J Comput Aided Mol Des (2010) 24:759–769
DOI 10.1007/s10822-010-9372-2

In silico analysis of the histaprodifen induced activation pathway of the guinea-pig histamine H_1-receptor

Andrea Straßer · Hans-Joachim Wittmann

Received: 7 January 2010 / Accepted: 28 June 2010 / Published online: 7 July 2010

Abstract The binding of (partial) agonists in the binding pocket of biogenic amine receptors induces a conformational change from the inactive to the active state of the receptors. There is only little knowledge about the binding pathways of ligands into binding pocket on molecular level. So far, it was not possible with molecular dynamic simulations to observe the ligand binding and receptor activation. Furthermore, there is nearly nothing known, in which state of ligand binding, the receptor gets activated. The aim of this study was to get more detailed insight into the process of ligand binding and receptor activation. With the recently developed *LigPath* algorithm, we scanned the potential energy surface of the binding process of dimeric histaprodifen, a partial agonist at the histamine H_1-receptor, into the guinea pig histamine H_1-receptor, taking also into account the receptor activation. The calculations exhibited large conformational changes of $Trp^{6.48}$ and $Phe^{6.55}$ during ligand binding and receptor activation. Additionally, conformational changes were also observed for $Phe^{6.52}$, $Tyr^{6.51}$ and $Phe^{6.44}$. Conformational changes of $Trp^{6.48}$ and $Phe^{6.52}$ are discussed in literature as rotamer toggle switch in context with receptor activation. Additionally, the calculations indicate that the binding of dimeric histaprodifen, accompanied by receptor activation is energetically preferred. In general, this study gives new, theoretical insights onto ligand binding and receptor activation on molecular level.

Keywords Guinea-pig Histamine H_1-Receptor · G protein-coupled receptor · Inactive conformation · Active conformation · Receptor activation · LigPath

A. Straßer (✉)
Department of Pharmaceutical and Medicinal Chemistry, Faculty of Chemistry and Pharmacy, University of Regensburg, Regensburg, Germany
e-mail: andrea.strasser@chemie.uni-regensburg.de

H.-J. Wittmann
Faculty of Chemistry and Pharmacy, University of Regensburg, Regensburg, Germany

Introduction

Biogenic amine receptors, belonging to the G protein-coupled receptors (GPCRs), consist of seven transmembrane domains, which are connected by intra- and extracellular loops [1, 2]. Compounds, showing affinity to these receptors bind in a pocket deep between the transmembrane domains of the receptor. Antagonists or inverse agonists stabilize an inactive conformation of the receptor, while partial agonists induce a conformational change of the receptor into an active state [3]. Several studies give hint that GPCRs exist in multiple inactive and active conformations which may be related to each other by dynamic equilibria [4–10]. The calculation of binding energy landscape of different ligands, binding to the β_2 adrenergic receptor, supports the thesis, that there exist ligand dependent receptor conformations [11]. Similar results were obtained for bovine rhodopsin [12]. A lot of experimental studies, like x-ray crystallography [13], site-directed mutagenesis [14–17], cysteine accessibility [18–20], fluorescence measurements [21], disulfide cross-linking [22, 23], zinc crosslinking [24], spin labelling [22, 25] and EPR spectroscopy [22] have been used to study the activation mechanism of GPCRs. Some theoretical studies with regard to ligand binding and unbinding are found in literature [26, 27]. However, there is only small knowledge about the binding process of ligands to a receptor and in

which state of ligand binding, the receptor gets activated. It can be suggested, that the ligands have to be recognized at the extracellular part of the receptor and subsequently penetrate into the binding pocket. MD simulations do not allow observing the whole process of ligand binding, since simulations with an adequate computation time are too short with regard to time scale. To get more insight into the binding process, the algorithm "*LigPath*" was developed [28, 29]. With this algorithm it is possible to scan the potential energy surface during the process of ligand binding and receptor activation. At the histamine H_1-receptor (H_1R), several studies to describe structural characteristics of the active state have been performed [17, 29–32]. The aim of this study was to analyze the binding process of dimeric histaprodifen, one of the largest H_1-receptor ligands known, to guinea-pig (gp) H_1R. Site directed mutagenesis studies revealed that the exchange of asparagine to serine at position 2.61 is responsible for the species differences in pharmacology of dimeric histaprodifen between hH_1R and gpH_1R [33]. Within further mutagenesis studies, the influence of other amino acids onto pharmacology of dimeric histaprodifen at hH_1R was analyzed [34, 35]. These studies revealed an essential influence of $Asp^{3.32}$ and $Phe^{6.52}$ onto binding of dimeric histaprodifen to hH_1R [34, 35]. Since dimeric histaprodifen acts as partial agonist at gpH_1R (efficacy of about 0.92) [36–38], a conformational change from the inactive state to the active state is expected. Thus, the question arises, (1) if an agonist penetrates into the inactive receptor and then the receptor gets activated or (2) if first the receptor gets activated and then the ligand penetrates into the binding pocket or (3), if ligand penetration is accompanied by receptor activation. The different pathways are schematically presented in Fig. 1. The aim of this study was to address this question.

Methods

Numbering of residues, the *rmsd*, the torsion angles χ_1 and χ_2 and rotable bonds

The amino acids of the transmembrane domains were numbered according the numbering scheme of Ballesteros [39]. The root mean square distance $rmsd^{(rec)}$ of each receptor state *s* (inactive or each state between inactive and active) with respect to the destination structure (active state of the receptor) is defined as given in Eq. 1.

$$rmsd^{(rec)} = \sqrt{\frac{\sum_{i=1}^{N(rec)} \left(\left(x_s^i - x_{active}^i\right)^2 + \left(y_s^i - y_{active}^i\right)^2 + \left(z_s^i - z_{active}^i\right)^2 \right)}{N(rec)}} \quad (1)$$

The $rmsd^{(rec)}$ is calculated including all *N(rec)* receptor atoms with the coordinates (x_s^i, y_s^i, z_s^i) for atom *i* in the state *s* and the coordinates $(x_{active}^i, y_{active}^i, z_{active}^i)$ for atom *i* in the destination structure. An analogue $rmsd^{(lig)}$ of the ligand with respect to its destination (in the binding pocket of the receptor) is defined as given in Eq. 2.

$$rmsd^{(lig)} = \sqrt{\frac{\sum_{i=1}^{N(lig)} \left(\left(x_s^i - x_{bound}^i\right)^2 + \left(y_s^i - y_{bound}^i\right)^2 + \left(z_s^i - z_{bound}^i\right)^2 \right)}{N(lig)}} \quad (2)$$

The $rmsd^{(lig)}$ is calculated including all *N(lig)* atoms of the ligand. (x_s^i, y_s^i, z_s^i) are the coordinates for each ligand atom in state *s* and $(x_{bound}^i, y_{bound}^i, z_{bound}^i)$ are the coordinates of each ligand atom in the destination structure (bound in the binding pocket of the active receptor).

The value $rmsd^{(rec+lig)}$ of the ligand-receptor-complex with respect to its destination structure is defined, as given in Eq. 3:

$$rmsd^{(rec+lig)} = \sqrt{\frac{\sum_{i=1}^{N(rec+lig)} \left(\left(x_s^i - x_{rec+lig}^i\right)^2 + \left(y_s^i - y_{rec+lig}^i\right)^2 + \left(z_s^i - z_{rec+lig}^i\right)^2 \right)}{N(rec+lig)}} \quad (3)$$

The $rmsd^{(rec+lig)}$ is calculated including all *N(rec + lig)* atoms of the receptor and ligand. (x_s^i, y_s^i, z_s^i) are the coordinates for each receptor and ligand atom in state s and $(x_{rec+lig}^i, y_{rec+lig}^i, z_{rec+lig}^i)$ are the coordinates for each receptor and ligand atom in the destination structure.

For a chain X–A–B–Y, the torsion angles χ_1 and χ_2 are calculated in the following manner: The torsion angle χ_1 or χ_2 is defined as the angle between the orthogonal projections of the bonds X–A and B–Y into the plane, bisecting A–B, when viewing from A to B. The resulting angle is considered to be positive, if the bond in front of the

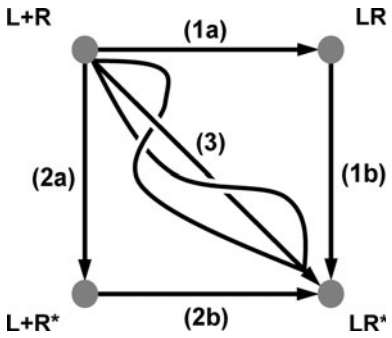

Fig. 1 Scheme for different ligand binding and receptor activation pathways: L + R: ligand L, not bound to the receptor and the inactive receptor R; LR: ligand bound in the binding pocket of the inactive receptor; L + R*: ligand L, not bound to the receptor and the active receptor R*; LR*: ligand bound in the binding pocket of the active receptor; three different pathways (1a)—(1b), (2a)—(2b) and (3) from the starting point (L + R) to the destination (LR*)

bisecting plane has to be rotated clockwise to eclipse the bond behind the bisecting plane. For χ_1, the following assignment was used: X=N, A=C_α, B=C_β and Y=C_γ. For χ_2, the assignment X=C_α, A=C_β, B=C_γ and Y=CD2 (Y=CD1 in case of $Tyr^{6.51}$) was used.

The rotable bonds of the amino acid side chains were already described [28]. The rotable bonds for the ligand dimeric histaprodifen are presented in Fig. 2.

A new version of the *LigPath*-algorithm

The *LigPath* algorithm is a combination of directional guiding, Monte-Carlo search and minimization [28]. In the previous versions of the *LigPath*-algorithm only the ligand penetration from the extracellular part into the binding pocket [28] *or* receptor activation [29] was implemented. In the new version of the *LigPath*-algorithm, ligand penetration *and* receptor activation was combined. Therefore, the starting structure can be defined by an inactive receptor with the ligand being located anywhere in the extracellular part and the destination structure may represent the active receptor with the ligand being located in the binding pocket. Thus, the pathway, from starting to destination structure is calculated by subsequent translational and rotational motions of ligand and receptor atoms in direction to destination structure. Compared to the previous versions of *LigPath* [28, 29], some modifications of parameters were introduced: Previously [28, 29], a parameter Δr was used. In the version for calculation of ligand penetration [28], Δr described the maximum step for translation of the ligand atoms in direction to its destination. In the version for calculation of receptor activation [29], Δr described the maximum step for translation of the receptor atoms in direction of their destination. Within this new version, the previous parameter Δr is substituted by Δr^{rec} and Δr^{lig}. Δr^{rec} or Δr^{lig} describe the maximum step for translation of a receptor or ligand atom in direction to its destination position. The translation of the receptor and ligand atoms is allowed to deviate $\Delta\Phi^{rec}$ or $\Delta\Phi^{lig}$ from the guiding line. Both parameters are introduced instead of $\Delta\Phi$. In the previous versions, $\Delta\Phi$ described the deviation from the guiding line for the ligand [28] or receptor atoms [29]. The maximum rotation of the rotable bonds of amino acid side chains or ligand is defined by $\Delta\varphi^{rec}$ and $\Delta\varphi^{lig}$.

Beginning with the starting structure, three groups with n new child structures for each group are generated by *LigPath* and energetically minimized (Fig. 3). In group I, only the ligand atoms are translated in direction of their destination position. This means, $\Delta r^{lig} \neq 0$ and $\Delta r^{rec} = 0$. In group II, only the receptor atoms are translated in direction to their destination position. This can be realized by $\Delta r^{rec} \neq 0$ and $\Delta r^{lig} = 0$. In group III, the ligand and the receptor atoms are translated in direction of their destination position. Therefore, parameters $\Delta r^{rec} \neq 0$ and $\Delta r^{lig} \neq 0$ are used. Both, $\Delta r^{rec} = \Delta r^{lig}$ or $\Delta r^{rec} \neq \Delta r^{lig}$ is possible. After minimizing each of the n children of each of the three groups, $3n$ new structures are received. The best child out of all $3n$ children is used as a new starting structure for the next generation. As already described for the previous version of *LigPath*, (Eq. 4) was used to determine the best child of a generation:

$$q_j = -\frac{E_j(i) - E(i-1)}{rmsd_j^{(rec-lig)}(i) - rmsd^{(rec-lig)}(i-1)} \tag{4}$$

The term $E_j(i)$ is the potential energy of the ligand-receptor-complex of child j in the actual generation i, whereas $E(i-1)$ defines the potential energy of the ligand-receptor-complex with smallest q of the previous generation

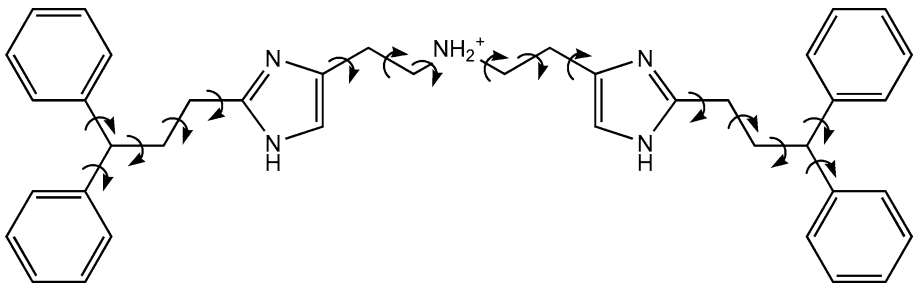

Fig. 2 Structure of dimeric histaprodifen. The arrows indicate the rotable bonds of dimeric histaprodifen

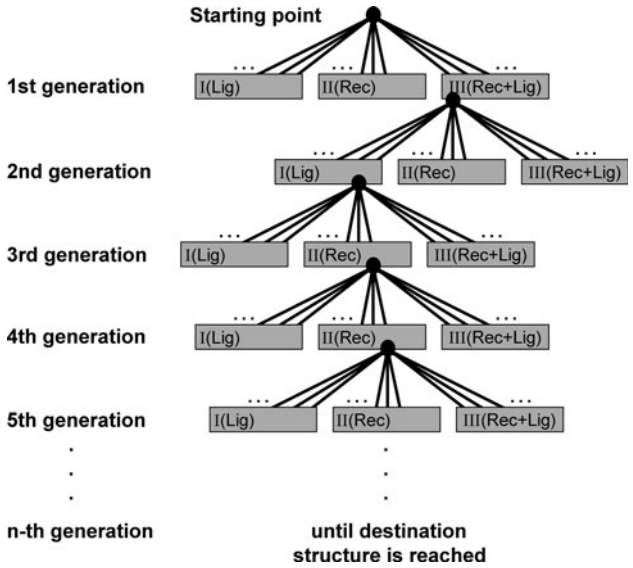

Fig. 3 Generation-child scheme of the *LigPath*-calculations. Based on a starting structure, new, minimized child structures are generated by the *LigPath*-algorithm. The children are divided into three groups I, II and III, as described under *Methods*. Each group contains n children. The best child of each generation is used as a starting structure for the next generation. The criterion for choice of the best child is described under *Methods*. The generation-child cycle is continued, until the destination structure is reached

$(i - 1)$. The child with smallest q_j was used as starting structure for the next generation.

The introduction of three child-groups I, II and III guarantees a non-restrained pathway calculation, because the ligand and receptor are allowed to reach their destination independent from each other. Thus, no distinct pathway from L + R → LR* is prejudiced (L + R is the free ligand and free inactive receptor, LR* is the active ligand-receptor-complex).

In the following a short summary of the parameters (Table 1) is given:

- *seed*: initialization of the random number generator
- *n*: number of children in each of the three groups I, II and III
- Δr^{rec}, Δr^{lig}: maximum stepsize for translation of a receptor or ligand atom
- $\Delta\phi^{rec}$, $\Delta\phi^{lig}$: maximum deviation from guiding line which is allowed for each atom; the guiding line is defined by the straight line between the coordinates of generation $(i - 1)$ and the destination coordinates; (i is the actual structure which has to be minimized)
- $\Delta\alpha$, $\Delta\beta$, $\Delta\gamma$: maximum angle for rotation of the ligand around x-, y- and z-axis; the ligand atom used as center of rotation is determined randomly in each step
- $\Delta\varphi^{rec}$, $\Delta\varphi^{lig}$: maximum angle for rotation of the rotable bond of the amino acid side chains and the ligand
- Δr_{coll}: interatomic distances below this limit are not allowed

All stepsizes and angles are determined randomly in such way that they do not exceed the maximum value allowed. A more detailed description of these parameters will be found elsewhere [28, 29].

Table 1 Initialization parameters for the different pathway calculations; the parameters n^{TM}, ΔTM_{xy} and $\Delta\Theta_z$ [28] were set to 0 because they were not used

	Systematic scan		Unconstrained pathway calculations			
	(1a), (2b)	(1b), (2a)	run 1	run 2	run 3	run 4
Seed	323496[a], 523[a], 8424[a]		323496	523	323496	523
n	1	1	2	2	2	2
Δr^{rec} [nm]	0.0	0.10	0.10	0.10	0.10	0.10
Δr^{lig} [nm]	0.05	0.0	0.05	0.05	0.10	0.10
$\Delta\phi^{rec}$ [°]	10	10	10	10	10	10
$\Delta\phi^{lig}$ [°]	20	20	20	20	20	20
$\Delta\alpha$, $\Delta\beta$, $\Delta\gamma$ [°]	2.5	2.5	2.5	2.5	2.5	2.5
$\Delta\varphi^{rec}$ [°]	10	10	5	5	10	10
$\Delta\varphi^{lig}$ [°]	10	10	5	5	10	10
Δr_{coll} [nm]	0.1	0.1	0.1	0.1	0.1	0.1

Parameters (more details are given in *A new version of the LigPath-algorithm* under *Methods*): *seed*: initialization of the random number generator; *n*: number of child structures in each group; Δr^{rec}, Δr^{lig}: maximum step of any receptor/ligand atom in direction to its destination position; $\Delta\phi^{rec}$, $\Delta\phi^{lig}$: the maximum deviation of the receptor/ligand atoms from the guinding line; $\Delta\alpha$, $\Delta\beta$, $\Delta\gamma$: maximum angle for rotation of the ligand around x-, y- and z-axis; $\Delta\varphi^{rec}$, $\Delta\varphi^{lig}$: maximum angle for rotation of the rotable bonds of the amino acid side chains and ligand; Δr_{coll}: interatomic distances below this limit are not allowed

[a] Each of the three values of the parameter *seed* was used once in each pathway calculations (1a), (1b), (2a) and (2b)

Calculation of potential energy surface and binding pathway

To calculate the ligand binding and receptor activation pathway L + R → LR* from the starting point L + R to the destination LR*, two different methods were used: 1) Systematic scan of the potential energy surface (Fig. 4a) and 2) a free, not restrained *LigPath* calculation (Fig. 4b).

Within the systematic scan of the potential energy surface, first the pathways (1a) and (2a) were calculated by adequate use of the *LigPath* parameters Δr^{lig} and Δr^{rec} (Table 1). The resulting structures of pathway calculations (1a) and (2a) were used as new starting structures for the subsequent pathway calculations (1b) and (2b) (Table 1). After that, the $rmsd^{(rec)}$-axis (0.0–0.45 nm) was divided into 9 intervals with a distinct length of 0.05 nm. The $rmsd^{(lig)}$-axis (0.0 nm to 2.5 nm) was divided into 10 intervals with a distinct length of 0.25 nm. Starting from the lattice points, located on (1a) or (2a), further pathway calculations, as indicated by the arrows (Fig. 4a) were performed. Each calculation (represented by an arrow, Fig. 4b) was performed three times, using different *seed*-values for initializing the random number generator. Thus, after complete calculation, three energy values were obtained for the lattice points, located on the border of the lattice (white boxes, Fig. 4a) and six values (black boxes, Fig. 4a) were obtained for the remaining lattice points. For subsequent construction of the potential energy surface, the mean potential energy was calculated at each lattice point. Based on these data, the potential energy surface was calculated with the software Maple 7 (Waterloo Maple Inc.), using the *surfdata*-command.

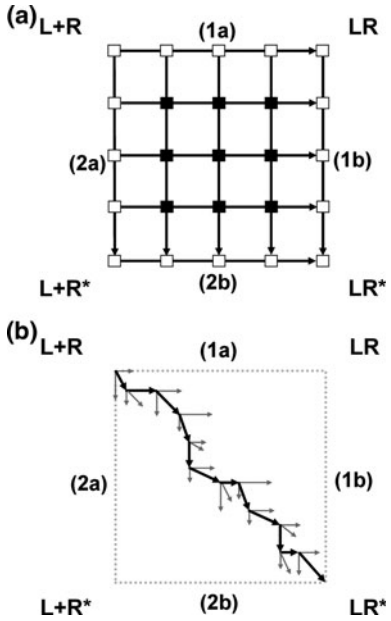

Fig. 4 **a** schematic presentation of the systematic surface scan; the *long black arrows* indicate the directions of pathway calculations; the *white* and *black boxes* present schematically the lattice points used for surface calculation. Three energy values were obtained for the *white boxes* on the border of the lattice, six energy values were obtained for the *black boxes*. **b** schematic presentation of a non-restrained *LigPath*-calculation; the three *arrows*, beginning in each case in the same origin represent the three groups of children in each generation; the *black arrows* represents the best child of each generation; the final point of each *black arrow* is the starting point for the next generation

The free, non-restrained *LigPath* calculation with L + R as starting and LR* as destination structure was performed four times, using the parameters, given in Table 1.

Starting and Destination Structure

The simulation box with the inactive gpH_1R (homology model based on the crystal structure of bovine rhodopsin (1F88) [40] was generated, as described [29]. This simulation box was modified by randomized putting of one positively charged dimeric histaprodifen into the extracellular part of the simulation box. The water molecules, clashing with the dimeric histaprodifen were removed. Additionally, one sodium ion was removed to conserve electroneutrality. Subsequently, a 2 ns MD simulation, using the simulation parameters, as described previously [29], was carried out. During the simulation, a high flexibility of the dimeric histaprodifen was observed. However, the MD simulation revealed snapshots with the dimeric histaprodifen located near to the extracellular surface of the receptor. One of these structures was energetically minimized and used as starting structure for the *LigPath* simulation.

The active dimeric histaprodifen—gpH_1R complex, embedded in a lipid bilayer and surrounded by extracellular and intracellular water, described previously [31] was energetically minimized and used as destination structure for the *LigPath* simulation. In a nutshell, a homology model of gpH_1R based on the crystal structure of the human adrenergic $h\beta_2R$ [41–43] was generated and embedded in a simulation box [31]. Afterwards, restrained MD simulations including experimental data of active GPCRs [44] were performed as described [29] in order to generate the active state model [31]. The active dimeric histaprodifen—gpH_1R complex was generated as described [31].

Summarizing, the starting structure corresponds to the dimeric histaprodifen on the extracellular part and the inactive gpH_1R: L + R (Fig. 1). The destination structure is represented by the dimeric histaprodifen in the binding pocket of the active gpH_1R: LR* (Fig. 1).

It should be noted, that there are only small differences between our inactive gpH_1R model, refined by MD simulations, and the crystal structure of $h\beta_2R$ [41–43], except with regard to the E2-loop. However, the difference in the

E2-loop is not astonishing, since there are two cysteine bridges in the hβ_2R, but only one in gpH$_1$R. The active state model, used within this study exhibits some differences with regard to the crystal structure of opsin in its G-protein-interacting conformation (3DQB) [45]. Only small differences were observed with regard to helix 3, but some larger differences with regard to the position of helix 5. In our model we have a similar outward movement of helix 6, as observed in 3DQB. However, we have decided not use a homology model based on the newer crystal structures, because of two reasons: The newer crystal structures are represented by adrenergic receptors or opsin. In contrast, we are modelling the H$_1$R. Because of the differences in amino acid sequences between H$_1$R and adrenergic receptors or opsin, we would expect differences in the structure of the receptor. Thus, we think it is just as well possible to use a model based on experimental data and additionally refined by MD simulations.

Parameters for calculation of different dimeric histaprodifen induced activation pathways at gpH$_1$R with the *LigPath*-algorithm

The aim of the study was the different ligand binding and receptor activation pathways by dimeric histaprodifen at gpH$_1$R. As indicated by Fig. 1, there may exist different ligand induced activation pathways of the gpH$_1$R: Pathway (1) is divided in the two parts (1a) and (1b) (Fig. 1): Firstly, the ligand penetrates into the binding pocket of the inactive receptor (1a) and secondly, the conformation of the receptor is changed from the inactive into the active state (1b). Pathway (2) is also divided into the two parts (2a) and (2b) (Fig. 1): There, the receptor is activated without ligand penetration (2a) and then the ligand penetrates into the binding pocket of the active receptor (2b). Pathways (1) and (2) describe extreme pathways for the process L + R $\rightleftharpoons$ LR*. As indicated by Fig. 1, other pathways (3) should also be taken into account. This third group of pathways would correspond to ligand binding accompanied by receptor activation. In our calculations we considered all three pathways. The different pathways were controlled by differences in the input parameters (Table 1). Each pathway calculation was performed three times with different *seed* parameters for initialization of the random number generator. Several calculations were performed with different input parameters given in Table 1. The rotable bonds of the amino acid side chains are defined, as given in [28]. The ffG53A6 force field was assigned to the protein [46]. All minimizations were performed with Gromacs 3.3.1 [47]. Within all minimizations, the whole simulation box, including lipid bilayer, intra- and extracellular water, was included. For the minimization steps during the *LigPath* calculation a maximum of 10000 iterations were allowed to reach a maximum force smaller than 10 kJ mol^{-1} nm^{-1} in the minimization. For energy minimization, the steepest decent method was used. For all minimizations, rectangular, periodic boundary conditions were used. The particle mesh Ewald (PME) method was applied to describe electrostatic interactions. The distances for coulomb cutoff and Lennard-Jones cutoff were set to 1.4 nm. The dielectric constant was set to 1. Thereby, all sites were allowed to be minimized, this means, no position constraints were set.

Results and Discussion

Energetical considerations

A mean potential surface FOR the binding process of dimeric histaprodifen and gpH$_1$R activation is shown in Fig. 5. Six corresponding structures are given in Fig. 6. The potential energy surface exhibits only small changes in potential energy for the process of ligand penetration from the extracellular part into the binding pocket of the inactive gpH$_1$R (corresponds to pathway (1a), Fig. 1). Furthermore, there are no significant differences in potential energy between L + R and LR. During the subsequent conformational change from LR to LR* (corresponds to pathway (1b), Fig. 1), a considerable energy barrier has to be crossed (Fig. 5). A comparison of potential energies of L + R and LR* shows that LR* has a significant smaller potential energy than L + R (Fig. 5). For pathway (2a) the potential energy surface exhibits a high energy barrier between the inactive and the active state of the receptor. Furthermore, the potential energy of the active gpH$_1$R without ligand in the binding pocket (L + R*) is energetically disfavoured, compared to the inactive gpH$_1$R (L + R). This suggests that the active conformation of gpH$_1$R is stabilized by dimeric histaprodifen, acting as partial agonist. This is in good accordance to the experimental result that the gpH$_1$R shows almost no constitutive activity [38]. The penetration of the dimeric histaprodifen into the activated gpH$_1$R (corresponds to pathway (2b), Fig. 1) leads to a continuous decrease in potential energy, until LR* is reached (Fig. 5). Thus, it can be summarized that there are high energy barriers in pathway (1), as well as in pathway (2) for L + R $\rightarrow$ LR*. In contrast, the mean minimum energy pathway indicates that the binding of dimeric histaprodifen, accompanied by receptor activation (pathway (3), Fig. 1) (Fig. 5) is the energetically favoured, compared to pathway (1) and (2). This result is supported by the non-restrained (non-directed) pathway calculation (grey, dotted area, Fig. 5). This means, that the course from L + R to LR* was calculated by *LigPath* without defining any restraints from the starting to destination structure.

Several non-restrained pathway calculations resulted in pathways, corresponding to pathway (3), but not in pathways, corresponding to pathways (1) or (2). The course of the pathway (3) is not linear from starting point to destination (grey, dotted area, Fig. 5). Distinct points on the energy surface (Fig. 5a–f) are correlated with the corresponding structures (Fig. 6a–f). Beginning from the starting point, in early steps, the $rmsd^{lig}$ of dimeric histaprodifen is reduced to about 2.1 nm, whereas only small changes are observed in receptor structure (Figs. 5a, b and 6b). This region corresponds to a local minimum on the potential energy surface. In the following steps, the ligand starts to penetrate significantly between the upper part of the transmembrane domains and only small changes of the receptor in direction to the active conformation are observed (Figs. 5c, 6c). However, compared to the starting structure (Fig. 6a) a straightening of TM VI was found. At a $rmsd^{lig}$ of about 1.5 nm subsequent penetration of the ligand between the transmembrane domains with strong structural changes of the receptor could be observed (Figs. 5d, 6d). However, in this region, the structural changes of the receptor are mainly found in the side chain conformations of the amino acids and not in the conformation of the transmembrane domains. During the following steps, the ligand shows further penetration into the binding pocket, whereas only small changes in receptor conformation were observed (Figs. 5e, 6e). In the phase from c to d (Fig. 5) there are, compared to the other areas of the potential energy surface, only small changes in potential energy. Between d and e (Fig. 5), a decrease in potential energy is observed. Within the remaining steps (e—destination, f), a significant decrease in potential energy is found. As the minimum energy path indicates, this decrease is related to the final steps of ligand binding and final changes in receptor conformation. Summarizing up, high potential energy barriers separate the inactive and active gpH_1R (Fig. 5). There is only a small corridor with low potential energy (Fig. 5, around d) between the inactive and active gpH_1R, which can be passed by dimeric histaprodifen during penetration from the extracellular part into the binding pocket, accompanied by receptor activation.

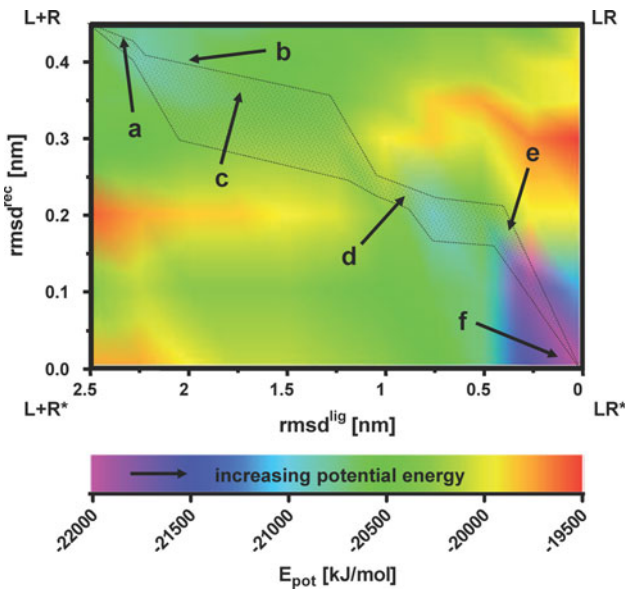

Fig. 5 Potential energy surface for penetration of dimeric histaprodifen into gpH_1R and activation of gpH_1R. In the *grey dotted* region the pathways without any restraints were found. Distinct regions of the energy surface are related according to their numbering scheme to the corresponding structures, shown in Fig. 6

A suggestion for a mechanistic model for ligand binding and activation of gpH_1R

In Fig. 7, the $rmsd^{lig}$ and $rmsd^{rec}$ with regard to the destination position is shown. Additionally, the angle α of the TM VI kink is given. All values represent one exemplary single pathway calculation (Table 1, run 1) and are given as function of the generation.

A strong decrease in $rmsd^{lig}$ from 2.5 to 1.0 nm within the first ~250 generations is found (Fig. 7). Within the same phase, the $rmsd^{rec}$ also shows a decrease from 0.45 nm to 0.25 nm (Fig. 7). Furthermore, a straightening of TM VI from 135° to 150° is revealed (Fig. 7). Thus, the straightening of TM VI, suggested to be important for receptor activation, takes place in an early phase of penetration of dimeric histaprodifen into gpH_1R, accompanied by strong structural changes within the receptor. From generations ~300 to ~1200, no changes in $rmsd^{lig}$ were observed. $rmsd^{rec}$ showed no changes from generation ~300 to ~800, but from ~800 to ~1200, a slight decrease was found. In the next phase, from ~1200 to ~1300, a strong decrease in $rmsd^{lig}$ as well as in $rmsd^{rec}$, but no further significant changes in straightening of TM VI was observed. In the last phase, from ~1300 to ~1950, only small changes in $rmsd^{lig}$, $rmsd^{rec}$ or α of TM VI were found. Generation ~1300 can be related with Fig. 5e. Thus, it can be suggested that from generations ~1300 to ~1900 only small changes in structure take place (Fig. 7) and the destination structure is achieved. However, a more detailed information about the changes within the receptor can be obtained by looking onto changes in the conformation of amino acid sided chains.

In Fig. 8, changes in dihedral angles of selected amino acid side chains $Phe^{6.55}$, $Phe^{6.52}$, $Tyr^{6.51}$, $Trp^{6.48}$ and $Phe^{6.44}$ are given. A detailed structural analysis of the binding pathway of dimeric histaprodifen reveals an aromatic channel, established by the amino acids Phe193 (E2-loop), Tyr194 (E2-loop), $Phe^{6.55}$ and $Tyr^{6.51}$. In early steps of ligand binding, an aromatic interaction between the

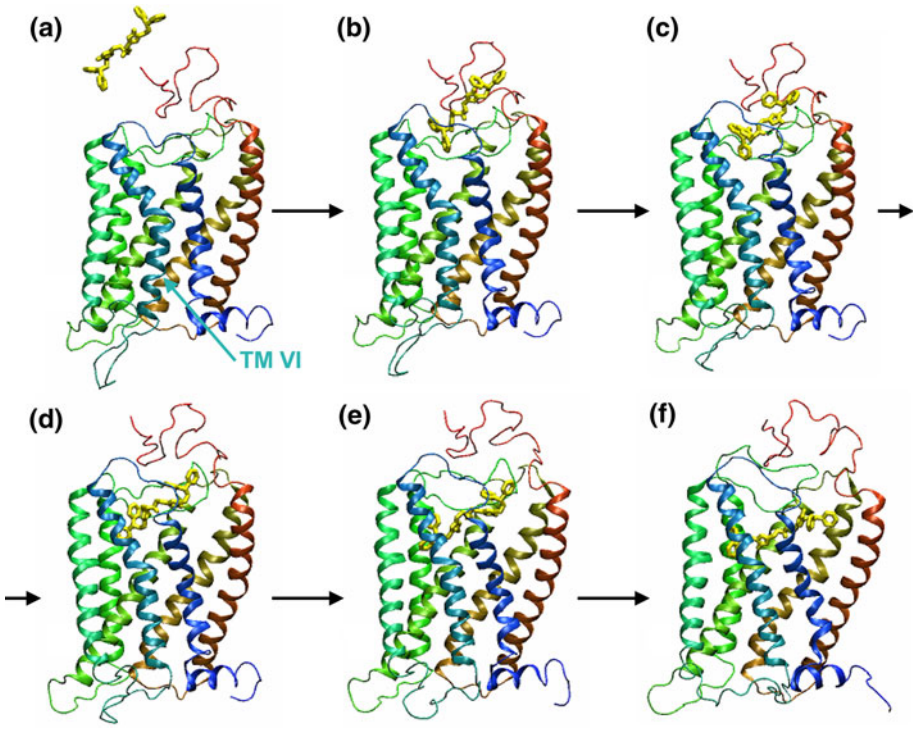

Fig. 6 Distinct structures for penetration of dimeric histaprodifen into the binding pocket of gpH_1R, accompanied by receptor activation. The ligand is shown in *yellow*. The structures are related according to their numbering scheme to distinct regions on the potential energy surface (Fig. 5)

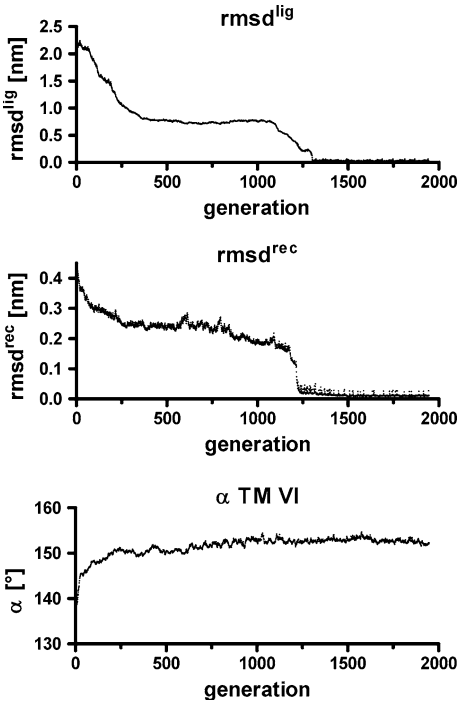

Fig. 7 Changes in $rmsd^{lig}$, $rmsd^{rec}$ and the angle α of TM VI for a representative, unconstrained *LigPath* run

diphenylpropyl moiety of the dimeric histaprodifen and Tyr194 is observed. The Tyr194 additionally establishes an aromatic interaction with Phe193. Further changes in the side chain conformation of Tyr194 and Phe193 open an aromatic channel on the receptor surface for the aromatic moiety of the ligand. The part of the dimeric histaprodifen, which is located in the binding pocket near TM V (Fig. 5f, left part) penetrates nearly complete into its part of the binding pocket within the first ~350 generations, enforcing several changes in dihedral angels of distinct amino acids (Fig. 8). The conformational changes in χ_1 and χ_2 of $Phe^{6.55}$, $Phe^{6.52}$ and $Tyr^{6.51}$ establish subsequently an aromatic channel for the diphenylpropyl moiety of dimeric histaprodifen. Site directed mutagenesis studies exhibited a strong influence of $Phe^{6.52}$ onto ligand binding [34, 35]. Within this phase, changes are also observed in χ_1 and χ_2 of $Trp^{6.48}$ and $Phe^{6.44}$ [29]. Strong changes in some dihedral angles take place at ~260 generations, maybe induced by sterical clashes. Here, χ_1 of $Phe^{6.55}$ switches from −90° to −180° and χ_2 from −120° to 90°. This switch is induced during penetration by the diphenylpropyl moiety of dimeric histaprodifen. A similar switch is observed for χ_1 and χ_2 of $Trp^{6.48}$. χ_1 switches from −40° to −180°, and χ_2 from −75° to −110°. This switch is correlated with the aromatic toggle switch, described for receptor activation [20, 48–50]. In Fig. 9, two representative structures immediate before and after the aromatic toggle switch are shown. These structures correspond to region d on the potential energy surface (Fig. 5). The penetration of the diphenylpropyl moiety of dimeric histaprodifen (Fig. 9(1)) leads to a clash with $Phe^{6.55}$. Consequently, $Phe^{6.55}$ undergoes a conformational change at generation ~260 (Figs. 8, 9(2)). Subsequently, conformational changes of $Tyr^{6.51}$, especially of χ_2 are found (Figs. 8, 9(3)). These changes would

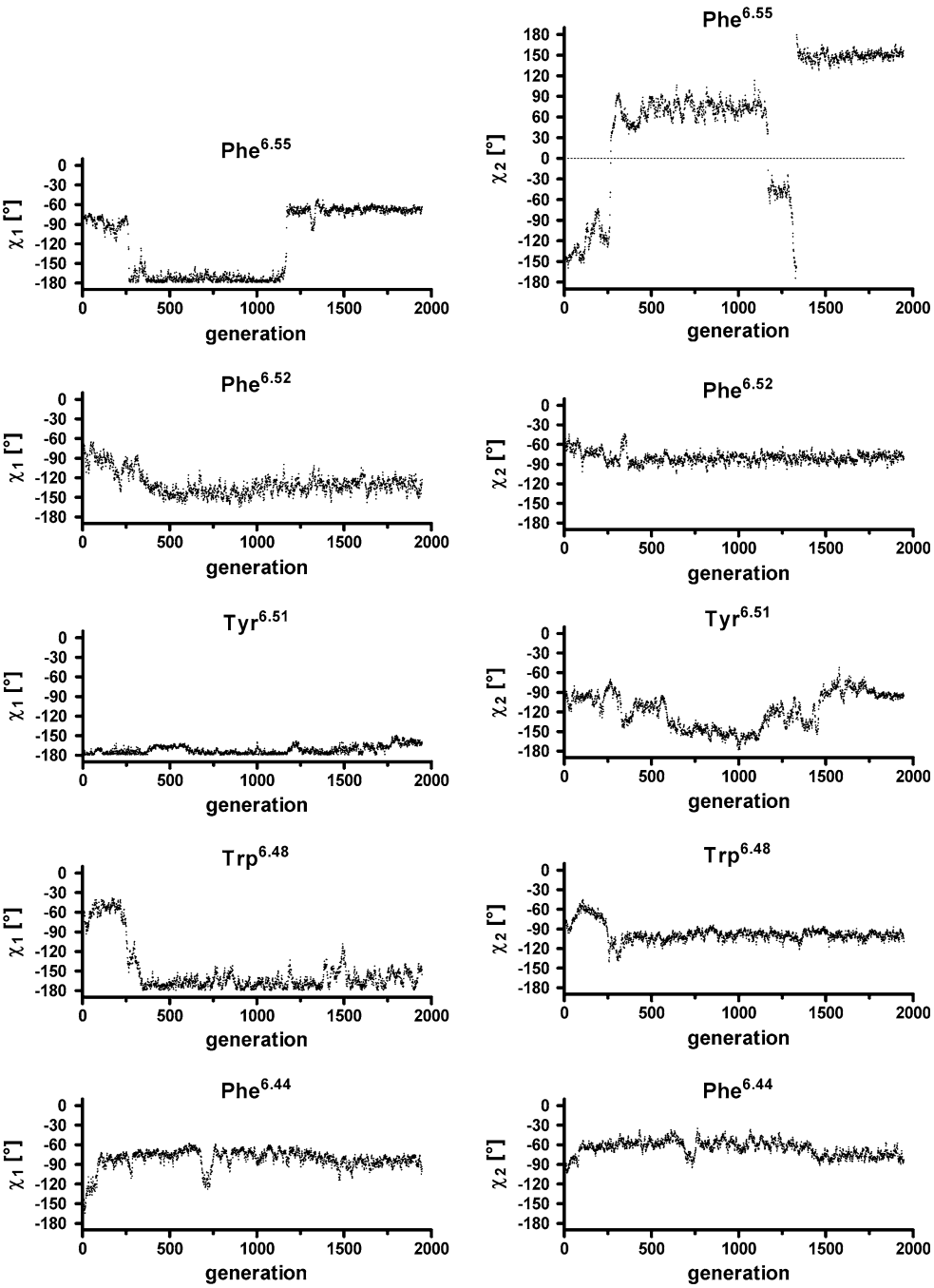

Fig. 8 Changes in dihedral angles χ_1 and χ_2 of distinct amino acid side chains

lead to a sterical clash with $Trp^{6.48}$ (Fig. 9). To avoid this clash, χ_1 and χ_2 of $Trp^{6.48}$ are observed to change up to generation ~325 (Figs. 8, 9(4)). Furthermore, $Phe^{6.52}$ undergoes a conformational change in χ_1 up to generation ~350 (Figs. 8, 9(5)). Within the generations ~325 to ~1100, no significant changes in dihedral angles were observed. At generation ~1100, strong changes in χ_1 and χ_2 of $Phe^{6.55}$ were observed, whereas smaller changes were observed for χ_2 of $Tyr^{6.51}$. These changes take place during the last phase of ligand penetration. The switch of χ_2 of $Phe^{6.55}$ leads at a χ_2 of ~90° to a nearly stack aromatic interaction between the phenyl moiety of $Phe^{6.55}$ and to the two corresponding phenyl moieties of dimeric histaprodifen. This switch is necessary to open the ligand a channel into the binding pocket. With a χ_2 of about −180° or 180° of the $Phe^{6.55}$, the phenyl moiety closes the channel into the binding pocket. The conformational change of χ_2 of $Tyr^{6.51}$ is correlated with the changes in χ_2 of $Phe^{6.55}$ since a switch of the $Tyr^{6.51}$ side chain is also needed to open the aromatic channel for the diphenylpropylmoiety of the

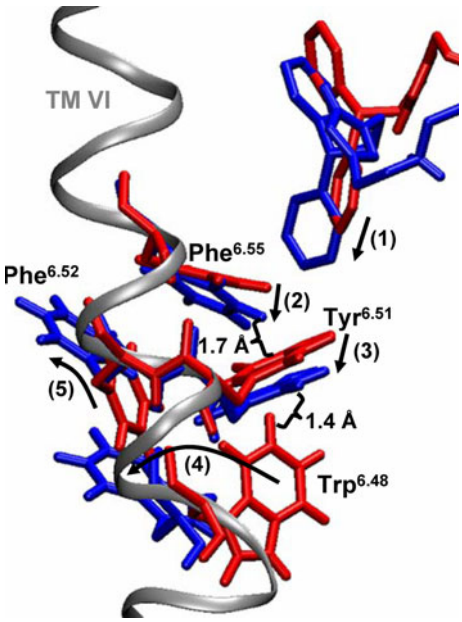

Fig. 9 Presentation of the rotamer toggle switch. Ligand and distinct amino acids some steps before (*red*) and some steps after (*blue*) the rotamer toggle switch

dimeric histaprodifen. In the last generations from ~1100 to ~1950, no further significant changes in the mentioned dihedral angles were observed.

Binding process of dimeric histaprodifen to gpH_1R and its activation

Batychara et al. [11, 12] used their software LITiCon to calculate several ligand-stabilized conformational states of bovine rhodopsin and the human β_2 adrenergic receptor. These studies give interesting and important insights into the ligand dependent stabilization of different receptor conformations. However, the calculations with LITiCon take only into account ligand-receptor complexes, with the ligand already being located in the binding pocket. In contrast, the *LigPath* algorithm, presented in this study, takes also into account the binding pathway of the ligand from the extracellular part into the binding pocket. Thus, *LigPath* is a suitable technique, which can be used to calculate the whole process of ligand binding and receptor activation.

The calculation of the potential energy surface for the binding process of dimeric histaprodifen to gpH_1R revealed that penetration of dimeric histaprodifen accompanied by successive activation of gpH_1R is the energetically preferred pathway. Up to now, there is no experimental technique available, to observe the structural changes during the binding process of a ligand completely on molecular level. Furthermore, the whole binding process of a partial agonist into the binding pocket and activation of the receptor could not yet be observed by in silico modelling techniques, like molecular dynamics. The reason for this is that sufficient simulation times to address this process are not yet realizable. Thus, the *LigPath*-algorithm is the first technique, which allows to get detailed insight into the process of ligand binding and receptor activation on molecular level. The modelling results, presented in this study may give new insight into the ligand binding process on molecular level. In future, experimental studies which may allow to observe the structural changes during the whole process of ligand binding have to be performed to proof the in silico results. However, until such experimental techniques will be available, the modelling studies will give hint to structural changes during the process of ligand binding and receptor activation.

Conclusion

Based on the results of this study, it can be proposed that the binding process of dimeric histaprodifen into the binding pocket of gpH_1R, accompanied by receptor activation is energetically favoured. The structural data of the pathway suggest that distinct amino acids, located within the penetration pathway of the ligand undergo conformational changes, opening a channel for the ligand from the extracellular part into the binding pocket. Furthermore, early steps of ligand penetration seem to induce a cascade of conformational changes of amino acid side chains in the pathway, resulting in a switch of the highly conserved $Trp^{6.48}$ with subsequent receptor activation. Thus, this study is one of the first studies, which presents a hypothesis about a ligand induced activation mechanism of GPCRs. However, in subsequent studies it has to be analyzed if there are ligand and receptor dependent differences in activation mechanism of GPCRs.

Acknowledgments We thank the DFG (Deutsche Forschungsgemeinschaft) for further support of this project.

References

1. Pierce KL, Premont RT, Lefkowitz RJ (2002) Nat Rev Mol Cell Biol 9:639–650
2. Kristiansen K (2004) Pharmacol Ther 103:21–80
3. Gether U, Kobilka BK (1998) J Biol Chem 273:17979–17982
4. Kobilka BK, Deupi X (2007) Trends Pharmacol Sci 28:397–406
5. Kobilka BK (2007) Biochim Biophys Acta 1768:794–807
6. Kobilka BK, Deupi X (2007) Trends Pharm Sci 28:397–406
7. Bartfai T, Benovic J, Bockaert J, Bond RA, Bouvier M, Christopoulos A, Civelli O, Devi LA, George SR, Inui A, Kobilka BK, Leurs R, Neubig R, Pin JP, Quirion R, Roques BP, Sakmar TP, Seifert R, Stenkam RE, Strange PG (2004) Nat Rev Drug Discov 3:577–626

8. Ghanouni P, Steenhuis JJ, Farrens DL, Kobilka BK (2001) Proc Natl Acad Sci USA 98:5997–6002
9. Xie SX, Ghorai P, Ye QZ, Buschauer A, Seifert R (2006) J Pharmacol Exp Ther 317:139–146
10. Wittmann HJ, Seifert R, Strasser A (2009) Mol Pharmacol 76:25–37
11. Bhattacharya S, Hall SE, Li H, Vaidehi N (2008) Biophys J 84:2027–2042
12. Bhattacharya S, Hall SE, Vaidehi N (2008) J Mol Biol 382:539–555
13. Schertler GF (2005) Curr Opin Struct Biol 15:408–415
14. Gabilondo AM, Cornelius K, Lohse MJ (1996) Eur J Pharmacol 307:243–250
15. Liapakis G, Ballesteros JA, Papchristou S, Chan WC, Chen X, Javitch JA (2000) J Biol chem 275:37779–37788
16. Chen S, Lin F, Xu M, Rie RP, Novotny J, Graham RM (2002) Biochemistry 41:6045–6053
17. Jongejan A, Bruysters M, Ballesteros JA, Haaksma E, Bakker RA, Pardo L, Leurs R (2005) Nat Chem Biol 1:98–103
18. Javitch JA, Fu D, Liapakis G, Chen J (1997) J Biol Chem 272:18546–18549
19. Rasmussen SGF, Jensen AD, Liapakis G, Ghanouni P, Javitch JA, Gether U (1999) Mol Pharmacol 56:175–184
20. Shi L, Liapakis G, Xu R, Guarnieri F, Ballesteros JA, Javitch JA (2002) J Biol Chem 277:40989–40996
21. Vilardaga JP, Steinmeyer R, Harms GS, Lohse MJ (2005) Nat Chem Biol 1:25–28
22. Farrens D, Altenbach C, Yan K, Hubbell W, Khorana HG (1996) Science 274:768–770
23. Ward SDC, Hamdan FF, Bloodworth LM, Wess J (2002) J Biol Chem 277:2247–2257
24. Dunham TD, Farrens DL (1999) J Biol Chem 274:1683–1690
25. Altenbach C, Klein-Seetharaman J, Cai K, Khorana HG, Hubbell WL (2001) Biochemistry 40:15493–15500
26. Shen J, Li W, Liu G, Tang Y, Jiang H (2009) J Phys Chem B 113:10436–10444
27. Wang T, Duan Y (2009) J Mol Biol 392:1102–1115
28. Straßer A, Wittmann HJ (2007) J Mol Model 13:209–218
29. Straßer A, Wittmann HJ (2007) J Comput Aided Mol Des 21:499–509
30. Bakker RA, Jongejan A, Sansuk K, Hacksell U, Timmerman H, Brann MR, Weiner DM, Pardo L, Leurs R (2008) Mol Pharmacol 73:94–103
31. Straßer A, Wittmann HJ, Seifert R (2008) J Pharmacol Exp Ther 326:783–791
32. Straßer A, Wittmann HJ, Kunze M, Elz S, Seifert R (2009) Mol Pharmacol 75:454–465
33. Bruysters M, Jongejan A, Gillard M, van de Manakker F, Bakker R, Chatelain P, Leurs R (2005) Mol Pharmacol 67:1045–1052
34. Jongejan A, Leurs R (2005) Arch Pharm Chem Life Sci 338:248–259
35. Bruysters M, Pertz HH, Teunissen A, Bakker RA, Gillard M, Chatelain P, Schunack W, Timmerman H, Leurs R (2004) 487:55–63
36. Menghin S, Pertz HH, Kramer K, Seifert R, Schunack W, Elz S (2003) J Med Chem 46:5458–5470
37. Seifert R, Wenzel Seifert K, Bürckstümmer T, Pertz HH, Schunack W, Dove S, Buschauer A, Elz S (2003) J Pharmacol Exp Ther 305:1104–1115
38. Straßer A, Striegl B, Wittmann HJ, Seifert R (2008) J Pharmacol Exp Ther 324:60–71
39. Ballesteros JA, Shi L, Javitch JA (2001) Mol Pharmacol 60:1–19
40. Palczewski K, Kumasaka T, Hori T, Behnke CA, Motoshima H, Fox BA, Le Trong I, Teller DC, Okada T, Stenkamp RE, Yamamoto M, Miyano M (2000) Science 289:739–745
41. Cherezov V, Rosenbaum DM, Hanson MA, Rasmussen SG, Thian FS, Kobilka TS, Choi HJ, Kuhn P, Weis WI, BK Kobilka, Stevens RC (2007) Science 318:1258–1265
42. Rasmussen SGF, Choi HJ, Rosenbaum DM, Kobilka TS, Thian FS, Edwards PC, Burghammer M, Ratnala VRP, Sanishvili R, Fischetti RF, Schertler GFX, Weis WI, Kobilka BR (2007) Nature 450:383–387
43. Rosenbaum DM, Cherezov V, Hanson MA, Rasmussen SGF, Thian FS, Kobilka BS, Choi HJ, Yao XJ, Weis WI, Stevens RC, Kobilka BK (2007) Science 318:1266–1273
44. Niv MY, Skrabanek L, Filizola M, Weinstein H (2006) J Comput Aided Mol Des 20:437–448
45. Scheerer P, Park JH, Hildebrand PW, Kim YJ, Krauss N, Choe HW, Hofmann KP, Ernst OP (2008) Nature 455:497–503
46. Oostenbrink C, Villa A, Mark AE, van Gunsteren WF (2004) J Comput Chem 25:1656–1676
47. van der Spool D, Lindahl E, Hess B, Groenhof G, Mark AE, Berendsen HJC (2005) J Comput Chem 26:1701–1718
48. Crocker E, Eilers M, Ahuja S, Hornak V, Hirshfeld A, Sheves M, Smith SO (2006) J Mol Biol 357:163–172
49. Singh R, Hurst DP, Barnett-Norris J, Lynch DL, Reggio PH, Guarnieri F (2002) J Peptide Res 60:357–370
50. Colson AO, Perlman JH, Jinsi-Parimoo A, Nussenzveig DR, Osman R, Gershengorn MC (1998) Mol Pharmacol 54:968–978