, such as β2-AR[78,79,119,120]. Fig. S8a shows the eigenvectors of cluster 1 for the **apo** receptor. The N-ter, ECL2 and ECL3 undergo large concerted unfolding-folding motions, with ECL2 exerting a hinge-bending motion on TM4 perpendicular to the plane of the membrane. For cluster 2 (Fig. S8b), the motions are like cluster 1, but this time accompanied of a large movement of the TM7-H8 loop that carries with it large movements of H8. For cluster 3 (Fig. S8c), large motions of N-ter and ECL2 are observed, with N-ter in the solvent. No correlation with

the intra-cellular regions of the receptor is observed. ECL2 is essentially disordered, although we can observe from time to time the presence of one or two one-turn helical segments.

Clusters of charged residues

Electrostatic interactions are of special importance for membrane proteins because of the low dielectric environment in membranes. We thus detected clusters of charged residues using a distance criterion of 12 Å between COMs of the side chains of amino acid residues with home-made scripts.

For the **antagonist-H3R** complex, there are two charged amino acid clusters that persist 90% of the time of the production trajectory. One cluster is in the extracellular region of the receptor and the other one in the intracellular region. The two clusters are represented in **Fig.** S9. The amino acid residues for the first cluster belong to TM3, 5, 6 and 7, and those of the second cluster to TM1, 3, 4, 5, 6, and H8. Both clusters involve several loop residues belonging to the N-ter, ICL1, ICL2 and ECL2 (located at the surface of the receptor and exposed to the extra-membrane solvent). For the **agonist-H3R** complex, two independent clusters persist while histamine is bound to the receptor (**Fig.** S10). They involve N-ter, TM3, ECL2, ECL3 and TM7 for the first cluster; and TM3, 4, 5 and 6, and ICL3 for the second. One residue of the first cluster, Asp 3.32, is in contact with the protonated amine (pKa ~9.4) of the aliphatic amino group of histamine, consistent with experimental data[121]. Four clusters of charged residues can be found for the **apo** receptor (**Fig.** S11). Amino acid residues from TM3 and ECL2 compose the first cluster; one residue from each TM3, 5 and 6 the second cluster; residues from TM3, 4, 5, 6, H8, ICL1 and ICL2 the third cluster; and residues from the N-ter, ECL1 and TM7 the fourth cluster. Many of these residues are also involved in long-lasting h-bonds, i.e. prevailing more than 70% of the time.

H-bonded networks between amino acid residues

For the **antagonist-H3R** complex, inter-secondary structure h-bonds with residence times greater than 70% connect ECL2 and ICL2 to TM3; N-ter to TM7; TM1 to H8; ECL2 to ECL2; and TM6 to TM7 (Table S1). For the **agonist-H3R** complex, ECL2 is H-bonded to the N-ter, TM3 and TM4; and the N-ter of TM1 to H8 (Table S1). Finally, for the **apo** receptor, only an h-bond between ECL2 and TM4 persists 78% of the time (Table S1). Thus,

it is interesting to notice that the antagonist-bound receptor shows the largest number of inter-secondary h-bonds, providing the structure of this complex with an added energetic and structural stability.

H-bonded water networks in the internal cavity

We determined the H-bonded networks established by water molecules in the cavity with occupancy greater than 75%. In all three systems, these occupancies are characterized by very low residence times, of the order of 20% maximum, indicating a high fluidity of water molecules. Thus, even if a site may be continually hydrated, it is not so by the same water molecule 4/5 of the time. Again, the observation of the itinerary of the water molecules indicates that these enter and exit the internal cavity of the receptor through the extra-cytoplasmic region of the receptor, no water molecules neither entering nor exiting the receptor through the intra-cytoplasmic region. The internal cavity of the receptor has the shape of a bent funnel with a flexible lid at the top composed of the N-ter and the ECLs, especially ECL2.

For the **antagonist-H3R** complex, residues that contribute with their side chains to H-bonded networks in the cavity zone are Asn 7.45; Asp 2.50 and 3.32; Glu 5.46 and 7.36; Ser 7.46; Thr 6.52; Tyr 2.61; and Trp 7.43 and 23.50 (**Fig.** S12a). Of these, Asp 2.50, Glu 5.46 and Ser 7.46 form two or more H-bonded bridges. The two largest networks involve five waters, Asp 2.50 and Ser 7.46 (Netwk1); and four waters, Glu 5.46 and Thr 6.52 (Netwk2, **Fig.** S12a), respectively. Finally, the interaction of CPX with two residues in the protein, Tyr 2.61 and Asp 3.32, is mediated by two water molecules for the former, and by one for the latter (Netwk3, **Fig.** S12a). Notice that TM1 and, of course TM4 and H8, do not participate in CPX binding. The only charged amino acids in the interior of the receptor are in the binding cavity and are Asp 2.50 and 3.32, and Glu 5.46 (**Fig.** S12a). **Fig. S12**b for the **agonist-H3R** complex shows two networks of h-bonds, the first of which includes histamine bonded to a water molecule and Asp 3.32. The two networks of the **apo** receptor involve, each one, two internal water molecules (**Fig.** S12c).

Hydrophobic and aromatic clusters

We report in Table S2 the hydrophobic clusters at 90% occupancy formed by at least three side chains for the antagonist-H3R complex, the agonist-H3R complex and the apo receptor, respectively. The cutoff used is of 6 Å between the COMs of each amino acid residue; the set of residues is composed of Ile, Leu, Val, Phe, Met, Cys, Trp, Pro and Ala.

For the **antagonist-H3R** complex, there are six clusters. One cluster, implying the N-ter and ECL1 contains 7 residues (Ala 23 from the N-ter, Phe 2.56, Cys 2.57, Leu 2.60, Trp 23.50 (ECL1), Leu 3.24 and Trp 3.28). The distribution of clusters is as follows: two clusters of three residues, two clusters of four residues, one cluster of six residues and one cluster of seven residues. For the **agonist-H3R** complex, there are six clusters. Finally, for the **apo** receptor, there are five clusters. These non-covalent interactions, in addition to the h-bonds described above, contribute also to stabilize the overall structure of the receptor.

As for the clusters containing aromatic side chains only, the **antagonist-H3R** complex shows a 3-residue cluster contributed; the **agonist-H3R** complex trajectory shows none; and the **apo** receptor show two clusters of 3 residues each (Table S3).

Moreover, we decided to detect the interactions between the aromatic side chains belonging to hydrophobic clusters to analyze their evolution during the MD simulation. For this purpose, we proceeded to use the COMs of the multi-atomic functional groups. The pairs of side chains studied were obtained from the hydrophobic clusters just described. For the **antagonist-H3R** complex, Fig. S13a shows an aromatic cluster composed of Phe 2.56, Tyr 2.64 of the C-ter of TM2, Trp 23.50 (ECL1) and 3.28, and Phe 23.52 (ECL1). This cluster remains stable throughout the trajectory and locks the conformation between the C-ter of TM2 and ECL1. The Tyr aromatic cycle is perpendicular to the Trp cycle, which in turn is perpendicular to the Phe cycle. This type of non-covalent interaction involving aromatic rings (π-to-π or π-stacking) has been previously documented[122]. The second cluster is composed of Tyr 7.53 of the NPXXY microdomain (C-ter of TM7), which establishes a long-lasting pi-pi interaction with the first residue of H8, Phe 8.50 (Fig. S13b). The distances between the COMs of the corresponding pairs are between, 4.5 and 8.5 Å. For the **histamine** complex, we found no such aromatic ring clusters. Lastly, the **apo** receptor presents two clusters. The first one is composed of Phe 6.44, Phe 7.39 and Trp 7.40; the second cluster is like the first cluster in the active complex (Not shown).

Mechanical properties of the helices

Global tilt

Fig. S14a, b, c to Fig. S21a, b, c show thus the three components -x, y and z- of the global helix tilt angle for all eight helices of the **antagonist-H3R** complex. The x- and y-components lie on the plane of the membrane whereas the z-component is perpendicular to them. As can be seen, the x- and y-components adopt values in the range 60-120° for all TM helices. The z-tilt of TM4 is 10°, making of this the most perpendicular helix to the plane of the membrane. The axis of H8, being quasi parallel to the membrane, shows a global z-tilt of 90°. Odd-numbered helices TM1, 3, 5 and 7 have their global z-tilt angles in the range 145-160°, whereas even-numbered helices TM2, 4 and 6 have it in the range 10-30°. This is expected since helices whose N- to C-ter directionality is directed towards the cytoplasmic space have an opposite directionality to those helices whose N- to C-ter sequence points to the extracellular space. The resulting differential z-tilt angle is of about 130° between the two subgroups and not of 180°, meaning that the change of orientation of one subgroup of helices with respect to the other subgroup is accompanied by rotations around the x- and y-axes. It can also be seen that, exception made of H8, the fluctuations in the values of the global helix tilt angles along the whole trajectory are rather small, implying a structural stability of the entire helical bundle. For the **agonist-H3R** complex, the mean values of the x- and y- tilt angles of helices TM1-7 are between 60 and 120° (Fig. S22a, b, c to Fig. S29a, b, c), with minor variations of the orientation of the helices with respect to the plane of the membrane. TM4 and TM6 are the most perpendicular to the plane of the membrane. As for the **apo** receptor, the values of the mean x- and y-tilt angles are also between 90 and 120° (Fig. S30a, b, c to Fig. S37a, b, c), like for the antagonist-H3R complex.

Thus, the average values of the x- and y-tilt angles for the three systems are pretty much in the same range (60-120°), with rather small dispersions around each average value. In this sense, the antagonist receptor and the apo form have in common a negative value for the average rotation of the eccentric TM4.

Global helix rotation

Fig. S14d to Fig. S21d show the values adopted by this variable during the MD simulation for TM1-7 and H8 of the **antagonist-H3R** complex. In general, a certain degree of fluctuation around the helix rotation average

values is found for all helices. The average rotation for TM1-3, 5-7 is between 0 and +/- 20°. That of the eccentric TM4 is of about -55°. Besides, the rotation of the latter undergoes a shift from -55 to -70° in the last 2/5 of the trajectory (Fig. S17d). TM6 undergoes a clear-cut transition (Fig. S19d) from ~5 to 30° in rotation also in the last 2/5 of the trajectory, making of this a concerted movement. For the **agonist-H3R** complex, the helix rotation value of TM4 is rather of 0° (Fig. S22d- Fig. S29d), which is reminiscent of the behavior of the global helix tilt angles. For the **apo** receptor, helix rotation is again "well-behaved" for all TM helices except for TM4, for which it is of about -40° at the end of the trajectory (Fig. S30d- Fig. S37d).

Local tilt and Global toggle switch

This is shown in Fig. S14e-Fig. S21e for each helix of the **antagonist-H3R** complex, in Fig. S22e-S29e for the first period of the **agonist-H3R** complex, and in Fig. S30e-S37e for the **apo** receptor. For all three systems, all TM helices show average values of 0-10° with little dispersion about these values.

Bend (Kink), Wobble, Face-shift (winding-unwinding) angles and Turn angle per residue (TPR)

The figures that follow show the values of the bend, the wobble and the face-shift angles around proline residues Pro 5.50, 6.50, and 7.50 of helices TM5, 6 and 7, respectively. As we can see from Fig. S38a for the **antagonist-H3R** complex, the values of bending for TM5 undergo several transitions during the trajectory but oscillate around 40° starting at about 630 ns. This implies a value of 140° between the helical fragments on each side of the proline kink. The wobble angle for TM5 in Fig. S38b decreases from about -70° to -120° at 560 ns to end-up close to +/-180°. The helix is under-wound as the face-shift angle is always positive, residing in the 37 to 75° range, although in the end it goes back to 37° (Fig. S38c). It is worthwhile noticing that the face-shift angle undergoes transitions at around 350, 490 and 560 ns, i.e. at the same frames in the trajectory as the bend and wobble angles undergo transitions, indicating the correlated behavior of these variables. After some oscillatory behavior in the 210-630 ns interval, the average kink angle of TM6 of the antagonist-H3R complex (Fig. S39a) attains a value of about 35°, just like in the beginning of the trajectory. The wobble angle attains values around -85° only, with deviations of +/- 75°, and the face-shift values of about 52° starting at 420 ns (Fig. S39b and Fig. S39c, respectively), indicating and under-wound helix. For TM7, the kink undergoes transitions in the 630-700 ns interval and adopts a value of 50° at the end of the trajectory (Fig. S40a). The wobble angle begins at -130° and finishes in -130°, visiting the +/-180° region (Fig. S40b). Finally, the face-shift values indicate

also an under-wound helix –final value around 50° (Fig. S40c). The kink angle of the **agonist-H3R** complex Pro 5.50 TM5 helix varies from 25 to 150°; the wobble operates a change between local average values around -130° et 100°, spanning the whole circle at times; the face shift angles span the whole circle (Fig. S41a, b, c). For Pro 6.50, the kink mean value is of ~135°, with values spanning the 0 to pi range at certain intervals. The wobble spans the whole circle also. The mean face-shift is ~-90°, indicating again an over-wound helix (Fig. S42a, b, c). The mean value of the bend angle for Pro 7.50 is about 80°, spanning a 20 to 150° range. The corresponding wobble angle spans the whole circle again, just like the face-shift (Fig. S43a, b, c). For the **apo** receptor, the bend angle of Pro 5.50 fluctuates slightly around 10°; the wobble, with a mean negative value is such that the post-proline axis moves towards the proline Cα. The face shift is close to 80°, indicating an underwinding of the helix (Fig. S44a, b, c). For Pro 6.50 of TM6, the bend angle fluctuates slightly around 40°, the behavior of the wobble shows variations in the -50 to -150° region. The face-shift values begin and end-up in 75° (Fig. S45a, b, c, respectively). For Pro 7.50, the mean kink value is around 35°, with little fluctuations during the trajectory. The wobble presents a value of -150°, and the face-shift values correspond to a slightly under-wound helix (50°), just like in TM5 and TM6 (Fig. S46a, b, c, respectively).

In summary, helices TM5 to TM7 of the agonist-H3R complex are overwound, whereas the same helices in the antagonist-H3R complex and the apo receptor are underwound. In addition, the agonist-H3R complex shows large variations in the bend angle values of TM5 to TM7, consonant with the global toggle switch activation model[123], in which a seesaw movement of TM6 around a pivot corresponding to Pro 6.50 takes place. On the contrary, the antagonist-H3R complex and the apo receptor show very narrow distributions about a given kink value. As for the wobble angle, we notice that all three helices span the whole 0-2pi zone for the agonist-H3R complex, whereas it is just TM5 of the apo receptor that shows this behavior. None of the helices of the antagonist-H3R complex spans the whole circle.

The overall winding-unwinding of a helix can be monitored using the turn angle per residue (TPR) and measuring its deviation from ideal values[113]. Fig. S14f-S21f show the TPR for TM1-7 and H8 of the **antagonist-H3R** complex. We can see that in all cases the average values fall close to the ideal values of an α-helix (99.4°), but that the dispersion around this value is higher for TM6, indicating more winding-unwinding, and thus compression-extension of the helix. These findings point to TM6 as having peculiar mechanical properties that

differentiate it from the other TM helices. For the **agonist-H3R** complex (Fig. S22f-S29f), the TPR values are close to the ideal value, TM2, 3, 4 6 and 7 showing fluctuations skewed towards less than 90° angle values. Finally, for the **apo** receptor, the average values are close to 99.4° and the fluctuations around it are rather of small magnitude. In conclusion, it is the agonist-H3R complex that shows larger oscillations of the TPR values as compared to the antagonist-H3R complex and the apo receptor (Fig. S30f- S37f).

Helix length (compression-extension)

To measure the variations in helix length, we took the difference between the maximum and minimum values of the length in the trajectory for each helix. The helices are listed in Table S4 for the **antagonist-H3R** complex with their number of residues, their average length, the average length per residue and the max-min difference in helix length for the equilibrated production interval of the trajectory (Fig. S47a-h). Thirty-six residues compose TM3. Its average length (51.4 Å) and the difference between the max and min length (8.3 Å) adopted during the trajectory are the highest, meaning that TM3 experiments the largest compression-extension movements. TM5, with just 24 residues, shows the second highest average length and max-min length difference (Table S4). For the **agonist-H3R** complex (Fig. S48a-h) and the **apo** receptor (Fig. S49a-h), again TM3 is the longest helix (Table S5 and Table S6). It shows the most compressibility for the agonist-H3R complex (Table S5), whereas the largest variation in helix length (8.5 Å), corresponds to TM6 for the apo receptor, followed by TM3 (7.6 Å). The longest helix for all three systems is TM3, measuring 51-52 Å, implying this helix extends beyond the physical limits of the membrane. It is followed in length by TM5 for the antagonist and agonist-H3R complexes, and by TM6 for the apo receptor. The smallest helix is TM7 for all three systems. For the understanding of the mechanism of action of signal transduction, it is interesting to observe that the largest variations in the helix length (compression-extension) are shown by TM3 for the antagonist and agonist-H3R complex (8.3 Å) and, instead, by TM6 for the apo receptor (8.5 Å). Finally, in all three systems, TM3 (36 residues), TM5 (24 residues) and TM6 (33-34 residues) are the helices whose average length is the largest and protrude into the aqueous phase of the intracellular side of the membrane.

Total helix displacement. Helix shift in the z-direction.

The largest value of the total helix displacement corresponds to TM4 (5.1 Å) of the **antagonist-H3R** complex, to TM3 (1.0 Å) of the **agonist-H3R** complex and to TM4 and TM6 (1.0 Å) of the **apo** receptor.

The helix shift in the Z direction, perpendicular to the plane of the membrane of the last frame of the trajectory with respect to the average structure for the **antagonist-H3R** complex ranges from -0.5 to 3.5 Å for the ensemble of the TM helices (**Fig. S50**a-h), with a large shift upward of 3.5 Å for TM4. Like TM4, TM2 is shifted upward, while the displacement of the rest of the helices is negligible. For the **agonist-H3R** complex (**Fig.** S51a-h), the helix shift range is -0.45 to 0.15 Å, with the largest shift downward for TM1. For the **apo** receptor (**Fig.** S52a-h), the absolute values of the z-shift of the last frame are much smaller than the ligand complexes, oscillating between -1.5 and 1.0 Å, but this time with TM6 showing the largest downward shift (-1.5 Å).

These last results suggest that TM4, even though eccentric with respect to the helix bundle formed by TM1-TM3, and TM5-TM7 may play a role in the inactivation mechanism of H3R through its total displacement on the plane of the membrane and its shift in the Z direction.

Switches for the Met side chains

Kofuku et al.[124], by monitoring the NMR signals, investigated the role of Met residue 82 (2.53) in antagonist- and partial agonist-bound states of the β2-AR, which are correlated with conformational changes of the transmembrane regions upon activation. The corresponding residue in H3R is a valine and none of its neighbors is a methionine; nevertheless, we decided to monitor the conformational states of all Met residues for the three systems. We found that in all three systems, all but one of the methionine residues remain in the t or g- states throughout the trajectory.

As opposed to the β2-AR then, the conformational states of the methionine residues remain unchanged between the apo receptor, and its antagonist- and agonist-bound states; there is thus no correlation of conformational changes of the methionines upon activation in H3R (Table S7).

Lipid binding sites on the receptor

We list in Table 5 the lipids attached to the receptor in the **antagonist** state. The residence time of the given lipid is listed in the first line of the "occupancy" column; the next lines show the residence times of the lipid with different amino acid residues of the receptor. We can see that one DPPC molecule (125) is permanently bound to the N-ter residues of TM6 Lys 6.32 and 6.35, Ser 6.36, Ile 6.39 and 6.43, and to the C-ter residues of TM7 (Leu 7.52 of the NPXXY motif and Leu 7.55; Cys 7.56 and His 78.00 belonging to loop TM7-TM8). This lipid

is part of the lower sheet of the membrane, on the luminal side. Another lipid binding site (lipid 110) is composed mostly by residues from TM3 (Val 3.40 and 3.44, Phe 3.41, Leu 3.45 and 3.52, and Tyr 3.48) and is an upper leaflet lipid. Lipid 161 lies proximal to residues from the N-ter of TM4, making it a lower layer lipid. An additional lipid in the upper leaflet (79) binds to two residues of the N-ter end of the receptor and to the N-ter of TM7. The first half of TM5 is occupied by an upper-leaflet lipid (31 or 33). A minor binding site is formed by the ICL3 loop and the N-ter of TM6 (lower-leaflet lipid 165). Lipid 43's site is made up of N-ter TM3 residues and C-ter TM4 residues. Furthermore, a site composed by about 12 residues from the C-ter of TM7 and the N-ter of H8 and the central region of TM1 binds a lower-leaflet lipid (149). An interhelical N-ter TM3 – C-ter TM4 site lodges an upper leaflet lipid. As can be seen from Table 6 for the **histamine** complex, lipid 2 is lodged in the middle section of TM6 and establishes one contact with one residue of H8; lipid 47 binds to the C-ter of TM5 and all along TM6; lipid 183 binds to the second half of TM2; lipid 7 interacts with H8; yet lipid 22 is in contact with ICL1, the last residue of TM1, and the first residue of TM2; lipid 59 is in contact with C-ter residues of TM5. In the case of the **apo** receptor, Table 7, the first lipid binds to the N-ter of TM1 and TM7, and to the C-ter of TM2; the 2nd lipid to the N-ter of TM3, to the C-ter of TM4; the 3^{d} lipid to the N-ter of TM1; the phosphate head of the 4d lipid to an Arg of the N-ter; the 5th lipid to the C-ter of TM6.

Table S8 shows the amino acid residues of the receptor grouped in four types -non-polar aliphatic, aromatic, uncharged polar and charged- and their number in contact with lipids for each of the three systems. First, the number of non-polar residues is the largest, followed by the aromatic residues, the polar and the charged types. The most frequent residue in the first group is leucine. In the second group, the number will vary according to the receptor complex, with Phe being the most abundant for the antagonist and agonist (section 1) systems, and Tyr and Trp for the apo system. In the third group, Ser and Thr for the **antagonist**- and **agonist**-H3R complexes are the most represented residues of the polar group. Finally, the antagonist-H3R complex is the system showing the most contacts of charged residues with lipids -eleven Arg and six Lys.

The average composition of transmembrane surface amino acid residues is 67, 71 and 67% hydrophobic (Phe, Met, Trp, Ile, Val, Leu, Pro, Ala) for the antagonist, agonist (section 1) and apo systems, respectively (red numbers in Table S9), with Leu being the most frequently exposed residue, representing 27-29% of all transmembrane exposed residues.

Acknowledgments

ECOS Nord-ANUIES Mexique and UMR-S 1204 INSERM/Université d'Evry-Val-d'Essonne/Université Paris-Saclay for visiting scholarships to LDHZ.

Institut Servier for a postdoctoral fellowship to LMMV.

Université d'Evry-Val d'Essonne/Université Paris Saclay for a visiting professorship to JCB.

Author contributions

Conceptualization, RCM; Methodology, RCM & LDHZ; Software, LDHZ; Formal analysis, RCM & LDHZ; Investigation RCM, LMMV, JCB & DP; Resources, JMA, PM; Writing - original draft, writing - review and editing, RCM, JMA; Visualization, LDHZ; Supervision, RCM; Project administration, RCM; Funding acquisition, RCM, JMA, PC.

Competing interests

None

REFERENCES

1. Fredriksson, R. The G-Protein-Coupled Receptors in the Human Genome Form Five Main Families. Phylogenetic Analysis, Paralogon Groups, and Fingerprints. *Mol. Pharmacol.* (2003). doi:10.1124/mol.63.6.1256
2. Lagerström, M. C. & Schiöth, H. B. Structural diversity of G protein-coupled receptors and significance for drug discovery. *Nat. Rev. Drug Discov.* **7**, 339–357 (2008).
3. Schiöth, H. B. & Lagerström, M. C. Structural diversity of g proteincoupled receptors and significance for drug discovery. *Nat. Rev. Drug Discov.* (2008). doi:10.1038/nrd2518
4. Fukui, H. *et al.* Molecular cloning of the human histamine H1 receptor gene. *Biochem. Biophys. Res. Commun.* (1994). doi:10.1006/bbrc.1994.1786
5. Gantz, I. *et al.* Molecular cloning of the human histamine H2 receptor. *Biochem. Biophys. Res. Commun.* **178**, 1386–1392 (1991).
6. Lovenberg, T. W. *et al.* Cloning and functional expression of the human histamine H3 receptor. *Mol. Pharmacol.* **55**, 1101–1107 (1999).
7. Arrang, J. M., Garbarg, M. & Schwartz, J. C. Auto-inhibition of brain histamine release mediated by a novel class (H3) of histamine receptor. *Nature* **302**, 832–837 (1983).
8. Zampeli, E. & Tiligada, E. The role of histamine H4 receptor in immune and inflammatory disorders. *Br. J. Pharmacol.* **157**, 24–33 (2009).
9. Brown, R. E., Stevens, D. R. & Haas, H. L. The physiology of brain histamine. *Prog. Neurobiol.* **63**, 637–672 (2001).
10. Ligneau, X. *et al.* Neurochemical and behavioral effects of ciproxifan, a potent histamine H3-receptor antagonist. *J. Pharmacol. Exp. Ther.* **287**, 658–66 (1998).
11. Pillot, C. *et al.* Ciproxifan, a histamine H3-receptor antagonist/inverse agonist, potentiates neurochemical and behavioral effects of haloperidol in the rat. *J. Neurosci. Off. J. Soc. Neurosci.* **22**, 7272–80 (2002).
12. Pillot, C., Héron, A., Schwartz, J.-C. & Arrang, J.-M. Ciproxifan, a histamine H3-receptor antagonist/inverse agonist, modulates the effects of methamphetamine on neuropeptide mRNA expression in rat striatum. *Eur. J. Neurosci.* **17**, 307–314 (2003).

13. Ligneau, X. *et al.* Distinct pharmacology of rat and human histamine H \textlesssub\textgreater3\textless/sub\textgreater receptors: role of two amino acids in the third transmembrane domain. *Br. J. Pharmacol.* **131**, 1247–1250 (2000).

14. Hill SJ, Ganellin CR, Timmerman H, Schwartz JC, Shankley NP, Young JM, Schunack W, Levi R, H. H. International Union of Pharmacology. XIII. Classification of histamine receptors. *Pharmacol. Rev.* **49**, 253–278 (1997).

15. Panula, P. & Nuutinen, S. The histaminergic network in the brain: basic organization and role in disease. *Nat. Rev. Neurosci.* **14**, 472–487 (2013).

16. Haas, H. L., Sergeeva, O. A. & Selbach, O. Histamine in the nervous system. *Physiol. Rev.* **88**, 1183–1241 (2008).

17. \Lazewska, D. & Kieć-Kononowicz, K. Recent advances in histamine H3 receptor antagonists/inverse agonists. *Expert Opin. Ther. Pat.* **20**, 1147–69 (2010).

18. Tiligada, E., Zampeli, E., Sander, K. & Stark, H. Histamine H3 and H4 receptors as novel drug targets. *Expert Opin. Investig. Drugs* **18**, 1519–31 (2009).

19. Sadek, B., Saad, A., Sadeq, A., Jalal, F. & Stark, H. Histamine H3 receptor as a potential target for cognitive symptoms in neuropsychiatric diseases. *Behav. Brain Res.* **312**, 415–30 (2016).

20. Axe, F. U., Bembenek, S. D. & Szalma, S. Three-dimensional models of histamine H3 receptor antagonist complexes and their pharmacophore. *J. Mol. Graph. Model.* (2006). doi:10.1016/j.jmgm.2005.10.005

21. Massink, A. *et al.* Sodium ion binding pocket mutations and adenosine A2A receptor function. *Mol. Pharmacol.* **87**, 305–13 (2015).

22. Schnell, D. & Seifert, R. Modulation of histamine H3 receptor function by monovalent ions. *Neurosci. Lett.* **472**, 114–118 (2010).

23. Uveges AJ, Kowal D, Zhang Y, Spangler TB, Dunlop J, Semus S, J. P. The Role of Transmembrane Helix 5 in Agonist Binding to the Human H3 Receptor. *J. Pharmacol. Exp. Ther.* **301**, 451–458 (2002).

24. Morisset, S. *et al.* High constitulive activity of native H3 receptors regulates histamine neurons in brain. *Nature* (2000). doi:10.1038/35048583

25. Gbahou, F. *et al.* Protean agonism at histamine H3 receptors in vitro and in vivo. *Proc. Natl. Acad. Sci.* (2003). doi:10.1073/pnas.1932276100

26. Palczewski, K. *et al.* Crystal Structure of Rhodopsin: A G Protein-Coupled Receptor. *Science* **289**, 739–745 (2000).

27. Rai, B. K., Tawa, G. J., Katz, A. H. & Humblet, C. Modeling G protein-coupled receptors for structure-based drug discovery using low-frequency normal modes for refinement of homology models: Application to H3 antagonists. *Proteins Struct. Funct. Bioinforma.* **78**, 457–473 (2010).

28. Levoin, N. *et al.* Refined Docking as a Valuable Tool for Lead Optimization: Application to Histamine H \textlesssub\textgreater3\textless/sub\textgreater Receptor Antagonists. *Arch. Pharm. (Weinheim)* **341**, 610–623 (2008).

29. Levoin, N. *et al.* Determination of the binding mode and interacting amino-acids for dibasic H3 receptor antagonists. *Bioorg. Med. Chem.* **21**, 4526–4529 (2013).

30. Kim, S.-K., Fristrup, P., Abrol, R. & Goddard, W. A. Structure-Based Prediction of Subtype Selectivity of Histamine H \textlesssub\textgreater3\textless/sub\textgreater Receptor Selective Antagonists in Clinical Trials. *J. Chem. Inf. Model.* **51**, 3262–3274 (2011).

31. Kuder, K. *et al.* Chlorophenoxy aminoalkyl derivatives as histamine H3R ligands and antiseizure agents. *Bioorg. Med. Chem.* **24**, 53–72 (2016).

32. Lorenzi, S. *et al.* Validation of a histamine H3 receptor model through structure-activity relationships for classical H3 antagonists. *Bioorg. Med. Chem.* (2005). doi:10.1016/j.bmc.2005.05.072

33. Dastmalchi, S., Hamzeh-Mivehroud, M., Ghafourian, T. & Hamzeiy, H. Molecular modeling of histamine H3 receptor and QSAR studies on arylbenzofuran derived H3 antagonists. *J. Mol. Graph. Model.* (2008). doi:10.1016/j.jmgm.2007.05.002

34. Stark, H. *et al.* Different antagonist binding properties of human and rat histamine H3 receptors. *Bioorg. Med. Chem. Lett.* **11**, 951–954 (2001).

35. Yao, B. . B. *et al.* No Title. *Neuropharmacology* **44**, (2003).

36. Jończyk, J., Malawska, B. & Bajda, M. Hybrid approach to structure modeling of the histamine H3 receptor: Multi-level assessment as a tool for model verification. *PLOS ONE* **12**, e0186108 (2017).

37. Morini, G. *et al.* Dibasic non-imidazole histamine H3 receptor antagonists with a rigid biphenyl scaffold. *Bioorg. Med. Chem. Lett.* (2006). doi:10.1016/j.bmcl.2006.04.092

38. Wittmann, H.-J., Seifert, R. & Strasser, A. Sodium binding to hH3R and hH4R — a molecular modeling study. *J. Mol. Model.* **20**, 2394 (2014).

39. Bissantz, C., Logean, A. & Rognan, D. High-throughput modeling of human g-protein coupled receptors: Amino acid sequence alignment, three-dimensional model building, and receptor library screening. *J. Chem. Inf. Comput. Sci.* (2004). doi:10.1021/ci034181a

40. Stansfeld, P. J. J. *et al.* MemProtMD: Automated Insertion of Membrane Protein Structures into Explicit Lipid Membranes. *Structure* **23**, 1350–1361 (2015).

41. Johnston, J. M. & Filizola, M. Showcasing modern molecular dynamics simulations of membrane proteins through G protein-coupled receptors. *Curr. Opin. Struct. Biol.* **21**, 552–8 (2011).

42. Plazinska, A., Kolinski, M., Wainer, I. W. & Jozwiak, K. Molecular interactions between fenoterol stereoisomers and derivatives and the β2-adrenergic receptor binding site studied by docking and molecular dynamics simulations. *J. Mol. Model.* (2013). doi:10.1007/s00894-013-1981-y

43. Huber, T., Botelho, A. V., Beyer, K. & Brown, M. F. Membrane model for the G-protein-coupled receptor rhodopsin: hydrophobic interface and dynamical structure. *Biophys. J.* **86**, 2078–2100 (2004).

44. Pitman, M. C., Suits, F., Gawrisch, K. & Feller, S. E. Molecular dynamics investigation of dynamical properties of phosphatidylethanolamine lipid bilayers. *J. Chem. Phys.* (2005). doi:10.1063/1.1899153

45. Pitman, M. C., Grossfield, A., Suits, F. & Feller, S. E. Role of cholesterol and polyunsaturated chains in lipid-protein interactions: Molecular dynamics simulation of rhodopsin in a realistic membrane environment. *J. Am. Chem. Soc.* (2005). doi:10.1021/ja042715y

46. Pitman, M. C., Suits, F., MacKerell, A. D. & Feller, S. E. Molecular-level organization of saturated and polyunsaturated fatty acids in a phosphatidylcholine bilayer containing cholesterol. *Biochemistry* (2004). doi:10.1021/bi048231w

47. Suits, F., Pitman, M. C. & Feller, S. E. Molecular dynamics investigation of the structural properties of phosphatidylethanolamine lipid bilayers. *J. Chem. Phys.* **122**, 244714 (2005).

48. Grossfield, A., Feller, S. E. & Pitman, M. C. A role for direct interactions in the modulation of rhodopsin by -3 polyunsaturated lipids. *Proc. Natl. Acad. Sci.* (2006). doi:10.1073/pnas.0508352103

49. Grossfield, A., Feller, S. E. & Pitman, M. C. Contribution of omega-3 fatty acids to the thermodynamics of membrane protein solvation. *J. Phys. Chem. B* (2006). doi:10.1021/jp060405r

50. Rubenstein, L. A., Zauhar, R. J. & Lanzara, R. G. Molecular dynamics of a biophysical model for β2-adrenergic and G protein-coupled receptor activation. *J. Mol. Graph. Model.* (2006). doi:10.1016/j.jmgm.2006.02.008

51. Kong, Y. & Karplus, M. The Signaling Pathway of Rhodopsin. *Structure* **15**, 611–623 (2007).

52. Espinoza-Fonseca, L. M., Pedretti, A. & Vistoli, G. Structure and dynamics of the full-length M1 muscarinic acetylcholine receptor studied by molecular dynamics simulations. *Arch. Biochem. Biophys.* (2008). doi:10.1016/j.abb.2007.09.002

53. Neri, M., Vanni, S., Tavernelli, I. & Rothlisberger, U. Role of aggregation in rhodopsin signal transduction. *Biochemistry* (2010). doi:10.1021/bi100478j

54. Stansfeld, P. J. & Sansom, M. S. P. Molecular simulation approaches to membrane proteins. *Struct. Lond. Engl. 1993* **19**, 1562–1572 (2011).

55. Shaw, D. E. *et al.* Anton, a special-purpose machine for molecular dynamics simulation. *Commun. ACM* **51**, 91 (2008).

56. Chavent, M., Duncan, A. L. & Sansom, M. S. P. *Molecular dynamics simulations of membrane proteins and their interactions: From nanoscale to mesoscale*. (2016). doi:10.1016/j.sbi.2016.06.007

57. Gantz, I. *et al.* Molecular basis for the interaction of histamine with the histamine H2 receptor. *J. Biol. Chem.* **267**, 20840–20843 (1992).

58. Shin, N. *et al.* Molecular modeling and site-specific mutagenesis of the histamine-binding site of the histamine H4 receptor. *Mol. Pharmacol.* **62**, 38–47 (2002).

59. West, R. E., Zweig, A., Granzow, R. T., Siegel, M. I. & Egan, R. W. Biexponential kinetics of (R)-alpha-[3H]methylhistamine binding to the rat brain H3 histamine receptor. *J. Neurochem.* **55**, 1612–6 (1990).

60. Arrang, J. M., Roy, J., Morgat, J. L., Schunack, W. & Schwartz, J. C. Histamine H3 receptor binding sites in rat brain membranes: modulations by guanine nucleotides and divalent cations. *Eur. J. Pharmacol.* **188**, 219–27 (1990).

61. Kappel, K., Miao, Y. & Andrew McCammon, J. Accelerated molecular dynamics simulations of ligand binding to a muscarinic G-protein-coupled receptor. *Q. Rev. Biophys.* (2015). doi:10.1017/S0033583515000153

62. Liu, W. *et al.* Structural Basis for Allosteric Regulation of GPCRs by Sodium Ions. *Science* **337**, 232–236 (2012).

63. Yuan, S. *et al.* The mechanism of ligand-induced activation or inhibition of μ- And κ-opioid receptors. *Angew. Chem. - Int. Ed.* (2015). doi:10.1002/anie.201501742

64. Holst, B., Rosenkilde, M. M., Elling, C. E., Frimurer, T. M. & Schwartz, T. W. MOLECULAR MECHANISM OF 7TM RECEPTOR ACTIVATION—A GLOBAL TOGGLE SWITCH MODEL. *Annu. Rev. Pharmacol. Toxicol.* (2006). doi:10.1146/annurev.pharmtox.46.120604.141218

65. Nygaard, R., Frimurer, T. M., Holst, B., Rosenkilde, M. M. & Schwartz, T. W. *Ligand binding and micro-switches in 7TM receptor structures*. (2009). doi:10.1016/j.tips.2009.02.006

66. Neale, C., Herce, H. D., Pomès, R. & García, A. E. Can Specific Protein-Lipid Interactions Stabilize an Active State of the Beta 2 Adrenergic Receptor? *Biophys. J.* **109**, 1652–62 (2015).

67. Nieto-Alamilla, G., Marquez-Gomez, R., Garcia-Galvez, A.-M., Morales-Figueroa, G.-E. & Arias-Montano, J.-A. The Histamine H3 Receptor: Structure, Pharmacology, and Function. *Mol. Pharmacol.* (2016). doi:10.1124/mol.116.104752

68. Dror, R. O., Jensen, M. Ø. & Shaw, D. E. Elucidating membrane protein function through long-timescale molecular dynamics simulation. *Conf. Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. IEEE Eng. Med. Biol. Soc. Annu. Conf.* **2009**, 2340–2342 (2009).

69. Nygaard, R. *et al.* The dynamic process of β2-adrenergic receptor activation. *Cell* (2013). doi:10.1016/j.cell.2013.01.008

70. Laskowski, R. A. & Swindells, M. B. LigPlot+: Multiple ligand-protein interaction diagrams for drug discovery. *J. Chem. Inf. Model.* (2011). doi:10.1021/ci200227u

71. Seifert, R. & Wenzel-Seifert, K. Unmasking different constitutive activity of four chemoattractant receptors using Na+ as universal stabilizer of the inactive (R) state. *Receptors.Channels* (2001).

72. Vilardaga, J. P., Bünemann, M., Krasell, C., Castro, M. & Lohse, M. J. Measurement of the millisecond activation switch of G protein-coupled receptors in living cells. *Nat. Biotechnol.* (2003). doi:10.1038/nbt838

73. Vilardaga, J.-P. Theme and variations on kinetics of GPCR activation/deactivation. *J. Recept. Signal Transduct.* **30**, 304–312 (2010).

74. Ciruela, F., Vilardaga, J. P. & Fernández-Dueñas, V. *Lighting up multiprotein complexes: Lessons from GPCR oligomerization*. (2010). doi:10.1016/j.tibtech.2010.05.002

75. Vilardaga, J. P., Steinmeyer, R., Harms, G. S. & Lohse, M. J. *Molecular Basis of Inverse Agonism in a G Protein-Coupled Receptor*. (2005). doi:10.1038/nchembio705

76. Kenakin, T. *Principles: Receptor theory in pharmacology*. **25**, (Elsevier Current Trends, 2004).

77. Kenakin, T. Theoretical Aspects of GPCR–Ligand Complex Pharmacology. *Chem. Rev.* **117**, 4–20 (2017).

78. Rosenbaum, D. M. *et al.* Structure and function of an irreversible agonist-β(2) adrenoceptor complex. *Nature* **469**, 236–40 (2011).

79. Rasmussen, S. G. F. *et al.* Crystal structure of the β2 adrenergic receptor-Gs protein complex. *Nature* **477**, 549–55 (2011).

80. Ghamari, N. *et al. Histamine H3 receptor ligands by hybrid virtual screening, docking, molecular dynamics simulations, and investigation of their biological effects*. (2019). doi:10.1111/cbdd.13471

81. Schlegel, B., Stark, H., Sippl, W. & Höltje, H.-D. Model of a specific human histamine H3 receptor (hH3R) binding pocket suitable for virtual drug design. *Inflamm. Res.* **54**, S50–S51 (2005).

82. Faraldo-Gómez, J. D. *et al.* Conformational sampling and dynamics of membrane proteins from 10-nanosecond computer simulations. *Proteins Struct. Funct. Genet*. (2004). doi:10.1002/prot.20257

83. Horst, R., Liu, J. J., Stevens, R. C. & Wüthrich, K. ??2-adrenergic receptor activation by agonists studied with 19F NMR spectroscopy. *Angew. Chem. - Int. Ed.* **52**, 10762–10765 (2013).

84. Peeters, M. C., Van Westen, G. J. P., Li, Q. & Ijzerman, A. P. *Importance of the extracellular loops in G protein-coupled receptors for ligand recognition and receptor activation*. (2011). doi:10.1016/j.tips.2010.10.001

85. Klco, J. M., Wiegand, C. B., Narzinski, K. & Baranski, T. J. Essential role for the second extracellular loop in C5a receptor activation. *Nat. Struct. Mol. Biol.* (2005). doi:10.1038/nsmb913

86. Scarselli, M., Li, B., Kim, S. K. & Wess, J. Multiple residues in the second extracellular loop are critical for M 3 muscarinic acetylcholine receptor activation. *J. Biol. Chem.* (2007). doi:10.1074/jbc.M610394200

87. Sykes DA, Dowling MR, C. S. Exploring the mechanism of agonist efficacy: a relationship between efficacy and agonist dissociation rate at the muscarinic M3 receptor. *Mol. Pharmacol.* **76**, 543–551 (2009).

88. Palsdottir, H. & Hunte, C. Lipids in membrane protein structures. *Biochim. Biophys. Acta BBA - Biomembr.* **1666**, 2–18 (2004).

89. Hite, R. K., Li, Z. & Walz, T. Principles of membrane protein interactions with annular lipids deduced from aquaporin-0 2D crystals. *EMBO J.* **29**, 1652–8 (2010).

90. Oates, J. & Watts, A. Uncovering the intimate relationship between lipids, cholesterol and GPCR activation. *Curr. Opin. Struct. Biol.* **21**, 802–807 (2011).

91. Contreras, F.-X., Ernst, A. M., Wieland, F. & Brügger, B. Specificity of intramembrane protein-lipid interactions. *Cold Spring Harb. Perspect. Biol.* **3**, a004705–a004705 (2011).

92. Hedger, G. & Sansom, M. S. P. Lipid interaction sites on channels, transporters and receptors: Recent insights from molecular dynamics simulations. *Biochim. Biophys. Acta* **1858**, 2390–2400 (2016).

93. Yeagle, P. L. Non-covalent binding of membrane lipids to membrane proteins. *Biochim. Biophys. Acta BBA - Biomembr.* **1838**, 1548–1559 (2014).

94. Cang, X. *et al.* Mapping the Functional Binding Sites of Cholesterol in β \textlesssub\textgreater2\textless/sub\textgreater -Adrenergic Receptor by Long-Time Molecular Dynamics Simulations. *J. Phys. Chem. B* **117**, 1085–1094 (2013).

95. Ballesteros, J. A. & 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). doi:10.1016/S1043-9471(05)80049-7

96. Isberg, V. *et al.* Generic GPCR residue numbers - aligning topology maps while minding the gaps. *Trends Pharmacol. Sci.* **36**, 22–31 (2015).

97. Webb, B. & Sali, A. Comparative protein structure modeling using MODELLER. *Curr. Protoc. Bioinforma.* (2016). doi:10.1002/cpbi.3

98. Ozgur, C., Doruker, P. & Akten, E. D. Investigation of allosteric coupling in human β2-adrenergic receptor in the presence of intracellular loop 3. *BMC Struct. Biol.* (2016). doi:10.1186/s12900-016-0061-9

99. Gbahou, F., Rouleau, A. & Arrang, J.-M. The histamine autoreceptor is a short isoform of the H3 receptor. *Br. J. Pharmacol.* **166**, 1860–1871 (2012).

100. Trott, O. & Olson, A. J. Software news and update AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J. Comput. Chem.* (2010). doi:10.1002/jcc.21334

101. Laskowski, R. A., MacArthur, M. W., Moss, D. S., Thornton, J. M. & IUCr. PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.* **26**, 283–291 (1993).

102. Vriend, G. WHAT IF: A molecular modeling and drug design program. *J. Mol. Graph.* (1990). doi:10.1016/0263-7855(90)80070-V

103. Jo, S., Lim, J. B., Klauda, J. B. & Im, W. CHARMM-GUI membrane builder for mixed bilayers and its application to yeast membranes. *Biophys. J.* (2009). doi:10.1016/j.bpj.2009.04.013

104. Wu, E. L. *et al. CHARMM-GUI membrane builder toward realistic biological membrane simulations*. (2014). doi:10.1002/jcc.23702

105. Jensen, J. H., Li, H., Robertson, A. D. & Molina, P. A. Prediction and rationalization of protein pK a values using QM and QM/MM methods. *J. Phys. Chem. A* (2005). doi:10.1021/jp051922x

106. Kucerka, N. *et al.* Lipid bilayer structure determined by the simultaneous analysis of neutron and X-ray scattering data. *Biophys. J.* **95**, 2356–2367 (2008).

107. Lobanov, M. Y. I., Bogatyreva, N. S. & Galzitskaya, O. V. Radius of gyration is indicator of compactness of protein structure. *Mol. Biol.* (2008). doi:10.1134/S0026893308040195

108. Amadei, A., Linssen, A. B. M. & Berendsen, H. J. C. Essential dynamics of proteins. *Proteins Struct. Funct. Bioinforma.* (1993). doi:10.1002/prot.340170408

109. Grossfield A, Feller SE, P. M. Convergence of molecular dynamics simulations of membrane proteins. **67**, 31–40 (2007).

110. Mezei, M. Simulaid: A simulation facilitator and analysis program. *J. Comput. Chem.* (2010). doi:10.1002/jcc.21551

111. Visiers, I., Braunheim, B. B. & Weinstein, H. Prokink: a protocol for numerical evaluation of helix distortions by proline. *Protein Eng. Des. Sel.* (2000). doi:10.1093/protein/13.9.603

112. Tsoulos, I. G. & Stavrakoudis, A. Eucb: A C++ program for molecular dynamics trajectory analysis. *Comput. Phys. Commun.* (2011). doi:10.1016/j.cpc.2010.11.032

113. Mezei, M. & Filizola, M. TRAJELIX: A computational tool for the geometric characterization of protein helices during molecular dynamics simulations. *J. Comput. Aided Mol. Des.* (2006). doi:10.1007/s10822-006-9039-1

114. Glykos, N. Software news and updates. CARMA: A molecular dynamics analysis program. *J. Comput. Chem.* **27**, 1765–1768 (2006).

115. Humphrey, W., Dalke, A. & Schulten, K. VMD: Visual molecular dynamics. *J. Mol. Graph.* (1996). doi:10.1016/0263-7855(96)00018-5

116. Schmidtke, P., Le Guilloux, V., Maupetit, J. & Tufféry, P. fpocket: Online tools for protein ensemble pocket detection and tracking. *Nucleic Acids Res.* (2010). doi:10.1093/nar/gkq383

117. Hernandez, M., Ghersi, D. & Sanchez, R. SITEHOUND-web: A server for ligand binding site identification in protein structures. *Nucleic Acids Res.* (2009). doi:10.1093/nar/gkp281

118. Chovancova, E. *et al.* CAVER 3.0: A Tool for the Analysis of Transport Pathways in Dynamic Protein Structures. *PLoS Comput. Biol.* (2012). doi:10.1371/journal.pcbi.1002708

119. Rasmussen, S. G. F. *et al.* Structure of a nanobody-stabilized active state of the β(2) adrenoceptor. *Nature* **469**, 175–80 (2011).

120. Weichert, D. *et al.* Covalent agonists for studying G protein-coupled receptor activation. *Proc. Natl. Acad. Sci. U. S. A.* **111**, 10744–10748 (2014).

121. Kiss, R. & Keserű, G. M. Structure-based discovery and binding site analysis of histamine receptor ligands. *Expert Opin. Drug Discov.* **11**, 1165–1185 (2016).

122. Neel, A. J., Hilton, M. J., Sigman, M. S. & Toste, F. D. *Exploiting non-covalent π interactions for catalyst design*. (2017). doi:10.1038/nature21701

123. Schwartz, T. W., Frimurer, T. M., Holst, B., Rosenkilde, M. M. & Elling, C. E. Molecular mechanism of 7TM receptor activation--a global toggle switch model. *Annu. Rev. Pharmacol. Toxicol.* **46**, 481–519 (2006).

124. Kofuku, Y. *et al.* Efficacy of the β 2-adrenergic receptor is determined by conformational equilibrium in the transmembrane region. *Nat. Commun.* (2012). doi:10.1038/ncomms2046

Figure legends

Figure 1. Chemical Structure of a) histamine and b) ciproxifan.

Figure 2. The global dynamics of the structures of the three systems identified by cartesian Principal Component Analysis. The panels a), b) and c) show a pseudo color representation of the distribution of the first two principal components PC1, PC2 obtained from a ~ 1 µs MD simulation for the antagonist, agonist and apo structures, respectively.

Figure 3. The global dynamics of the structures of the three systems identified by cartesian Principal Component Analysis. The panels a), b) and c) show a pseudo color representation of the distribution of the principal components PC1, PC3 obtained from a ~ 1 µs MD simulation for the antagonist, agonist and apo structures, respectively.

Figure 4. The global dynamics of the structures of the three systems identified by cartesian Principal Component Analysis. The panels a), b) and c) shows a pseudo color representation of the distribution principal components PC2, PC3 obtained from a ~ 1 µs MD simulation for the antagonist, agonist and apo structures, respectively.

Figure 5. The χ_1 torsion for the side chain of residues (a) Trp6.48 and (b) Phe5.47. (c) concomitant transitions of Chi1 torsions for the same residues.

Figure 6. Schematic representation of the DRF motif ($^{3.49}$AspArgPhe$^{3.51}$), with a bifurcated h-bond originating from Arg3.50. (a) side view; (b) cytoplasmic view.

Figure 7. Schematic representation of a small hydrophobic cluster formed by Ala6.40 and Tyr7.53.

Figure 8. 2D-binding mode of ciproxifan within the H3 receptor. Polar interactions between the amino acids, water and ciproxifan ligand (green line: H-bond) can be appreciated. The a) to f) panels show the most relevant configurations and conformations of CPX throughout the simulation.

Figure 9. 2D-binding mode of histamine within the H3 receptor. Polar interactions between the amino acids, water and histamine ligand (green line: H-bond, magenta line: hydrophobic interaction, light green: π staking)

can be appreciated. The three panels correspond to sections 1-3 in the trajectory and show the most relevant configurations and conformations of histamine throughout the simulation.

Figure 10. 2D-binding major mode of two DPPC lipids to the H3 receptor in the antagonist state showing the interacting amino acid residues.

Figure 11. 3D representation of the average binding mode of DPPC lipids to the H3 receptor in the antagonist state. Above, the extracellular face; below, the intracellular face; the ciproxifan ligand is visualized in CPK code at the center of these.

Figure 12. 2D-binding major mode of DPPC lipids to the H3 receptor in the agonist state showing the interacting amino acid residues. (a) Two amino acids with h-bonds to the lipid. Multiple DPPC interactions can be observed. (b) Several water molecules (in red) are present, in conjunction with seven amino acids.

Figure 13. Lipid-receptor interactions for the three systems. The images are representatives from the whole production phase of the simulation. The upper part shows the thickness of the lipid layer; the lower part view shows just the lipids without the H3 receptor and allows to see the lipids in the backward part of the membrane.

Figure 14. 3D superposition of the average structure of each system. In yellow, the antagonist-H3R complex; in blue the agonist-H3R complex; in red the apo receptor.

Supplemental figure legends

Figure S1. Time-dependent Root Mean Square Deviation (RMSD) of antagonist (yellow), agonist (red) and apo (blue) structures. The observed fluctuations suggest a good packing quality of each model.

Figure S2. The 2D RMSD matrix. Panels a) to c) correspond to the antagonist- and agonist-H3R complexes, and to the apo receptor, respectively.

Figure S3. The 2D contact maps. Panels a) to c) correspond to the antagonist- and agonist-H3R complexes, and to the apo receptor, respectively.

Figure S4. Radius of gyration (Rg) of antagonist complex (yellow), agonist complex (red), and apo (blue) structures.

Figure S5. Root mean square fluctuation (RMSF) of antagonist complex (yellow), agonist complex (red), and apo (blue) structures.

Figure S6. Clustering of the antagonist state structures. Visualization of the movements of the first principal component PC1 in the four clusters, from blue to green depicting low to high atomic displacements. Panels a) cluster C1, b) cluster C2, c) cluster C3, and d) cluster C4 as determined with the Carma package.

Figure S7. Clustering of the agonist state structures. Visualization of the first principal component PC1 of the only agonist cluster C1 as determined with the Carma package. Color scale from blue to green depicts low to high atomic displacements.

Figure S8. Clustering of the apo state structure. Visualization of the first principal component PC1 of the three clusters C1, C2 and C3 as determined by the Carma package. Color scale from blue to green depict low to high atomic displacements. Panels, a) cluster C1, b) cluster C2, and c) cluster C3.

Figure S9. Networks of charged residues for the antagonist-H3R complex. The two charged amino acids networks stabilize the structure of the receptor, important for the formation of an electrostatic clamp.

Figure S10. Networks of charged residues for the agonist-H3R complex. In the two first sections at least one cluster interacts with the ligand. The first section in the trajectory shows two inside electrostatic networks, whereas the second section shows three.

Figure S11. Networks of charged residues for the apo receptor. The four charged amino acids networks stabilize the apo structure. The first three networks involve less than six residues.

Figure S12 a) H-bonded networks for the antagonist-H3R complex. The seven networks involve several internal water molecules and clusters. Other water molecules mediate the interactions between charged and polar amino acid residues. b) H-bonded networks for the agonist-H3R complex. The first network shows a water molecule mediating the interaction between the histamine ligand and Trp7.43. c) H-bonded networks for the

apo receptor. As for the antagonist and agonist-H3R complexes, the water molecules mediate interactions between residues and are polyvalent.

Figure S13. Hydrophobic clusters formed by aromatic amino acids in the case of the antagonist-H3R complex. a) stacking interactions coming from TM2, TM3 and EC1; b) stacking interaction between two aromatic residues belonging to TM7 and H8.

Figure S14a to S21a. The global helix X tilt angle for all eight TM helices of the antagonist-H3R complex, respectively.

Figure S14b to S21b. The global helix Y tilt angle for all eight TM helices of the antagonist-H3R complex, respectively.

Figure S14c to S21c. The global helix Z tilt angle for all eight TM helices of the antagonist-H3R complex, respectively.

Figure S14d to S21d. The helix rotation for all eight TM helices of the antagonist-H3R complex, respectively.

Figure S14e to S21e. The local helix tilt angle for all eight TM helices of the antagonist-H3R complex, respectively.

Figure S14f to S21f. The turn angle per residue (TPR) for all eight TM helices of the antagonist-H3R complex, respectively.

Figure S22a to S29a. The global helix X tilt angle for all eight TM helices of the agonist-H3R complex.

Figure S22b to S29b. The global helix Y tilt angle for all eight TM helices of the agonist-H3R complex

Figure S22c to S29c. The global helix Z tilt angle for all eight TM helices of the agonist-H3R complex

Figure S22d to S29d. The helix rotation for all eight TM helices of the agonist-H3R complex

Figure S22e to S29e. Idem for the agonist-H3R complex

Figure S22f to S29f. Idem for the agonist-H3R complex

Figure S30a to S37a. Idem for the apo receptor.

Figure S30b to S37b. Idem for the apo receptor

Figure S30c to S37c. Idem for the apo receptor

Figure S30d to S37d. Idem for the apo receptor

Figure S30e to S37e. Idem for the apo receptor

Figure S30f to S37f. Idem for the apo receptor

Figure S38. a) bend, b) wobble and c) face-shift angles of the antagonist-H3R complex around proline residue Pro 5.50 (TM5).

Figure S39. idem for Pro 6.50 (TM6).

Figure S40 idem for Pro 7.50 (TM7).

Figure S41. a) bend, b) wobble and c) face-shift angles of the agonist-H3R complex around proline residue Pro 5.50 (TM5).

Figure S42. idem for Pro 6.50 (TM6).

Figure S43. idem for Pro 7.50 (TM7).

Figure S44. a) bend, b) wobble and c) face-shift angles of the apo receptor around proline residue Pro 5.50 (TM5).

Figure S45. idem for Pro 6.50 (TM6).

Figure S46. idem for Pro 7.50 (TM7).

Figure S47. Helix length and radius of fitted circle for the antagonist-H3R complex. In red color is the radius of the fitted circle, and in blue the helix length. Figures a to h correspond to helices TM1-TM7 and H8, respectively.

Figure S48. Helix length and radius of fitted circle for the agonist-H3R complex. In red color is the radius of the fitted circle, and in blue the helix length. Figures a to h correspond to helices TM1-TM7 and H8, respectively.

Figure S49. Helix length and radius of fitted circle for the apo receptor. In red color is the radius of the fitted circle, and in blue the helix length. Figures a to h correspond to helices TM1-TM7 and H8, respectively.

Figure S50. Total helix displacement (in blue) and Helix shift the Z direction (in red) for the antagonist-H3R complex. Figures a to h correspond to helices TM1-TM7 and H8, respectively.

Figure S51. idem for the agonist-H3R complex.

Figure S52. idem for the apo receptor.

Table legends

Table 1. Amino acid residues involved in the formation of the receptor cavity for the antagonist, agonist and apo systems.

Table 2. Selected inter-residue distances and ionic locks.

Table 3. Interactions between the ciproxifan antagonist/inverse agonist and the receptor.

Table 4. Interactions between the histamine agonist and the receptor for the four segments of the production trajectory.

Table 5. Lipid-binding sites on the inactive-state receptor (antagonist-bound).

Table 6. Lipid-binding sites on the active-state receptor (agonist-bound).

Table 7. Lipid-binding sites on the constitutive-state receptor.

Supplemental table legends

Table S1. H-bonding residue pairs for residence times greater than 70% for the antagonist, agonist and apo systems, respectively.

Table S2. Hydrophobic clusters at 90% occupancy formed by at least three side chains for the antagonist-H3R complex, the agonist-H3R complex and the apo receptor, respectively.

Table S3. Aromatic residue clusters at 90% occupancy formed by at least three side chains for the antagonist-H3R complex, the agonist-H3R complex and the apo receptor, respectively.

Table S4. Number of residues, average length, average length per residue and difference between the maximum and minimum values of the length in the trajectory for each helix of the antagonist-H3R complex.

Table S5. Number of residues, average length, average length per residue and difference between the maximum and minimum values of the length in the trajectory for each helix of the agonist-H3R complex.

Table S6. Number of residues, average length, average length per residue and difference between the maximum and minimum values of the length in the trajectory for each helix of the apo receptor.

Table S7. Conformational states of Met residues for the three systems -antagonist, agonist and apo.

Table S8. Residue population in contact with the lipids grouped according to amino acid types (non-polar aliphatic, uncharged polar, aromatic, and charged) for the antagonist, agonist and apo systems, respectively.

Table S9 Average composition of transmembrane surface amino acid residues.

FIGURES

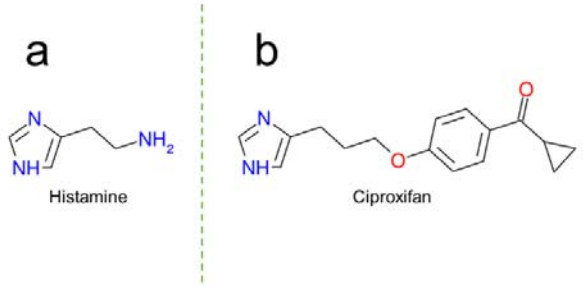


Fig. 1a-b

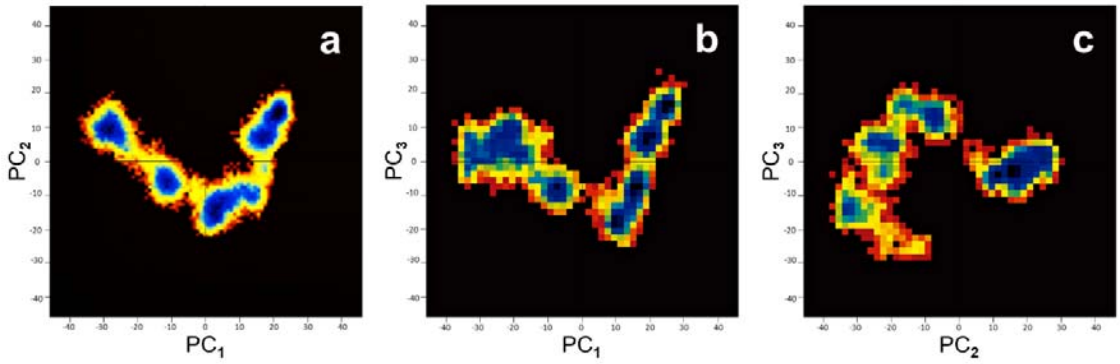


Fig.2a-c

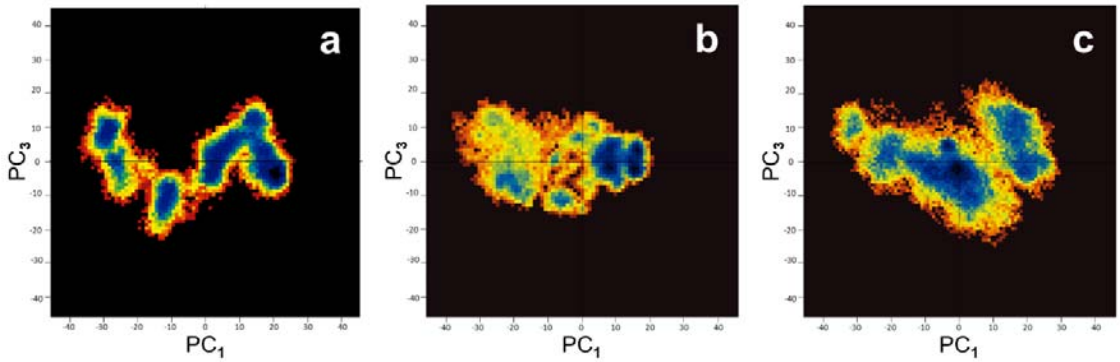


Fig.3a-c

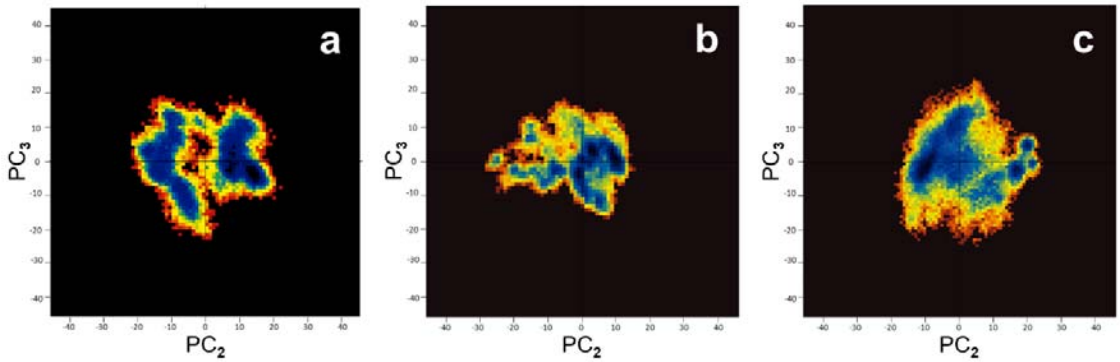
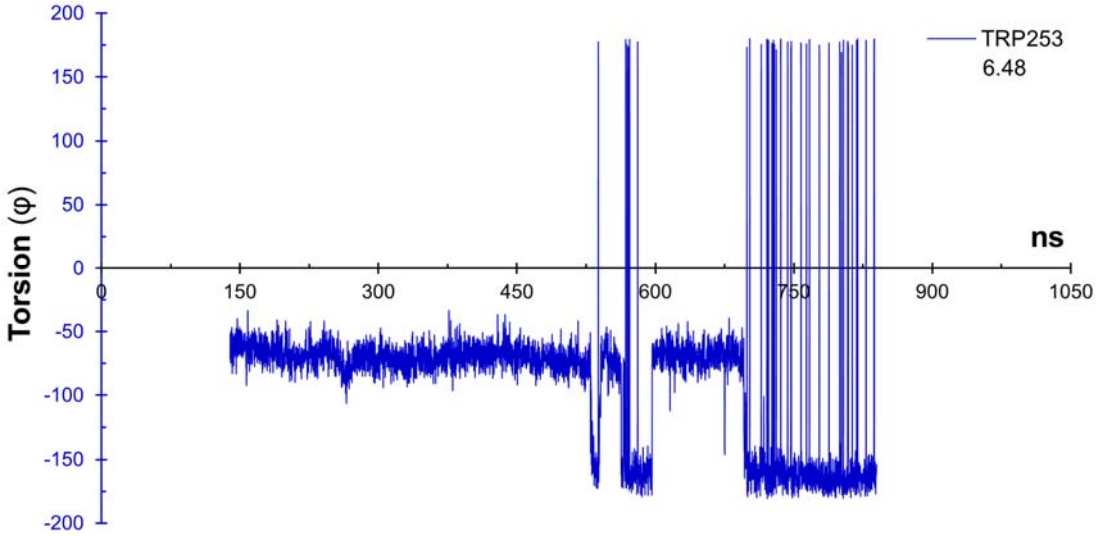
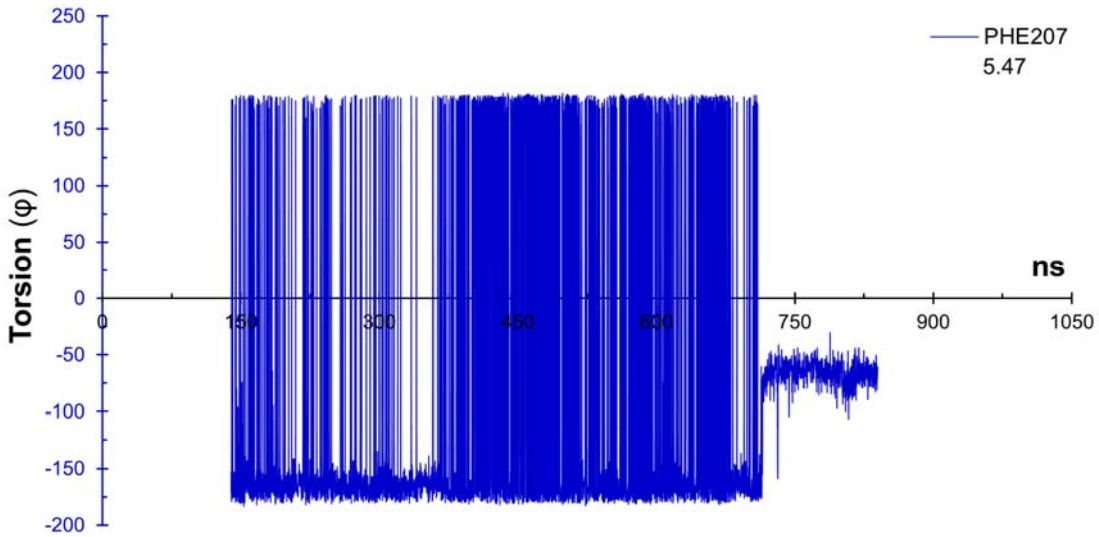

a
b
c
PC3
PC2

Fig.4a-c

a


TRP253
6.48
Torsion (φ)
ns
200
150
100
50
0
-50
-100
-150
-200
0
150
300
450
600
750
900
1050

b


PHE207
5.47
Torsion (φ)
ns
250
200
150
100
50
0
-50
-100
-150
-200
0
150
300
450
600
750
900
1050

c

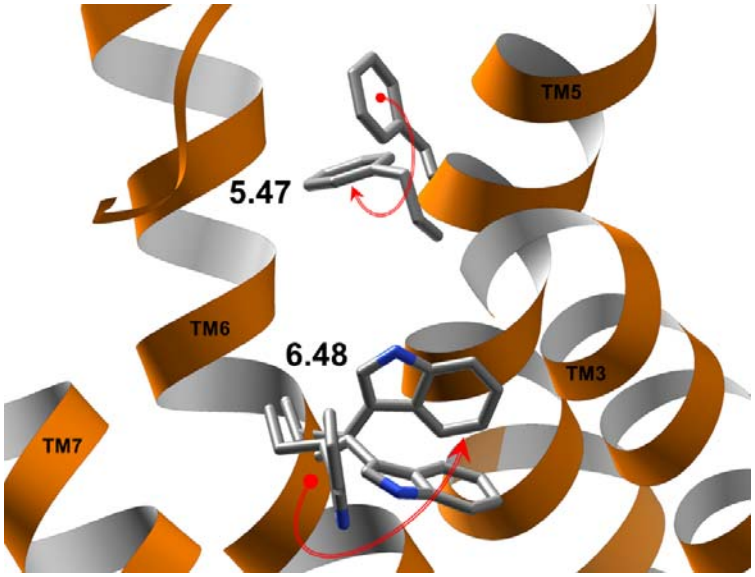

TM5
5.47
TM6
6.48
TM3
TM7

Fig. 5a-c

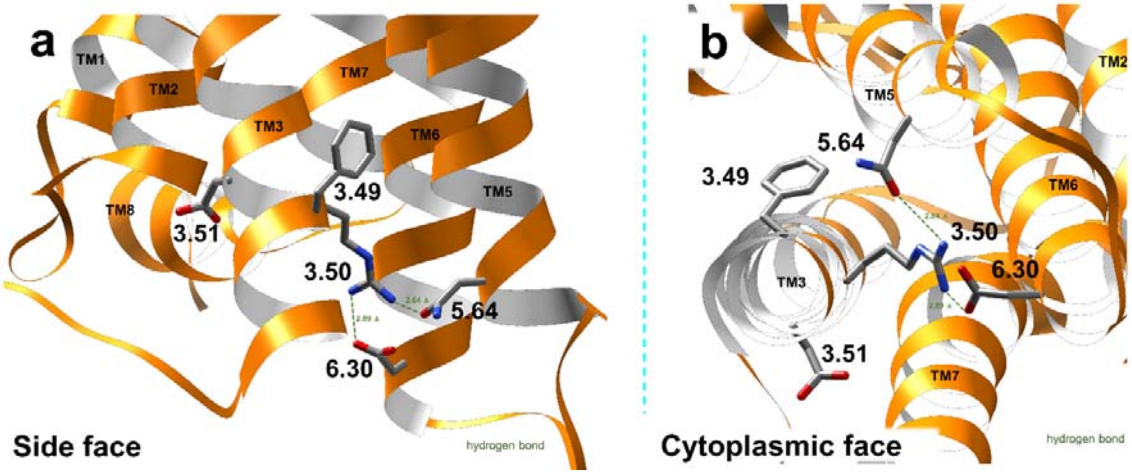

a
TM1
TM2
TM7
TM3
TM6
TM5
TM8
3.49
3.51
3.50
5.64
6.30
Side face
hydrogen bond
b
TM2
TM5
5.64
3.49
TM6
3.50
6.30
TM3
3.51
TM7
Cytoplasmic face
hydrogen bond

Fig. 6a-b

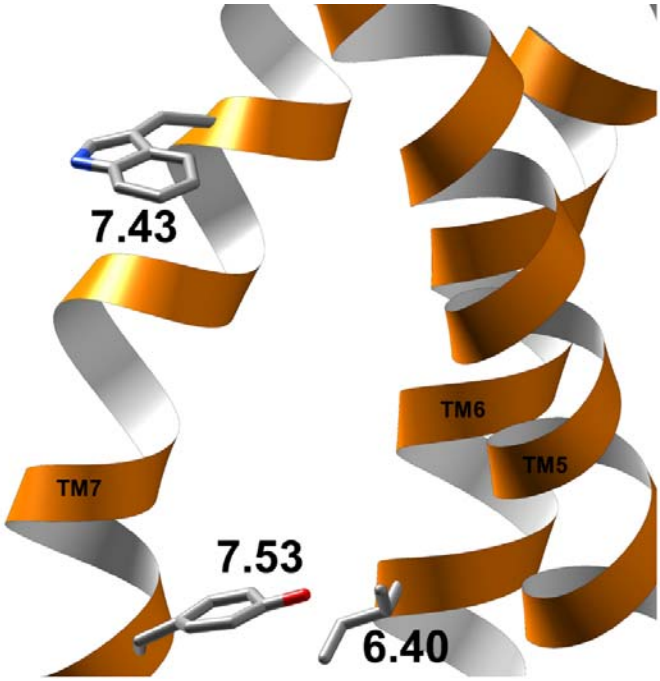
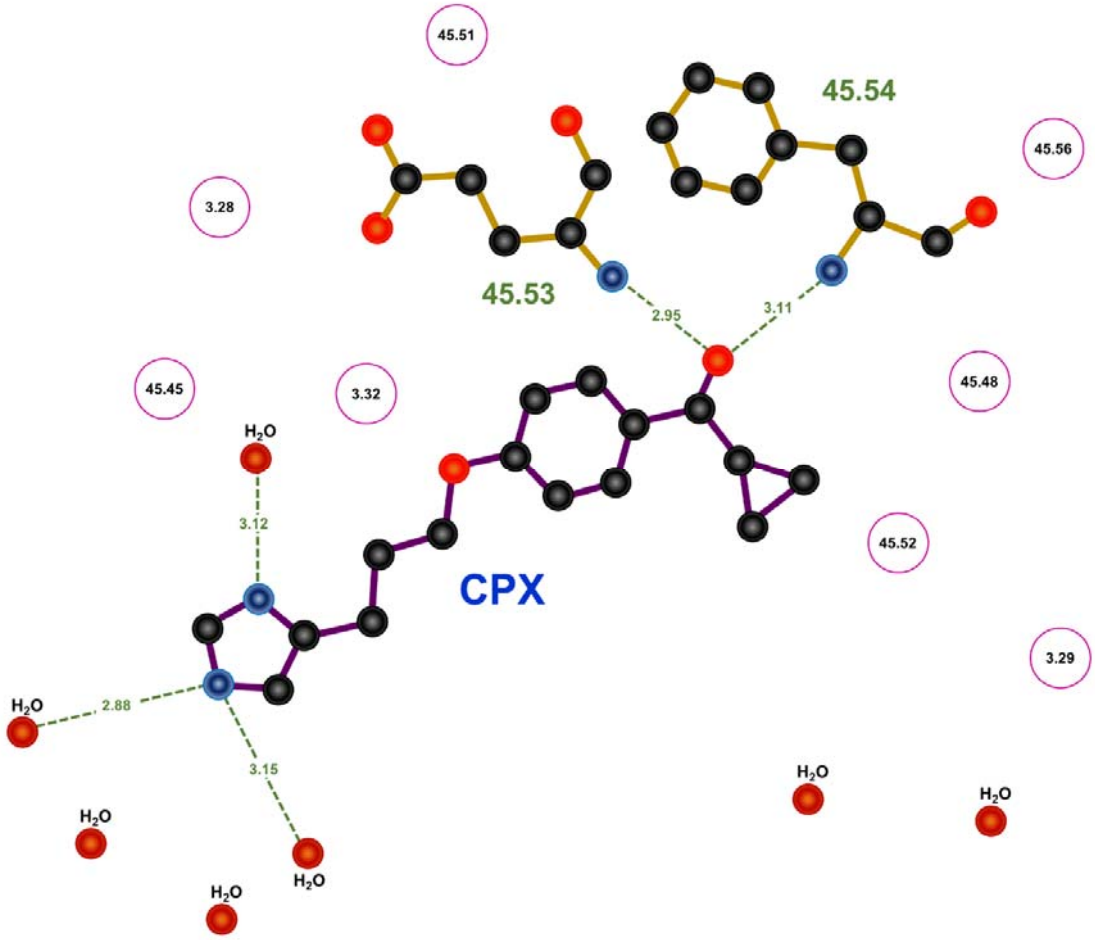
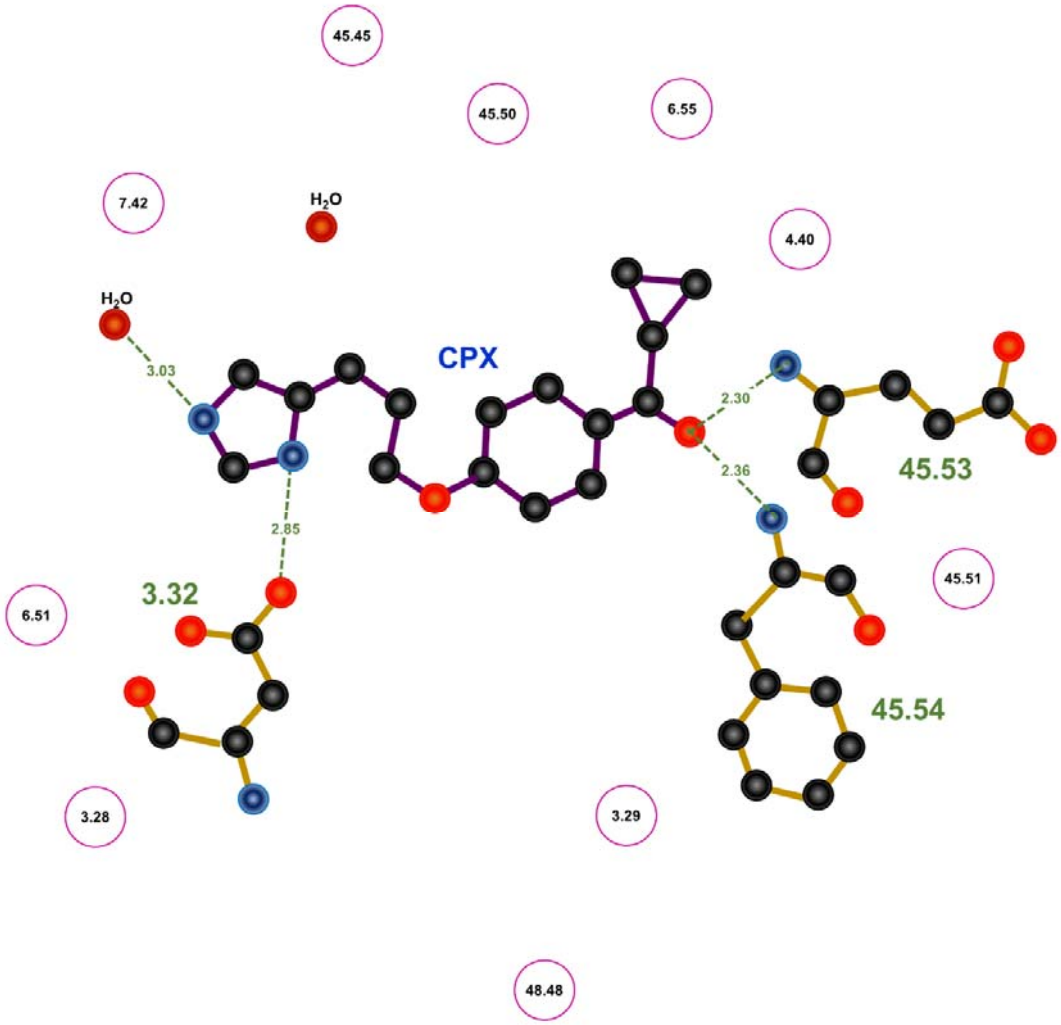

7.43
TM6
TM5
TM7
7.53
6.40

Fig. 7

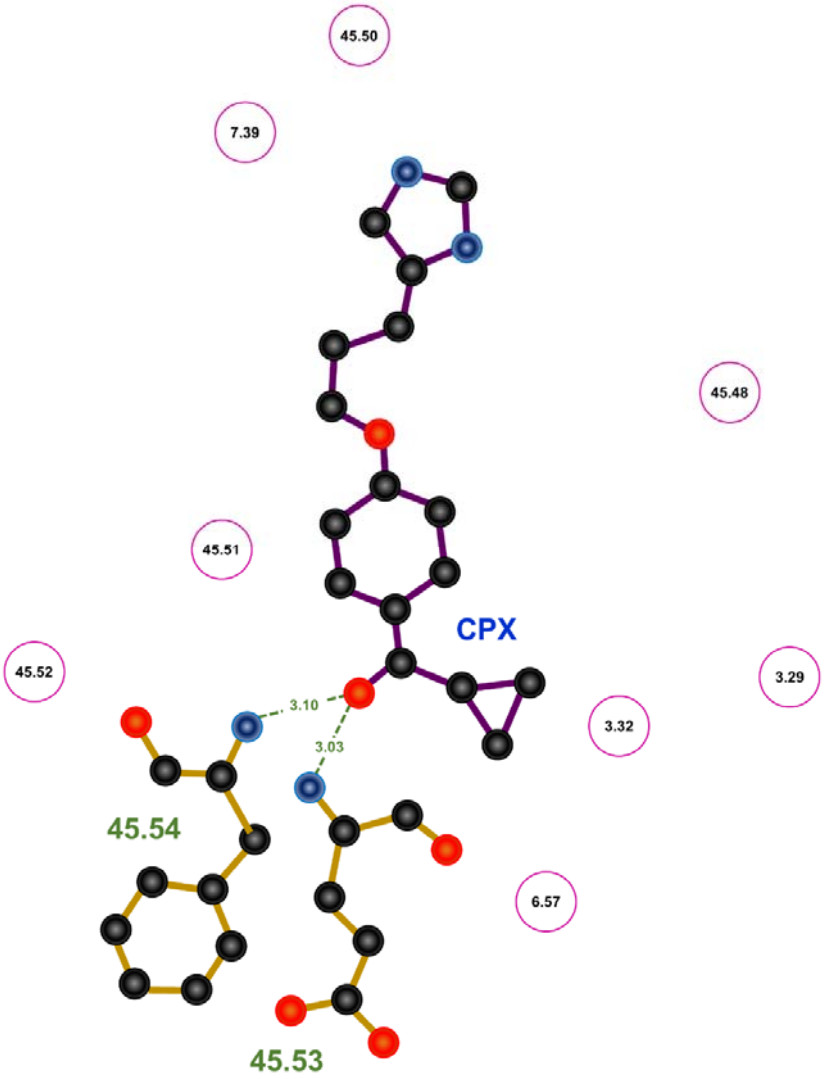
a


45.51
45.54
45.56
3.28
45.53
2.95
3.11
45.45
3.32
45.48
H2O
3.12
45.52
CPX
3.29
H2O
2.88
3.15
H2O
H2O
H2O
H2O
H2O

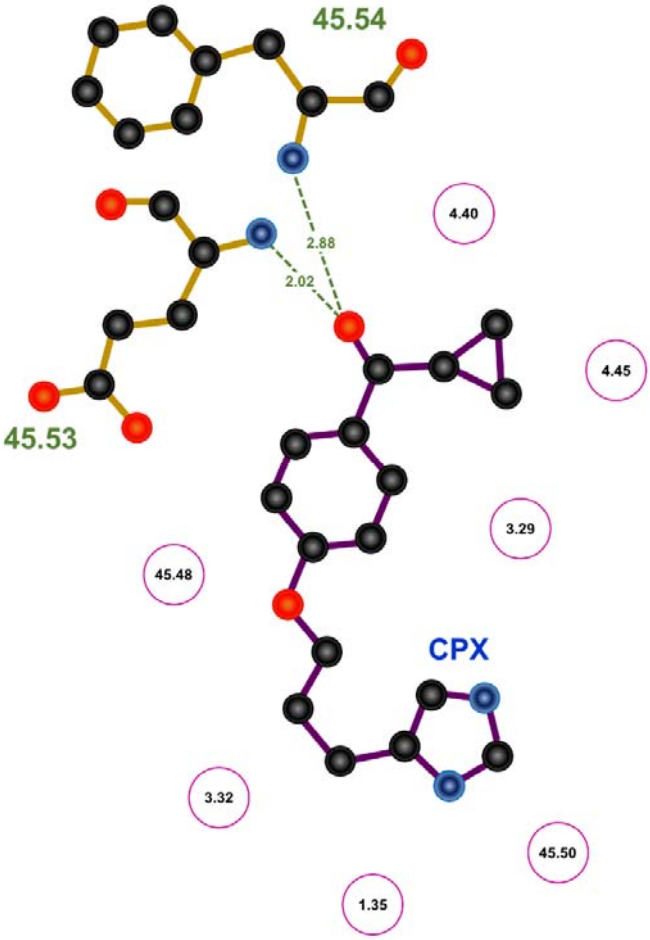
b


45.45
45.50
6.55
7.42
H_2O
4.40
H_2O
3.03
CPX
2.30
2.36
45.53
2.85
3.32
45.51
6.51
45.54
3.28
3.29
48.48

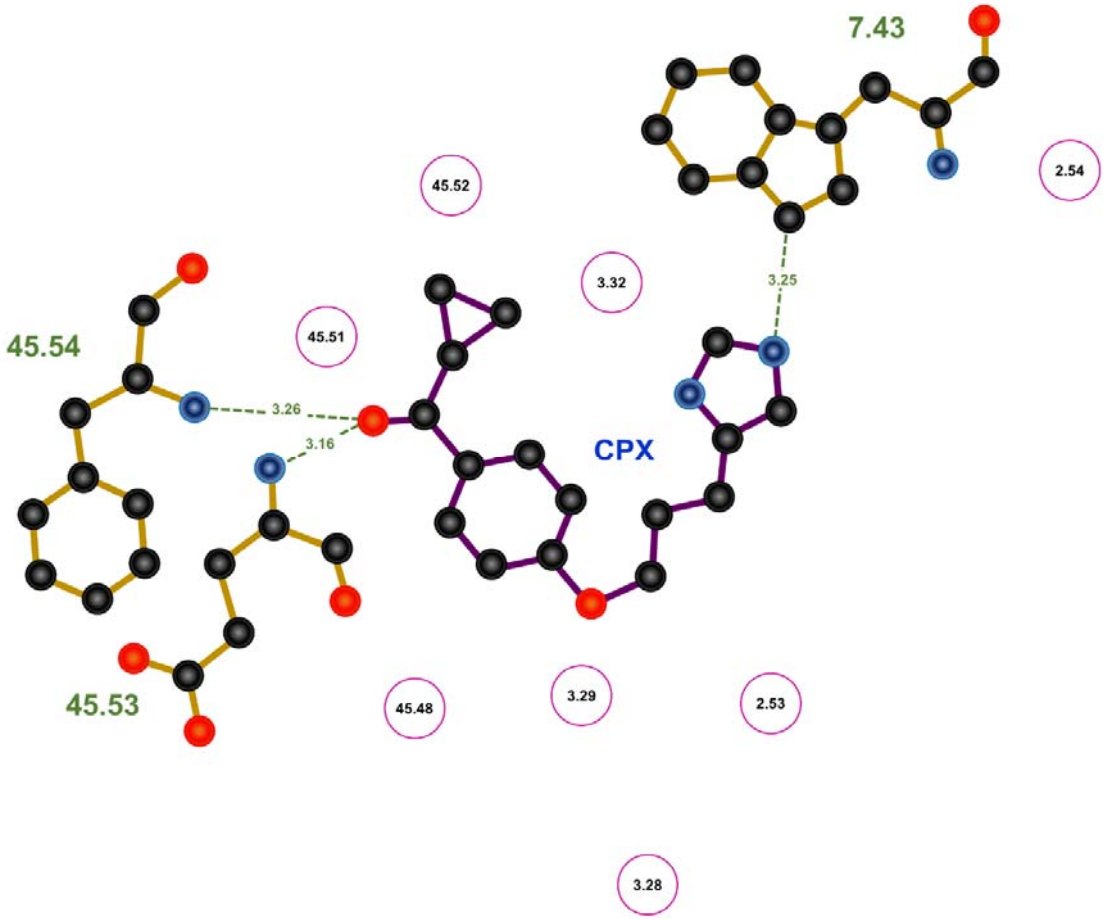
c



d



e


7.43
45.52
2.54
3.32
3.25
45.54
45.51
3.26
3.16
CPX
45.53
45.48
3.29
2.53
3.28

f

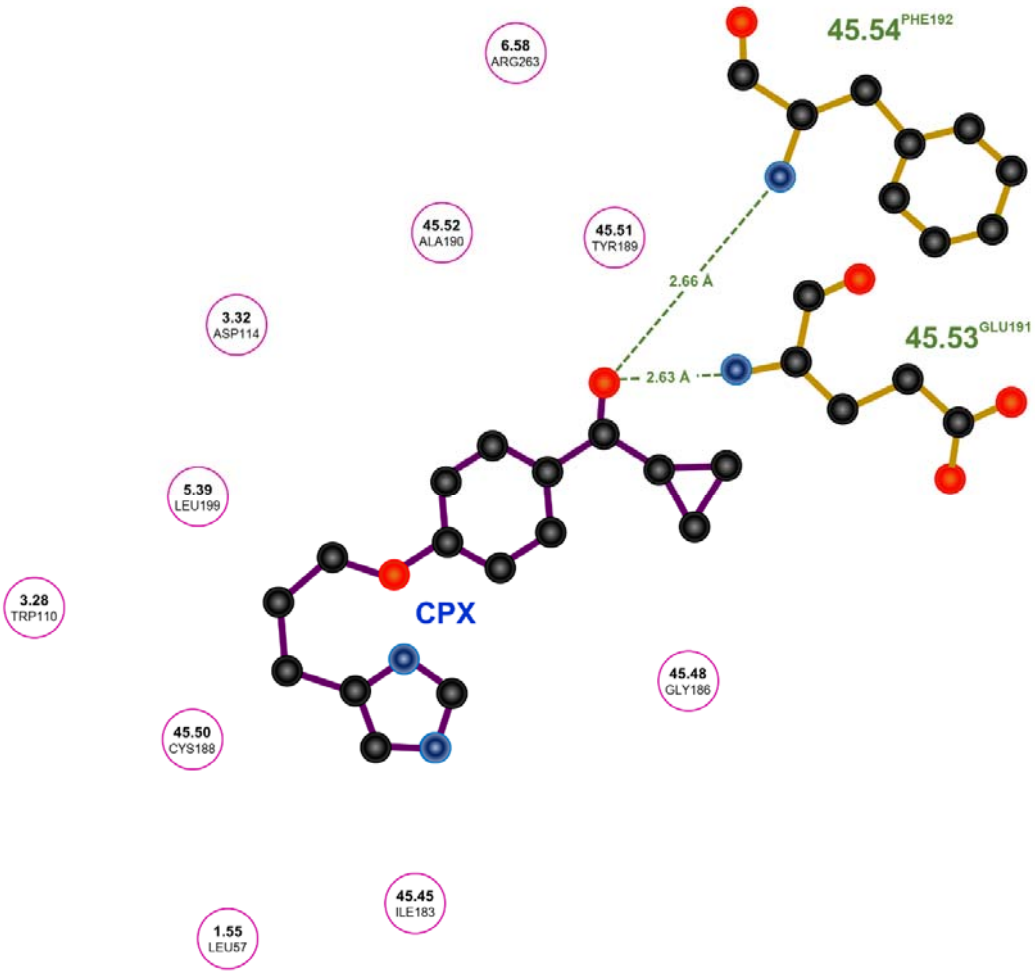

45.54 PHE192
6.58
ARG263
45.52
ALA190
45.51
TYR189
2.66 Å
3.32
ASP114
45.53 GLU191
2.63 Å
5.39
LEU199
3.28
TRP110
CPX
45.48
GLY186
45.50
CYS188
45.45
ILE183
1.55
LEU57

Fig. 8a-f

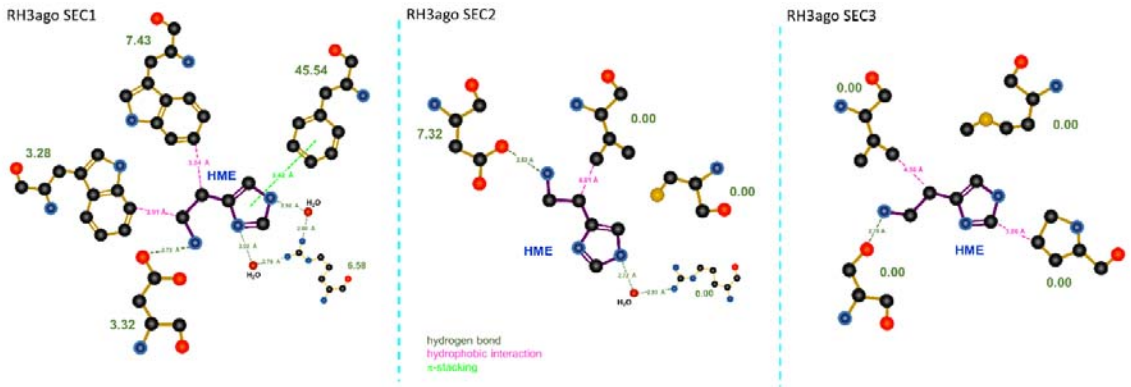

RH3ago SEC1
7.43
45.54
3.28
HME
3.32
6.58
RH3ago SEC2
7.32
0.00
0.00
HME
0.00
hydrogen bond
hydrophobic interaction
π-stacking
RH3ago SEC3
0.00
0.00
HME
0.00
0.00

Fig. 9

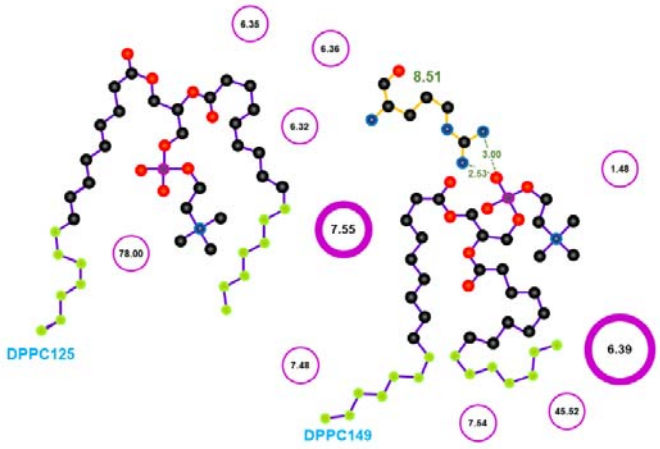

6.35
6.36
8.51
6.32
1.48
7.55
78.00
DPPC125
7.48
6.39
DPPC149
7.64
45.52

Fig. 10

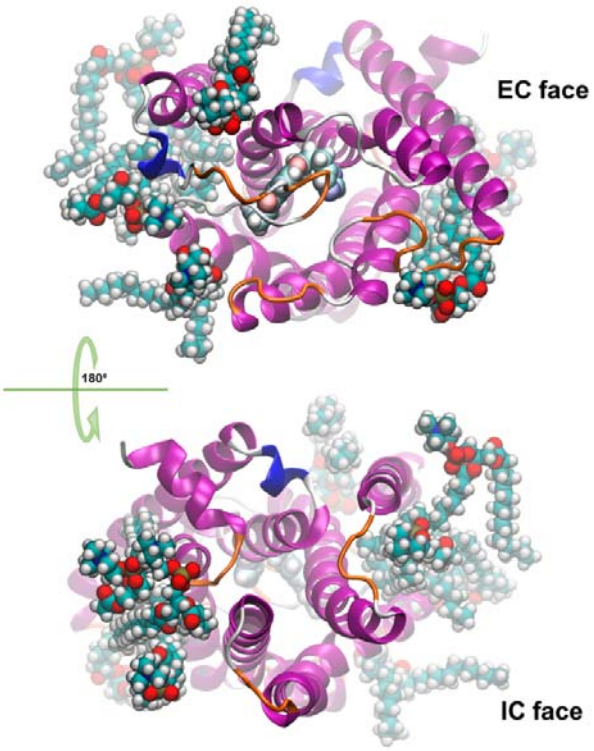
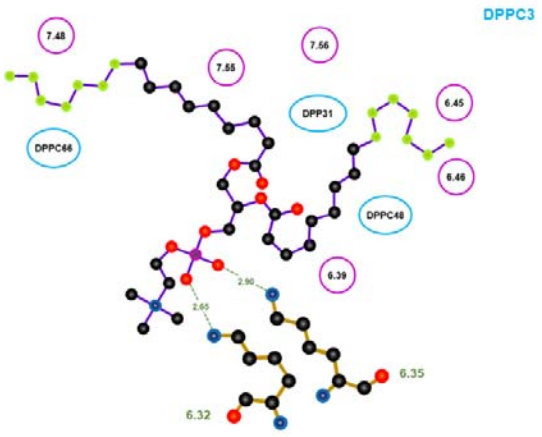


Fig. 11

a


DPPC3
7.48
7.55
7.56
DPP31
6.45
DPPC46
6.46
DPPC40
6.39
2.90
2.65
6.35
6.32

b

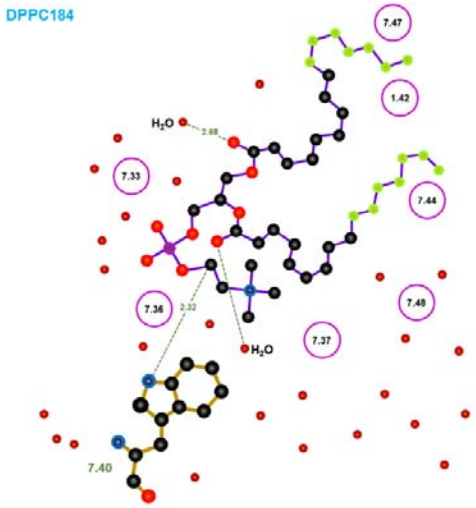

DPPC184
7.47
1.42
H_2O
2.68
7.33
7.44
7.36
2.32
7.48
7.37
H_2O
7.40

Fig. 12a-b

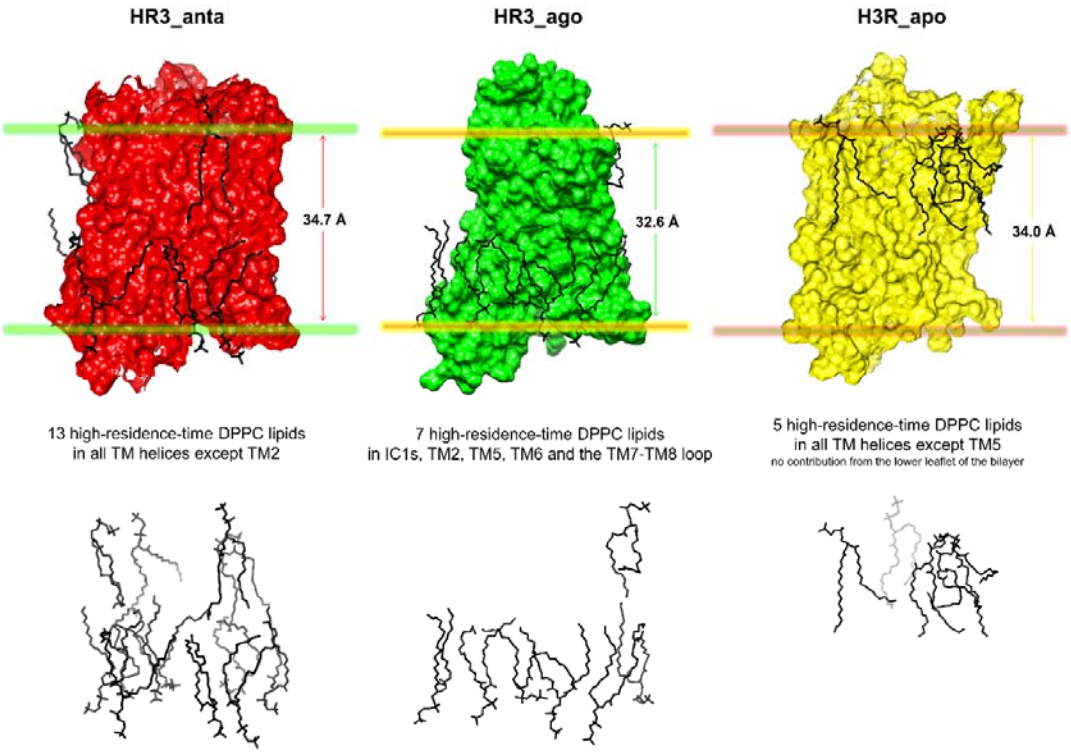

HR3_anta
HR3_ago
H3R_apo
34.7 Å
32.6 Å
34.0 Å
13 high-residence-time DPPC lipids
in all TM helices except TM2
7 high-residence-time DPPC lipids
in IC1s, TM2, TM5, TM6 and the TM7-TM8 loop
5 high-residence-time DPPC lipids
in all TM helices except TM5
no contribution from the lower leaflet of the bilayer

Fig. 13

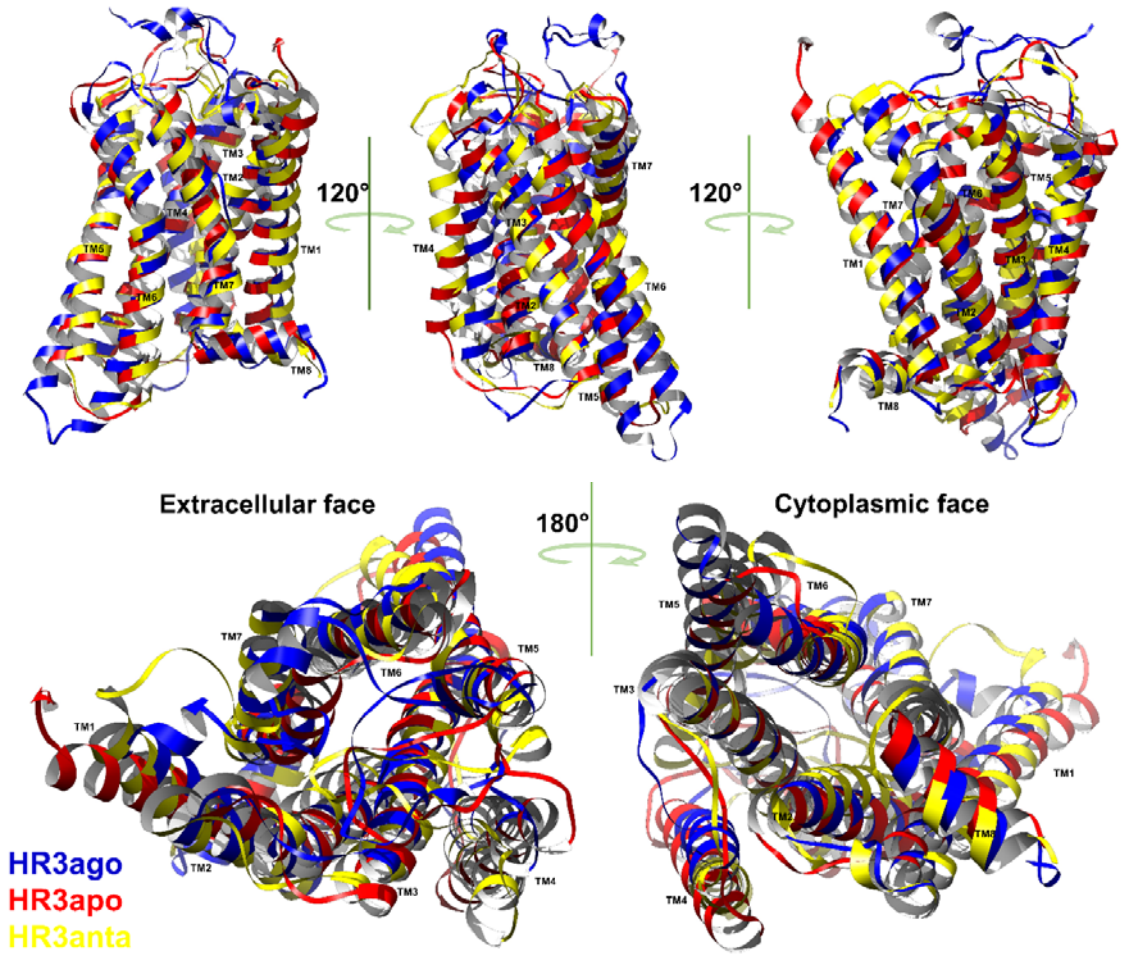

120°
120°
Extracellular face
180°
Cytoplasmic face
HR3ago
HR3apo
HR3anta

Fig. 14

TABLE 1

	no. waters
H3Rapo Bulk=	48,9516
H3Ranta Bulk=	61,3404
H3Rago Bulk=	
Bulk-sec1=	35,1545
Bulk-sec2=	19,8312
Bulk-sec3=	19,0037
Bulk-sec4=	24,9105

N-ter

H3Ranta	H3Rago	H3Rapo

TM1

H3Ranta		H3Rago		H3Rapo	
		MET41	1,39		
		LEU44	1,42		

ICL1

H3Ranta	H3Rago	H3Rapo

TM2

H3Ranta		H3Rago		H3Rapo	
		ASN70	2,4		
SER79	2,49			SER79	2,49
ASP80	2,5				
PHE81	2,51				
VAL83	2,53	VAL83	2,53	VAL83	2,53
		CYS87	2,57		
ILE88	2,58	ILE88	2,58	ILE88	2,58
		LEU90	2,6		
TYR91	2,61	TYR91	2,61	TYR91	2,61
TYR91	2,64	TYR91	2,61		

ECL1

H3Ranta	H3Rago	H3Rapo

TM3

H3Ranta		H3Rago		H3Rapo	
CYS107	3,25	CYS107	3,25		
TRP110	3,28				
LEU111	3,29	LEU111	3,29	LEU111	3,29
ASP114	3,32	ASP114	3,32	ASP114	3,32
TYR115	3,33	TYR115	3,33	TYR115	3,33
LEU117	3,35			LEU117	3,35
CYS118	3,36				
SER121	3,39			SER121	3,39
VAL122	3,4			VAL122	3,4

ICL2

H3Ranta	H3Rago	H3Rapo

TM4

H3Ranta	H3Rago	H3Rapo

ECL2

H3Ranta		H3Rago		H3Rapo	
		GLU175	45,37		
		TYR189	45,51		
		ALA190	45,52		
GLU191	45,53	GLU191	45,53		
PHE192	45,54	PHE192	45,54	PHE192	45,54
PHE193	45,55				
TYR194	45,56			TYR194	45,56

TM5

H3Ranta		H3Rago		H3Rapo	
		PHE198	5,38		
LEU199	5,39			LEU199	5,39
ALA202	5,42				
SER203	5,43				
GLU206	5,46				
PHE207	5,47			PHE207	5,47

ICL3

H3Ranta	H3Rago	H3Rapo

TM6

H3Ranta		H3Rago		H3Rapo	
		LEU249	6,37		
TRP253	6,48				
TYR256	6,51			TYR256	6,51
MET260	6,55			MET260	6,55
ARG263	6,58	TYR263	6,51	ARG263	6,58

ECL3

H3Ranta	H3Rago	H3Rapo

TM7

H3Ranta		H3Rago		H3Rapo	
TYR276	7,35			TYR276	7,35
GLU277	7,36				
SER279	7,38				
PHE280	7,39			PHE280	7,39
LEU283	2,42	TYR283	7,35	LEU283	2,42
TRP284	7,43	GLU284	7,36	TRP284	7,43
ASN286	7,45			ASN286	7,45
SER287	7,46	PHE287	7,39	SER287	7,46
ASN290	7,49			ASN290	7,49
		TRP291	7,43		
		SER291	7,49		

ICL4

H3Ranta	H3Rago	H3Rapo

H8

H3Ranta	H3Rago	H3Rapo

C-ter

H3Ranta	H3Rago	H3Rapo

TABLE 2

Anta		Ago		Apo	
MET41:TRP281	1.39:7.40	MET41:TRP288	1.39:7.40	MET41:TYR91	1.39:2.61
				MET41:TRP281	1.39:7.40
MET41:TRP284	1.39:7.43			MET41:TRP284	1.39:7.43
				ARG150:PRO291	1.50:7.50
ASN69:ASP131	2.39:3.49			MET56:PHE81	1.54:2.51
ASN69:ARG132	2.39:3.50			ASN69:ASP131	2.39:3.49
				ASN69:ARG132	2.39:3.50
				ASP80:PRO291	2.50:7.50
ARG132:ASP131	3.50:3.49			ASP114:TRP284	3.32:7.43
ARG132:ASP235	3.50:6.30				
PHE151:TRP253	5.51:6.48				
				PHE207:TRP253	5.47:6.48
MET260:TYR256	6.55:6.51	MET267:PHE208	6.55:5.48	TRP253:SER287	6.48:7.46
		MET267:TYR115	6.55:3.33	MET260:TYR256	6.55:6.51
MET260:PHE280	6.55:7.39	MET267:TYR263	6.55:6.51		
		TYR301:PHE308	7.53:8.50		
				TYR294:PHE301	7.53:8.50
				TYR294:PHE305	7.53:8.54
	8		6		13

TABLE 3

H3Ranta		
resid		esidence time
PHE192	45,54	100,00%
CYS188	45,5	100,00%
ALA190	45,52	100,00%
LEU111	3,29	100,00%
TYR189	45,51	99,00%
TRP110	3,28	99,00%
ASP114	3,32	99,00%
CYS87	2,57	97,00%
GLY186	45,48	96,00%
GLU191	45,53	96,00%
CYS107	3,25	93,00%
VAL83	2,53	83,00%
PHE280	7,39	77,00%

TABLE 4

HR3ago		
Sec1		
resid		residence time
ASP114	3,32	97,20%
TRP110	3,28	94,60%
TRP291	7,43	88,10%
PHE192	45,54	76,90%
TYR283	7,35	61,50%
TYR91	2,61	59,00%
TYR263	6,51	57,60%
ARG270	6,58	56,40%
ALA190	45,52	54,40%
LEU11	3,29	54,20%
TRP100	23,50	24,30%
TYR115	3,33	24,00%
PHE287	7,39	22,70%
GLU191	45,53	22,30%
CYS87	2,57	12,90%
LEU266	6,54	12,80%
CYS107	3,25	10,80%
VLA83	2,53	7,70%
LEU90	2,60	7,20%
LEU290	7,42	6,10%
TYR94	2,64	6,00%
TRP260	6,48	5,40%
SER294	7,46	5,10%
PHE86	2,56	5,00%
LEU199	5,39	4,80%
ARG27	0,00	4,70%
CYS118	3,36	4,50%
TYR189	45,51	4,10%
PHE29	0,00	3,00%
ALA25	0,00	2,30%
LYS108	3,26	2,10%
GLY28	0,00	2,00%
MET267	6,55	1,60%
ASP280	7,32	1,30%
PHE102	23,52	1,10%

TABLE 5

	HR3anta																																	
	name	sidence tim	name	sidence tim	name	sidence tim	name	sidence tim	name	sidence tim	name	sidence tim	name	sidence tim	name	sidence tim	name	sidence tim	name	sidence tim	name	sidence tim	name	sidence tim	name	sidence tim	name	sidence tim	name	sidence tim
DPPC	125	99,98%	110	99,98%	161	99,94%	79	99,78%	33	99,64%	105	98,13%	184	85,85%	159	94,83%	43	82,67%	31	82,79%	149	82,48%	131	81,48%	166	81,43%
ALANINE AND FOLATE	LEU 196	7,55	PHE 123	3,41	ILE 158	4,48	PHE 128	8,84	THR 197	5,37	ASN 185	45,57	ILE 124	6,35	THR 124	6,35	VAL 88	5,58	ILE 169	3,37	TRP 196	5,56	VAL 182	7,51	PHE 126	5,56	VAL 153	4,43	SER 174	5,56
	ILE 144	6,59	LEU 145	3,45	ARG 154	4,44	TRP 181	7,49	ASN 185	45,57	ARG 176	5,62	THR 122	5,62	PHE 183	3,31	MET 56	5,54	THR 167	4,57	ILE 184	5,49	PRO 181	7,54	THR 137	3,55				
	LYS 148	6,35	LEU 137	3,45	LYS 155	4,45	SER 174	7,33	PHE 188	5,38	ARG 131	6,26	ASP 131	6,26	ILE 182	5,32	ILE 57	5,55	ILE 165	3,23	THR 184	5,44	PRO 185	7,54	PHE 183	3,51				
	LYS 159	7,56	THR 138	3,48	PHE 133	3,41	TRP 174	7,37	PHE 186	5,36	TRP 186	5,36	ILE 152	6,45	THR 137	3,35			ALA 170	4,68	PHE 188	5,48	ARG 182	8,52	ILE 132	5,52				
	LYS 137	6,32	VAL 126	3,44	ARG 152	4,41	ILE 178	8,88	ILE 165	4,55	SER 147	6,42	SER 147	6,42	ASP 138	3,36			ILE 171	4,61	HIS 197	57,03	SER 124	1,45						
	HIS 198	78,84	ILE 158	3,52	ARG 152	4,48	SER 138	8,88	ILE 166	4,56	ILE 144	6,39	ILE 144	6,39	PHE 188	3,48			ARG 164	3,22	ALA 196	6,59	GLY 152	1,49						
	ILE 148	6,43	ARG 152	4,41	ALA 119	3,37	ALA 193	7,44	ILE 162	4,52	ILE 125	5,65	ILE 125	5,65	ILE 164	3,32			LYS 148	3,26	SER 193	5,63	ILE 154	1,52						
	SER 141	6,39	VAL 122	3,48	ILE 152	4,52	SER 182	7,41	ILE 182	5,41	SER 148	6,38	ALA 143	6,38	PHE 187	3,37			ILE 166	3,24	ILE 157	5,47	PHE 185	8,54						
	ILE 145	7,52	ALA 113	3,37	SER 138	3,38	SER 174	7,36	THR 188	4,59	GLN 126	5,66	LYS 159	6,31					SER 173	4,63	ILE 159	5,38	ILE 159	7,55						
	ARG 182	8,51	ARG 138	3,56	VAL 159	4,49	TRP 181	8,31					ASN 126	5,66					TRP 184	4,54			ALA 189	7,47						
	VAL 189	7,48	THR 141	3,41	ILE 116	3,34	ILE 148	1,38					ASN 118	5,58					PHE 184	4,54			THR 148	1,46						
	ILE 152	6,46	ILE 132	3,52			ARG 127	8,88											ALA 145	4,53			VAL 189	7,48						
							ALA 132	8,88											VAL 153	3,31			ILE 158	1,48						
							ASP 173	7,32											GLU 175	4,37			ILE 141	1,42						

TABLE 6

H3Rago

	sec1																			
	name		residence time	name		residence time	name		residence time	name		residence time	name		residence time	name		residence time	name	residence time
DPPC	2		100,00%	47		98,43%	183		94,05%	7		93,16%	65		88,81%	22		85,25%	59	85,15%
amino acid residue	ILE251	6,39	87,11%	TYR222	5,62	98,43%	VAL192	2,62	77,96%	VAL293	7,51	59,14%	LEU303	7,55	85,18%	ARG66	12,51	85,21%	TYR222	5,62	80,87%
	LYS247	6,35	81,72%	ALA250	6,38	98,18%	ILE188	2,58	75,50%	ARG309	8,51	53,02%	ARG309	8,51	76,62%	VAL160	1,58	85,21%	GLN226	5,66	65,76%
	ILE255	6,43	81,36%	ASN218	5,58	98,07%	PRO93	2,63	75,20%	THR313	8,55	52,78%	VAL299	7,51	59,36%	MET56	1,56	84,93%	LEU219	5,59	57,52%
	LEU303	7,55	77,37%	LYS247	6,35	95,25%	LEU96	2,66	72,26%				TYR306	0,00	58,67%	PHE59	1,57	84,81%			
	LEU258	6,46	70,21%	THR215	5,55	95,16%	PRO89	2,59	62,70%							LEU74	2,46	84,65%			
	ARG309	8,51	69,49%	LYS243	6,31	93,09%	THR97	2,67	60,71%							PHE81	2,51	81,07%			
	CYS304	7,56	68,37%	LEU219	5,59	93,06%										ILE78	2,48	80,02%			
	LEU300	7,52	67,95%	ALA246	6,34	91,71%										SER63	1,48	79,50%			
	SER254	6,42	65,25%	VAL253	6,41	90,88%										LEU57	1,55	69,68%			
	ALA250	6,38	62,41%	SER256	6,42	88,47%										PHE71	2,41	64,01%			
	SER248	6,36	61,75%	GLY257	6,45	86,60%										ALA61	1,59	61,09%			
	LYS244	6,32	58,71%	ILE225	5,65	84,01%										LEU82	2,52	59,72%			
	VAL296	7,48	58,45%	VAL214	5,54	83,60%										ASP62	1,60	52,81%			
	TYR306	0,00	49,86%	ILE221	5,61	81,21%															
				PHE211	5,51	79,03%															
				TRP260	6,48	78,62%															
				LEU258	6,46	57,06%															
				PHE256	6,44	51,17%															
				ALA261	6,49	50,16%															

TABLE 7

	H3Rapo														
	name		ecidence tim	name		ecidence tim	name		ecidence tim	name		ecidence tim	name		ecidence time
DPPC	184		100,00%	151		98,25%	133		96,42%	155		88,47%	128		82,44%
amino acid residue	TRP 281	7,40	100,00%	LEU109	3,27	98,74%	TRP33	1,31	73,40%	ARG27	0,00	79,35%	LEU258	6,53	80,12%
	GLU277	7,36	99,83%	LYS108	3,26	97,76%	LEU40	1,38	72,31%	PHE29	0,00	53,83%	ILE262	6,57	73,91%
	TYR274	7,33	99,27%	VAL112	3,3	95,42%	LEU37	1,35	59,57%				PRO255	6,50	65,46%
	ALA285	7,44	98,93%	TYR167	4,57	94,57%	TYR274	7,33	58,91%				TRP275	7,34	64,76%
	THR278	7,37	96,32%	VAL113	3,31	92,42%							CYS266	6,61	55,80%
	LEU44	1,42	95,96%	LEU116	3,34	92,05%							ALA254	6,49	53,68%
	LEU37	1,35	95,44%	ALA163	4,53	91,62%							LEU259	6,54	53,05%
	VAL95	2,65	95,14%	GLY105	3,23	89,95%									
	LEU40	1,38	90,32%	TRP160	4,5	82,08%									
	TYR91	2,61	90,18%	VAL159	4,49	82,04%									
	LEU282	7,41	88,36%	ARG104	3,22	78,15%									
	GLY98	2,68	86,11%	ALA170	4,6	66,44%									
	ALA288	7,47	74,47%	LEU82	2,52	66,34%									
	VAL289	7,48	73,43%	PHE164	4,54	61,74%									
	ASP273	7,32	70,32%	TYR176	45,38	61,73%									
	TRP33	1,31	65,56%	ILE171	4,61	61,59%									
	TYR94	2,64	63,54%	MET156	4,46	59,43%									
	LEU96	2,66	61,77%	LEU117	3,35	56,34%									
	LEU43	1,41	57,98%												
	TRP284	7,43	53,40%												
	THR97	2,67	52,68%												