JCTC Journal of Chemical Theory and Computation

Article

pubs.acs.org/JCTC

Evidence for a Transient Additional Ligand Binding Site in the TAS2R46 Bitter Taste Receptor

Massimo Sandal,[†] Maik Behrens,[‡] Anne Brockhoff,[‡] Francesco Musiani,[§,∥] Alejandro Giorgetti,*[,†,⊥] Paolo Carloni,*[,†,#,○] and Wolfgang Meyerhof[‡]

[†]Computational Biophysics, German Research School for Simulation Sciences, 52425 Jülich, Germany

[‡]Department of Molecular Genetics, German Institute of Human Nutrition Potsdam-Rehbruecke (DIfE), Arthur-Scheunert-Allee 114-116, 14558 Nuthetal, Germany

[§]Scuola Internazionale Superiore di Studi Avanzati (SISSA/ISAS), Via Bonomea 265, 34151 Trieste, Italy

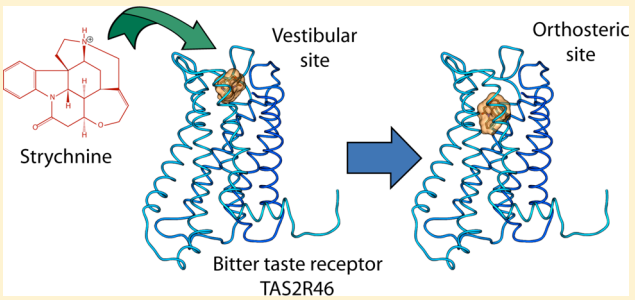
[∥]Laboratory of Bioinorganic Chemistry, Department of Pharmacy and Biotechnology (FaBIT), University of Bologna, Viale Giuseppe Fanin 40, I-40127, Bologna, Italy

[⊥]Department of Biotechnology, University of Verona, Ca' Vignal 1, Strada Le Grazie 15, 37134 Verona, Italy

[#]Computational Biomedicine, Institute for Advanced Simulation IAS-5, and [○]Computational Biomedicine, Institute of Neuroscience and Medicine INM-9,Forschungszentrum Jülich, 52425 Jülich, Germany

Supporting Information

ABSTRACT: Most human G protein coupled receptors (GPCRs) are activated by small molecules binding to their 7-transmembrane (7-TM) helix bundle. They belong to basally diverging branches: the 25 bitter taste 2 receptors and most members of the very large rhodopsin-like/class A GPCRs subfamily. Some members of the class A branch have been suggested to feature not only an orthosteric agonist-binding site but also a more extracellular or "vestibular" site, involved in the binding process. Here we use a hybrid molecular mechanics/coarse-grained (MM/CG) molecular dynamics approach on a widely studied bitter taste receptor (TAS2R46) receptor in complex with its agonist strychnine. Three ∼1 μs molecular simulation trajectories find two sites hosting the agonist, which together elucidate experimental data measured previously and in this work. This mechanism shares similarities with the one suggested for the evolutionarily distant class A GPCRs. It might be instrumental for the remarkably broad but specific spectrum of agonists of these chemosensory receptors.



INTRODUCTION

Human G protein coupled receptors (GPCRs) are the largest transmembrane receptor superfamily in the human genome, with about 800 known members.[1] They are of great pharmaceutical importance, being currently targeted by as many as 30% of marketed drugs.[2] Despite their extremely high sequence diversity, GPCRs share a typical 7-transmembrane (7-TM) helix bundle architecture. Most of them (ca. 70%) are activated by small (nonpeptidic) ligands binding inside their 7-TM bundle in the so-called orthosteric binding site. That the orthosteric site may act thermodynamically as a proper binding site is suggested by location of agonists in crystal structures[3−6] and by molecular dynamics studies.[7,8] This vast subset of GPCRs consists of two basally diverging branches of the family's phylogenetic tree: the 25 bitter taste 2 receptors (TAS2Rs[9,10]) and most (ca. 540 out of 701) rhodopsin-like/class A GPCRs.[1] Members of two (out of five) subfamilies of the latter branch (the muscarinic[8,11] and the β2-adrenergic (B_2AR),[7] receptors belonging to the α subfamily;[1] the opioid receptors[12] belonging to the γ subfamily[1]) have been suggested to feature also a "vestibular binding site", en route to the orthosteric site. The vestibular site, along with the extracellular loops (especially the extracellular loop 2) may be involved in ligand selectivity.[8,11−14] To our knowledge, no member of the TAS2Rs branch has been proposed to feature a vestibular site. Yet bioinformatics and molecular docking-based protein/ligand predictions guided by mutagenesis experiments[15−25] have suggested the presence of an "access control" within the extracellular opening of one of the most studied receptors of the family, the human TAS2R46.[17] Here we ask ourselves the question on whether this "access control" within the channel leading to the orthosteric site could make use of a vestibular site as found in class A GPCRs members.

When structural information on a protein is lacking and particularly when the sequence identity of a target protein with the template is low, approaches beyond simple homology modeling and rigid docking may greatly improve structural

Received: May 21, 2015

DOI: 10.1021/acs.jctc.5b00472
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

predictions of protein/ligand complexes. These include the combination of homology modeling with flexible docking techniques (the so-called ensemble docking approach)[26] or the use of molecular simulation.[16,27−30] Within the latter framework, we have here used an apt hybrid molecular mechanics/coarse grained (MM/CG) molecular dynamics approach, developed by us, on a homology model of the TAS2R46 receptor in complex with its potent agonist strychnine [submicromolar EC_{50} (0.43 ± 0.02 μM)].[31] This approach was shown to be useful particularly when the sequence identity is so low (like here, ∼11%) that the location of the side chains in the predicted model is likely to be wrong.

In fact, by limiting the atomistic representation to a limited region of interest and constraining the rest of the protein with a simplified Go model, we allow side chains in the region of interest to relax much more quickly than in an all-atom simulation. We have previously shown that our approach, while considering the MM region, does not introduce significant differences with all-atom simulations.[16,27] An all-atom simulation, on the other hand, might maintain the homology modeling poor prediction of side chains rotamers for a much longer time scale. MM/CG simulations have been repeatedly shown to be highly predictive for GPCR structural determinants starting from distant homology modeling structures.[15,16,32] Specifically, MM/CG-based structural ensembles turned out (i) to reproduce crystallographic binding poses of B_2AR ligands starting from a purposefully wrong initial pose, docked in a homology model of the receptor;[16,27] (ii) to be consistent with extensive experimental mutagenesis data on agonists binding to the bitter taste receptor TAS2R38,[16] which has a sequence identity with the most apt structural template (B_2AR) as low as 13%.[16] This is not dissimilar from that of TAS2R46 (11% with the template dopamine D3 receptor, PDB ID: 3PBL). In contrast with previous state-of-the-art structural predictions based on bioinformatics and molecular docking, the simulated structural ensemble described here turns out to be consistent with mutagenesis experiments performed here and previously.[17] Based on these encouraging results, we next predicted the effect of additional mutations, which were further confirmed by experiments performed here. Building on the overall consistency between experiment and simulations, we suggest that the access channel of the receptor does indeed exploit a site hosting the agonist prior it reaches its binding site inside the receptor. This site is remarkably similar to the vestibular site identified in members of other GPCR branches.

METHODS

In Vitro Mutagenesis of TAS2R46. Site-directed mutagenesis of TAS2R46 cDNAs was carried out by PCR-mediated recombination.[17,33] Briefly, a two-step PCR protocol was applied to generate two subfragments. The 5′ subfragments were amplified with the vector-specific CMV forward primer and the reverse mutagenesis primer, the corresponding 3′ subfragments were obtained using the complementary forward mutagenesis primer together with the vector-specific BGH reverse primer. PCR conditions were the following: 5 min, 95 °C; 15 cycles of 1 min annealing [annealing temperatures were calculated using the formula $T \approx T_{melt} - 3\ °C - (3\ °C \times$ mismatching bp)], 0.5−2.5 min 72 °C, 30 s 95 °C, followed by 5 min annealing, 10 min 72 °C. For the subsequent PCR-reaction, both purified subfragments were mixed and amplified using CMV forward primer and BGH reverse primer. PCR conditions were the following: 5 min, 95 °C; 15 cycles of 2 min, 54 °C, 3 min, 72 °C, 30 s, 95 °C, followed by 5 min, 54 °C, 10 min 72 °C. Finally, the constructs were cloned and sequenced. For subcloning the vector pcDNA5/FRT-sst3-MCS-hsv was used resulting in the addition of an N-terminal sst3-tag and a C-terminal hsv-tag.

Functional Expression. For functional expression analyses the receptor cDNAs were transfected into HEK 293T cells stably expressing the G protein chimera $G_{\alpha 16}gust_{44}$, incubated for 22 h, loaded with the calcium sensitive dye Fluo-4 AM, and washed with C1 solution (130 mM NaCl, 5 mM KCl, 10 mM Hepes, 2 mM $CaCl_2$, 10 mM glucose, pH 7.4). Test substances were diluted in C1 solution. Changes in cytosolic calcium levels were recorded in a fluorometric imaging plate reader (Molecular Devices). Data were collected from at least two independent experiments carried out in duplicates. Fluorescence signals were corrected for responses of mock-transfected cells and normalized to background fluorescence. Dose−response relations and EC_{50} values were calculated in SigmaPlot (SPSS) by nonlinear regression using the following equation:

$$y = \{(a - d)/[1 + (x/EC_{50})^{nH}] + d\} \quad (1)$$

where y is the observed value, a is the maximum observed value, d is the lowest observed value, and nH is the Hill coefficient, or the largest absolute value of the curve slope. Receptor expression was evaluated as in[6] and not significantly altered in the mutant cells.

Computations. The structure of human TAS2R46 (UniProtKB ID: P59540) was initially predicted by means of our GOMoDo server.[34] Human TAS2R sequences were aligned with PROMALS[35] to generate a hidden Markov model to guide the template/target alignment, using the HHSearch suite.[36] GOMoDo then generated, using MODELLER 9v10,[37] 100 models for selected structurally solved GPCR templates (PDB IDs: 1U19, 2LNL, 2R4E, 2RH1, 2Z73, 3EML, 3ODU, 3PBL, 3RZE, 3SN6, 3UON, 3V2Y, 3VW7, 4AMJ, 4DKL, 4EA3, 4EIY, 4GRV, 4IAR, 4JKV, 4L6R, 4LDE, 4N6H, and 4PXZ). Models were evaluated on the basis of: (i) MODELLER normalized DOPE and GA341 heuristic scores; (ii) structural parameters (Ramachandran plot quality, solvent exposed surface, etc.) obtained by the VADAR server;[38] (iii) consistency with experimental mutagenesis data. We only discarded a priori the rhodopsin X-ray structures (PDB IDs: 1U19 and 2Z73) since they feature a binding cavity structure poorly representative of soluble small ligands. Among the remaining templates, the one yielding the best models as identified by both quality scores and consistence with experimental data was the inactive state human dopamine D3 receptor (PDB ID: 3PBL,[39] sequence identity 11% with TAS2R46). The alignment was then further slightly refined by hand (final alignment in Supporting Information (SI) 1); new models were regenerated with MODELLER 9v10 and re-evaluated as above. Two rounds of MODELLER loop optimization were then performed with soft dihedral constrains on residues L71 and N150 to induce the side chains to point toward the binding site. The coordinates of the model are available in PDB format as Supporting Information file ct5b00472_si_002.pdb.

Strychnine has a pK_a of 8.25.[40] It was therefore considered protonated, with a total charge +1. The molecule was docked on the final protein model using HADDOCK 2.1.[41] The strychnine structure and its parameters for HADDOCK docking, including charges, were generated using PRODRG.[42] For molecular dynamics, RESP charges were derived using

DOI: 10.1021/acs.jctc.5b00472
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

structures optimized with HF/6-31G(d), obtained with Gaussian03, while bonded parameters were obtained using the PRODRG server, as described previously.[15,16,27] Cavity residues in the model were identified using FPOCKET[43] and used as both passive and active restrains for HADDOCK (see SI 2). We have previously shown that this approach can lead to reasonable initial predictions of ligand/receptor complexes based on homology models.[34] For each ligand, 1000 structures were generated for initial rigid docking, and the top scoring 200 complexes underwent further flexible minimization/refinement and then energy minimization in explicit water. Docked structures were then clustered according to root-mean-square deviation (RMSD).

The lowest (best) scoring structure of the most populated cluster was then used as a starting point for MM/CG simulations. Ligand parametrization and system preparation were performed with the implementation described in refs.[16,27,44] This is based on a customized version of GROMACS 4.5.[45]

The system was split in a MM part, which includes the GPCR agonist and the residues in the binding region (see SI 3) and a CG part, containing the protein frame, and in an interface region (I), defined between the MM and CG regions. The MM and the I regions were described by the GROMOS96 force field.[46] The CG part is described using a Go-like potential.[27,47] Residues at a distance less or equal than 6 Å from the MM boundary and not belonging to the MM part are belonging to the I region. Hydration at the active site was accounted by including a droplet of ca. 4300 SPC water molecules[48] around the MM region, above the membrane layer. The presence of the lipid bilayer was modeled by introducing a wall located at 2.0 Å from the proteins Cα atoms. Two planar walls coincide with the height of the lipid heads, while two hemispheric walls cap the extracellular and cytoplasmic ends of the protein. Within the two walls describing the bilayer (which is about 31 Å thick) another wall follows the initial boundary between protein and membrane. This wall has weakly attractive van der Waals-like interactions with Trp, Phe, and Tyr side chains, that function as "anchors" preventing the protein to exit the membrane. A 16 Å cutoff was used for electrostatics, van der Waals and Go-like interactions. The SHAKE algorithm was used to keep fixed the distance of bonds containing hydrogen(s).[49] Stochastic dynamics simulations were performed at a 300 K. The stochastic dynamics inverse friction constant was set at a value of 0.4 ps. The obtained system was used to run three 1 μs-long MM/CG simulations starting from the same initial geometry and differing only for the initial velocities.

Trajectory Analysis. The final trajectories of each replica were analyzed separately using the standard GROMACS 4.5 tools (*g_rms, g_clust, g_mindist*). For ligand RMSD analysis, all the protein backbone was realigned for each frame, and then the ligand conformations were clustered using the GROMOS clustering algorithm,[50] with a 1 Å RSMD cutoff. Analysis of ligand–receptor contacts was performed using UCSF Chimera[51] and GROMACS 4.5 tools. To analyze the main clusters, the top two representative structures of each trajectory were then automatically analyzed to count ligand–protein and ligand–solvent contacts and hydrogen bonds. Contacts were defined as atom–atom couples whose van der Waals-radius spheres are closer than 0.4 Å. Hydrogen bonds are defined by the criteria described in ref 52 relaxed by a 0.75 Å distance tolerance and a 45° angle tolerance, to allow for molecular motions around the central cluster structure. This analysis was further refined by checking the persistence of contacts during the trajectory. Residues which contact the agonist more than 75% of the time after equilibration are considered as stable contacts even if they do not appear as such in the single isolated cluster structure. The contact definition includes the residue backbone atoms. Water occupancy analysis was performed using UCSF Chimera, with a grid of 1 Å^3 voxels, analyzing all waters up to 5 Å from MM atoms.[51]

Bioinformatics Sequence Analysis. To have a consistent mapping between class A receptors and TAS2Rs, which belong to a separate branch of the GPCR superfamily, we use the GPCRDB generic number positions[53] which in turn generalized the classical Ballesteros–Weinstein numbering.[54] Specifically, we map the GPCRDB generic numbering from our template, the human dopamine D3 receptor,[39] to the TAS2R46 homology model. See Table S1 and the snake plot in Figure S1 for details of the TAS2R46 numbering scheme and the TAS2R46 topology according to our model. Generic GPCR numbers are indicated throughout the text in a superscript after the residue sequence number, e.g. T180$^{5.43}$.

For bioinformatics conservation analysis, sequences of human GPCR families (TAS2Rs, opioid, muscarinic, β-adrenergic) were retrieved from UniProtKB (http://www.uniprot.org) and aligned using PROMALS.[35] The alignment was then mapped on the TAS2R46 model or on a reference PDB structure by means of UCSF Chimera:[51] (i) muscarinic receptors: human M2 muscarinic acetylcholine receptor crystal structure (PDB ID: 3UON);[55] (ii) beta-adrenergic receptors: human B_2AR receptor crystal structure (PDB ID: 3PDS);[5] (iii) dopaminergic receptors: human dopamine D3 receptor crystal structure (PDB ID: 3PBL).[39] To define conservation going from the extracellular to intracellular side of the 7-TM helix bundle, a "water level" was defined in the intracellular half of the receptors, corresponding to generic GPCR numbering positions: 1.50, 2.47, 3.42, 4.49, 5.54, 6.40, and 7.49. Residues in each helix were then indexed according to their relative position going from extracellular to intracellular in the following fashion:

$$\text{Towards extracellular loops} \leftarrow \ldots -5, -4, -3, -2, -1, 0,$$
$$+1, +2, +3, +4, +5 \ldots \rightarrow \text{Towards intracellular loops}$$

Following this index, residues in all helices ≤ -12 were considered as in the position of a putative vestibular site (V group) and residues between -12 and -4 were considered to belong to the orthosteric site (O group).

We then performed PROMALS alignments of the following human GPCR subfamilies: TAS2Rs, muscarinic receptors, beta-adrenergic receptors, and dopaminergic receptors. The conservation of each residue was calculated as the percentage of the most prevalent residue type at the corresponding position in the alignment. We then generated the frequency histogram of the conservation values for residues in the V group and O group.

RESULTS AND DISCUSSION

Receptor Architecture. We used state-of-the-art bioinformatics and advanced flexible docking protocols to construct an initial model of the agonist/receptor complex (see Methods). This initial complex then underwent several independent, microsecond-long MM/CG simulations differing only for the initial velocities (Figure S2 in the SI). During the simulations,

DOI: 10.1021/acs.jctc.5b00472
J. Chem. Theory Comput. XXXX, XXX, XXX–XXX

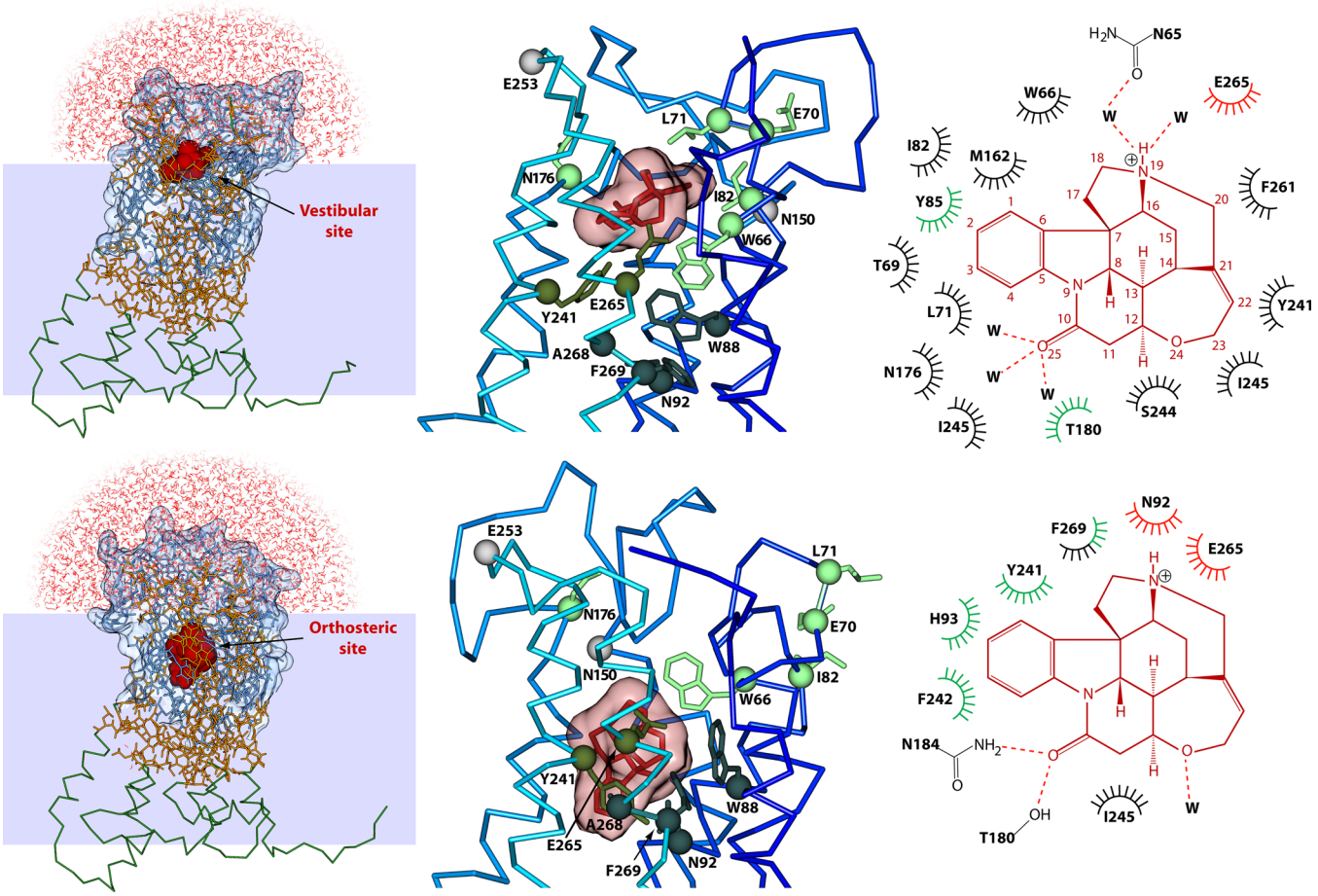


Figure 1. MM/CG calculations on TAS2R46. The top row shows the pose in the vestibular site, the bottom row shows that in the orthosteric site. (left) Scheme of the MM/CG setup. The MM part consists of strychnine (dark red spheres), protein residues close to the extracellular side (see SI 3), and water molecules (lines colored accordingly to the atom type). Only a few of the latter are represented for the sake of clarity. The CG regions include the Cα (green tubes). Residues joining the MM and CG regions are represented by orange sticks. (center) Close up of the binding poses. Strychnine is represented by dark red sticks and by its molecular surface. The protein backbone is colored from dark blue in the proximity of the N-terminus to light blue at the C-terminus. α carbons of residues studied by mutagenesis are represented as light green, olive green, and slate gray when the corresponding residues interact with the agonists in the vestibular site, in both sites, and in the orthosteric site, respectively. Only two residues previously identified as important for TAS2R46 activation by strychnine ($N150^{ECL2}$ and $E253^{ECL3}$, gray spheres; see also Tables S3 and S4) do not interact with the agonists in the 1 μs simulations: it is possible that their role becomes apparent at larger time scales because of the high flexibility of the ECL2 and ECL3 loops where these residues are located or because these residues influence the overall folding of the extracellular loops and therefore rather exert indirect effects on strychnine interaction. (right) Schematic representation of strychnine binding poses (*pose* 1 and *pose* 2 in the main text). Schemes represent H-bonds (red dashed lines), electrostatic (red spiked arcs), hydrophobic (black spiked arcs), π-stacking (green spiked arcs) agonist/receptor interactions.

the ligands quickly abandoned the docked pose. This suggests that the starting pose obtained by flexible docking, while plausible when visually inspected, might not be correct. Two conformational ensembles corresponding to two distinct strychnine poses are identified already after 0.1 μs (Figure 1). This relatively rapid convergence, which we previously observed,[16,27] may be due, at least in part, to the fact that our system has far fewer degrees of freedom than the correspondent one using an all-atom representation. Hence we expect that convergence is reached in a shorter time scale than in all-atom simulation.

Both poses, once settled, are stable on the microsecond time scale of the simulation (see Figure S2). They occupy distinct regions in the receptor: Taking as a reference the central atom 8 of strychnine (Figure 1, right panels), the locations of the two binding sites are about 7 Å distant, along a vector almost parallel to the *Z*-axis.

Pose 1 assumes a conformation roughly parallel to the membrane slab. The agonist is here solvated, contacting up to more than a dozen water molecules (Figure 1 and Table S2). The site hosting the agonist is relatively close to the extracellular region; by analogy we name it vestibular site as those found in Class A GPCRs[12] [We also recovered, in another trajectory, a similar pose in the same vestibular site, which is however less compatible with experimental data. We discuss it in SI 9.].

Pose 2 shares similarities to the binding pose from the bioinformatics models reported previously,[17] but also it features several additional residues involved in agonist binding (Figure 1, right panel) along with about four water molecules (see Table S2). The perpendicular orientation, the deeper location, and the presence of a lower solvation, contrast with the parallel orientation and extensive hydration of the ligand in the vestibular site. The region in which the agonist is accommodated basically overlaps with the orthosteric site

DOI: 10.1021/acs.jctc.5b00472
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

across Class A GPCR/small agonist complex X-ray structures (see Figure S3) [Structural information on a variety of agonist/GPCR complexes shows small-molecule agonists bound to the orthosteric binding site. These include the purinergic receptor $P2Y_{12}$/2-MeSADP,[3] serotonin receptor 5-HT_1/ergotamine,[4] B_2AR/FAUC50,[5] and adenosine A_{2A} receptor/adenosine.[6]].

Water in the Access Channel. When the agonist occupies the vestibular site in pose 1, a "water wire" of 5−6 water molecules within the TAS2R46 extracellular access channel connects the agonist, the vestibular, and the orthosteric site, contacting TM 3, 4, 5, 6, and 7 (see Figure S4 A). A similar feature has been observed in several Class A GPCRs X-ray structures[56] and dubbed the "extracellular water cluster".[57] These water molecules occupy the receptor for more than 75% of the time-scale investigated, exchanging places between them every ∼100−300 ns. They interact with several residues identified here below and in ref 17 as important for binding ($Y85^{3.29}$, $H93^{3.37}$, $T180^{5.43}$, $Y241^{6.51}$, and $Y271^{7.45}$) (superscript here and following are residue numbering according to the generic GPCR numbering system; see Methods). Another hydration site is located between TM helix 2 and extra cellular loop 1 (ECL), around residue $T69^{2.64}$. Waters there exchange on a time scale of ca. 10 ns. Several positions contacted by waters in in our model ($A89^{3.33}$, $H93^{3.37}$, $Y241^{6.51}$, and $Y271^{7.45}$) correspond to those contacted by structural waters in known GPCR crystallographic structures such as rhodopsin, the β-adrenergic receptor, and the A2A adenosin receptor.[56,58]

When the agonist occupies the orthosteric site in pose 2, about five water molecules occupy the extracellular opening at any moment but persisting for 40% of the time scale investigated or less. The most persistent water occupancy is located between $A89^{3.33}$ and $T180^{5.43}$ (see Figure S4 B); at any moment 1−2 water molecules, exchanging with the bulk solvent on a time scale <10 ns, also contact $W66^{2.61}$. Residues in the orthosteric site involved in the "water wire" described above are mostly dehydrated here. Hence, dehydration seems a key step for agonist binding in the orthosteric site.

Comparison with Bioinformatics Predictions. Both poses differ distinctly from bioinformatics/docking-based predictions performed here and in ref 17. In particular, in the initial configuration obtained here using state-of-the-art bioinformatics combined with standard docking, before MM/CG, the agonist was able to contact only a few of the residues which experimentally play a key role for agonist binding, namely $L71^{ECL1}$, $N92^{3.36}$, $N176^{5.39}$, $Y241^{6.51}$, and $E265^{7.39}$. This might be due, at least in part, to the fact that bioinformatics models based on templates with low sequence identity are deep in the "twilight zone" of homology modeling. This challenges the accuracy of the results obtained by modeling and docking alone, and point to the requirement of molecular dynamics.[59,60] Ligand−receptor contact patterns in both poses, as emerging from present calculations, are presented in Table S3.

Validation of Simulations. We now compare our simulations with reported experiments measuring the impact of mutations on the intracellular calcium fluorescence imaging.[17] Correlation of the agonist concentration with intracellular calcium release is here used to estimate the agonist efficacy for activation as well as potency. Efficacy is measured by the saturation (maximal) receptor response, while potency is measured by the half-maximal effective concentration (EC_{50}) of agonist (see SI 11 for details). The mutated residues (Figure S6 and Table S4) were selected as important for TAS2R46 specificity by comparison to structurally highly related receptors with little or no agonist overlap.[17]

Mutations in the Vestibular Site. The experimentally investigated mutation $L71^{ECL1}F$ on the ECL1 loop (Figure 1, center) can destabilize strychnine binding by reducing the volume of the vestibular space. This provides an explanation for the experimentally observed increase of the EC_{50}-concentration from the wild type EC_{50} of 0.43 ± 0.02 μM, without affecting maximal receptor activation (Figure 2). Among the mutations experimentally studied in this and previous works (see Figure 2, Figure S6, and Table S4) are also $W66^{2.61}A$ and $I82^{3.26}T$ (Figure 1, center). $W66^{2.61}A$ is experimentally disruptive for activation by strychnine.[17] In our model it modifies the volume of the vestibular site, thereby weakening agonist interaction by loss of hydrophobic interactions, and destabilizes an interaction with $E265^{7.39}$. In this respect, the $W66^{2.61}$/$E265^{7.39}$ role in shaping the binding site is reminiscent of that of $W201^{5.45}$/$S260^{6.52}$ in TAS2R38.[16] Mutation $I82^{3.26}T$, which we characterized experimentally as leading to a reduced response (see Table S4), decreases the hydrophobic nature of the vestibular site. It also might lead to spurious interactions with $N148^{4.61}$ and $N150^{ECL2}$. Mutations that change the chemical nature of residue $N176^{5.39}$ (Figure 1, center; Figure 2; and Table S4), at the border between ECL2 and TM5, lead both to moderate effects of increased ($N176^{5.39}D$) or decreased ($N176^{5.39}A$) maximum receptor activation, with little effect on EC_{50} (Figure 2). $N176^{5.39}Y$ instead shows only a very limited increase in efficacy.[17] Our simulations show that $N176^{5.39}$, while contacting the agonist in the vestibular site, forms a hydrogen bond network with $W154^{ECL2}$ and $N148^{ECL2}$, both located in ECL2 (see Figure S7). ECL2 has been shown to impact on receptor activation across several GPCRs,[14,61−63] as it may act as a "damper" of receptor activation.[64] $N176^{5.39}D$ charged side chain strengthens the hydrogen bond with $W154^{ECL2}$, yielding a higher efficacy (see Table S4). The null effect of $W154^{ECL2}R^{17}$ is also consistent with this model, since it can maintain the hydrogen bonding.

$E70^{ECL1}$, $N150^{ECL2}$, and $E253^{ECL3}$ (Figure 1, central panel) have been experimentally shown to affect receptor EC_{50} values in a way consistent with affecting the binding process.[17] The three positions are located in extracellular loops (Figure 1) and hence their mutations may affect the molecular recognition process in a time scale beyond the one that can be simulated here.

Mutations Shared between Both Sites (Residues $W88^{3.32}$, $Y241^{6.51}$, $E265^{7.39}$). The massive effect of the $E265^{7.39}K$ mutation (ref 17; see Table S4 for data) can be explained by the electrostatics repulsion with the positive charge of strychnine. The large effect of $E265^{7.39}Q$ and the less significant effect of $E265^{7.39}D$ (ref 17 see Table S4) are clearly explained in this way as well. $Y241^{6.51}$ [This position is correctly identified as 6.51; it was misidentified as 6.50 in ref 17.] and $W88^{3.32}$ roles are most readily explained in the orthosteric binding pose (see below).

Mutations in the Orthosteric Site. $Y241^{6.51}F$ experimental curves (Figure 2, Table S4) might support a role of this residue prevalently in activation:[17] The $Y241^{6.51}F$ mutant strongly reduces the maximum activation level while increasing EC_{50} moderately. When strychnine occupies the orthosteric site, $Y241^{6.51}$ hydroxyl forms a stable H-bond (76% of time) with the $N92^{3.36}$ side chain (see Figure S8) and a staggered π-stacking interaction with the strychnine aromatic ring. Strychnine is positioned so to contact both side chains,

DOI: 10.1021/acs.jctc.5b00472
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

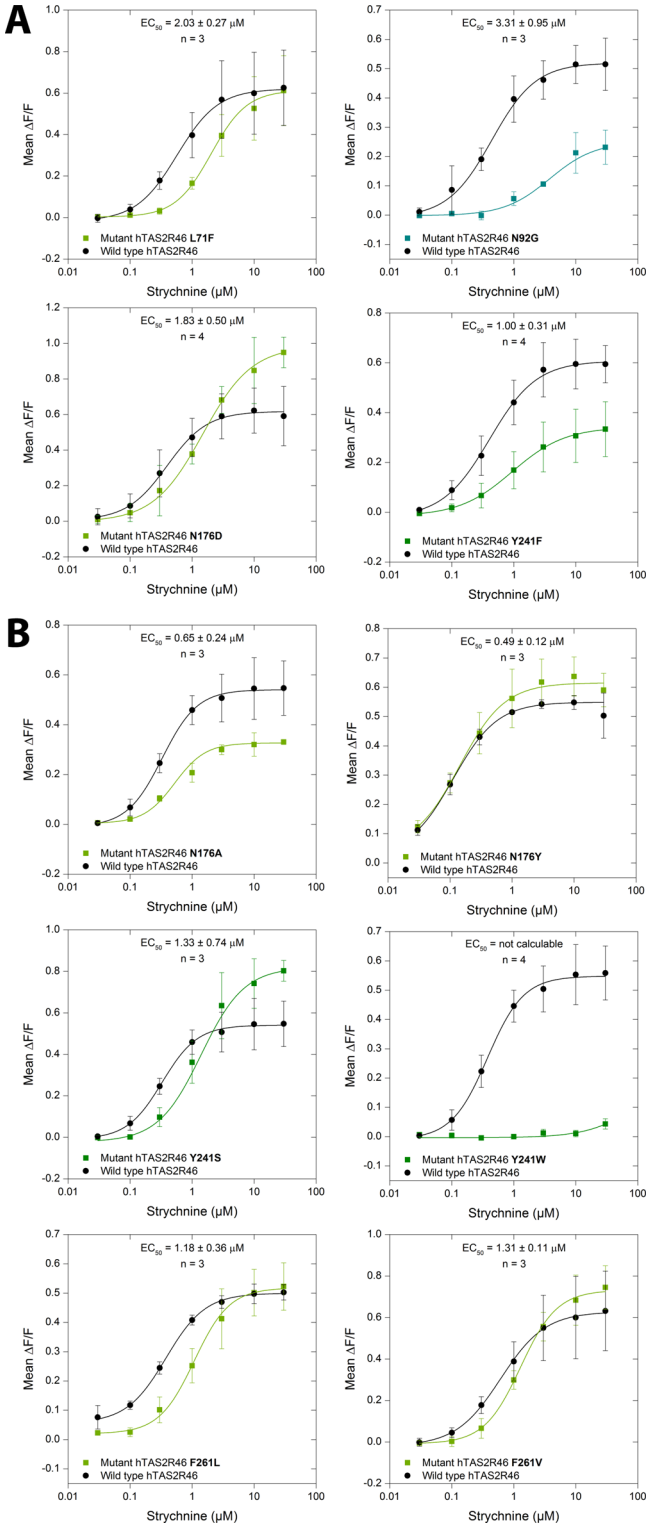


Figure 2. Mean calcium imaging concentration response curves of TAS2R46 mutants in response to strychnine. The wild type EC_{50} is 0.43 ± 0.02 μM.[17] For each panel, the black curve shows the wild type response, while the line in color shows the mutant response, following the same color code of Figure 1. (**A**) Curves for mutants already discussed in ref 17. (**B**) Curves for mutants designed in the present work.

therefore the mutants lose one or the other interaction. The corresponding positions to Y241$^{6.51}$ and especially N92$^{3.36}$ are highly conserved in bitter taste receptors, with 64% and 84% conservation respectively among human TAS2Rs (see Figure S9 for an alignment of the human TAS2R family), supporting a role biased more toward activation than ligand selectivity. The importance and conservation of position 3.36 has been also evidenced experimentally in other human TAS2Rs (see Table 1). Consistently, the Y241$^{6.51}$F mutant shows a lower maximum activation level (see Table S4). This model explains also the strongly reduced sensitivity and decreased response magnitude, as observed form the experimental curves (Figure 2), in the N92$^{3.36}$G mutant (Figure 2, Table S4, and refs 17 and 65). Here the mutation causes the loss of the H-bond with Y241$^{6.51}$ and loss of contact with the ligand. The conserved requirement of this H-bond for bitter taste receptor activation is also consistent with studies made on N92$^{3.36}$ in other closely related TAS2Rs such as TAS2R43, where N92$^{3.36}$D, while reducing agonist binding, increased basal activity while N92$^{3.36}$A and N92$^{3.36}$G abolished activity.[65]

W88$^{3.32}$ (Figure 1, central panel) another highly characterized and conserved residue across TAS2Rs (Figure S7) involved in activation in related TAS2Rs such as TAS2R43 and TAS2R30[65] (see Table 1), contacts the agonist in all trajectories. Here, this residue forms a cation-π interaction with the agonist in the orthosteric site. It also shapes the site by forming $\pi-\pi$ interactions with F269$^{7.43}$ and W66$^{2.61}$.

The orthosteric A268$^{7.42}$R mutation (Figure 1, central panel) is known experimentally to abolish TAS2R46 sensitivity to strychnine almost completely[17] (see Table S3). In our model it is expected to be detrimental to the volume of the site and to cause electrostatic repulsions with strychnine. The slightly decreased EC_{50} of A268$^{7.42}$G further supports a steric effect.[17] Consistently, we performed a set of MM/CG simulations of the A268$^{7.42}$R mutant which show that the ligand is expelled from the orthosteric site in less than 0.1 μs (see Figures S10 and S11).

Novel Mutations. We have performed several mutations in three critical residues, to check the robustness of our model and dissect the role of intramolecular interactions.

To investigate the role of intramolecular hydrogen bonding interactions within ECL2 for activation, we generated two additional mutants of residue N176$^{5.39}$ (see Figure 1, central panel). N176$^{5.39}$A mutant disrupts the hydrogen bond network, resulting in wild-type-like EC_{50} but lower total activation efficacy (Figure 2), while N176$^{5.39}$Y instead either maintains the network by the hydroxyl or substitutes the interaction with an aromatic stacking with W154^{ECL2}, thus remaining close to wild type activity.

Residue F261$^{7.35}$ is predicted here to be involved in the formation of hydrophobic contacts in the vestibular site (Figure 1). F261$^{7.35}$V and F261$^{7.35}$L mutant (Figure 2) should reduce the hydrophobic surface, without altering neither the charge nor polarity of the vestibular site. The mutants have indeed, a significant yet not dramatic effect, which in our model is explained by the partial loss of hydrophobic strychnine-protein interactions.

To investigate the role of Y241$^{6.51}$/N92$^{3.36}$ interactions, we have performed two additional mutations on the Y241 residue. The Y241$^{6.51}$W mutant is expected to disrupt the interaction with N92$^{3.36}$ side chain. It is also sterically unfavorable. Consistently, the mutant receptor cannot be activated by strychnine (Figure 2). Vice versa, Y241$^{6.51}$S shows a higher EC_{50}—possibly due to a reduced interaction with the agonist—but a higher activation level, consistent with the importance of the hydrogen bond for activation (Figure 2; Table S4).

DOI: 10.1021/acs.jctc.5b00472
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

Table 1. Comparison of Residues Involved in Contacts with Strychnine in TAS2R46 According to Simulations Performed Here, Along with Those Experimentally Shown to Be Involved in Agonist Selectivity and Receptor Activation Across TAS2Rs[a]

helix position	TAS2R46 simulations and experiments (this work)	TAS2R38 ref 16	TAS2R10 ref 18	TAS2R16 ref 19	TAS2R43 refs 17, 65	TAS2R31[b] refs 17, 65	TAS2R30[c] ref 65	TAS2R1 ref 21	TAS2R4 ref 23
1.46									G28
1.50								N24	
1.53								I27	
2.54								R55	
2.60	N65								
2.61	W66								
2.64	T69								
2.65	E70							N66	
3.26	I82				T82	T82			
3.29	Y85		S85						
3.30	N86								
3.31						V87			
3.32	W88	W99	W88			W88	W88		
3.33	A89	M100	V89	E86					
3.35					I91	T91			
3.36	N92	N103	N92	N89	N92	G92	N92		
3.37	H93		Q93						
3.40	N96			F93					
3.41				W94					
3.45								N89	
3.50								W94	
3.55								L99	
4.60	I147								
5.38	S175								
5.39	N176			Q177		D176			
5.40			Q175						
5.42		F197							
5.43	T180		L178	H181					
5.47	N184								
5.50								L197	
5.53								S200	
5.54								L201	
6.51	Y241	S259	Y239						
6.52	F242	S260		F240					
6.54	S244								
6.55	I245	F264		I243					
6.58	S248								
6.59	V249								
7.35	F261								
7.39	E265		M263		K265	K265			
7.42	A268		T266		R268	R268			
7.43	F269								
7.46								H273	
7.47									S285
7.50								L277	
7.51								I278	

[a]Studies on TAS2R1 focused mostly on the TM region, while studies on TAS2R4 mostly focused on the TM and the intracellular side of the receptors. In bold, positions identified in this work which have experimental counterparts in other human TAS2Rs. [b]Also known as TAS2R44. [c]Also known as TAS2R47.

In conclusion, our model appears to be validated by experiments performed here and in ref 17 and to offer mechanistic explanations of most experimental mutagenesis effects. In particular, it suggests that residues involved in agonist binding and/or activation are located in two distinct cavities hosting the agonist (see Figure 1, center column, and Figure S3). In particular, some residues experimentally known to be involved in ligand binding and/or receptor activation ($L71^{ECL1}$, $I82^{3.26}$, $N176^{5.39}$; see Figure 1 and Figure S6) interact with strychnine only in the vestibular site (see Table S3). The ligand interacts with experimentally studied residues $Y241^{6.51}$, $I245^{6.55}$ and $E265^{7.39}$ both in the vestibular and in the orthosteric binding poses (see Table S3).

Any attempt to push all of relevant residues to contact the agonist in a single binding pose turn out not to be feasible (see SI 18). Hence, not all mutagenesis effects can be explained simply by the disruption of intermolecular interactions within a

DOI: 10.1021/acs.jctc.5b00472
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

single site if only one simple 1:1 agonist/receptor complex is formed.

We therefore suggest, within the limitations of the molecular modeling performed here (see SI 11), that members of both evolutionary branches of GPCRs targeted by small molecules (Rhodopsin-like/class A and TAS2Rs) might feature a two-site architecture.

Two-Site Architecture Across TAS2Rs and GPCRs. We next investigate if the two binding site architecture may be present across TAS2Rs other than TAS2R46. Bioinformatics calculations show that 14 out of the 27 functional positions herein predicted on TAS2R46 helices are functionally conserved across two or more members of this GPCRs subfamily (Table 1). Four of them (Generic GPCR numbering positions 2.65, 3.26, 3.29, and 5.39) are located distinctly in the vestibular site while five of them (3.33, 3.36, 3.37, 3.40, and 7.42) are located distinctly in the orthosteric site. The remaining (5.43, 6.51, 6.52, 6.55, 7.39) are ligand/receptor contacts common to both poses in the two binding sites. This suggests that several TAS2Rs may share a two-sites architecture, similar to what is indicated by several lines of evidence for the opioid,[12] the muscarinic,[8,11] and the β-adrenergic[7] class A members. This suggestion is corroborated by an analysis of ligand/proteins contact patterns across class A GPCRs and TAS2Rs (Table 2). These are conserved in both orthosteric [Generic GPCR numbering positions 3.32 (W88 in TAS2R46), 3.33 (A89), 3.36 (N92), 6.51 (Y241), 6.52 (F242), 6.55 (I245), and 7.39 (E265)] and vestibular sites [2.64 (T69), 7.35 (F261)] (see Table 3).

Table 2. Comparison Between Strychnine/TAS2R46 Contacts and Corresponding Contacts in the Different X-ray Structures of GPCRs in Complex with Agonists or Antagonists: 19 Aminergic Receptors, 8 Opsins, and 8 Nucleosidic, Analyzed in Reference 77[a]

Generic GPCR numeration	TAS2R46 sequence	Strychnine	GPCR conservation	Generic GPCR numeration	TAS2R46 sequence	Strychnine	GPCR conservation
2.60	65			ECL2	161		
2.61	66			ECL2	162		
2.64	69			5.38	175		
ECL1	70			5.39	176		
ECL1	71			5.42	179		
ECL1	73			5.43	180		
ECL1	75			5.46	183		
ECL1	77			5.47	184		
3.26	82			6.44	234		
3.28	84			6.48	238		
3.29	85			6.51	241		
3.30	86			6.52	242		
3.32	88			6.54	244		
3.33	89			6.55	245		
3.36	92			6.58	248		
3.37	93			ECL3	253		
3.40	96			7.35	261		
4.60	147			7.39	265		
ECL2	150			7.42	268		
ECL2	154			7.43	269		

[a](white) No contact; (light gray) contact in a minority of trajectories/conserved in one GPCR subfamily; (dark grey) contact in a majority of trajectories/conserved in two GPCR subfamilies; (black) contact in all trajectories/conserved in all GPCR subfamilies. Functionally relevant residues identified from experiments in TAS2R46 (represented as spheres in Figure 1) are labeled in red.

Table 3. Contacts Exclusive to the Vestibular Region Conserved in Three Receptors Where Experimental Data ($\surd$) and/or Molecular Simulation (*) Unveiled a Vestibular Site Involved in Agonist Recognition

generic GPCR numbering position	TAS2R46 (this work)	M2 muscarinic acetylcholine receptor[11]	β2 adrenergic receptor[7]
2.64	*	$\surd$ and *	
5.39	$\surd$ and *		*
6.58	*		*
6.59	*		*
7.35	$\surd$ and *	$\surd$ and *	*

We next assume (as first suggested by Granier and Kobilka)[12] that, for a given receptor subclass, a low residue conservation value in a receptor site correlates with a role in shaping divergent agonist specificities among members of the subclass. Conversely, high residue conservation would imply a conserved role of the site, independently from ligand specificity. This reasonable hypothesis has been used successfully to predict residues conveying substrate-specificity across enzyme families.[66] Let us then identify the residues that could be, in principle, involved in the vestibular binding site (V group hereafter) [We do not assume that the vestibular site exists functionally for all these receptors, or that it is significantly involved in recognition of every possible ligand in these receptors. The V designation implies only that, *if* a distinct vestibular site is present, it might, by analogy, include residues in group V.] and those in the orthosteric binding site (O group) in the TAS2R, opioid, muscarinic and β-adrenergic subfamilies (see Methods for definition) [Opioid, muscarinic, and β-adrenergic subfamilies belong to Class A. TAS2Rs are a separate branch of the GPCR superfamily, close either to the Frizzled family1 or remotely related to Class A receptors.[67] However, X-ray studies of GPCRs belonging to different classes show that the general 7-TM structure and orientation of GPCRs is highly conserved even between highly diverging families.[68−70] This suggests that one may obtain useful insights by comparing the structural determinants across GPCRs belonging to different classes, as in this case.[68,69]]. Next, let us calculate the normalized residue conservation histograms in the two groups across the subfamilies that these residues belong to. Even if TAS2Rs do not belong to class A, the general conservation of the 7-TM architecture, confirmed by crystal structures of GPCRs belonging to different classes, allows us to compare them.

The conservation histograms of TAS2Rs for the two groups are similar to each other and differ largely from those of Class A receptors (turquoise and red histograms in Figure 3). They are peaked at low values, between 20% and 40%. This suggests a selectivity role of both the orthosteric binding site and (if present as such in a given receptor) of the vestibular site, consistently with the large number of TAS2Rs members and agonists (25 and more than 100,[71] respectively). In particular TAS2R46, which is suggested here to feature a vestibular site, is activated by as many as 28 natural agonists.[31]

In Class A receptors, conservation histograms of the V group and, more, of the O group are peaked at larger values than those of TAS2Rs (turquoise and red histograms in Figure 3). This suggests that, in these cases, both sites have a less clear selectivity role, and if anything selection is prevalently due to the residues in the V group. This is consistent with the paucity of receptors and of natural agonists of the families: the human

DOI: 10.1021/acs.jctc.5b00472
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

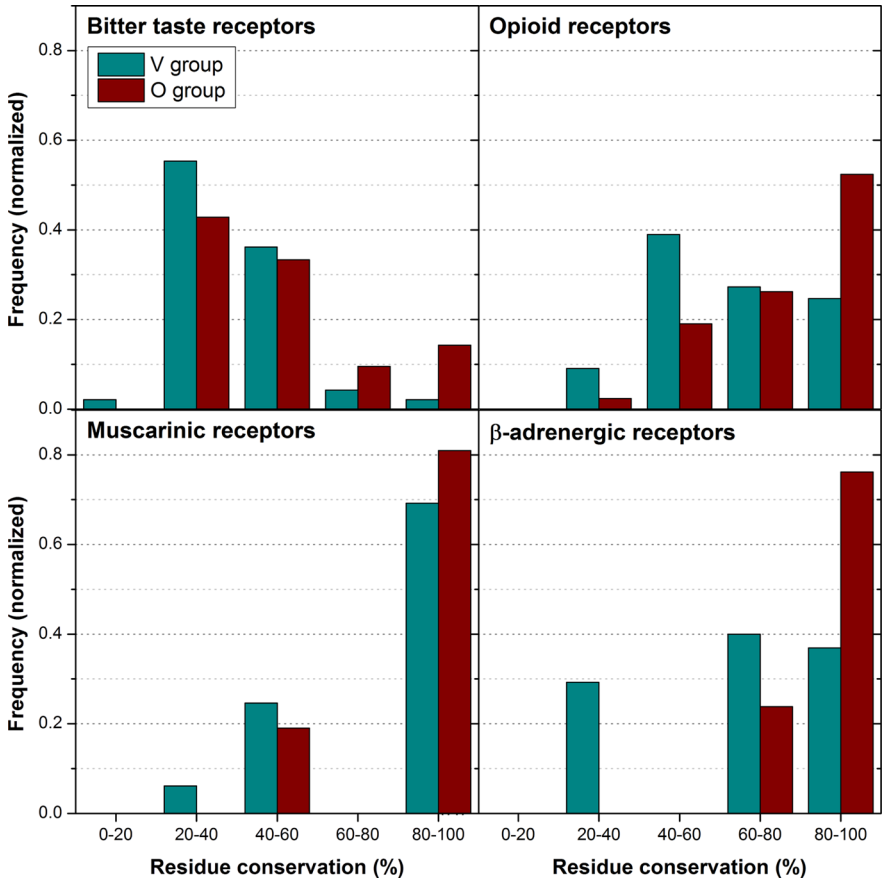


Figure 3. Conservation histograms of residues in V group (turquoise) and O group (red) across human GPCRs subfamilies where the vestibular site has been identified. See the main text for the definition of both groups.

opioid, β-adrenergic and muscarinic receptor subfamilies comprise 4, 3, and 5 members, targeting 15,[72] 2, and 1 natural agonists, respectively.

CONCLUDING REMARKS

Based on our combined in silico/in vitro study, we suggest that a typical agonist to TAS2R46, a prototypical broad-spectrum chemosensory GPCR, features a two cavities recognition process. Moreover we suggest, on the basis of bioinformatics analysis, that both cavities could be involved in receptor selectivity. This seems to be apt to chemosensory function. A single binding site may not contain enough information, encoded as potential side chain/ligand contacts, to discriminate between chemically similar compounds while also accepting a wide spectrum of chemically different compounds, as it happens in bitter taste receptors.[20,31,71,73,74] It is conceivable that a similar mechanism could be found on other chemosensory receptors, such as olfactory receptors. Our modeling and simulation approach is very general and it could be extended to them as well, provided that enough mutagenesis experimental data is available to validate the resulting models.

Our current and previous[17] analysis along with previous simulation studies,[7,8] point to the presence of two cavities and ligand/protein contact pattern similarities across members of GPCRs targeted by small agonists in their 7-TM bundle—a large majority of GPCRs across at least two major and distant evolutionary branches of the GPCR superfamily.[75] Evidence is emerging that both branches evolved from an ancestral small molecule-activated GPCR family (the cyclic AMP receptor family) in early eukaryotes.[76] We thus conjecture that the general binding strategy of GPCRs targeted by small ligands, including the existence of a vestibular site, might predate the radiation of GPCRs into separate subfamilies.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jctc.5b00472.

Supporting Information 1−18 show alignments and parameters used for the modeling and docking procedures, model numbering and topology, details of system evolution during simulation including hydration, details of the second vestibular pose, residue−ligand contact map, discussion of experimental limitations, location and effect of experimentally mutated residues, details of intramolecular interactions, details of the A268R simulation, and effects of attempts to force experimentally identified residues into a single binding pose (PDF)

Coordinates of the TAS2R46 homology model (PDB)

AUTHOR INFORMATION

Corresponding Authors

*E-mail: alejandro.giorgetti@univr.it (A.G.).
*E-mail: p.carloni@fz-juelich.de (P.C.).

DOI: 10.1021/acs.jctc.5b00472
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

Author Contributions

All authors conceived and designed the experiments and analyzed the data. : M.S., A.B., and M.B. performed the experiments and the simulations. M.S., F.M., P.C., W.M. contributed reagents/materials/analysis tools. M.S., F.M., M.B., W.M., A.G., and P.C. wrote the paper.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We acknowledge the Jülich-Aachen Research Alliance−High Performance Computing time grants JARA0023 and JARA0082. P.C. and A.G. acknowledge a long-standing collaboration with the E. Illy Foundation for chemical senses. F.M. was financed by Programma Operativo del Fondo Sociale Europeo 2007/2013 of Regione Autonoma Friuli Venezia Giulia and by CIRMMP (Consorzio Interuniversitario di Risonanze Magnetiche di Metallo-Proteine).

■ REFERENCES

(1) Fredriksson, R.; Lagerström, M. C.; et al. The G-protein-coupled receptors in the human genome form five main families. Phylogenetic analysis, paralogon groups, and fingerprints. *Mol. Pharmacol.* **2003**, *63*, 1256−1272.
(2) Overington, J. P.; Al-Lazikani, B.; et al. How many drug targets are there? *Nat. Rev. Drug Discovery* **2006**, *5*, 993−996.
(3) Zhang, J.; Zhang, K.; et al. Agonist-bound structure of the human P2Y12 receptor. *Nature* **2014**, *509*, 119−22.
(4) Wang, C.; Jiang, Y.; et al. Structural basis for molecular recognition at serotonin receptors. *Science* **2013**, *340*, 610−614.
(5) Rosenbaum, D. M.; Zhang, C.; et al. Structure and function of an irreversible agonist-beta(2) adrenoceptor complex. *Nature* **2011**, *469*, 236−40.
(6) Lebon, G.; Warne, T.; et al. Agonist-bound adenosine A2A receptor structures reveal common features of GPCR activation. *Nature* **2011**, *474*, 521−5.
(7) Dror, R. O.; Pan, A. C.; et al. Pathway and mechanism of drug binding to G-protein-coupled receptors. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108*, 13118−13123.
(8) Kruse, A. C.; Hu, J.; et al. Structure and dynamics of the M3 muscarinic acetylcholine receptor. *Nature* **2012**, *482*, 552−556.
(9) Breslin, P. A. An evolutionary perspective on food and human taste. *Curr. Biol.* **2013**, *23*, R409−R418.
(10) Meyerhof, W.; Behrens, M.; et al. Human bitter taste perception. *Chem. Senses* **2005**, *30* (Suppl1), i14−i15.
(11) Dror, R. O.; Green, H. F.; et al. Structural basis for modulation of a G-protein-coupled receptor by allosteric drugs. *Nature* **2013**, *503*, 295−299.
(12) Granier, S.; Kobilka, B. A new era of GPCR structural and chemical biology. *Nat. Chem. Biol.* **2012**, *8*, 670−673.
(13) Granier, S.; Manglik, A.; et al. Structure of the delta-opioid receptor bound to naltrindole. *Nature* **2012**, *485*, 400−4.
(14) Wheatley, M.; Wootten, D.; et al. Lifting the lid on GPCRs: the role of extracellular loops. *Br. J. Pharmacol.* **2012**, *165*, 1688−1703.
(15) Biarnes, X.; Marchiori, A.; et al. Insights into the binding of Phenyltiocarbamide (PTC) agonist to its target human TAS2R38 bitter receptor. *PLoS One* **2010**, *5*, e12394.
(16) Marchiori, A.; Capece, L.; et al. Coarse-Grained/Molecular Mechanics of the TAS2R38 Bitter Taste Receptor: Experimentally-Validated Detailed Structural Prediction of Agonist Binding. *PLoS One* **2013**, *8*, e64675.
(17) Brockhoff, A.; Behrens, M.; et al. Structural requirements of bitter taste receptor activation. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107*, 11110−11115.
(18) Born, S.; Levit, A.; et al. The human bitter taste receptor TAS2R10 is tailored to accommodate numerous diverse ligands. *J. Neurosci.* **2013**, *33*, 201−213.
(19) Sakurai, T.; Misaka, T.; et al. Characterization of the beta-D-glucopyranoside binding site of the human bitter taste receptor hTAS2R16. *J. Biol. Chem.* **2010**, *285*, 28373−28378.
(20) Sakurai, T.; Misaka, T.; et al. The human bitter taste receptor, hTAS2R16, discriminates slight differences in the configuration of disaccharides. *Biochem. Biophys. Res. Commun.* **2010**, *402*, 595−601.
(21) Singh, N.; Pydi, S. P.; et al. Structural basis of activation of bitter taste receptor T2R1 and comparison with Class A G-protein-coupled receptors (GPCRs). *J. Biol. Chem.* **2011**, *286*, 36032−36041.
(22) Upadhyaya, J.; Pydi, S. P.; et al. Bitter taste receptor T2R1 is activated by dipeptides and tripeptides. *Biochem. Biophys. Res. Commun.* **2010**, *398*, 331−335.
(23) Pydi, S. P.; Bhullar, R. P.; et al. Constitutively active mutant gives novel insights into the mechanism of bitter taste receptor activation. *J. Neurochem.* **2012**, *122*, 537−544.
(24) Prasad Pydi, S.; Singh, N.; et al. The third intracellular loop plays a critical role in bitter taste receptor activation. *Biochim. Biophys. Acta, Biomembr.* **2014**, *1838*, 231−236.
(25) Pydi, S. P.; Sobotkiewicz, T.; et al. Amino acid derivatives as bitter taste receptor (T2R) blockers. *J. Biol. Chem.* **2014**, *289*, 25054−25066.
(26) Novoa, E. M.; Pouplana, L. R. d.; et al. Ensemble docking from homology models. *J. Chem. Theory Comput.* **2010**, *6*, 2547−2557.
(27) Leguebe, M.; Nguyen, C.; et al. Hybrid molecular mechanics/coarse-grained simulations for structural prediction of G-protein coupled receptor/ligand complexes. *PLoS One* **2012**, *7*, e47332.
(28) Holyoake, J.; Caulfeild, V.; et al. Modeling, docking, and simulation of the major facilitator superfamily. *Biophys. J.* **2006**, *91*, L84−L86.
(29) Tai, K.; Stansfeld, P. J.; et al. Ion-blocking sites of the Kir2. 1 channel revealed by multiscale modeling. *Biochemistry* **2009**, *48*, 8758−8763.
(30) Norimatsu, Y.; Ivetac, A.; et al. Locating a plausible binding site for an open-channel blocker, GlyH-101, in the pore of the cystic fibrosis transmembrane conductance regulator. *Mol. Pharmacol.* **2012**, *82*, 1042−1055.
(31) Brockhoff, A.; Behrens, M.; et al. Broad tuning of the human bitter taste receptor hTAS2R46 to various sesquiterpene lactones, clerodane and labdane diterpenoids, strychnine, and denatonium. *J. Agric. Food Chem.* **2007**, *55*, 6236−6243.
(32) Musiani, F.; Rossetti, G.; et al. Chemosensorial G-proteins-Coupled Receptors: A Perspective from Computational Methods. *Adv. Exp. Med. Biol.* **2014**, *805*, 441−457.
(33) Burger, H.; Weiser, B.; Fang, G.; et al. PCR-mediated recombination: a general method applied to construct chimeric infectious molecular clones of plasma-derived HIV-1 RNA. *Nat. Med.* **1999**, *5*, 239−242.
(34) Sandal, M.; Duy, T. P.; et al. GOMoDo: A GPCRs Online Modeling and Docking Webserver. *PLoS One* **2013**, *8*, e74092.
(35) Pei, J.; Kim, B. H.; et al. PROMALS web server for accurate multiple protein sequence alignments. *Nucleic Acids Res.* **2007**, *35*, W649−W652.
(36) Hildebrand, A.; Remmert, M.; et al. Fast and accurate automatic structure prediction with HHpred. *Proteins: Struct., Funct., Genet.* **2009**, *77* (Suppl 9), 128−132.
(37) Eswar, N.; Eramian, D.; et al. Protein structure modeling with MODELLER. *Methods Mol. Biol.* **2008**, *426*, 145−159.
(38) Willard, L.; Ranjan, A.; et al. VADAR: a web server for quantitative evaluation of protein structure quality. *Nucleic Acids Res.* **2003**, *31*, 3316−3319.
(39) Chien, E. Y.; Liu, W.; et al. Structure of the human dopamine D3 receptor in complex with a D2/D3 selective antagonist. *Science* **2010**, *330*, 1091−1095.
(40) Everett, A. J.; Openshaw, H. T.; et al. 221. The constitution of aspidospermine. Part III. Reactivity at the nitrogen atoms, and biogenetic considerations. *J. Chem. Soc.* **1957**, 1120−1123.
(41) Dominguez, C.; Boelens, R.; et al. HADDOCK: a protein-protein docking approach based on biochemical or biophysical information. *J. Am. Chem. Soc.* **2003**, *125*, 1731−1737.

DOI: 10.1021/acs.jctc.5b00472
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

(42) Schuttelkopf, A. W.; van Aalten, D. M. PRODRG: a tool for high-throughput crystallography of protein-ligand complexes. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2004**, *60*, 1355−1363.
(43) Le Guilloux, V.; Schmidtke, P.; et al. Fpocket: an open source platform for ligand pocket detection. *BMC Bioinf.* **2009**, *10*, 168.
(44) Neri, M.; Anselmi, C.; et al. Coarse-grained model of proteins incorporating atomistic detail of the active site. *Phys. Rev. Lett.* **2005**, *95*, 218102.
(45) Pronk, S.; Pall, S.; et al. GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit. *Bioinformatics* **2013**, *29*, 845−854.
(46) Scott, W. R. P.; Hünenberger, P. H.; et al. The GROMOS Biomolecular Simulation Program Package. *J. Phys. Chem. A* **1999**, *103*, 3596−3607.
(47) Go, N.; Abe, H. Noninteracting local-structure model of folding and unfolding transition in globular proteins. I. Formulation. *Biopolymers* **1981**, *20*, 991−1011.
(48) Berendsen, H. J.; Postma, J., et al. Interaction models for water in relation to protein hydration. In *Intermolecular forces*; Pullman, B., Ed.; Springer, 1981; pp 331−342.
(49) Ryckaert, J.-P.; Ciccotti, G.; et al. Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J. Comput. Phys.* **1977**, *23*, 327−341.
(50) Daura, X.; Gademann, K.; et al. Peptide Folding: When Simulation Meets Experiment. *Angew. Chem., Int. Ed.* **1999**, *38*, 236−240.
(51) Pettersen, E. F.; Goddard, T. D.; et al. UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.* **2004**, *25*, 1605−1612.
(52) Mills, J.; Dean, P. M. Three-dimensional hydrogen-bond geometry and probability information from a crystal survey. *J. Comput.-Aided Mol. Des.* **1996**, *10*, 607−622.
(53) Isberg, V.; de Graaf, C.; et al. Generic GPCR residue numbers−aligning topology maps while minding the gaps. *Trends Pharmacol. Sci.* **2015**, *36*, 22−31.
(54) Ballesteros, J. A.; Weinstein, H. Integrated methods for the construction of three-dimensional models and computational probing of structure-function relations in G protein-coupled receptors. In *Methods in Neurosciences*; Sealfon, S. C., Ed.; Academic Press, 1995; Vol. *25*, pp 366−428.
(55) Haga, K.; Kruse, A. C.; et al. Structure of the human M2 muscarinic acetylcholine receptor bound to an antagonist. *Nature* **2012**, *482*, 547−551.
(56) Angel, T. E.; Chance, M. R.; et al. Conserved waters mediate structural and functional activation of family A (rhodopsin-like) G protein-coupled receptors. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106*, 8555−8560.
(57) Liu, W.; Chun, E.; et al. Structural basis for allosteric regulation of GPCRs by sodium ions. *Science* **2012**, *337*, 232−236.
(58) Yuan, S.; Filipek, S.; et al. Activation of G-protein-coupled receptors correlates with the formation of a continuous internal water pathway. *Nat. Commun.* **2014**, *5*, 4733.
(59) Kufareva, I.; Rueda, M.; et al. Status of GPCR modeling and docking as reflected by community-wide GPCR Dock 2010 assessment. *Structure* **2011**, *19*, 1108−1126.
(60) Kufareva, I.; Katritch, V.; et al. Advances in GPCR modeling evaluated by the GPCR Dock 2013 assessment: meeting new challenges. *Structure* **2014**, *22*, 1120−39.
(61) Klco, J. M.; Wiegand, C. B.; et al. Essential role for the second extracellular loop in C5a receptor activation. *Nat. Struct. Mol. Biol.* **2005**, *12*, 320−326.
(62) Peeters, M. C.; Wisse, L. E.; et al. The role of the second and third extracellular loops of the adenosine A1 receptor in activation and allosteric modulation. *Biochem. Pharmacol.* **2012**, *84*, 76−87.
(63) Unal, H.; Jagannathan, R.; et al. Long range effect of mutations on specific conformational changes in the extracellular loop 2 of angiotensin II type 1 receptor. *J. Biol. Chem.* **2013**, *288*, 540−551.
(64) Massotte, D.; Kieffer, B. L. The second extracellular loop: a damper for G protein-coupled receptors? *Nat. Struct. Mol. Biol.* **2005**, *12*, 287−288.
(65) Pronin, A. N.; Tang, H.; et al. Identification of ligands for two human bitter T2R receptors. *Chem. Senses* **2004**, *29*, 583−593.
(66) Yu, G.-X.; Park, B.-H.; et al. In silico discovery of enzyme−substrate specificity-determining residue clusters. *J. Mol. Biol.* **2005**, *352*, 1105−1117.
(67) Nordstrom, K. J.; Sallman Almen, M.; et al. Independent HHsearch, Needleman–Wunsch-based, and motif analyses reveal the overall hierarchy for most of the G protein-coupled receptor families. *Mol. Biol. Evol.* **2011**, *28*, 2471−2480.
(68) Siu, F. Y.; He, M.; et al. Structure of the human glucagon class B G-protein-coupled receptor. *Nature* **2013**, *499*, 444−9.
(69) Wang, C.; Wu, H.; et al. Structure of the human smoothened receptor bound to an antitumour agent. *Nature* **2013**, *497*, 338−43.
(70) Wu, H.; Wang, C.; et al. Structure of a class C GPCR metabotropic glutamate receptor 1 bound to an allosteric modulator. *Science* **2014**, *344*, 58−64.
(71) Meyerhof, W.; Batram, C.; et al. The molecular receptive ranges of human TAS2R bitter taste receptors. *Chem. Senses* **2010**, *35*, 157−170.
(72) Waldhoer, M.; Bartlett, S. E.; et al. Opioid receptors. *Annu. Rev. Biochem.* **2004**, *73*, 953−990.
(73) Behrens, M.; Brockhoff, A.; et al. The human taste receptor hTAS2R14 responds to a variety of different bitter compounds. *Biochem. Biophys. Res. Commun.* **2004**, *319*, 479−485.
(74) Behrens, M.; Brockhoff, A.; et al. The human bitter taste receptor hTAS2R50 is activated by the two natural bitter terpenoids andrographolide and amarogentin. *J. Agric. Food Chem.* **2009**, *57*, 9860−9866.
(75) Lagerstrom, M. C.; Schioth, H. B. Structural diversity of G protein-coupled receptors and significance for drug discovery. *Nat. Rev. Drug Discovery* **2008**, *7*, 339−57.
(76) Krishnan, A.; Almen, M. S.; et al. The origin of GPCRs: identification of mammalian like Rhodopsin, Adhesion, Glutamate and Frizzled GPCRs in fungi. *PLoS One* **2012**, *7*, e29817.
(77) Venkatakrishnan, A. J.; Deupi, X.; et al. Molecular signatures of G-protein-coupled receptors. *Nature* **2013**, *494*, 185−194.

DOI: 10.1021/acs.jctc.5b00472
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX