Basal Histamine H₄ Receptor Activation: Agonist Mimicry by the Diphenylalanine Motif

David Wifling,[b] Christopher Pfleger,[c] Jonas Kaindl,[a] Passainte Ibrahim,[a] Ralf C. Kling,[a] Armin Buschauer †,[b] Holger Gohlke †,*[c, d] and Timothy Clark †,*[a]

In memory of our friend and colleague Armin Buschauer

Abstract: Histamine H₄ receptor (H₄R) orthologues are G-protein-coupled receptors (GPCRs) that exhibit species-dependent basal activity. In contrast to the basally inactive mouse H₄R (mH₄R), human H₄R (hH₄R) shows a high degree of basal activity. We have performed long-timescale molecular dynamics simulations and rigidity analyses on wild-type hH₄R, the experimentally characterized hH₄R variants S₁₇₉M, F₁₆₉V, F₁₆₉V+S₁₇₉M, F₁₆₈A, and on mH₄R to investigate the molecular nature of the differential basal activity. H₄R variant-dependent differences between essential motifs of GPCR activation and structural stabilities correlate with experimentally determined basal activities and provide a molecular explanation for the differences in basal activation. Strikingly, during the MD simulations, F₁₆₉ dips into the orthosteric binding pocket only in the case of hH₄R, thus adopting the role of an agonist and contributing to the stabilization of the active state. The results shed new light on the molecular mechanism of basal H₄R activation that are of importance for other GPCRs.

Introduction

GPCR crystal structures,[¹] coupled with molecular dynamics (MD) simulations, site-directed mutagenesis, and NMR studies, have led to increased knowledge about the activation mechanism of GPCRs.[²] Agonists that bind to the orthosteric binding site located between transmembrane helices (TM) III, V, VI and VII[³] favor the active receptor state, eliciting the maximal possible response and inducing G-protein- and/or β-arrestin-mediated signaling.[⁴] Inverse agonists exert the opposite effect to agonists and favor the inactive receptor state. Neutral antagonists do not affect the basal activity. However, little is known, both on the pharmacological and molecular levels, about basal (constitutive) GPCR activation, which describes the activity of a GPCR in the absence of a ligand.[⁵] The existence of basal activity suggests that conformational transitions leading to the active receptor state may occur spontaneously. Such dynamic oscillations between active and inactive conformations have also been postulated as a mechanism for partial activation,[⁶] suggesting a mechanistic link between ligand-independent and -dependent GPCR activation. Alternatively, a partially active receptor may adopt a static conformation that is intermediate between active and inactive.[⁷] Furthermore, single-molecule studies have indicated that differential ligand dwell-times and/or influences on nucleotide exchange rates in the G-protein add further layers of complexity to GPCR activation.[⁸] Under-
we characterized the basal activity of a series of hH 4R variants, Chem. Eur. J. 2019. The histamine H 4 receptor (H 4R), a class A GPCR expressed mainly on immune cells, plays a fundamental role in processes such as cytokine release and chemotaxis. H 4R offers a unique opportunity towards understanding the mechanisms and pathways of GPCR activation in a ligand-independent manner. H 4R orthologues represent ideal naturally occurring GPCRs with different degrees of basal activity. In contrast to mouse H 4R (mH 4R), human H 4R (hH 4R) shows a high level of basal activity (Figure 1). In a previous molecular pharmacological study, we characterized the basal activity of a series of hH 4R variants, comprising hH 4R-S179M, hH 4R-F169V, hH 4R-F169V+S179M, and hH 4R-F168A (Figure 1). While the hH 4R-S179M variant exhibits a basal activity similar to that of hH 4R, replacing F169 with valine significantly decreased the basal activity. The basal activity of the hH 4R-S179M, hH 4R-F169V, hH 4R-F169V+S179M, and hH 4R-F168A variants in different regions shown in Figure 1 and, hence, provide a molecular explanation for basal activity.

| Site-directed mutagenesis | H 4R variant | FIA(%) | FIV(%) | SIM(%) |
|--------------------------|--------------|--------|--------|--------|
| H 4R                     | F168         | E169   | S173   |
| hH 4R-S179M              | F168         | E169   | M179   |
| hH 4R-F169V              | F168         | V169   | S173   |
| hH 4R-F169V+S179M        | F168         | V169   | M179   |
| hH 4R-F168A              | A168         | E169   | S173   |
| mH 4R                    | E170         | V171   | M181   |

**Results and Discussion**

**Dipping of F169 into the orthosteric binding pocket**

ECL2 has been proposed to be involved in ligand binding, selectivity, recognition, and to act as a gatekeeper for ligand entry. However, little is known about its contribution to basal GPCR activity. Among others, the M 2 muscarinic receptor (M 2R), the β 2-adrenergic receptor (β 2AR), the histamine H 3 receptor and hH 4R, which all show basal activity, share the same F 45.54–E 45.55 motif in ECL2 (F168–F169 for hH 4R). Furthermore, 15 other GPCRs feature analogous F–Y, Y–F or Y–Y motifs in place of F–F. Remarkably, our MD simulations of hH 4R revealed that F169 dips into the orthosteric binding pocket and mainly interacts with the surrounding hydrophobic and aromatic residues W90, L91, Y95, P166, W172, L175, L326, Y340, and F344 (Figure 2). The dipping of F169 was observed in four out of seven (approximately 60%) of our wild-type hH 4R simulations but not at all for other variants (simulations performed in singlet). Strikingly, F169 occupies the same area of the binding pocket as assumed for the imidazole moiety of histamine. It may thus act as a surrogate ligand in the highly basally active hH 4R. Interestingly, GPCR crystal structures show two different orientations of the diphenylalanine (F–Y) in case of H 4R motif: the two amino acids either point to opposite directions (M 2R) or the same directions compared with hH 4R simulations. Hence, though our simulations suggest the dipping of F169 into the binding pocket and the direction of F169 towards the cytoplasm, the two amino acids could potentially swap their orientations on longer timescales.

**Contraction of the orthosteric binding pocket**

In all MD simulations, the binding-pocket dimension between TM III and VII, represented by the distances between the C 8 carbon atoms of the aromatic amino acids (F, Y, W, and L) in the orthosteric binding pocket, decreased significantly (Figure 3). The reduction of the orthosteric binding pocket is particularly evident in the simulations with the H 4R variants containing mutations in either F168 or F169 and/or S179. Further amino acids found to be involved in basal H 4R activation are shown as ball and sticks. Amino acid numbering corresponds to hH 4R.
atoms of D94 (hH4R variants) or Q349 (mH4R), was initially approximately 11 Å, close to the corresponding distance in the inactive hH4R structure used for homology modelling. This distance increases by a maximum of approximately 3 Å for the basally inactive hH4R variants F169V, F169V+S179M and F168A, and for mH4R (Figure 3b and Table 1). In contrast, it remained nearly constant for the hH4R-S179M variant and decreased by 2 to 3 Å for wild-type hH4R, both of which exhibit high basal activity. This distance is approximately 11 Å in aminergic GPCR crystal structures (Table S1 in the Supporting Information), both with bound inverse agonist and antagonist ((10.9±0.2) Å; N=30) and bound agonist ((10.9±0.5) Å; N=20). Thus, in our simulations, TM III approaches TM VII more closely in apo-hH4R (≈8 Å) than previously observed in any ligand-bound GPCR or found for the other H4R variants. The decreased distance between TM III and TM VII enables D94, which is involved in ligand binding and receptor activation, to form additional interactions to residues in TM VII. The simulations of hH4R, its variants S179M (basally active) and F169V+S179M (weakly basally active) exhibited hydrogen bonds between D94, R341 and W348 in addition to a water-mediated contact to Q347 (Figures 3a, 4, and S1a and c in the Supporting Information). By contrast, in all variants with lower basal activity and mH4R, D94 is involved in a hydrogen-bond network with S68 (and with W90 in hH4R-F169V+S179M) (Figures 3d, 4, and S1b–d in the Supporting Information). A network comparable to the latter has been described for the crystal structures of M1 and M4 receptor-antagonist complexes, and may be an indicator of the inactive receptor conformation of H4R. In a further hy-

Figure 2. H4R variant-dependent locations and orientations of key amino acids for basal H4R activation: F168, F169, S/M179 for a) hH4R and the hH4R variants b) S179M, c) F169V, d) F169V+S179M, e) F168A, and f) F170, V171, M181 of mH4R. Wild-type and mutated residues of pre-aligned H4R variant structures (a–f), showing the respective predominating cluster, are illustrated as sticks and semi-transparent surface representation. To enhance comparability, the hH4R cartoon is highlighted in red (semi-transparent) for all H4R variants (a–f). Carbon atoms of aromatic and hydrophobic hH4R residues, capable of interacting with F169, are colored in gray (a).

Figure 3. Structural differences within the orthosteric binding pocket that relate to differences in basal activity. Insights into a) hH4R and d) mH4R binding pockets, represented by the corresponding predominant cluster (for cluster sizes see Table S2 in the Supporting Information). Residues involved in a hydrogen-bond network with D94 (a,d) or E182 (a) or E184 (d) in at least one H4R variant are highlighted in dark purple and green, respectively. W316 (a) or W318 (d), and the semi-transparent surface are shaded in red (a) or magenta (d). b) Time-evolution of distances between Cα atoms of D94 and Q347 (hH4R variants) or Q349 (mH4R), c) Time course of W316 or W318 (hH4R variants) or W318 (mH4R) θ2 rotational angles. For comparison, the corresponding distances of the inactive state hH4R (b) and the θ2 torsional angles of both the inactive state M1R and the active state β2AR (c) are illustrated with straight lines. Colors referring to the H4R variants in (b,c) are specified in the legend.
drogen-bonding network, all H₄R variants showed differential, highly variant-specific interactions between the key residues\(^{16,18}\) E182\(^\mathrm{i}\) (E184\(^\mathrm{j}\) in mH₄R) and Y95\(^\mathrm{j}\) or T178\(^\mathrm{k}\) (Figures 3a, d, 4, and S1 in the Supporting Information). While E182\(^\mathrm{i}\) established hydrogen-bond contacts to both Y95\(^\mathrm{j}\) and T178\(^\mathrm{k}\) in hH₄R and its variants S179M and F169V+S179M, it only displayed a single interaction to T178\(^\mathrm{k}\) in hH₄R-F169V, and Y95\(^\mathrm{j}\) in hH₄R-F168A and mH₄R. Consequently, the results emphasize that the key amino acids D\(^\mathrm{l}\) and E\(^\mathrm{k}\), which are usually involved in ligand-dependent H₄R activation, formed intra-receptor interactions in the absence of a ligand. Depending on the amino acid composition, H₄R variant-specific active receptor conformations are stabilized to a greater or lesser extent.

The rotamer toggle switch contributes to basal H₄R activation

The binding pocket connects to structures closer to the cytoplasm, so that interactions formed within the binding pocket affect distal rearrangements of the receptor. In this context, the proposed W\(^\mathrm{48}\) rotamer toggle switch has been suggested\(^{20}\) to drive the P\(^\mathrm{50}\) kink, and thus the outward movement of TM VI at the intracellular side. Analysis of rotameric W\(^\mathrm{48}\) states of aminicergic GPCR crystal structures (Table S1 in the Supporting Information) resulted in \(\chi_2\) torsional angles ranging from 29.7 to 113.7° for antagonist/inverse agonist-bound structures (\(94.7\pm 29.5\), \(N = 30\), from 80.1 to 86.1° for ligand-free structures (\(82.2\pm 2.8\), \(N = 3\) and from 97.3 to 124.4° for agonist-bound structures (\(112.4\pm 6.5\), \(N = 20\). In the hH₄R simulations, this angle (\(117.3\pm 13.6\)) was comparable to that of active-state aminicergic structures such as, for example, \(\beta_2\)AR\(^\mathrm{15a}\) (115.4°) (Table 1, Figure 3c). Simulations of both mH₄R and hH₄R-F169V gave values (21.5±26.4) and (69.7±14.7)° comparable to those of inactive-state aminicergic structures such as, for example, M₃ \(\mathrm{R}17\) (52.2°). The corresponding angles of the other H₄R variants ranged between these two values. These findings thus support the hypothesis that particularly apo-hH₄R features more highly populated active states than the other H₄R variants. Hence, although the rotamer toggle switch has been discussed controversially in literature\(^{20,21}\), our results suggest that it is involved in basal H₄R activation.

| Parameter | hH₄R | hH₄R-S179M | hH₄R-F169V | hH₄R-F169V+S179M | hH₄R-F168A | mH₄R |
|-----------|------|------------|------------|------------------|------------|--------|
| distance\(^\mathrm{b}\) (C \(_\mathrm{j}\)) D94\(^\mathrm{l}\)-Q347\(^\mathrm{q}\) | 9.5±0.7 | 10.7±0.5 | 10.8±0.8 | 11.2±1.2 | 12.1±0.5 | 12.7±0.8 |
| torsional angle\(^\mathrm{b}\) for \(\chi_2\) W316\(^\mathrm{d}\) | 117.3±13.6 | 89.5±14.9 | 69.7±14.7 | 85.3±10.7 | 94.5±11.4 | 71.5±26.4 |
| torsional angle\(^\mathrm{b}\) for \(\chi_2\) Y358\(^\mathrm{d}\) | 102.8±21.1 | 87.3±31.4 | 129.7±30.3 | 126.9±20.2 | 105.8±20.4 | 139.1±20.7 |
| distance\(^\mathrm{b}\) (C \(_\mathrm{j}\)) R112\(^\mathrm{d}\)-A299\(^\mathrm{m}\) | 10.1±0.6 | 9.1±0.6 | 9.4±0.4 | 9.3±0.5 | 9.5±0.4 | 8.5±0.5 |

\(^{16}\) Amino acid names and numbers correspond to hH₄R. \(^{18}\) inactive-state hH₄R \(\beta_2\)AR\(^\mathrm{15}\); 10.9, inactive-state \(\beta_2\)AR\(^\mathrm{15}\); 10.9, active-state \(\beta_2\)AR\(^\mathrm{15}\); 11.4 Å. \(^{16}\) inactive-state hH₁R\(^\mathrm{16}\): 10.9, inactive-state hH₁R\(^\mathrm{16}\); 10.9, active-state hH₁R\(^\mathrm{16}\); 11.4 Å. \(^{16}\) Inactive-state hH₁R\(^\mathrm{16}\): 128.7, inactive-state hH₁R\(^\mathrm{16}\); 138.4, active-state hH₁R\(^\mathrm{16}\); 86.3°. For torsion angle calculation, a periodicity of 180° was considered. \(^{15a}\) inactive-state hH₄R\(^\mathrm{15}\); 8.0, inactive-state hH₄R\(^\mathrm{15}\); 11.1, active-state hH₄R\(^\mathrm{15}\); 17.2 Å.

D\(^\mathrm{50}\), whose fundamental role for agonist-dependent GPCR activation and allosteric sodium ion binding has been demonstrated\(^{22}\), is located in the central core of the receptor. Hydrogen bonds between D\(^\mathrm{61}\) and S101 were more pronounced for both the highly basally active hH₄R and its hH₄R-S179M variant than for the other H₄R variants (see Figures 4, 5a, b, and S2 in the Supporting Information; for further interactions formed by D\(^\mathrm{50}\), please also see Figure S3a, Supporting Information). Consequently, the hydrogen-bond network of the
D2.50 cluster is expanded, and thus, TM III could undergo concerted helix movements with TM I, II and VII. The fact that active-state crystal structures, [15b, 17b, 23] unlike inactive [15a, 17a, 24] ones, usually show this hydrogen bond confirms its important role for basal GPCR activation. Y358 7.53, which is part of the NPxxY motif and lies adjacent to D2.50, acts as a further key player in GPCR activation. [20] It conducts the effect of the tyrosine toggle switch after the so-called hydrophobic barrier has been opened and thus, the water-mediated hydrogen bonding network is expanded towards the DRY motif. [20] Consistently with our interpretation of the simulations, Y358 7.53 (Y360 7.53 in mH4R) revealed overall \( \phi \psi \) dihedral angles comparable to those of the inactive [15a] and active [15b] state \( \beta_{2}AR \) (Table 1, Figure 5 c) in mH4R and dhH4R, respectively. The mean \( \phi \psi \) angles increased in the order hH4R-S179M < hH4R-F168A < hH4R-F169V+S179M < hH4R-F169V < hH4R-F168A < hH4R-F169V+S179M (Table 1). Consequently, in hH4R and its variants S179M and F168A, Y7.53 prevalently occupies a conformation that corresponds to a (basally) activated receptor state. In contrast, Y7.53 conformations of the hH4R variants F169V and F169V+S179M are comparable to that of mH4R and thus reflect an inactive receptor state. Driven by such conformational changes, Y7.53 only formed a prevalent hydrogen bond with D61 2.50 in mH4R (Figures 4, 5a, b and S2 in the Supporting Information). Water-mediated interactions, however, were present both in the other H4R variants (Figure 5a, Supporting Information) and identified in both inactive and active aminergic GPCR crystal structures. [24, 25] This result further supports the hypothesis that, of the H4R variants investigated, the inactive receptor state is most predominant in mH4R.

Outward movement of TM VI at the intracellular side

R3.50 has been proposed to play a key role in G-protein activation. [26] It forms the ionic lock with E6.30 that has been found in many GPCRs such as, for example, \( \beta_{2}AR \) or H1R, and was suggested to play a role in the stabilization of the inactive receptor state. [20] In contrast, H4R orthologues cannot form the ionic lock because E6.30 is replaced by non-acidic amino acids (hH4R: Ala, mH4R: Gly). [3a, 27] The simulations suggest that R112 3.50 forms more prevalent hydrogen-bond contacts with D111 3.49 in the H4R-S5.43 variants (hH4R, hH4R-F169V, hH4R-F168A) than H4R-M5.43 variants (hH4R-S179M, hH4R-F169V+S179M, mH4R) (Figures 4, 6a, b, and S4 in the Supporting Information). However, only the latter showed hydrogen bonds between R3.50 and S6.36 that connect TM III and TM VI; these may replace the ionic-
As a measure of GPCR activation, the substitution of S5.43 by Met leads to structural rearrangements that diminish the hydrogen-bond formation between D3.49 and R3.50 and simultaneously enhance contacts between R3.50 and E6.30 (see Figures 4, 6a, b, and S4 in the Supporting Information; for further interactions formed by D3.49, see Figure S3b, Supporting Information). This result is in agreement with site-directed mutagenesis studies, which show that the missing E6.30 alone is not the sole reason for the basal activity of hH4R.

Depending on the amino acid composition of the H4R variant, all aforementioned motifs in concert contribute more or less to helical rearrangements at the intracellular face. In particular, TM VI is bent around D9.32 as part of the FxxCxWxP2.04 motif, and thus forms a kink at this position. As a result, TM VI moves significantly outward upon receptor activation, enabling the binding of G-proteins into the gap thus formed, and downstream G-protein and/or β-arrestin-mediated signaling.

As a measure of GPCR activation, initial mean Cα–Cα distances of around 8.0 Å (as in the hH4R template) between R112.50 and A298.60 (G300.60 in mH4R) increased by a maximum of approximately 3 Å for WT-hH4R during the last quarter of the simulation (Figure 6c). In all other H4R variants the distance increase does not exceed 1.5 Å. This is consistent with the highest hH4R basal activity of all the H4R variants investigated, though not valid for hH4R-S179M with the second highest basal activity. Comparison of the R3.50–E6.30 distance of inactive (1.1 Å) with that of active state (1.7 Å) (see Table 1) revealed an outward movement of 6 Å during activation, suggesting that hH4R (∼3 Å) switched from an inactive to a partially active receptor conformation during MD simulation. Such an intermediate receptor conformation has also been suggested by Dror et al.

Mechanism of low-level activation

To both gain insights into the mechanism of low-level activation and to validate the reproducibility of our simulation results on basal hH4R activation, we performed six additional replica MD simulations of wild-type hH4R (Figure 7). During the last 500 ns of the in total seven hH4R simulations, the dipping of F169.55 into the binding pocket was observed in four simulations (1–4). Strikingly, the two of those four simulations that showed the most predominant dipping of F169.55 into the binding pocket (>70%) and either a binding pocket contraction (1) or only a subtle dilatation (2) also showed the most pronounced TM VI outward movement (>2.5 Å). By contrast, the other two simulations that either revealed lower dipping occupancies of F169.55 (4 or 5) or higher binding pocket dilata- tions (3) showed a markedly reduced outward movement of TM VI (<2.0 Å). The three remaining simulations that showed no dipping of F169.55 (5–7) resembled the H4R variants with a lower degree of basal activity (S179M, F169V, F169V+S179M or F168A) with respect to binding pocket contraction (<0.2 Å) and TM VI outward movement (0.9–1.3 Å). We thus observed two different activation states in the seven hH4R MD simulations, suggesting that hH4R switches between on (1, 2) and off (3–7) states in favor of adopting partially active conformations.

This is fully consistent with the concept of basal activity since basally active receptors transition spontaneously between inactive and active-like conformations in the absence of a ligand. Note, however, that although the dipping and TM VI outward movement were reproducible and predominant for...
hH4R, we cannot entirely exclude their occurrences for other variants on a statistical basis. Indeed, such structural changes would have been expected for the still highly basally active hH4R-S179M variant.

**Shift in structural stability occurs with changing basal activity**

To complement the analyses of differences in structural dynamics, we used Constraint Network Analysis (CNA) to compare the structural stability of WT-hH4R, its four variants S179M, F169V, F169V+S179M, F168A, and mH4R. In CNA, biomolecules are modeled as a constraint network, in which atoms are represented as nodes, and covalent and non-covalent binding interactions as constraints. The constraint network is efficiently decomposed into rigid clusters and connecting flexible hinge regions by applying rigidity theory. The inherent long-range aspect of rigidity percolation (i.e., whether a region is flexible or rigid depends on distant structural details) makes rigidity analysis an attractive tool for studying changes in structural stability due to distant influences. Biomolecules display a hierarchy of structural stability that reflects the modularity of their structure. To identify this hierarchy, a “constraint dilution trajectory” of network states, obtained from an initial network topology by successively removing non-covalent constraints, is analyzed (see Methods). Figure 8a shows a succession of five network states of hH4R and mH4R along such a constraint dilution trajectory, starting from the homology models (see Methods). In particular, regions known to be important in GPCR activation such as helices TM V and TM VI, become flexible earlier in hH4R than in mH4R. This qualitative analysis suggests that the lower structural stability of hH4R favors adaptations of these helices that allow the G-protein to bind and, hence, that hH4R can be basally active, in contrast to mH4R.

To obtain more detailed information about structural stability in hH4R, the four hH4R variants and mH4R, we applied CNA to conformational ensembles of the last 500 ns of the 2 μs MD trajectories. We then computed the difference of the residue-wise chemical potential energy [Eq. (1)], which reflects changes in biomolecular stability between the system being considered and hH4R (Figures 8b, S8, and S9 in the Supporting Information).

\[
\Delta E_{\text{CNA}} = E_{\text{CNA, hH4R}} - E_{\text{CNA, hH4R variants/mH4R}} \tag{1}
\]

\( \Delta E_{\text{CNA}} \) has been shown to be a robust local stability measure for predicting the thermal stability of proteins. From the \( \Delta E_{\text{CNA}} \) values, we identified regions in the receptors that correlate with the rank-ordering of the systems with respect to basal activity (Figure 8b): The structural stability of the systems increases with increasing levels of basal activity for the regions enclosing residues F54–P70 in helix TM II as well as W90–T92 and S109–D111 in helix TM III. In contrast, the structural stability of the systems decreases with increasing levels of basal activity for residues V184–A317 in TM V and as well as A306–G310–G310–A317 in TM VI. Notably, the analyses thus reveal focused, sequentially continuous but topologically distant residue clusters in the receptors in which changes in structural stability correlate with basal activity, rather than a broad distribution of stability changes across the structures.

Furthermore, the analyses show that, while regions critical for and known to move during GPCR activation (TM V and TM VI) become less structurally stable with increasing levels of basal activity, regions that are rather immobile during GPCR activation (TM II and TM III) become more structurally stable. Therefore, a local shift in structural stability within the GPCR occurs with changing basal activity, rather than a global, uni-
form change, so that overall structural stability is maintained in each case.

Interestingly, analyzing residues in TM VI adjacent to the G-protein binding site (L296–F312), which undergoes the largest structural changes during activation, in terms of the sum of their $\Delta E_{\text{CNA}}$ values yields a measure that significantly and very well correlates ($R^2 = 0.94, p = 0.001$) with the level of basal activity (cf. Figure 9; a significant and good correlation is also found if the entire, 2 µs-long trajectories are used for analysis (Figure S10 in the Supporting Information); likewise, a significant and very good ($R^2 = 0.94, p = 0.001$) correlation is found if the last 500 ns of six additional replica MD simulations for hH4R are evaluated (Figures 9a and S11 in the Supporting Information); both findings demonstrate the robustness of the CNA results: The hH4R-S179M variant shows only a minor stabilization of $-0.80 \text{ kcal mol}^{-1}$ compared to hH4R, which is in agreement with the very similar basal activities of the two systems. A stronger stabilization is found for the hH4R-F169V variant ($-9.48 \text{ kcal mol}^{-1}$), and the strongest effect ($-14.20, -20.94, \text{ and } -17.57 \text{ kcal mol}^{-1}$, respectively) for hH4R-F169V+S179M, hH4R-F168A, and mH4R. The findings are again in agreement with hH4R-F169V+S179M and hH4R-F168A showing a decrease in basal activity comparable to that of mH4R.19,10

**Predictive value of the parameters**

Table 2 summarizes further geometrical features observed in the simulations in relation to the basal activity of the six H4R variants. Although the binding pocket contraction is not as clear for the hH4R-S179M mutant as for the wild-type receptor, the D32–Q34 distance is quite predictive. Both direct and water-mediated hydrogen bonds between TM III and VII indicate activation, although they are also found for the inactive hH4R-F169V+S179M variant. The hydrogen bond between D943.32 and S68 2.57 or W90 3.28, on the other hand, is only found in the inactive variants and mH4R, making its absence indicative of activation. The remaining hydrogen-bond networks in the binding site and the W168 toggle switch are not diagnostic. More distally located motifs such as the D2.50–S3.39 hydrogen bond, the Y7.53 toggle switch, and the R3.50–A6.30 distance are quite predictive, but not the D2.50–Y7.53 hydrogen bond. As an alternative measure, structural stabilities of the lower part of TM VI (close to the G-protein binding site) can accurately predict basal activities (Figure 9b).

**Conclusion**

Prior molecular pharmacological (site-directed mutagenesis) studies have demonstrated the importance of the diphenylalanine (F168–F169) motif for basal H4R activation: whereas hH4R-F169V showed intermediate basal activity, the basal activities of both hH4R-F169V+S179M and hH4R-F168A were even comparable to that of the basally inactive mH4R. Therefore, the question arose, by which molecular mechanism the diphenylalanine motif contributes to basal H4R activation.

Our simulations unveiled the dipping of F169 into the orthosteric binding pocket and its conformational stabilization by the neighboring F168 as the trigger for high basal activity. The dipping was observed in approximately 60% of the hH4R simulations but not at all for other variants (Table 2). hH4R, therefore, appears to be unique among the receptor variants studied. Strikingly, our hH4R simulations showed an overlap of the pharmacophoric groups of F169 with those of the hH4R ligands such as histamine. F169 apparently resembles a ligand with an intrinsic activity ranging between that of a neutral antagonist and a full agonist. Depending on the ability of F169 to interact with the surrounding hydrophobic and aromatic amino acids and, thus, to contract the binding pocket, TM VI at the intracellular side moves outward to a greater or lesser extent (Table 2). In this respect, hH4R prefers switching between on and off states rather than stabilizing partially active conformations. Basal activity thus increases to

---

**Figure 9.** Correlation between structural stability of TM VI and experimental basal activity. a) The violin plot shows the distribution of structural stability in TM VI enclosing the residues L296–F312. The relative change in structural stability ($\Delta E_{\text{CNA}}$) for each species compared to hH4R is shown on the top. The statistical significance between pairs of mean $E_{\text{CNA}}$ values (dashed line in increasing order is **: $p < 0.01$, ***: $p < 0.001$, ****: $p < 0.005$). For hH4R, open circles depict the mean $\Delta E_{\text{CNA}}$ values of six additional replica MD simulations. b) The structural stability of TM VI correlates significantly and very well with experimental basal activity. Additionallly, the mean $\Delta E_{\text{CNA}}$ value over all seven hH4R replicas is shown, relative to the $\Delta E_{\text{CNA}}$ value of the one hH4R trajectory shown as violin plot in Figure 9a; error bars denote the SEM. The basal activity in percentage (0 (basally inactive) to 100% (basally active)) was transformed into the logit scale.
the extent that the basal equilibrium between inactive and active states is shifted towards the active state. An agonist binding to such a pre-activated receptor can subsequently only shift an intermediate to an active receptor state. In the absence of basal activity, by contrast, an agonist could shift an inactive to an active state.

Complementarily, results from rigidity analysis provide an excellent measure for describing the degree of basal activity in H4R. First, we observed focused, sequentially continuous but excellent measure for describing the degree of basal activity in H4R, its four variants, and mH4R on GPCR-activating motifs.

### Table 2. Summary of the impact of wild-type hH4R, its four variants, and mH4R on GPCR-activating motifs.

| Parameter                        | hH4R | S179M | F169V | F169V-S179M | F168A | mH4R |
|----------------------------------|------|-------|-------|-------------|-------|------|
| basal activity [%]               | 100  | 86    | 45    | 20          | 0     | 0    |
| binding pocket                   |      |       |       |             |       |      |
| dipping of F169 into the binding pocket | ✓    | ✓     | ✓     | ✓           | ✓     | ✓    |
| distance (C_{169}–D394 [%])      | 0    | 38    | 39    | 54          | 80    | 100  |
| TM III-VIII direct hydrogen bonds | ✓    | ✓     | ✓     | ✓           | ✓     | ✓    |
| TM III-VII water-mediated hydrogen bonds | ✓    | ✓     | ✓     | ✓           | ✓     | ✓    |
| hydrophobic bond D94–S68 [%]     | ✓    | ✓     | ✓     | ✓           | ✓     | ✓    |
| hydrophobic bond Y95–E182 [%]    | 97   | 24    | 0     | 79          | 100   | 64   |
| hydrophobic bond T178–E182 [%]   | 24   | 37    | 90    | 100         | 0     | 4    |
| torsional angle for W316 [%]     | 100  | 42    | 0     | 33          | 52    | 4    |
| central region                   |      |       |       |             |       |      |
| hydrophobic bond D61–S50 [%]     | 100  | 58    | 0     | 2           | 1     | 19   |
| hydrophobic bond D61–Y358 [%]    |      |       |       |             |       |      |
| torsional angle for Y358 [%]     | 30   | 0     | 82    | 77          | 36    | 100  |
| G-protein binding site           |      |       |       |             |       |      |
| structural stability L296–F312 [%]| 0    | 45    | 2     | 13          | 0     | 100  |
| structural stability L296–F312 [%]| 100  | 32    | 57    | 48          | 61    | 0    |

[a] Effects of H4R variants on GPCR activating motifs are either illustrated with ticks (applicable) and crosses (not applicable) or in a range between 0% (lowest value) and 100% (highest value). The percentage values were calculated from mean distances, torsional angles, or from occupancy over the simulation time values (hydrogen bonds) as outlined in the Computational methods section and Table S3 in the Supporting Information. [b] Amino acid naming and numbering correspond to hH4R. [c] Basal activities (hH4R = 100, mH4R = 0%) were calculated from experimentally determined intrinsic activities of thioperamide.

### Computational Methods

#### Homology modelling

A previous hH4R homology model,[3a] based on the inactive state hH4R (PDB ID: 3RZE) was used. It includes all extracellular (ECL) and intracellular (ICL) loops.[25] Models of the other H4R variants (hH4R variants S179M, F169V, F169V+S179M, F168A, and mH4R) were prepared with the modeling suite SYBYL-X 1.3 (Tripos Inc., St. Louis, MO USA) using the hH4R homology model as template. For the preparation of the mH4R model, the ECL2 loop upstream of C164–C165 was remodeled using the loop search module within SYBYL-X 1.3. All models showed the disulfide bond between C87 and C164, and C164 and C165 (hH4R variants) or C166 and C167 (mH4R). In place of ICL3, eight alanine residues were introduced. All residues were simulated in their dominant protonation state at pH 7. The N and C termini were positively (NH$_3^+$) and negatively (COO$^-$) charged.

#### Molecular dynamics simulations

Topologies and coordinates of the solvated (explicit water) and ionized (net charge of the entire system was zero) wild-type or mutant hH4R models, and the mH4R model, were prepared using the pdb2gmx module within GROMACS 4.5.[20] The systems were submitted to both steepest-descent and conjugated gradient energy minimization (without restraints) to remove bad van der Waals contacts of the amino acid side chains. The solvated GPCRs were then equilibrated in the NPT ensemble for 10 ns, applying harmonic restraints of 1000 kJ mol$^{-1}$ nm$^{-2}$ to protein main-chain atoms.
For membrane simulations, a pre-equilibrated dioleoylphosphatidylcholine (DOPC) lipid bilayer containing 72 DOPC lipids was used. This system was enlarged in the x and y directions, resulting in a lipid bilayer consisting of 253 DOPC molecules. This extended system was subjected to energy minimization (no restraints) and successively, equilibration runs were performed in the NVT (100 ps; weak harmonic restraints of 1000 kJ mol\(^{-1}\) nm\(^{-2}\) applied on DOPC atoms) and NPT (10 ns; no restraints) ensembles. Each receptor was inserted into this augmented, fully hydrated and equilibrated DOPC lipid bilayer using g_memb. The net charge of the simulation systems was neutralized by addition of 5 (hH4R variants) and 11 (mH4R) \(\text{Cl}^+\) ions. The systems additionally contained around 217 DOPC and 28,000 water molecules, accounting for a total of around 120,000 atoms in a box spanning approximately 90×91×143 Å. Consequently, these systems were energy minimized (without restraints) and equilibrated in the NPT ensemble for 10 ns (weak harmonic restraints of 1000 kJ mol\(^{-1}\) nm\(^{-2}\) applied to protein main-chain atoms). Production molecular-dynamics (MD) simulations (2 µs) of each H4R variant were started from the final frame of the equilibration runs. Data were collected every 100 ps. Although in other cases GPCR activation/deactivation typically demands much longer than 2 µs of simulation or enhanced sampling, we usually observe an induction period of approximately 500 ns with the AMBER force field, suggesting that 2 µs simulations are adequate for GPCR simulations in a membrane. The nature of basal activity might furthermore contribute to lower timescales needed for basal compared to ligand-dependent GPCR activation.

All minimization and MD simulation steps were carried out with GROMACS 4.5-4.6 (GROMACS 2016/2018 in the case of six replicate hH4R simulations). Initial velocities of all MD simulation runs were randomly assigned to the atoms. The SPC/E water model, the Amber ff99SB protein force field, and for DOPC molecules the GAFF force field were employed. A temperature of 310 K was maintained using a temperature bath and a time constant of 0.1 ps. Protein, DOPC, and water (including \(\text{Cl}^-\) ions) atoms were coupled separately. The Berendsen barostat with a time constant of 5 ps and a pressure of 100 bar was applied to maintain pressure of 1 bar and to ensure membrane properties in agreement with experiment. The volume compressibility was chosen to be \(4.5 \times 10^{-5} \text{bar}^{-1}\). The nonbonded cutoff was set to 10 Å and long-range electrostatic interactions were computed using the Particle Mesh Ewald (PME) method with an interpolation order of 4, an Ewald tolerance of \(1 \times 10^{-5}\) and FFT grid spacing of 1.2 Å. Periodic boundary conditions were applied in all directions. A cutoff of 14 Å was used for short-range van der Waals interactions. Bonds involving hydrogen atoms were constrained using LINCS, enabling a time step of 2 fs.

**Induced-fit docking**

“Flexible” docking of histamine to the hH4R model was essentially performed as described for muscarinic acetylcholine receptors in Pegoli et al. by using the Protein Preparation Wizard, and the LipPrep and Induced-fit docking modules (Schrödinger LLC, Portland, OR, USA).

**Data analysis**

Data were analyzed by using GROMACS 4.5 (g_cluster, g_density, g_rms, g_rmsf) and GROMACS 5 (gmx distance) analysis tools every nanosecond. The gromos clustering method was applied for cluster analysis of the entire (2 µs) trajectories, setting an RMSD cutoff of 2.5 Å. Hydrogen bonds and torsional angles were analyzed using the CPPTRAJ module of Amber 14 (University of California, San Francisco, CA, USA). Hydrogen bond and y plots (distance, torsion) were visualized by using the programing language R and the packages devEM, plot3d, Plotsix, Peptides and CircStats. For the purpose of comparison, distances and torsional angles of aminergic GPCRs (Table S1 in the Supporting Information) were calculated using the UCSF Chimera package, version 1.11.2, and are given as mean ± standard deviation. % values given in Table 2 were calculated as follows [Eq. (2)] from mean distances, torsional angles (Table 1), or occupancy over the simulation time values in the case of hydrogen bonds (Figure 4). For minimum and maximum values, please see Table S3 in the Supporting Information. Figure 7 was prepared with the Prism 5.01 software (GraphPad, San Diego, CA, USA), and all molecular figures with PyMOL Molecular Graphics system, version 1.8.2.1 (Schrödinger LLC, Portland, OR, USA).

\[
\% \text{ value} = \frac{\text{value} - \text{value}_{\text{min}}(\text{H}_4 \text{R variants})}{\text{value}_{\text{max}}(\text{H}_4 \text{R variants}) - \text{value}_{\text{min}}(\text{H}_4 \text{R variants})} \times 100
\]

(2)

**Rigidity analysis**

Rigidity analyses were performed using the CNA software package. CNA was applied on ensembles of network topologies generated from conformational ensembles. Each network of covalent and non-covalent (hydrogen bonds including salt bridges and hydrophobic tethers) interactions was constructed with the FIRST (Floppy Inclusions and Rigid Substructure Topography) software (version 6.2), to which CNA is a front- and back-end. The strengths of hydrogen bonds (including salt bridges) were assigned by the energy \(E_{\text{hb}}\) computed by FIRST. Hydrophobic interactions between carbon or sulfur atoms were taken into account if the distance between these atoms was less than the sum of their van der Waals radii (C: 1.7 Å, S: 1.8 Å) plus \(D_{\text{ion}} = 0.25 \text{ Å}\). In order to elucidate the hierarchy of structural stability in a biomolecule, a “constraint dilution trajectory” of network states \(\alpha\) was analyzed, which was generated by successively removing hydrogen bond constraints (including salt bridges) in the order of increasing strength. Thus, only those hydrogen bonds are retained in a network of state \(\alpha\) that have an energy \(E_{\text{hb}} \leq E_{\text{hb}}(\alpha)\). Altered biomolecular stability along a constraint dilution trajectory was quantified based on neighbor stability maps \(r_{ij,\text{neighbor}}\) with \(i, j\) being residue numbers [Eq. (3)]

\[
r_{ij} = \min(\{E_{\text{hb}}(\{x \in C^{\alpha-} : R_i < R_j < c\})\})
\]

(3)

Here, only short-range rigid contacts were considered that have at least one pair of heavy atoms of the residue pair \(R_{ij, \text{rho}}\) separated by a distance \(\leq 4.5 \text{ Å}\). A rigid contact \(r_c\) between pairs of residues ceases to exist when both residues stop sharing the same rigid cluster \(c\) of a set of rigid cluster \(C^{\alpha-}\) (Figure S7, Supporting Information). The double sum [Eq. (4)]

\[
E_{\text{CNA}} = \sum_{i=1}^{\alpha} \sum_{j=1}^{\alpha} r_{ij, \text{neighbor}}
\]

then represents the chemical potential energy \(E_{\text{CNA}}\) due to non-covalent bonding, obtained from the coarse-grained, residue-wise network representation of the underlying biomolecular structure. A per-residue decomposition of Equation (4), which is the chemical potential energy of residue \(i\) obtained by summation...
over all n short-range rigid contacts the residue is involved in (Figure S7 in the Supporting Information), is computed according to Equation (5):

\[ E_{CNA} = \frac{1}{3} \sum_{j=1}^{n} f_{j} \cdot \text{neighbor} \]  

(5)

Conformational ensembles used as input for CNA were extracted from MD trajectories of 2 μs length starting from hH4R, the four hH4R variants S179M, F169V, F169V+S179M, and F168A, as well as mH4R. The first 1.5 μs of each trajectory were considered equilibration phase of the system, and conformational ensembles for CNA were extracted from the last 500 ns. To estimate the uncertainty in the CNA computations, we split the last 500 ns of each trajectory into five individual sets of 100 ns length. Because the correlation time for decay of fluctuations of \( E_{CNA} \) (Eq. (4)) for all trajectories is about 0.1 ns, the five individual sets are statistically independent, and the standard error of the mean (SEM) is calculated from the standard deviation (SD) of the five mean \( E_{CNA} \) values [Eq. (5)] according to Equation (6):

\[ \text{SEM} = \frac{\text{SD}}{\sqrt{5}} \]  

(6)

To probe for a drift in the \( E_{CNA} \) values over the simulation time, we computed \( E_{CNA} \) for intervals of 100 ns along the whole trajectory of 2 μs length for residues in TM VI, lying next to the orthosteric binding site and the G-protein binding site; these values show no correlation with simulation time, indicating that there is no drift (Figure S12, Supporting Information).

To probe for the robustness of the \( E_{CNA} \) values for residues in TM VI, lying next to the orthosteric binding site and the G-protein binding site, we also analyzed entire, 2 μs-long MD trajectories (Figure S10 in the Supporting Information) as well as the last 500 ns of six additional replica MD simulations for hH4R (Figures 9 and S11 in the Supporting Information).

Statistical analysis

For each system, the average \( E_{CNA} \) values and the corresponding SEMs from Eq. 6 were calculated using NumPy[66]. According to the law of error propagation, the total SEM for differences in structural stability between wild-type hH4R and the four hH4R variants or mH4R was computed according to Equation (7):

\[ \text{SEM}_{\text{total}} = \sqrt{\text{SEM}_{\text{het,nt}}^2 + \text{SEM}_{\text{het,nt variant/mH4R}}^2} \]  

(7)

We used SciPy[66] for performing the following statistical analyses: (I) A two-sided Welch’s t-test to probe if \( \Delta E_{CNA} \) values are significantly different from zero (p < 0.05) and if independent samples of \( E_{CNA} \) values have significantly different average values; (II) Pearson’s and Spearman’s rank correlation between \( \Delta E_{CNA} \) values and (the rank of) the basal activity of each hH4R variant and mH4R. For computing the Pearson correlation, we summed \( \Delta E_{CNA} \) values for the residues in TM VI (L296–S312)[44] and then averaged over the last 500 ns for each species. The basal activity scale, ranging from 0% (inactive) to 100% (active), was transformed into the logit scale, with an adjustment of 2.5% for the lower and upper end of the scale. For computing the rank correlation, we calculated a running average over the \( \Delta E_{CNA} \) values with respect to \( i \), with a window size of ten and a step size of one.

Acknowledgements

This work was supported by the Deutsche Forschungsgemeinschaft as part of the Research Training Group GRK1910 “Medicinal Chemistry of Selective GPCR Ligands” and, in part, the Research Unit FOR2518 “Dynlon” (project P7, GO 1367/2-1). H.G. is grateful for computational support and infrastructure provided by the "Zentrum für Informations- und Medientechnologie" (ZIM) at the Heinrich Heine University Düsseldorf and the computing time provided by the John von Neumann Institute for Computing (NIC) to H.G. on the supercomputer JUWELS at Jülich Supercomputing Centre (JSC) (user ID: HKF7).

Conflict of interest

The authors declare no conflict of interest.

Keywords: basal activation · computational chemistry · GPCR · molecular dynamics · rigidity analysis

[1] a) J. A. Venkatakrishnan, X. Deupi, G. Lebon, C. G. Tate, G. F. Schertler, M. M. Babu, Nature 2013, 494, 185–194; b) K. Palczewski, T. Kumakura, T. Hori, C. A. Behnke, H. Motoshima, B. A. Fox, I. Le Trong, D. C. Teller, T. Okada, R. E. Stenkamp, M. Yamamoto, M. Miyano, Science 2000, 289, 739–745.
[2] a) B. K. Kobilka, Biochim. Biophys. Acta Biomembr. 2007, 1768, 794–807; b) R. O. Dörer, D. H. Arlow, P. Maragakis, T. J. Mildorf, A. C. Pan, H. Xu, D. W. Borhani, D. E. Shaw, Proc. Natl. Acad. Sci. USA 2011, 108, 18664–18669; c) R. Nygaard, Y. Zou, D. O. Dörer, T. J. Mildorf, D. H. Arlow, A. Manguik, A. C. Pan, C. W. Liu, J. J. Fung, M. S. Bokoch, F. S. Thian, T. S. Kobilka, D. E. Shaw, L. Mueller, R. S. Prosser, B. K. Kobilka, Cell 2013, 152, 532–542; d) E. C. Hulme, Trends Pharmacol. Sci. 2013, 34, 67–84.
[3] a) D. Wilfing, K. Loffel, Ü. Nordemann, A. Strasser, B. Berndt, S. Dove, R. Seifert, A. Buschauer, Br. J. Pharmacol. 2015, 172, 785–798; b) J. Nishimori, E. Coates, N. J. Murgolo, K. L. Morse, M. Bayne, C. D. Strader, F. J. Monsla, Jr., Mol. Pharmacol. 2002, 62, 38–47.
[4] T. Kenakin, FASEB J. 2001, 15, 598–611.
[5] R. J. Lefkowitz, S. Cotecchia, P. Samama, T. Costa, Trends Pharmacol. Sci. 1993, 14, 303–307.
[6] G. G. Gregorio, M. Masureel, D. Hilger, D. S. Tery, M. Juette, H. Zhao, Z. Zhou, J. M. Perez-Aguilar, M. Hauge, S. Mathiasen, J. A. Javitch, H. Wein- stein, B. K. Kobilka, S. C. Blanchard, Nature 2017, 547, 68–73.
[7] R. Lamichhane, J. J. Liu, G. Pijewalsicz, K. L. White, E. van der Schans, V. Katritch, R. C. Stevens, K. Wüthrich, D. P. Millar, Proc. Natl. Acad. Sci. USA 2015, 112, 14254–14259.
[8] E. Zampeli, E. Tiligada, Br. J. Pharmacol. 2009, 157, 24–33.
[9] D. Schnell, I. Brunskole, K. Ladova, E. H. Schneider, P. Igel, S. Dove, A. Buschauer, R. Seifert, Nouven-Schniedeberg’s Arch. Pharmacol. 2011, 382, 457–470.
[10] D. Wilfing, G. Berndt, S. Dove, A. Buschauer, PLoS One 2015, 10, e0117185.
[11] V. Isegb, C. de Graaf, A. Bortolato, V. Cherezov, V. Katritch, F. H. Marshall, S. Mordalski, J.-P. Pin, R. C. Stevens, G. Friend, D. E. Gloriam, Trends Pharmacol. Sci. 2015, 36, 22–31.
[12] a) M. Wheatley, D. Wootten, M. T. Conner, J. Simms, R. Kendrick, R. T. Logan, D. R. Poyner, J. Barwell, Br. J. Pharmacol. 2012, 165, 1688–1703; b) M. C. Peeters, G. J. van Westen, Q. Li, A. P. IJzerman, Trends Pharmacol. Sci. 2011, 32, 35–42.
[13] H. D. Lim, A. Jongejan, R. A. Bakker, E. Haaksma, I. J. de Esch, R. Leurs, J. Pharmacol. Exp. Ther. 2008, 327, 88–96.
[14] J. A. Ballesteros, H. Weinstein, Methods Neurosci. 1995, 25, 366–428.
[15] a) D. Wacker, G. Fenelli, M. A. Brown, V. Katritch, R. Abagyan, V. Chere- zov, R. C. Stevens, J. Am. Chem. Soc. 2010, 132, 11443–11445; b) S. G. Rasmussen, H. J. Choi, J. J. Fung, E. Pardon, P. Casarosa, P. S. Chae, B. T. Devree, D. M. Rosenbaum, F. S. Thian, T. S. Kobilka, A. Schnapp, I. Ko-
