

Title	Structure of the human histamine H1 receptor complex with doxepin.
Author(s)	Shimamura, Tatsuro; Shiroishi, Mitsunori; Weyand, Simone; Tsujimoto, Hirokazu; Winter, Graeme; Katritch, Vsevolod; Abagyan, Ruben; Cherezov, Vadim; Liu, Wei; Han, Gye Won; Kobayashi, Takuya; Stevens, Raymond C; Iwata, So
Citation	Nature (2011), 475(7354): 65-70
Issue Date	2011-07-07
URL	<a href="http://hdl.handle.net/2433/156845">http://hdl.handle.net/2433/156845</a>
Right	© 2011 Nature Publishing Group, a division of Macmillan Publishers Limited.
Type	Journal Article
Textversion	author

**Title: Structure of the human histamine H<sub>1</sub> receptor in complex with doxepin.**

**Authors**

Tatsuro Shimamura<sup>1,2,3\*</sup>, Mitsunori Shiroishi<sup>1,2,4\*</sup>, Simone Weyand<sup>1,5,6</sup>, Hirokazu Tsujimoto<sup>1,2</sup>, Graeme Winter<sup>6</sup>, Vsevolod Katritch<sup>7</sup>, Ruben Abagyan<sup>7</sup>, Vadim Cherezov<sup>3</sup>, Wei Liu<sup>3</sup>, Gye Won Han<sup>3</sup>, Takuya Kobayashi<sup>1,2‡</sup>, Raymond C. Stevens<sup>3‡</sup> and So Iwata<sup>1,2,5,6,8‡</sup>

1. Human Receptor Crystallography Project, ERATO, Japan Science and Technology Agency, Yoshidakonoe-cho, Sakyo-ku, Kyoto 606-8501, Japan.

2. Department of Cell Biology, Graduate School of Medicine, Kyoto University, Yoshidakonoe-cho, Sakyo-Ku, Kyoto 606-8501, Japan.

3. Department of Molecular Biology, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, CA 92037, USA.

4. Graduate School of Pharmaceutical Sciences, Kyushu University, 3-1-1 Maidashi, Higashi-ku, Fukuoka 812-8582, Japan.

5. Division of Molecular Biosciences, Membrane Protein Crystallography Group, Imperial College, London SW7 2AZ, UK.

6. Diamond Light Source, Harwell Science and Innovation Campus, Chilton, Didcot, Oxfordshire OX11 0DE, UK.

7. Skaggs School of Pharmacy and Pharmaceutical Sciences and San Diego Supercomputer Center, University of California, San Diego, La Jolla, CA 92093, USA.

8. Systems and Structural Biology Center, RIKEN, 1-7-22 Suehiro-cho Tsurumi-ku, Yokohama 230-0045

Japan.

**\*These authors contributed equally to this work.**

**‡ To whom correspondence should be addressed. E-mails: [stevens@scripps.edu](mailto:stevens@scripps.edu) (R.C.S.);**

**[so\\_iwata@me.com](mailto:so_iwata@me.com) (S.I.); [t-coba@mfour.med.kyoto-u.ac.jp](mailto:t-coba@mfour.med.kyoto-u.ac.jp) (T.K.)**

**Running title:** Crystal structure of human histamine H<sub>1</sub> receptor

## **Summary**

Histamine H<sub>1</sub> receptor (H<sub>1</sub>R) antagonists (H<sub>1</sub>R-antagonists) are very effective drugs inhibiting the H<sub>1</sub>R action and alleviating the symptoms of allergic reactions. The crystal structure of H<sub>1</sub>R-T4 lysozyme fusion-protein in complex with doxepin, a first-generation H<sub>1</sub>R-antagonist, allows us to characterize its ligand-binding pocket in detail. Doxepin sits much deeper in the pocket than the antagonists in other aminergic G protein coupled receptor (GPCR) structures and directly interacts with the highly conserved Trp428<sup>6,48</sup>, a key residue in GPCR activation. This well-conserved pocket and its mostly hydrophobic nature contribute to low selectivity of doxepin and other first-generation compounds causing considerable side-reactions. The pocket is associated with an anion-binding region occupied by a phosphate molecule. Docking of various second-generation H<sub>1</sub>R-antagonists reveals that the unique carboxyl-group present in this class of compounds interacts with Lys191<sup>5,39</sup> and/or Lys179<sup>ECL2</sup>, both of which form part of the anion-binding region and are not conserved in other aminergic receptors.

Histamine is a biogenic amine and an important mediator in various physiological and pathophysiological conditions such as gastric acid secretion, arousal state, locomotor activity, feeding, drinking, allergy and inflammation<sup>1,2,3</sup>. Histamine exerts its effects through the activation of four distinct histamine receptors (H<sub>1</sub>, H<sub>2</sub>, H<sub>3</sub> and H<sub>4</sub>) that belong to the G protein-coupled receptor (GPCR) superfamily. Histamine H<sub>1</sub> receptor (H<sub>1</sub>R), was originally cloned from bovine adrenal medulla by Yamashita *et al.* in 1992<sup>4</sup>, is now known to be expressed in various tissues including airway, retina, intestinal and vascular smooth muscle, brain, and lymphocytes<sup>2</sup>. In type I hypersensitivity allergic reactions, H<sub>1</sub>R is activated by histamine released from mast cells, which are stimulated by various antigens<sup>5</sup>. Many studies have been performed to develop H<sub>1</sub>R-antagonists, also known generally as antihistamines. Many of these compounds have inverse agonism activity and inhibit the action of histamine on H<sub>1</sub>R to alleviate the symptoms of the allergic reactions, making H<sub>1</sub>R one of the most validated drug targets judging from the number of drugs approved<sup>6</sup>. H<sub>1</sub>R displays constitutive activity, and H<sub>1</sub>R-antagonists generally act as inverse agonists for H<sub>1</sub>R<sup>7,8</sup>. Development of H<sub>1</sub>R-antagonists has progressed through two generations. First-generation drugs such as pyrilamine and doxepin (known under many brand-names such as Sinequan<sup>TM</sup> and Adapin<sup>TM</sup>, Supplementary Fig. 1) are effective H<sub>1</sub>R-antagonists. These compounds are, however, known to show considerable side effects such as sedation and appetite increment, dry mouth, hypertension, tachycardia and arrhythmia, because of penetration across the blood-brain barrier (BBB), and low receptor selectivity. These H<sub>1</sub>R-antagonists can bind not only to H<sub>1</sub>R but also to other aminergic GPCRs, monoamine transporters and cardiac ion channels.

Second-generation drugs such as cetirizine, terfenadine and olopatadine (Supplementary Figure 1) are less sedating and in general have fewer side effects. The improved pharmacology of the second-generation zwitterionic drugs can be attributed to a new carboxylic moiety, in combination with the protonated-amine, which significantly reduces brain permeability, although residual CNS effects are still reported<sup>9</sup>. The introduction of the carboxyl moiety also improves the H<sub>1</sub>R selectivity of these compounds, but certain second-generation H<sub>1</sub>R antagonists, such as terfenadine and astemizole, still show cardiotoxicity because of the interaction with cardiac potassium channels<sup>10,11</sup>. Therefore, structural details of the H<sub>1</sub>R agonist binding-pocket will be highly beneficial for guiding rational design of new H<sub>1</sub>R antagonists that do not penetrate the BBB while maintaining H<sub>1</sub>R selectivity.

A first-generation H<sub>1</sub>R-antagonist, doxepin, can cause many types of side effects due to its antagonistic effects on H<sub>2</sub>R<sup>12</sup>, serotonin 5-HT<sub>2</sub>,  $\alpha_1$ -adrenergic, and muscarinic acetylcholine receptors<sup>13</sup> in addition to the inhibition of the reuptake of serotonin and norepinephrine<sup>14</sup>. Although GPCR homology models have been successfully used for the design and discovery of novel GPCR ligands (for reviews, see refs 15 and 16), reliable receptor structures are essential to understand ligand selectivity at the molecular level. Recently determined GPCR structures have enabled structure-based approaches to modeling ligand interactions in the binding pocket<sup>17,18,19,20,21,22,23</sup> and are already yielding novel chemotypes predicted by virtual screening of large chemical libraries<sup>24,25</sup>. Here, we report the 3.1 Å resolution structure of the H<sub>1</sub>R-T4 lysozyme fusion protein (H<sub>1</sub>R-T4L) complex with doxepin. The crystal structure reveals the atomic details

of doxepin binding and its inverse agonistic activity. The H<sub>1</sub>R crystal structure and the models of second-generation H<sub>1</sub>R antagonists will be highly beneficial for guiding rational design of ligands that do not penetrate the BBB while maintaining H<sub>1</sub> selectivity. For example, one can use the H<sub>1</sub>R structure to design optimally bound carboxyl substitutes for other existing first-generation H<sub>1</sub>R antagonists. It can also guide further medicinal chemistry efforts for optimal replacement of the carboxyl group by other determinants of low BBB permeability.

### **Overall architecture of H<sub>1</sub>R**

H<sub>1</sub>R constructs were screened for expression and stability as green fluorescent protein (GFP) fusions in *Saccharomyces cerevisiae*<sup>26</sup>. The most stable construct was expressed in *Pichia pastoris*, purified in detergent solution and crystallized in lipidic cubic phase supplemented with cholesterol. In the H<sub>1</sub>R construct, we replaced most of the third cytoplasmic loop (ICL3) (Gln222-Gly404) with T4-lysozyme<sup>27</sup> and truncated the N-terminal 19 residues (Met1-Lys19). H<sub>1</sub>R-T4L showed similar binding affinities for H<sub>1</sub>R-antagonists and for histamine as the wild type H<sub>1</sub>R expressed in yeast cells (Supplementary Table 1) and in COS-7 cells<sup>28</sup>. The structure of H<sub>1</sub>R-T4L was determined in complex with the H<sub>1</sub>R-antagonist doxepin at 3.1 Å resolution (Supplementary Table 2).

The main fold of the H<sub>1</sub>R consists of the canonical 7 transmembrane bundle of  $\alpha$ -helices (Fig. 1a), which is structurally most similar to the aminergic receptors:  $\beta_2$ -adrenergic ( $\beta_2$ AR)<sup>18</sup>,  $\beta_1$ -adrenergic

( $\beta_1$ AR)<sup>19</sup> and dopamine D3 (D3R)<sup>23</sup> receptors, while having larger deviations from the more phylogenetically distant rhodopsin<sup>17,21</sup>, A<sub>2A</sub> adenosine receptor (A<sub>2A</sub>AR)<sup>20</sup> and CXCR4<sup>22</sup> (Supplementary Table 3). H<sub>1</sub>R also shares the common motifs with other GPCRs including D(E)RY in helix III, CWxP in helix VI and NPxxY in helix VII, as well as a disulfide bond connecting extracellular loop 2 (ECL2) with the extracellular end of helix III (Cys100<sup>3,25</sup> to Cys180). Like most other GPCR structures, H<sub>1</sub>R also has a short amphiphilic  $\alpha$ -helix (helix VIII) following helix VII, although H<sub>1</sub>R lacks the palmitoylation site at the end of helix VIII found in many other GPCRs<sup>29</sup>.

Previous GPCR structures revealed that not only the residues in the transmembrane segments but also those in the loops are critical for ligand specificity<sup>17,18,19,20,21,22,23</sup>. ECL2 connecting helices IV and V is attached to helix III through a disulfide bond between Cys180 in ECL2 and Cys100<sup>3,25</sup> in helix III. Seven residues (Phe168-Val174) before the disulfide are not included in the structure as they did not have interpretable densities. A section of ECL2, between the disulfide bridge and the extracellular end of helix V, is particularly important because it is located at the entrance to the ligand-binding pocket. This section of ECL2 contains 7 amino acids in H<sub>1</sub>R, as compared to 5 in  $\beta_2$ AR, 4 in D3R, and 8 in A<sub>2A</sub>AR. The extra length of this ECL2 section is apparently accommodated by the increased distance between the extracellular ends of helices III and V by  $\sim 1.5$  Å and  $\sim 3.1$  Å when compared to  $\beta_2$ AR and D3R, respectively (Fig. 1b,c). This creates more space within the ligand-binding pocket, which can now accommodate the larger second-generation H<sub>1</sub>R-antagonists as discussed below.



Some unique features are also observed in the transmembrane segments. A conserved Pro161<sup>4,59</sup> induced kink in helix IV forms a tight i-3 helical turn instead of i-4 as in  $\beta_2$ AR and D3R (Fig. 2a). This tighter turn allows accommodation of a bulky Trp side chain at position 4.56, which seems essential for ligand specificity of aminergic GPCRs because this position is occupied by Ser in  $\beta_2$ AR and D3R, and the mutations of this Trp in the guinea pig H<sub>1</sub>R to Ala, Met and Phe reduce the affinity against the antagonist pyrilamine<sup>30</sup>.

The "ionic lock", a salt bridge between Arg<sup>3,50</sup> in the conserved D(E)R<sup>3,50</sup>Y motif and Asp/Glu<sup>6,30</sup>, which is suggested to stabilize the inactive conformation, was observed in rhodopsin structures<sup>17,21</sup> and dopamine D3R<sup>23</sup>, but broken in all the other GPCRs<sup>18,19,20,22</sup>. In H<sub>1</sub>R, Arg125<sup>3,50</sup> of the D(E)R<sup>3,50</sup>Y motif does not form a salt bridge either with Glu410<sup>6,30</sup> or with Asp124<sup>3,49</sup>. Instead, the side chain of Arg125<sup>3,50</sup> adopts in a new conformer relative to previous structures forming a hydrogen bond to Gln416<sup>6,36</sup> in helix VI (Fig. 2b). Different structures of the "ionic lock" regions of the receptors could be caused by modifications of ICL3. Otherwise it might be related to the different levels of constitutive activities of the receptors.

### **Doxepin isomers and conformers**

The H<sub>1</sub>R structure was determined with an H<sub>1</sub>R-antagonist doxepin (Supplementary Fig. 1) that shows very high affinity for H<sub>1</sub>R ( $K_i = 2.3$ - $2.4$  nM, Supplementary Table 1) as well as for several other aminergic

receptors<sup>12,13</sup>. The doxepin used in this study contains a mixture of E- and Z- isomers, and each isomer can take two distinct rotational conformers of the dibenzo[b,e]oxepin ring, resulting in 4 distinct conformers (conformers 1 to 4, Supplementary Fig. 2). Two conformers, one E-isomer (conformer 1) and one Z-isomer (conformer 4) fit the electron density better than the other two (Supplementary Fig. 3). This result is also consistent with the  $R_{\text{free}}$  and the averaged B-factor values for each conformer (Supplementary Table 5). We used the 1:1 mixtures of the E- and the Z- isomers in the refinement. The two conformers are indistinguishable at this resolution and have nearly identical interactions with the binding pocket, so in the following sections we will refer to the E-isomer unless noted otherwise.

### **Ligand binding pocket**

Doxepin binds in a pocket mainly defined by the side chains of helices III, V and VI (Fig. 3a,b). Asp107<sup>3,32</sup>, a strictly conserved residue in aminergic receptors (Supplementary Table 4), and forms an anchor salt bridge with the amine moiety of the ligand. This interaction has been reported to be essential for the binding of H<sub>1</sub>R-antagonists as well as agonists by the mutational studies<sup>31,32,33</sup>. This amine moiety is connected via a flexible carbon chain to the tricyclic dibenzo[b,e]oxepin ring in a hydrophobic pocket comprised of the side chains of helices III, V and VI. The tricyclic ring of doxepin sits much deeper (by ~5 Å) in the binding pocket than the ligands in the other non-rhodopsin GPCR structures (Fig. 3c). The ligand is surrounded mainly by highly conserved residues among aminergic receptors including Ile115<sup>3,40</sup>, Trp428<sup>6,48</sup>,

Phe424<sup>6.44</sup> and Phe432<sup>6.52</sup>, whereas the non-conserved residues Trp158<sup>4.56</sup> and Asn198<sup>5.46</sup> in the pocket make only minor hydrophobic interactions with doxepin (Fig. 3a,b). The importance of a large side chain at position 6.52 has been suggested for the binding of pyrilamine.<sup>30,33</sup> Thr112<sup>3.37</sup> can form a hydrogen bond to the oxygen atom of the E-isomer (but not the Z- isomer) of doxepin as shown in Fig. 3a,b. A suboptimal geometry and bifurcated nature of this H-bond suggest that it does not contribute significantly to binding affinity. This observation is also consistent with binding of the isomer-specific doxepin analogue olopatadine, as described below. This well-conserved pocket and its mostly hydrophobic nature should contribute to low selectivity of doxepin and other first-generation H<sub>1</sub>R-antagonists<sup>13,32</sup>. Moreover, because of its deep binding position, doxepin does not interact with ECL2, whose highly variable primary and tertiary structures are known to contribute to binding specificity of GPCR ligands<sup>34</sup>.

A novel feature of the H<sub>1</sub>R-doxepin complex is the existence of an anion-binding site at the entrance to the ligand-binding pocket (Fig. 3d). A phosphate ion, which is present at a high concentration in the crystallization buffer (300mM ammonium phosphate), is modeled into the observed strong density in the site. This model is supported by the fact that PO<sub>4</sub><sup>3-</sup> affects the binding of some ligands and the stability of H<sub>1</sub>R (Supplementary Tables 1 and 6). The phosphate ion is coordinated by Lys179 in ECL2, Lys191<sup>5.39</sup>, Tyr431<sup>6.51</sup> and His450<sup>7.35</sup>; all of which, except for Tyr431<sup>6.51</sup> are unique to H<sub>1</sub>R (Supplementary Table 4). This encasement of the ligand in the pocket combined with an ionic interaction between the phosphate ion and the tertiary amine of doxepin (N-O distance 4.8 Å) suggest that PO<sub>4</sub><sup>3-</sup> may serve as a positive modulator

of ligand-binding. This hypothesis has been validated by comparing thermostability (Supplementary Table 4) and ligand affinity (Supplementary Table 1) in buffers with and without  $\text{PO}_4^{3-}$ . Thermostability of the receptor is increased in the presence of phosphate for all ligands except for cetirizine, which likely prevents the phosphate-binding according to the modeling study as discussed below. The phosphate effect is observed at a phosphate concentration as low as 1.5 mM suggesting its physiological relevance. The affinity of histamine and pyrilamine to the receptors also increased in the presence of phosphate.

### **H<sub>1</sub> selectivity of the second-generation H<sub>1</sub>R-antagonists**

Supplementary Figure 1 lists the first- and second-generations of H<sub>1</sub>R-antagonists. It has been shown that the second-generation H<sub>1</sub>R-antagonists are much more specific to H<sub>1</sub>R and show much lower affinity to the other aminergic receptors<sup>32,35</sup>. H<sub>1</sub>R-antagonist specificity has been previously analyzed using H<sub>1</sub>R homology models based on the bacteriorhodopsin structures in combination with the H<sub>1</sub>R antagonist pharmacophore model and mutational studies (for example, refs 30, 36 and 37). These studies have successfully determined some residues important for the selectivity including Lys191<sup>5,39</sup>, however, contributions of the ECL residues have not been examined because these loops could not be modeled accurately based on the bacteriorhodopsin structure. Our H<sub>1</sub>R X-ray structure with the ECL loops should significantly improve the understanding of the H<sub>1</sub>R-antagonist selectivity. Using flexible ligand-receptor docking<sup>38,39</sup> in the ICM molecular modeling package<sup>40</sup> (see also Methods), we have studied the H<sub>1</sub>R

selectivity for representative second-generation zwitterionic H<sub>1</sub>R-antagonists: olopatadine, acrivastine, R-cetirizine (levocetirizine, Xyzal<sup>TM</sup>) and fexofenadine (Allegra<sup>TM</sup>) (Fig. 4). Olopatadine (Fig. 4a) is a close doxepin analogue with a methyl-carboxyl substitution in one of its benzene rings. Its binding mode closely resembles doxepin, while the carboxyl group extends out of the pocket toward the extracellular space and interacts with Lys191<sup>5,39</sup> and Tyr108<sup>3,33</sup> without displacing the phosphate ion. These additional interactions can explain a reduced effect of the mutation of the conserved Asp107<sup>3,32</sup> to Ala on olopatadine binding (14 fold for olopatadine as compared to 280 fold for doxepin)<sup>32,41</sup>. The orientation of the carboxyl moiety in the ECL region dictates that the oxygen atom of the oxepin ring is in a position where it cannot form a H-bond with Thr112<sup>3,37</sup>. Although the marketed drug is only the Z-isomer, both olopatadine Z- and E-isomers show similar H<sub>1</sub>R affinities<sup>41</sup>, which is also in line with our docking results for the E-isomer (not shown).

Acrivastin (Fig. 4b) has a different chemical scaffold with a carboxyl group in its pyridine ring. Its longer carbon chain positions the carboxyl group higher in the ECL region, where it can form salt bridges to both Lys191<sup>5,39</sup> and Lys179 (ECL2) amine moieties. R-cetirizine (Fig. 4c) has its carboxylic moiety attached directly to a piperazine amino group. The conformational modeling suggests that the carboxyl moiety can reach towards the ECL region forming salt bridges to Lys191<sup>5,39</sup> and to Lys179. Finally, fexofenadine (R-isomer, Fig. 4d) has the most extended carboxyl-containing substituent, which reaches outside of the binding cavity and forms a salt bridge to Lys191<sup>5,39</sup>.

Modeling of the second-generation H<sub>1</sub>R-antagonist binding to H<sub>1</sub>R suggests that no significant protein backbone rearrangements are required to accommodate these diverse ligands. Instead, the enhanced H<sub>1</sub>R selectivity of these compounds<sup>32,35</sup> can be explained by the specific interaction of the carboxyl group with Lys residues in the ECL region, unique to H<sub>1</sub>R. The result also shows a good agreement with earlier modeling and site-directed mutagenesis studies. Lys191<sup>5,39</sup> is known to be important for increasing affinity for some of these ligands<sup>30,42,43</sup>, whereas the involvement of Lys179 was suggested in the modeling study of 8R-lisuride into the ligand-binding pocket<sup>44</sup>. Our modeling results also suggest that olopatadine is the only second-generation compound studied here for which the carboxyl moiety does not interfere with phosphate binding. The results are also supported by the fact that the presence of the phosphate ion increased the thermal stability of the H<sub>1</sub>R-doxepin or H<sub>1</sub>R-olopatadine complex, whereas it does not affect the stability of the H<sub>1</sub>R-cetirizine complex (Supplementary Table 6).

### **Stabilizing inactive H<sub>1</sub>R conformation by H<sub>1</sub>R-antagonists**

H<sub>1</sub>R-antagonists act as highly effective inverse agonists of H<sub>1</sub>R, which reduce basal activity of the receptor and therefore are expected to interfere with the key molecular switches involved in the GPCR activation mechanism. One of the switches is represented by Trp<sup>6,48</sup> of the conserved CWxP<sup>6,50</sup> motif, which helps to stabilize rhodopsin in its inactive dark state through a direct interaction with retinal. The recently published structure of the active-state A<sub>2A</sub>AR<sup>45</sup> also showed that Trp<sup>6,48</sup> participates in the activation-related

conformational changes, where a small ligand induced shift of Trp<sup>6.48</sup> was observed in concert with the large movement of the intracellular part of helix VI. In other receptors, the role of Trp<sup>6.48</sup> is less obvious, *e.g.* it lacks direct ligand interactions with either inverse agonists or full agonists of  $\beta_2$ AR<sup>46</sup>. It is interesting to note that in the H<sub>1</sub>R structure, like in inactive rhodopsin, the H<sub>1</sub>R-antagonist doxepin does make extensive hydrophobic interactions with the Trp428<sup>6.48</sup> rings, which is unique among the known non-rhodopsin GPCR structures and could stabilize the hydrophobic packing around helix VI (Fig. 3c). Another important ligand-induced switch described in  $\beta_2$ AR is activation-related contraction of the extracellular ligand-binding pocket<sup>36</sup>. Because the natural agonist histamine is much smaller than bulky H<sub>1</sub>R-antagonists, some contraction of the binding pocket is likely to accompany ligand induced H<sub>1</sub>R activation. Bulky compounds, capable of blocking both activation-related contraction of the pocket and the Trp428<sup>6.48</sup> switch would be very efficient in locking H<sub>1</sub>R in an inactive conformation, which is likely to explain as much as 78% reduction of H<sub>1</sub>R basal activity by some H<sub>1</sub>R-antagonists<sup>8</sup>.

## Methods summary

H<sub>1</sub>R-T4L was expressed in yeast *Pichia pastoris*. Ligand binding assays were performed as described in Methods. *Pichia pastoris* membranes were solubilized using 1% (w/v) n-dodecyl- $\beta$ -D-maltopyranoside and 0.2% (w/v) cholesteryl hemisuccinate, and purified by immobilized metal ion affinity chromatography (IMAC). After IMAC, the C-terminal GFP was cleaved by Tobacco Etch virus (TEV) protease. Then the sample mixture was passed through IMAC to remove the cleaved His-tagged GFP and TEV protease. Receptor crystallization was performed by lipidic cubic phase (LCP) method. The protein-LCP mixture contained 40% (w/w) receptor solution, 54% (w/w) monoolein, and 6% (w/w) cholesterol. Crystals were grown in 40-50 nl protein-laden LCP boluses overlaid by 0.8  $\mu$ l of precipitant solution (26-30% (v/v) PEG400, 300 mM ammonium phosphate, 10 mM MgCl<sub>2</sub>, 100 mM Na-citrate pH 4.5 and 1 mM doxepin) at 20 °C. Crystals were harvested directly from LCP matrix and flash frozen in liquid nitrogen. X-ray diffraction data were collected at 100 K with a beam size of 10 x 10 microns on the microfocus beamline I24 at the Diamond Light Source (UK). Data collection, processing, structure solution and refinement are described in Methods.



## Figure legends

**Fig. 1. Structure of H<sub>1</sub>R complex with doxepin.** (a) Ribbon representation of the H<sub>1</sub>R structure.

Doxepin is shown as yellow spheres whereas the phosphate ion as spheres with carbon, and oxygen atoms colored orange and red, respectively. Disulfide bonds are shown as yellow sticks and Trp428 and Asp107 as pink sticks. Three conserved motifs D(E)R<sup>3.50</sup>Y, CWxP<sup>6.50</sup> and NP<sup>7.50</sup>xxY are highlighted in blue. (b) Superimposition of the H<sub>1</sub>R (green) and β<sub>2</sub>AR (cyan) structures. (c) Same as (b) but with the D3R structure colored magenta.

**Fig. 2. Comparison of the structures of H<sub>1</sub>R, β<sub>2</sub>AR and D3R.** (a) Variations in backbone structures of the helix IV proline-induced kink (H<sub>1</sub>R: green, β<sub>2</sub>AR: cyan, D3R: magenta). The side chain of Trp158<sup>4.56</sup> and Pro161<sup>4.59</sup> of H<sub>1</sub>R and the equivalent residues of β<sub>2</sub>AR (Ser165<sup>4.57</sup> and Pro168<sup>4.60</sup>) and of D3R (Ser165<sup>4.57</sup> and Pro167<sup>4.59</sup>) are also shown. (b) Variations in the D(E)RY motif structures of H<sub>1</sub>R, β<sub>2</sub>AR and D3R colored in green, cyan and magenta, respectively. Side chains of Asp124<sup>3.49</sup>, Arg125<sup>3.50</sup>, Glu410<sup>6.30</sup> of H<sub>1</sub>R and the equivalents of β<sub>2</sub>AR and D3R are represented as stick models. For H<sub>1</sub>R, Gln416<sup>6.36</sup>, which forms a hydrogen bond with Arg125<sup>3.50</sup>, are also shown. Possible hydrogen bonds are indicated by dotted lines.

**Fig. 3. Binding interactions of doxepin.** (a) Doxepin in the ligand binding pocket is shown as sticks with

cyan carbon atoms, whereas the contact residues within 4 Å of doxepin are shown as sticks with grey carbon atoms. Nitrogen and oxygen atoms are colored blue and red, respectively. An E-doxepin isomer is used for all the figures. Possible hydrogen bonds/salt bridges are indicated by dashed lines. (b) Schematic representation of doxepin binding interactions, shown as blue dotted lines for the salt bridges/hydrogen bonds and black for hydrophobic interactions. (c) Ligand binding positions in the structures of non-rhodopsin GPCRs. Carbon atoms of doxepin (H<sub>1</sub>R) are shown in grey, carazolol (β<sub>2</sub>AR) in cyan, eticlopride (D3R) in magenta and ZM241385 (A<sub>2A</sub>AR) in grey. Doxepin occupies the deepest position among them and directly interacts with Trp428<sup>6,48</sup> in the conserved CWXP motif. (d) Structure of the anion-binding region at the entrance of the ligand-binding pocket. A phosphate ion was modeled into the observed electron density in the region. Nitrogen and oxygen atoms are colored blue and red, respectively. Possible hydrogen bonds/salt bridges are indicated as blue dotted lines.

**Fig.4 Interactions of second-generation selective H<sub>1</sub>R-antagonists with the H<sub>1</sub>R ligand-binding pocket.**

Conformation of each complex was predicted by global optimization of the ligand in the all-atom flexible H<sub>1</sub>R model<sup>38,39,40</sup> based on the structure of H<sub>1</sub>R-doxepin complex. Carbon atoms for (a) Z-olopatadine co-bound with PO<sub>4</sub><sup>3-</sup> ion, (b) acrivastine, (c) R-cetirizine (levocetirizine), and (d) fexofenadine are colored magenta. Nitrogen and oxygen atoms are colored blue and red, respectively. Ligand contact residues of H<sub>1</sub>R are shown with grey carbon atoms; parts of helices III, IV and ECL2 are not displayed for clarity. Hydrogen

bonds between ligands and contact side chains are shown in cyan. These figures were prepared with ICM molecular modeling package (Molsoft LLC).

## **Acknowledgements**

This work was supported by the ERATO Human Receptor Crystallography Project from the Japan Science and Technology Agency and by the Targeted Proteins Research Program of MEXT (S.I.), Japan; NIH Common Fund grant P50 GM073197 for technology development (R.C.S.) and NIH PSI:BiologY grant U54 GM094618 (R.C.S, V.C., V.K. and R.A.); R.A. was also partly funded by NIH R01 GM071872.

The work was also partly funded by the Biotechnology and Biological Sciences Research Council (BBSRC) BB/G023425/1 (S.I.), the Mochida Memorial Foundation for Medical and Pharmaceutical Research (to T. S. and T.K.), Takeda Scientific Foundation (M.S.) and the Sumitomo Foundation (T. K.). A part of the work was performed in the Membrane Protein Laboratory funded by the Wellcome Trust (grant 062164/ Z/00/Z) at the Diamond Light Source Limited and at The Scripps Research Institute. We thank D. Axford for help with data collection at the Diamond Light Source Limited, H. Wu for help with the preparation of Supplementary Figure 1 and Q. Xu for help on validation on data processing and A. Walker for assistance with manuscript preparation. The authors acknowledge Y. Zheng (The Ohio State University) and M. Caffrey, Trinity College (Dublin, Ireland), for the generous loan of the in meso robot (built with support from the National Institutes of Health [GM075915], the National Science Foundation [IIS0308078], and Science Foundation Ireland [02-IN1-B266]).

## **Author contributions**

T.S. purified and crystallized the receptor in LCP, optimized crystallization conditions, grew crystals for data collection, solved and refined the structure, and prepared the manuscript. M.S. designed, characterized and screened the constructs, purified the receptor, and prepared the manuscript. S.W. and S.I. collected the data and processed diffraction data with G.W.H. H.T. expressed the receptor, prepared the membrane, and performed the ligand-binding assay. V.K. and R.A. performed flexible ligand-receptor docking, and prepared the manuscript. V.C. assisted the crystallization in LCP and prepared the manuscript. W.L. performed the thermal stability assay and assisted the crystallization in LCP. G.W.H refined the structure and assisted preparing the manuscript. T.K. designed the receptor production strategy and assisted preparing the manuscript. R.C.S. and S.I. were responsible for the overall project strategy and management and wrote the manuscript.

**Author information.** The coordinates and the structure factors will be available from the Protein Data Bank upon the publication.

## References

- 1 Schwartz, J. C., Arrang, J. M., Garbarg, M., Pollard, H. & Ruat, M. Histaminergic transmission in the mammalian brain. *Physiol Rev* **71**, 1-51 (1991).
- 2 Hill, S. J. Distribution, properties, and functional characteristics of three classes of histamine receptor. *Pharmacol Rev* **42**, 45-83 (1990).
- 3 Hill, S. J. *et al.* International Union of Pharmacology. XIII. Classification of histamine receptors. *Pharmacol Rev* **49**, 253-278 (1997).
- 4 Yamashita, M. *et al.* Expression cloning of a cDNA encoding the bovine histamine H1 receptor. *Proc Natl Acad Sci USA* **88**, 11515-11519 (1991).
- 5 Simons, F. E. Advances in H1-antihistamines. *N Engl J Med* **351**, 2203-2217 (2004).
- 6 Overington, J. P., Al-Lazikani, B. & Hopkins, A. L. How many drug targets are there? *Nat Rev Drug Discov* **5**, 993-996 (2006).
- 7 Bakker, R. A., Wieland, K., Timmerman, H. & Leurs, R. Constitutive activity of the histamine H(1) receptor reveals inverse agonism of histamine H(1) receptor antagonists. *Eur J Pharmacol* **387**, R5-7 (2000).
- 8 Bakker, R. A., Schoonus, S. B., Smit, M. J., Timmerman, H. & Leurs, R. Histamine H(1)-receptor activation of nuclear factor-kappa B: roles for G beta gamma- and G alpha(q/11)-subunits in constitutive and agonist-mediated signaling. *Mol Pharmacol* **60**, 1133-1142 (2001).
- 9 Tashiro, M. *et al.* Dose dependency of brain histamine H(1) receptor occupancy following oral administration of cetirizine hydrochloride measured using PET with [<sup>11</sup>C]doxepin. *Hum Psychopharmacol* **24**, 540-548 (2009).
- 10 Woosley, R. L., Chen, Y., Freiman, J. P. & Gillis, R. A. Mechanism of the cardiotoxic actions of terfenadine. *JAMA* **269**, 1532-1536 (1993).
- 11 Yap, Y. G. & Camm, A. J. Potential cardiac toxicity of H1-antihistamines. *Clin Allergy Immunol* **17**, 389-419 (2002).
- 12 Okamura, N. *et al.* Functional neuroimaging of cognition impaired by a classical antihistamine, d-chlorpheniramine. *Br J Pharmacol* **129**, 115-123 (2000).
- 13 Cusack, B., Nelson, A. & Richelson, E. Binding of antidepressants to human brain receptors: focus on newer generation compounds. *Psychopharmacology (Berl)* **114**, 559-565 (1994).
- 14 Sarker, S. *et al.* The high-affinity binding site for tricyclic antidepressants resides in the outer vestibule of the serotonin transporter. *Mol Pharmacol* **78**, 1026-1035 (2010).
- 15 Klabunde, T. & Hessler, G. Drug design strategies for targeting G-protein-coupled receptors. *Chembiochem* **3**, 928-944 (2002).
- 16 de Graaf, C. & Rognan, D. Customizing G Protein-coupled receptor models for structure-based virtual screening. *Curr Pharm Des* **15**, 4026-4048 (2009).
- 17 Palczewski, K. *et al.* Crystal structure of rhodopsin: A G protein-coupled receptor. *Science* **289**,

739-745 (2000).

18 Cherezov, V. *et al.* High-resolution crystal structure of an engineered human beta2-adrenergic G protein-coupled receptor. *Science* **318**, 1258-1265 (2007).

19 Warne, T. *et al.* Structure of a beta(1)-adrenergic G-protein-coupled receptor. *Nature* **454**, 486-491 (2008).

20 Jaakola, V. P. *et al.* The 2.6 angstrom crystal structure of a human A2A adenosine receptor bound to an antagonist. *Science* **322**, 1211-1217 (2008).

21 Murakami, M. & Kouyama, T. Crystal structure of squid rhodopsin. *Nature* **453**, 363-367 (2008).

22 Wu, B. *et al.* Structures of the CXCR4 chemokine GPCR with small-molecule and cyclic peptide antagonists. *Science* **330**, 1066-1071 (2010).

23 Chien, E. Y. *et al.* Structure of the human dopamine D3 receptor in complex with a D2/D3 selective antagonist. *Science* **330**, 1091-1095 (2010).

24 Kolb, P. *et al.* Structure-based discovery of {beta}2-adrenergic receptor ligands. *Proc Natl Acad Sci U S A* **106**, 6843-6848 (2009).

25 Katritch, V. *et al.* Structure-based discovery of novel chemotypes for adenosine A(2A) receptor antagonists. *J Med Chem* **53**, 1799-1809 (2010).

26 Newstead, S., Kim, H., von Heijne, G., Iwata, S. & Drew, D. High-throughput fluorescent-based optimization of eukaryotic membrane protein overexpression and purification in *Saccharomyces cerevisiae*. *Proc Natl Acad Sci U S A* **104** (2007).

27 Rosenbaum, D. M. *et al.* GPCR engineering yields high-resolution structural insights into beta2-adrenergic receptor function. *Science* **318**, 1266-1273 (2007).

28 Ratnala, V. R. *et al.* Large-scale overproduction, functional purification and ligand affinities of the His-tagged human histamine H1 receptor. *Eur J Biochem* **271**, 2636-2646 (2004).

29 Qanbar, R. & Bouvier, M. Role of palmitoylation/depalmitoylation reactions in G-protein-coupled receptor function. *Pharmacol Ther* **97**, 1-33 (2003).

30 Wieland, K. *et al.* Mutational analysis of the antagonist-binding site of the histamine H(1) receptor. *J Biol Chem* **274**, 29994-30000 (1999).

31 Ohta, K. *et al.* Site-directed mutagenesis of the histamine H1 receptor: roles of aspartic acid107, asparagine198 and threonine194. *Biochem Biophys Res Commun* **203**, 1096-1101 (1994).

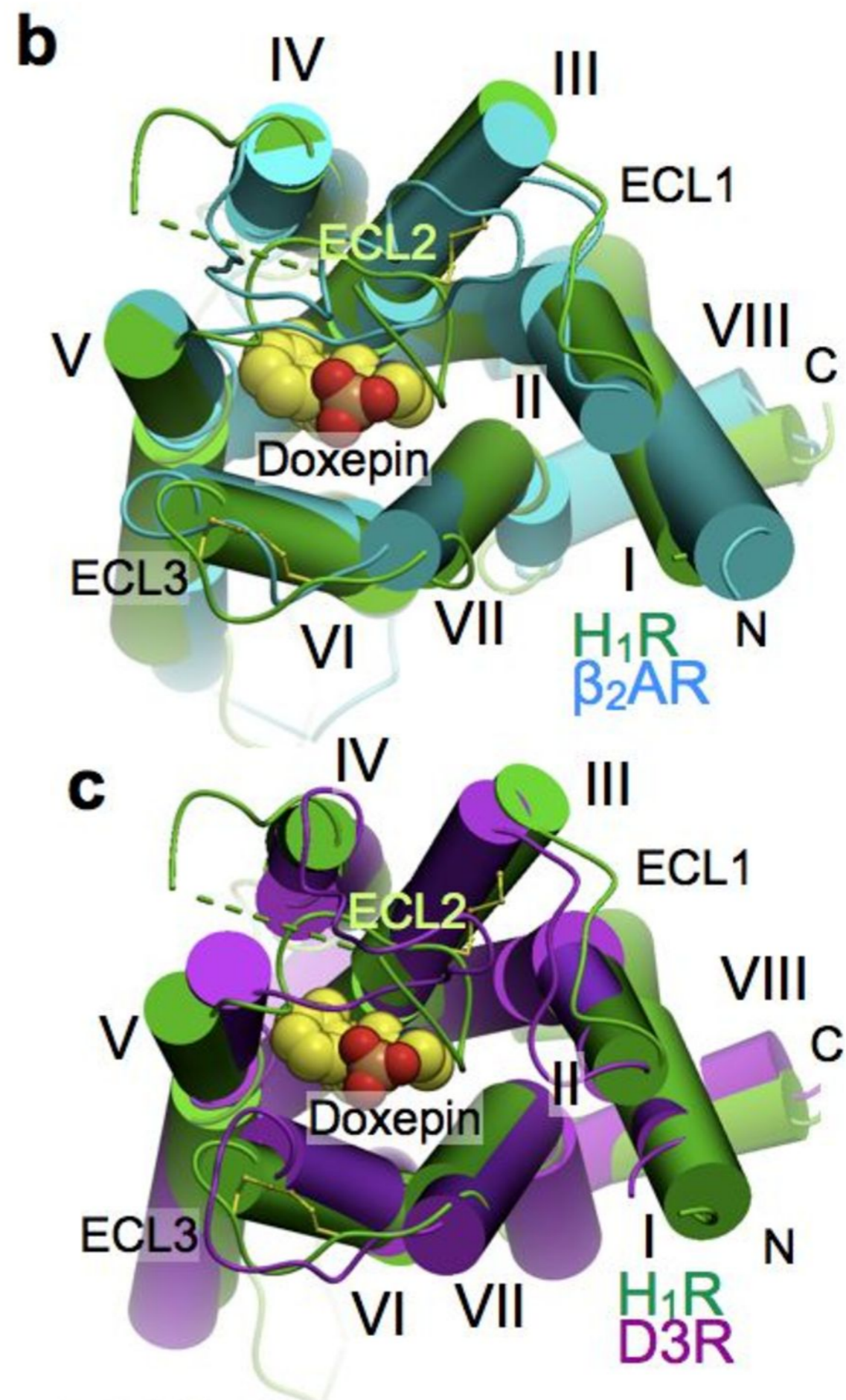
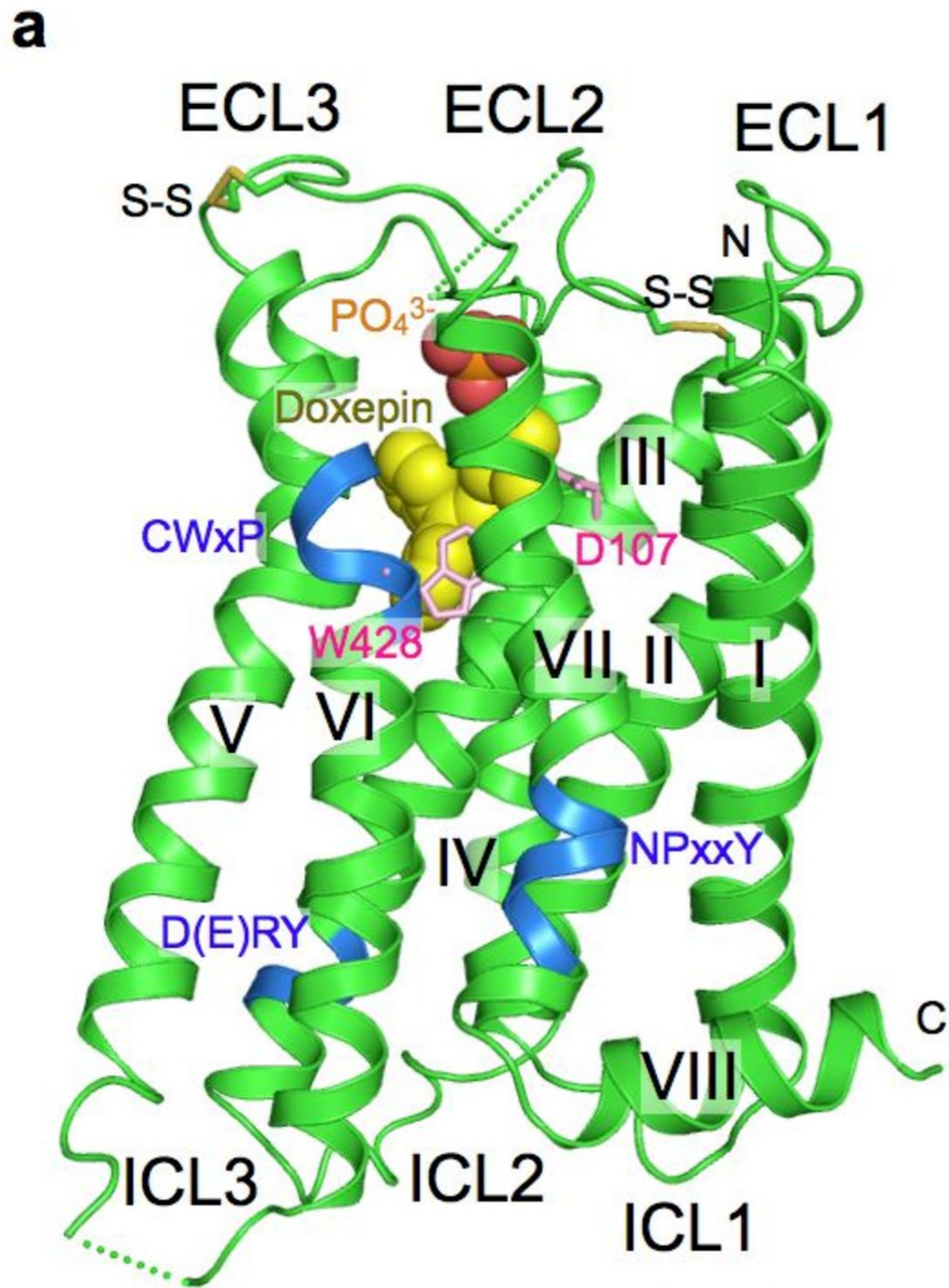
32 Nonaka, H. *et al.* Unique binding pocket for KW-4679 in the histamine H1 receptor. *Eur J Pharmacol* **345**, 111-117 (1998).

33 Bruysters, M. *et al.* Mutational analysis of the histamine H1-receptor binding pocket of histaprodifens. *Eur J Pharmacol* **487**, 55-63 (2004).

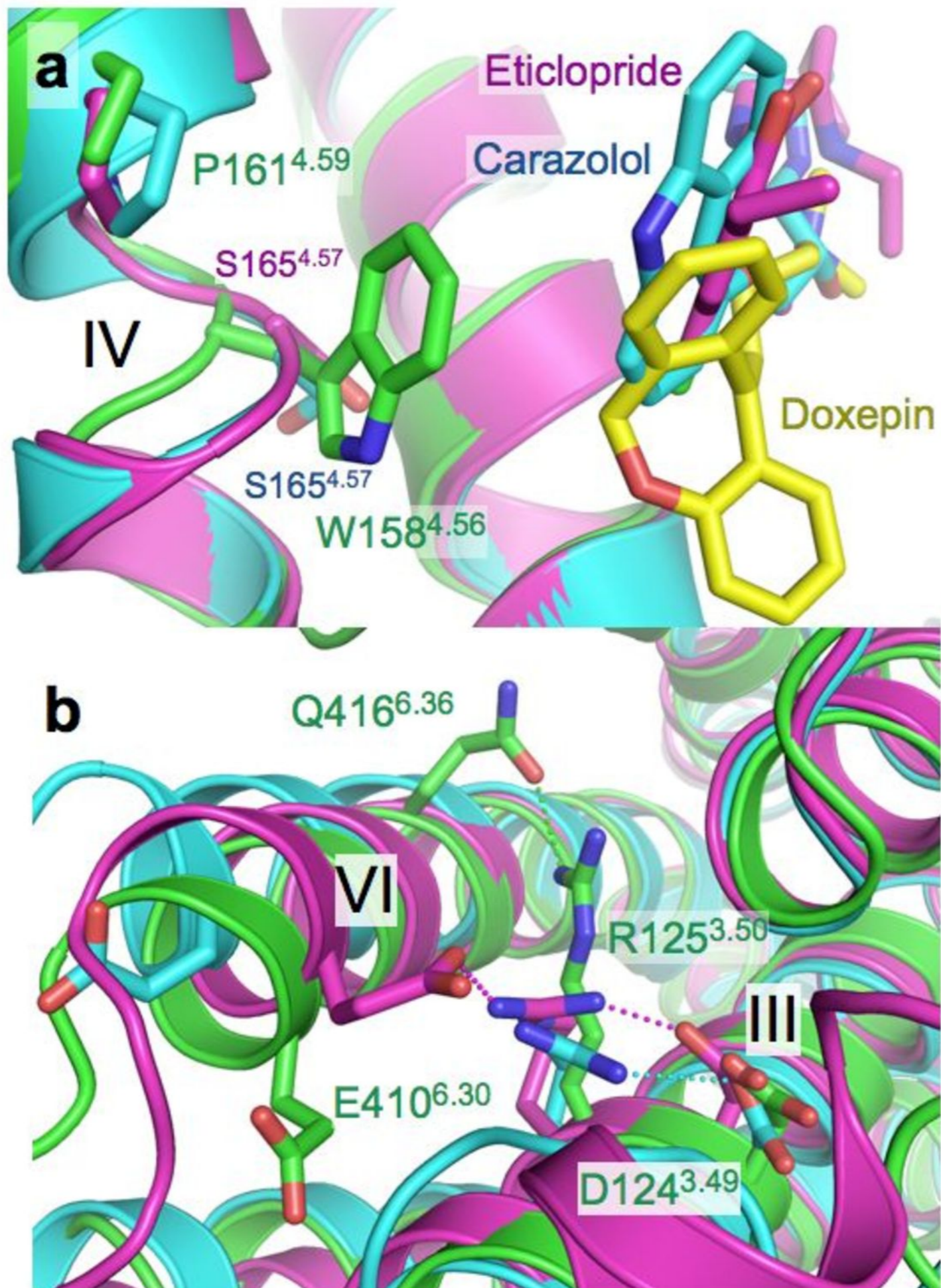
34 Peeters, M. C., van Westen, G. J., Li, Q. & Ijzerman, A. P. Importance of the extracellular loops in G protein-coupled receptors for ligand recognition and receptor activation. *Trends Pharmacol Sci* **32**, 35-42 (2011).

- 35 Gillard, M. *et al.* H1 antagonists: receptor affinity versus selectivity. *Inflamm Res* **52 Suppl 1**, S49-50 (2003).
- 36 Kiss, R., Kovari, Z. & Keseru, G. M. Homology modelling and binding site mapping of the human histamine H1 receptor. *Eur J Med Chem* **39**, 959-967 (2004).
- 37 Jongejan, A. & Leurs, R. Delineation of receptor-ligand interactions at the human histamine H1 receptor by a combined approach of site-directed mutagenesis and computational techniques - or - how to bind the H1 receptor. *Arch Pharm (Weinheim)* **338**, 248-259 (2005).
- 38 Totrov, M. & Abagyan, R. Flexible protein-ligand docking by global energy optimization in internal coordinates. *Proteins Suppl 1*, 215-220 (1997).
- 39 Katritch, V. *et al.* Analysis of full and partial agonists binding to beta2-adrenergic receptor suggests a role of transmembrane helix V in agonist-specific conformational changes. *J Mol Recognit* **22**, 307-318 (2009).
- 40 Katritch, V., Kufareva, I. & Abagyan, R. Structure based prediction of subtype-selectivity for adenosine receptor antagonists. *Neuropharmacology* **60**, 108-115 (2011).
- 41 Matsumoto, Y., Funahashi, J., Mori, K., Hayashi, K. & Yano, H. The noncompetitive antagonism of histamine H1 receptors expressed in Chinese hamster ovary cells by olopatadine hydrochloride: its potency and molecular mechanism. *Pharmacology* **81**, 266-274 (2008).
- 42 Gillard, M., Van Der Perren, C., Moguilevsky, N., Massingham, R. & Chatelain, P. Binding characteristics of cetirizine and levocetirizine to human H(1) histamine receptors: contribution of Lys(191) and Thr(194). *Mol Pharmacol* **61**, 391-399 (2002).
- 43 Leurs, R., Smit, M. J., Meeder, R., Ter Laak, A. M. & Timmerman, H. Lysine200 located in the fifth transmembrane domain of the histamine H1 receptor interacts with histamine but not with all H1 agonists. *Biochem Biophys Res Commun* **214**, 110-117 (1995).
- 44 Bakker, R. A. *et al.* 8R-lisuride is a potent stereospecific histamine H1-receptor partial agonist. *Mol Pharmacol* **65**, 538-549 (2004).
- 45 Xu, F. *et al.* Structure of an agonist-bound human A2A adenosine receptor. *Science* **332**, 322-327 (2011).
- 46 Rasmussen, S. G. *et al.* Structure of a nanobody-stabilized active state of the beta(2) adrenoceptor. *Nature* **469**, 175-180 (2011).

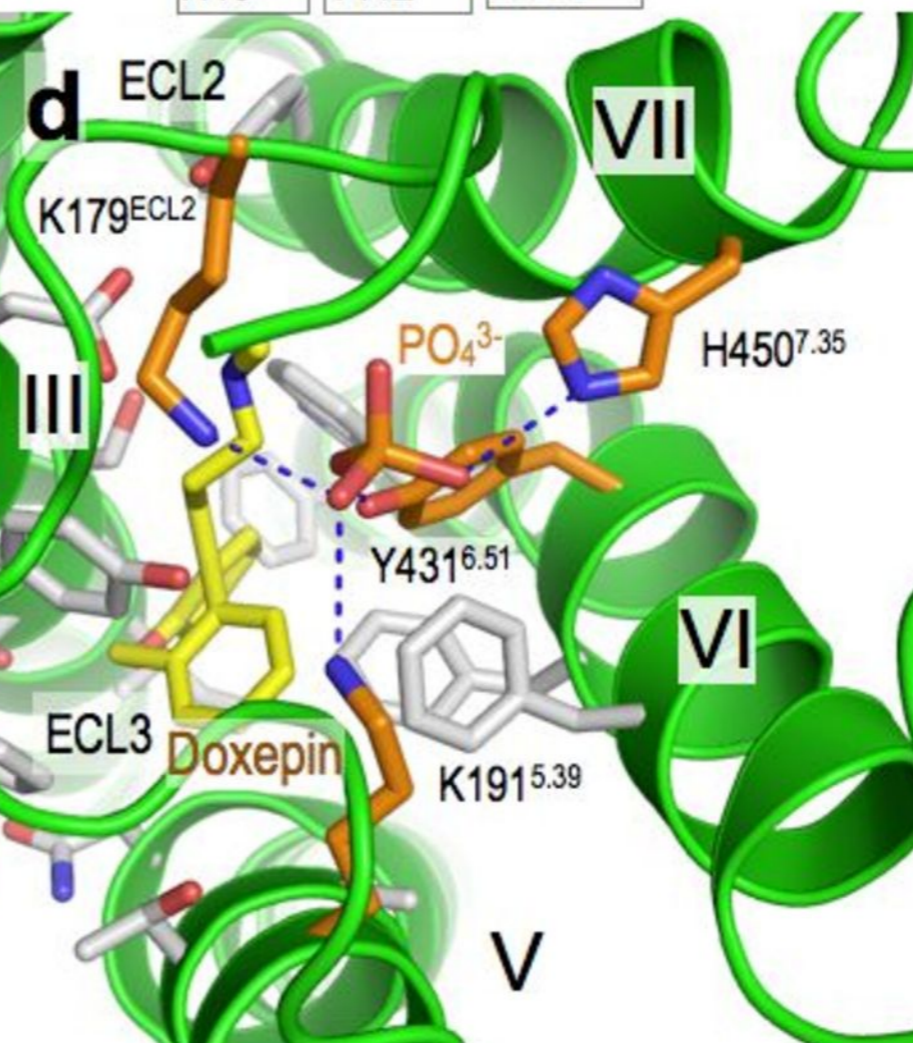
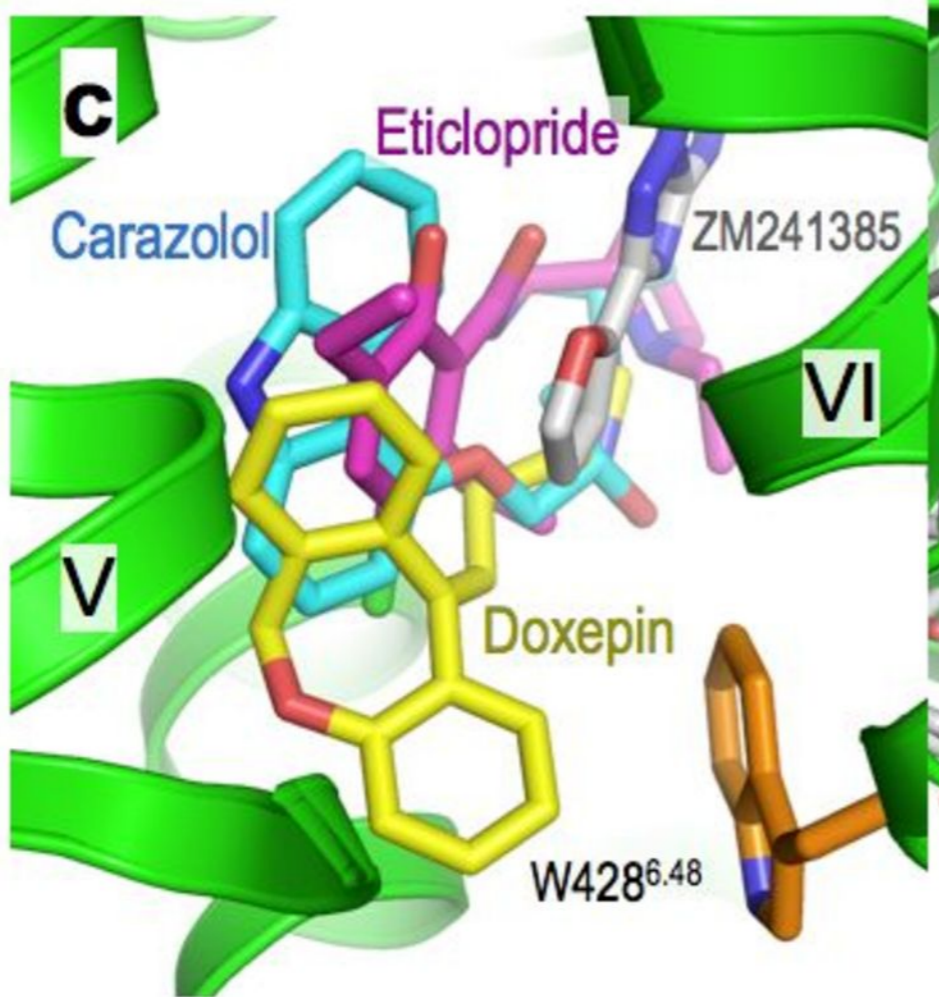
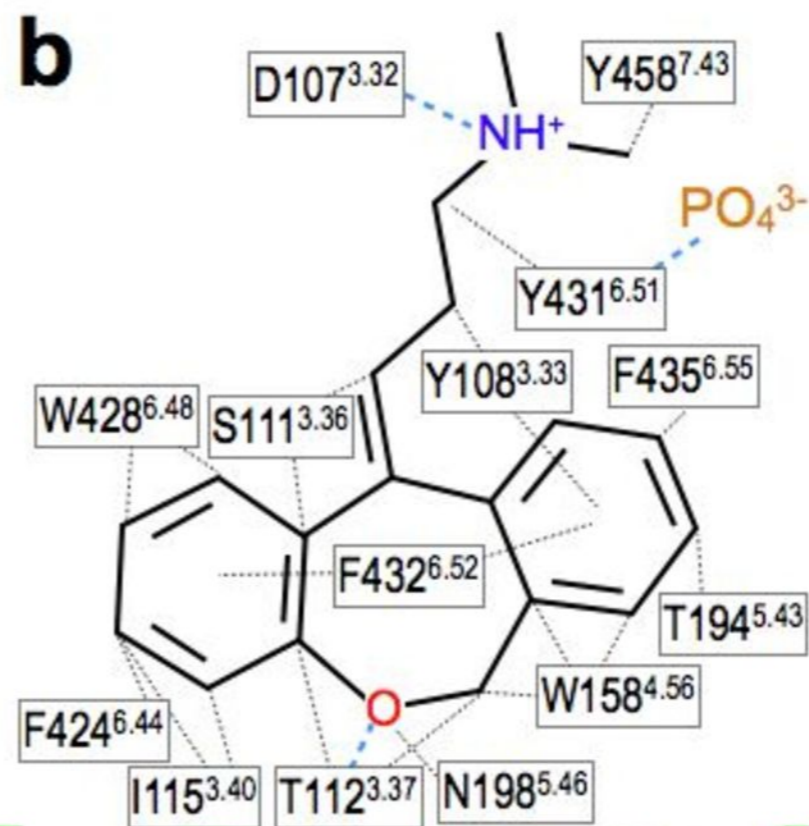
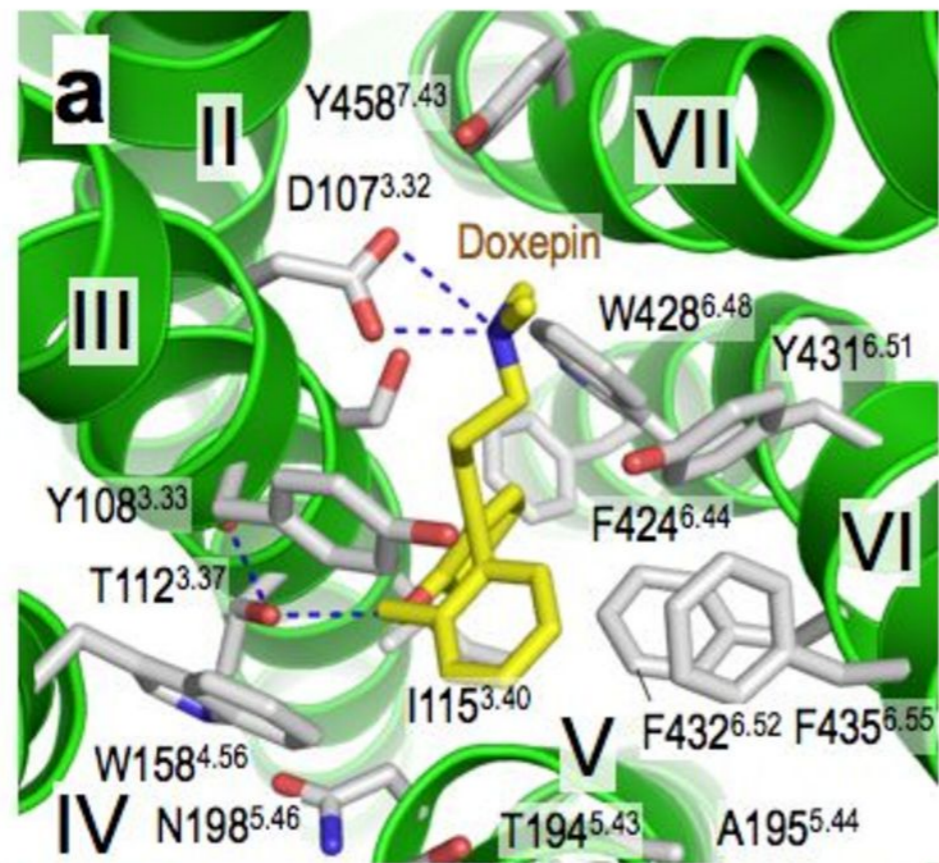


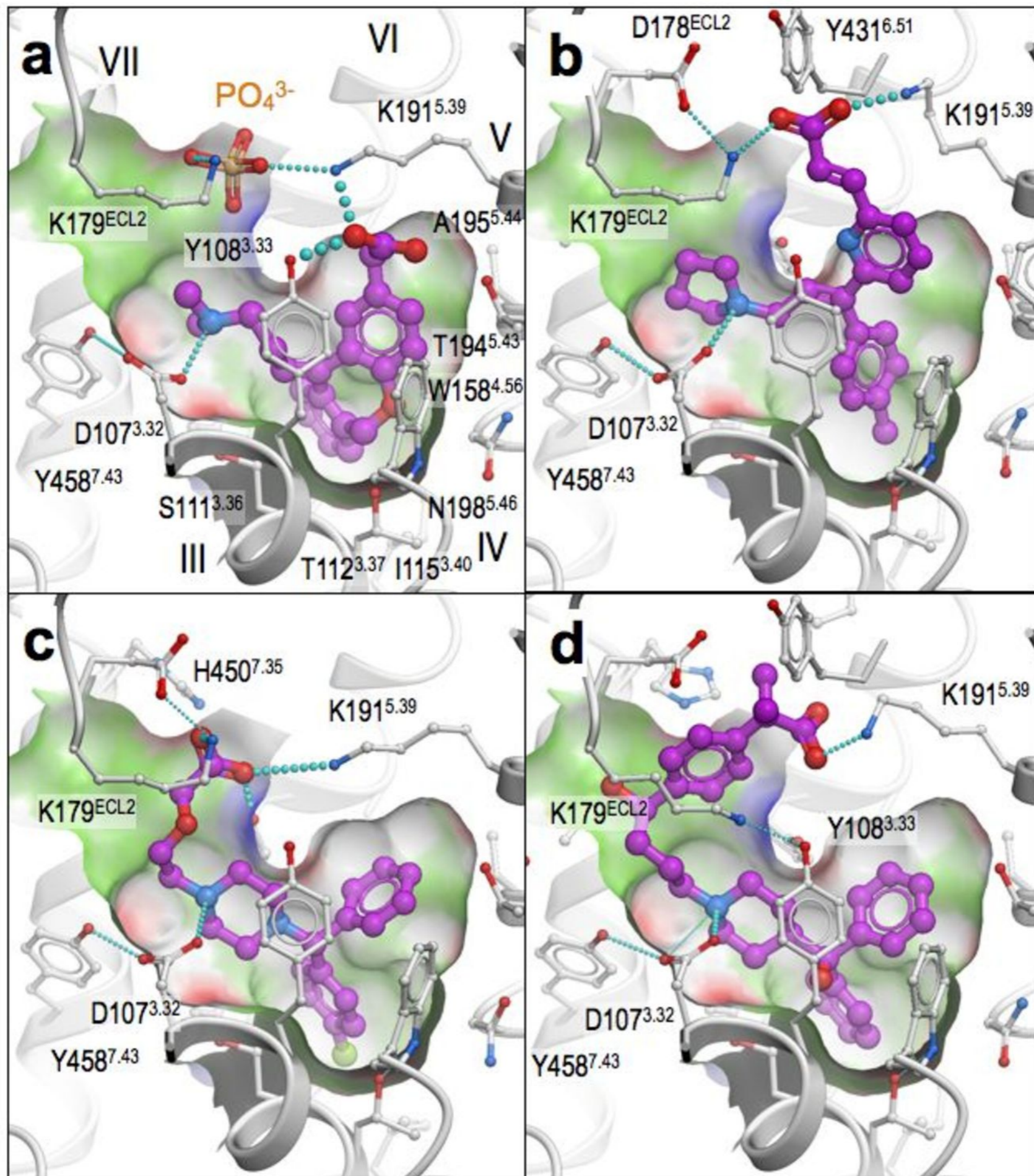


Shimamura *et al.* Fig.1



Shimamura *et al.*,  
Fig.2





Shimamura  
*et al.*, Fig.4

## **Methods**

### **Construction of the H<sub>1</sub>R expression vectors for *Pichia pastoris***

The coding sequence of the full-length human histamine H<sub>1</sub> receptor (H<sub>1</sub>R-fl), in which N-linked glycosylation sites (Asn5 and Asn18) were mutated to glutamines, was synthesized with optimization of codon usage for *P. pastoris* (TAKARA bio Inc.), and cloned into the pPIC9K expression vector (Invitrogen). The H<sub>1</sub>R-T4L construct with N-terminal 19 residues deletion and insertion of cystein-less (C54T, C97A) T4 lysozyme into the third intracellular loop was generated by yeast homologous recombination technique in *Saccharomyces cerevisiae* with the SmaI linearized plasmid pDDGFP2<sup>1</sup> and three PCR products with ~30 bp overlapped sequences. The three fragments were individually generated by standard PCR techniques with the indicated primers. The generated plasmid integrating H<sub>1</sub>R-T4L followed by TEV cleavage sequence (ENLYFQG), yeast enhanced GFP and octa-histidine tag (H<sub>1</sub>R-T4L-GFP) was isolated from *S. cerevisiae*. Coding regions of the H<sub>1</sub>R-T4L-GFP fusions were amplified by PCR using a forward primer containing a *Bam*HI site (5'-CTA GAA CTA GTG GAT CCA CCA TG-3') and a reverse primer containing an *Eco*RI site (5'-GCT TGA TAT CGA ATT CCT GCA GTT AAT G-3'). The PCR products were digested with BamHI and EcoRI, and subcloned into the pPIC9K vector.

### **Expression and membrane preparation**

The PmeI linearized pPIC9K expression vector integrating H<sub>1</sub>R-fl-GFP or H<sub>1</sub>R-T4L-GFP was then transformed into the *P. pastoris* SMD1163 strain by electroporation (2000 V, 25 mF, and 600 Ω) using a Gene Pulser I (Bio-Rad). Clone selection was performed on the YPD-agar plate containing 0.1 mg/ml geneticine. A single colony of *P. pastoris* transformant was inoculated into BMGY medium [1% (w/v) yeast extract, 2% (w/v) peptone, 1.34% (w/v) yeast nitrogen base without amino acids, 0.00004% (w/v) biotin, 1% (w/v) glycerol, 0.1 M phosphate buffer at pH 6.0] at 30 °C with shaking at 250 rpm until an OD<sub>600</sub> of 2–6 was reached. The cells were harvested by centrifugation. To induce expression, the cell pellet was resuspended to an OD<sub>600</sub> of 1.0 in BMMY medium [1% (w/v) yeast extract,

2% (w/v) peptone, 1.34% (w/v) yeast nitrogen base without amino acids, 0.00004% (w/v) biotin, 0.5% (v/v) methanol, 0.1 M phosphate buffer at pH 7.0] containing 2.5% (v/v) DMSO at 30 °C. Cells were harvested within 20 to 24 hours after induction, and stored at -80 °C. Yeast cells were disrupted with 0.5 mm glass beads in a buffer containing 50 mM HEPES, pH 7.5, 120 mM NaCl, 5%(v/v) glycerol, 2 mM EDTA and EDTA-free protein inhibitor cocktail (Roche). Undisrupted cells and cell debris were separated by centrifugation at 3000 x g, and yeast membrane were collected by ultracentrifugation at 100,000 x g for 30 min at 4 °C. Washing of the membranes was performed by repeating dounce homogenation and centrifugation in a high salt buffer containing 10 mM HEPES, pH 7.5, 1 M NaCl, 10 mM MgCl<sub>2</sub>, 20 mM KCl and EDTA-free protease inhibitor cocktail. Prepared membranes were resuspended in a buffer containing 50 mM HEPES pH 7.5, 120 mM NaCl, 20% (v/v) glycerol and EDTA-free protease inhibitor cocktail, and snap-frozen in liquid nitrogen and stored at -80 °C until use. Membrane proteins were quantified using the bicinchoninic acid method (Pierce).

### **Purification of H<sub>1</sub>R-T4L**

Membrane suspension containing H<sub>1</sub>R-T4L-GFP was thawed and incubated on ice for 30 min in the presence of 5 mM doxepin, 10 mg/ml iodoacetamide, and EDTA-free protease inhibitor cocktail (Roche). The membrane suspension was poured into the buffer containing 20 mM HEPES pH 7.5, 500 mM NaCl, 1% (w/v) n-dodecyl-β-D-maltopyranoside (DDM, Anatrace), 0.2% (w/v) cholesteryl hemisuccinate (CHS, Sigma), 20% (v/v) glycerol and 2-3 mg/ml membrane, and stirred gently at 4 °C for 1-2 hours. The unsolubilized material was separated by centrifugation at 100,000 × g for 30 min. The supernatant was incubated with TALON IMAC resin (Clontech) overnight. The resin was washed with twenty column volumes of 20 mM HEPES pH7.5, 500 mM NaCl, 10% glycerol, 0.025% (w/v) DDM, 0.005% (w/v) CHS, 100 μM doxepin and 20 mM imidazole. The protein was eluted with 4 column volumes of 20 mM HEPES pH 7.5, 500 mM NaCl, 10% glycerol, 0.025% (w/v) DDM, 0.005% (w/v) CHS, 500 μM doxepin and 200 mM imidazole. The eluted fractions were concentrated to 2.5 ml with a 100 kDa molecular weight cut-off AmiconUltra

(Millipore). Imidazole was removed using PD-10 column (GE healthcare). Then the protein was loaded onto the Ni-Sepharose high performance resin (GE healthcare) (1.5 ml resin for ~10 mg of protein). The resin was washed with 20 column volume of 20 mM HEPES pH 7.5, 500 mM NaCl, 10% glycerol, 0.025% (w/v) DDM, 0.005% (w/v) CHS, 500  $\mu$ M doxepin and 20 mM imidazole. The sample was eluted with three column volumes of 20 mM HEPES pH 7.5, 500 mM NaCl, 10% glycerol, 0.025% (w/v) DDM, 0.005% (w/v) CHS, 1 mM doxepin and 500 mM imidazole. Imidazole was removed using PD-10 column (GE healthcare). The protein was processed overnight with His-tagged TEV protease (expressed and purified in house). TEV protease and the cleaved His-tagged GFP were removed by passing the sample through the Ni-Sepharose high performance resin. The receptor was concentrated to 30-40 mg/ml with a 100 kDa molecular weight cut-off Vivaspin concentrator (Vivascience). Protein purity and monodispersity were tested by SDS-PAGE and by size-exclusion chromatography using Superdex 200 (GE healthcare).

### **Lipidic cubic phase crystallization**

Lipidic cubic phase (LCP) crystallization trials were performed using an *in meso* crystallization robot as previously described<sup>2</sup>. 96-well glass sandwich plates were filled with 40-50 nl protein-laden LCP boluses overlaid by 0.8  $\mu$ l of precipitant solution in each well and sealed with a glass cover-slip. The protein-LCP mixture contained 40% (w/w) receptor solution, 54% (w/w) monoolein, and 6% (w/w) cholesterol. Crystallization set-ups were performed at room temperature (20-22 °C). Plates were incubated and imaged at 20 °C using an automated incubator/imager (RockImager 1000, Formulatrix). Crystals were obtained in 26-30% (v/v) PEG400, 300 mM ammonium phosphate, 10mM MgCl<sub>2</sub>, 100 mM Na-citrate pH 4.5 and 1 mM doxepin (Sigma) (Supplementary Figure 5). Crystals were harvested directly from LCP matrix using MiTeGen micromounts and were flash-frozen in liquid nitrogen without additional cryoprotectant.

### **Data collection and refinement**

X-ray diffraction data were collected at 100 K with a wavelength of 0.97780 Å and with a

beamsize of 10 x 10 microns on the microfocus beamline I24 at the Diamond Light Source (UK) with a Pilatus 6M detector. Each loop was subjected to a grid scanning <sup>3</sup> in order to locate the crystals, which are invisible in the lipidic cubic phase once they are mounted. The exact locations and dimensions of the chosen crystals were determined by further grid scanning with a smaller search area. Data collection was carried out by collecting several overlapping wedges of data from adjacent positions within a single crystal. The data were processed initially with xia2 <sup>4</sup> using Mosflm <sup>5</sup> and Scala <sup>6</sup> with the merging statistics used to determine an optimum subset of measurements to merge. The final data set consisted of data from five of the eight positions recorded, giving a total of 75 degrees of data. These data were then remerged with Scala to give the final data set summarized in Table S2. The space group was determined to be *I*422 with one molecule in the asymmetric unit. Diffraction data were slightly anisotropic, extending to 2.9 Å in the *c*\* direction and 3.1 Å in the *a*\* and *b*\* directions. The structure factors up to 3.1 Å resolution were anisotropically scaled by PHASER <sup>7</sup> and then used for the subsequent molecular replacement and refinement. The structure was determined by molecular replacement with the program PHASER <sup>7</sup> using two independent search models (polyalanine of the 7 TM  $\alpha$ -helices, and T4L) from  $\beta_2$ AR (PDB ID: 2RH1) structure. We chose  $\beta_2$ AR as a model structure because it has the highest homology of transmembrane helices with H<sub>1</sub>R (41.7%) among the human GPCR structures. For the initial map calculation after molecular replacement, however, we used a  $\beta_2$ AR model without side chains, loops, ligand, lipids and any solvents, therefore the final H<sub>1</sub>R structure is not biased to the  $\beta_2$ AR structure. This is supported by low  $R_{\text{work}}$  and  $R_{\text{free}}$  values (Supplementary Table 2). All refinements were performed with REFMAC5 <sup>8</sup> and autoBUSTER <sup>9</sup> followed by manual examination and rebuilding of the refined coordinates in the program Coot <sup>10</sup>. The non-lysozyme portion contains higher B-factors (116Å<sup>2</sup>) due to fewer contacts as compared to T4 lysozyme (36 Å<sup>2</sup>). Calculation of the surface area buried by crystal contacts also explains this. For the non-lysozyme portion, only 8% (1,225 Å<sup>2</sup>) of 15,689 Å<sup>2</sup> solvent accessible surface area is buried by crystal contacts. In contrast, for the T4 lysozyme portion, 32% (2,733 Å<sup>2</sup>) of the solvent accessible area (8,648 Å<sup>2</sup>) is buried by crystal interactions. Supplementary Figure 4 also shows there are strong interactions between



T4 lysozyme domains, but relatively fewer between non-lysozyme domains throughout the crystal packing. Although the average B-factor of the non-lysozyme domain is high as compared to T4 lysozyme, as shown in the supplementary Figures 3 and 4, electron densities were clear for unambiguous model building. The H<sub>1</sub>R 8 N-terminal residues (Thr20-Leu27), 2 C-terminal residues (Arg486-Ser487), and 7 residues (Phe168-Val174) in the second extracellular loop (ECL2) are not included in the structure as they did not have interpretable densities.

Strong and spherical electron densities (about 4 sigma) were found in the anion-binding region in the Fo-Fc omit map. We excluded the presence of a water molecule in this region due to strong residual positive Fo-Fc densities when we modeled it as a water molecule. The coordination geometry in the highly electropositive environment surrounded by His450<sup>7,35</sup>, Lys179 and Lys191<sup>5,39</sup> implied that either a phosphate or sulfate ion could be modeled. Since ammonium phosphate was added to our crystallization buffer, we modeled it as a phosphate ion. The average B-factors of the phosphate ion and the interacting atoms are 177 Å<sup>2</sup> and 154 Å<sup>2</sup>, respectively.

### **Ligand binding assays**

For the saturation binding experiment, yeast membrane suspensions containing H<sub>1</sub>R-fl-GFP (20 µg) or H<sub>1</sub>R-T4L-GFP (5 µg) were incubated with increasing concentrations of [<sup>3</sup>H] pyrilamine (from 0.15 to 40 nM) in a total assay volume of 200 µl for 1 h at 25 °C. In order to investigate effect of phosphate on the ligand binding, assays were performed in PBS buffer pH 7.4 (138 mM NaCl, 8.1 mM Na<sub>2</sub>HPO<sub>4</sub>, 27 mM KCl, 1.8 mM KH<sub>2</sub>PO<sub>4</sub>) or in the HEPES buffer containing 20 mM HEPES pH 7.5 and 150 mM NaCl. Nonspecific binding was determined in the presence of 1000-times excess unlabeled pyrilamine. Membranes were trapped on Whatman GF/B filters pre-soaked in 0.3% polyethylenimine, and unbound radioligands were washed with 9 ml of the PBS or the HEPES buffers. The retained radioactivity was measured on an LCS-5100 liquid scintillation counter (ALOKA) in a Clearzol I scintillation liquid (Nakarai, Japan). Data were analyzed by non-linear curve-fitting with a rectangular hyperbola function using the Prism 4.0 software (GraphPad) to determine

dissociation constant ( $K_d$ ).

For competition binding assays, yeast membrane suspensions containing H<sub>1</sub>R-fl-GFP or H<sub>1</sub>R-T4L-GFP were incubated with 4 nM or 20 nM [<sup>3</sup>H]pyrilamine in the PBS buffer or the HEPES buffer in the presence of 10 nM to 100 mM histamine hydrochloride or 0.001 nM to 1 μM doxepin, or 0.01 nM to 10 μM cetirizine, pyrilamine, olopatadine and fexofenadine. Data were analyzed by non-linear curve fitting with a sigmoidal function using the Prism 4.0 to determine the half maximal inhibitory concentrations (IC<sub>50</sub>). All data shown calculated based on more than three independent experiments. Inhibition constant  $K_i$  was calculated based on the equation  $K_i = IC_{50}/(1+L/K_d)$ , where L is the concentration of [<sup>3</sup>H]pyrilamine with the dissociation constant  $K_d$ .

### **Thermal stability assay**

N-[4-(7-diethylamino-4-methyl-3-coumarinyl)phenyl]maleimide (CPM) dye was purchased from Invitrogen and dissolved in DMSO (Sigma) at 4 mg/ml as the stock solution for future use. The stock solution was kept at -80°C and was diluted 1:40 in dye dilution solution (10 mM buffer, 500 mM NaCl, 10% glycerol, 0.025% DDM and 0.005% CHS) before use. The thermal denaturation assay was performed with total volume of 200 μl sample in a quartz fluorometer cuvette (Starna Cells, Inc., Atascadero, CA). H<sub>1</sub>R (4 μg) was diluted in the appropriate buffer solution to a final volume of 200 μl. Five micro litter of the diluted dye was added to the protein solution and it was incubated for 30 min at 4°C. The mixed solution was transferred to the cuvette and the data were collected by a Cary Eclipse spectrofluorometer (Varian, USA) with a temperature ramping rate at 1°C/min. The excitation wavelength was 387 nm and the emission wavelength was 463 nm. All assays were performed over a temperature range starting from 20°C and 80°C. The stability data were processed with GraphPad Prism program (GraphPadPrism, Graphpad Software, San Diego, CA, USA). In order to determine the melting temperature (T<sub>m</sub>), a Boltzmann sigmoidal equation was used to fit to the data.

### **Flexible Ligand-Receptor Docking**

Docking of ligands was performed using all-atom flexible receptor docking algorithm in ICM-Pro molecular modeling package<sup>11</sup> as described previously<sup>12-13</sup>. The initial H<sub>1</sub>R receptor model was generated in ICM by building hydrogen atoms for the crystal structure of H<sub>1</sub>R. Internal coordinate (torsion) movements were allowed in the side chains of the binding pocket, defined as residues within 8Å distance of doxepin in the H<sub>1</sub>R-doxepin complex. Other side chains and backbone of the protein were kept as in the crystal structure. An initial conformation for each of the ligands was generated by Cartesian optimization of the ligand model in MMFF force field. Docking was performed by placing the ligand in a random position within 5Å from the binding pocket and global optimization of the complex conformational energy. The global energy of the complex was calculated as a sum of van der Waals, electrostatic, hydrogen-bonding and torsion stress terms. Stochastic global energy optimization of the complex was performed using the ICM Monte Carlo (MC) procedure with minimization<sup>14</sup>. To facilitate side chain rotamer switches in flexible H<sub>1</sub>R models, the first 10<sup>6</sup> steps of the MC procedure used “soft” vdW potentials and high MC temperature, followed by another 10<sup>6</sup> steps with “exact” vdW method and gradually decreasing temperature. A harmonic “distance restraint” has been applied between amino group of the ligand and carboxyl of Asp107 side chain in the initial 10<sup>6</sup> steps to facilitate formation of the known salt bridge interaction between these two groups. At least 10 independent runs of the docking procedure were performed for each H<sub>1</sub>R-ligand. The docking results were considered “consistent” when at least 80% of the individual runs resulted in conformations clustered within a RMSD of <0.5 Å to the overall best energy pose of the ligand.

## References

1. Newstead, S., Kim, H., von Heijne, G., Iwata, S., Drew, D. High-throughput fluorescent-based optimization of eukaryotic membrane protein overexpression and purification in *Saccharomyces cerevisiae*. *Proc Natl Acad Sci U S A.*, **104**, 13936-13941 (2007).
2. Cherezov, V., Peddi, A., Muthusubramaniam, L., Zheng, Y. F., Caffrey, M. A robotic system for crystallizing membrane and soluble proteins in lipidic mesophases. *Acta Crystallogr. D* **60**, 1795-1807 (2004).
3. Aishima, J., Owen, R. L., Axford, D., Shepherd, E., Winter, G., Levik, K., Gibbons, P., Ashton, A., Evans, G. High-speed crystal detection and characterization using a fast-readout

- detector. *Acta Crystallogr. D* **66**, 1032-1035 (2010).
4. Winter, G. xia2: an expert system for macromolecular crystallography data reduction. *J Appl Cryst.* **43**, 186–190 (2010).
  5. Leslie, A. G. W. Recent Changes to the MOSFLM Package for Processing Film and Image Plate Data., *Joint CCP4 and ESF-EACMB Newsletter on Protein Crystallography*, No. 26 (1992).
  6. Evans, P. Scaling and assessment of data quality. *Acta Crystallogr. D* **62**, 72-82 (2006).
  7. McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., Read, R. J. Phaser crystallographic software. *J Appl Crystallogr.* **40**, 658-674 (2007).
  8. Skubak, P., Murshudov, G. N., Pannu, N. S. Direct incorporation of experimental phase information in model refinement. *Acta Crystallogr. D* **60**, 2196-2201 (2004).
  9. Murshudov, G. N., Vagin, A. A. and Dodson, E. J. Refinement of Macromolecular Structures by the Maximum-Likelihood Method. *Acta Crystallogr. D* **53**, 240-255 (1997).
  10. Emsley, P., Lohkamp, B., Scott, W. G., Cowtan, K. Features and Development of Coot. *Acta Crystallogr. D* **66**, 486-501 (2010).
  11. ICM Manual v.3.0 (MolSoft LLC, La Jolla, CA, 2011).
  12. Totrov, M. & R., A. Derivation of sensitive discrimination potential for virtual ligand screening. (*RECOMB 99*) Lyon France, *ACM Press.* , 312–317. (1999).
  13. Katritch, V. *et al.* Analysis of full and partial agonists binding to beta2-adrenergic receptor suggests a role of transmembrane helix V in agonist-specific conformational changes. *J Mol Recognit* **22**, 307-318, doi:10.1002/jmr.949 (2009).
  14. Abagyan, R. & Totrov, M. Biased probability Monte Carlo conformational searches and electrostatic calculations for peptides and proteins. *J Mol Biol* **235**, 983-1002 (1994).