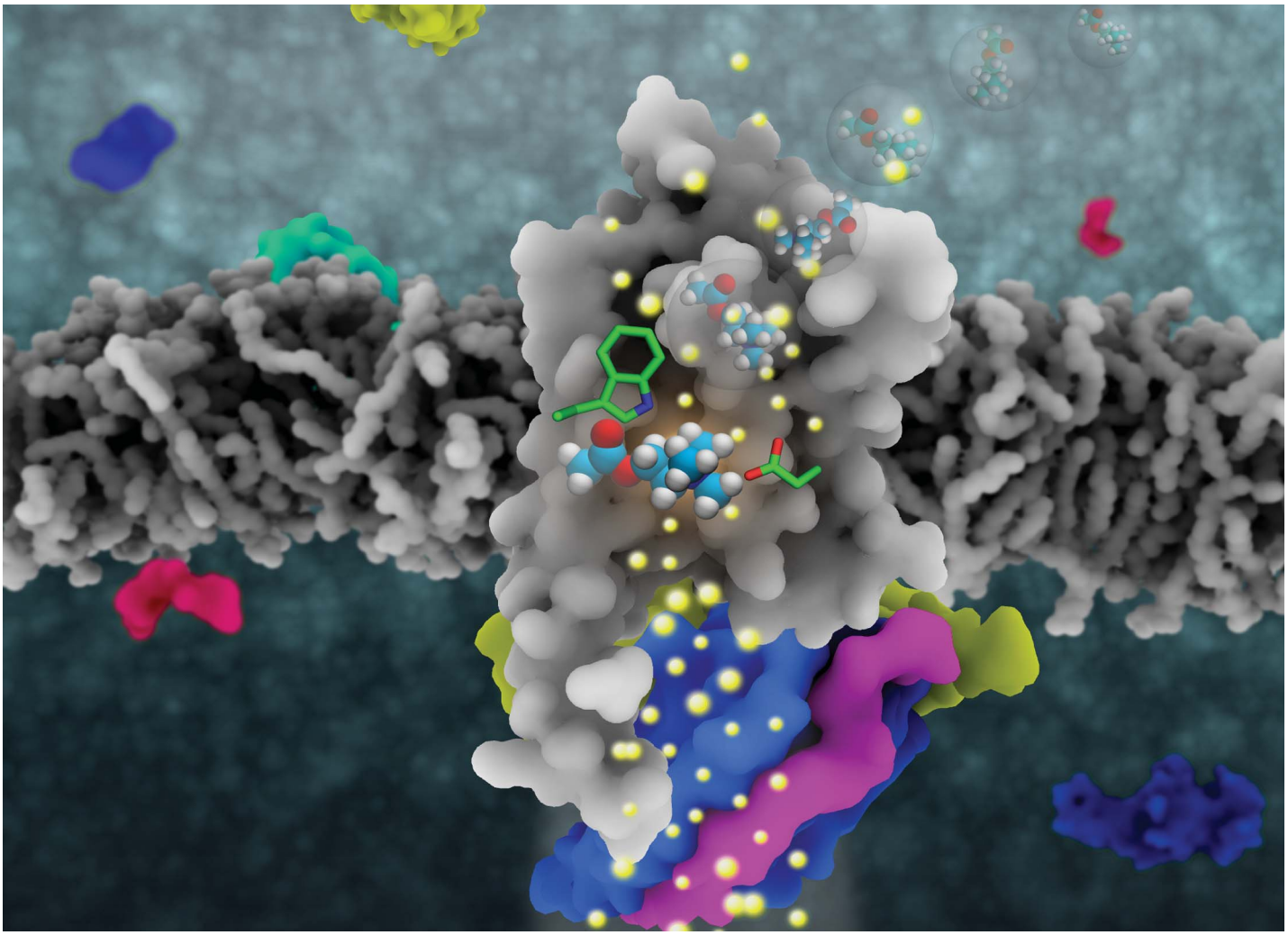

Showcasing research from Professor Horst Vogel's laboratory, Institute of Chemical Sciences and Engineering, École Polytechnique Fédérale de Lausanne, Lausanne, Switzerland.

Exploring a new ligand binding site of G protein-coupled receptors

A new ligand binding pocket next to the highly conserved $D^{2.50}$ residue of GPCRs has been discovered, using all-atom long time scale molecular dynamics simulations. This has been confirmed by follow-up mutagenesis studies. This newly discovered binding pocket provides new opportunities for modern drug discovery.

As featured in:

See Horst Vogel, Zhi-Jie Liu, Shuguang Yuan *et al.*, *Chem. Sci.*, 2018, **9**, 6480.

ROYAL SOCIETY OF CHEMISTRY

rsc.li/chemical-science

Registered charity number: 207890

Chemical Science

EDGE ARTICLE

View Article Online
View Journal | View Issue

Check for updates

Cite this: *Chem. Sci.*, 2018, **9**, 6480

All publication charges for this article have been paid for by the Royal Society of Chemistry

Exploring a new ligand binding site of G protein-coupled receptors†

H. C. Stephen Chan,‡[a] Jingjing Wang,‡[b] Krzysztof Palczewski,[c] Slawomir Filipek,[d] Horst Vogel,*[e] Zhi-Jie Liu*[b] and Shuguang Yuan *[e]

Identifying a target ligand binding site is an important step for structure-based rational drug design as shown here for G protein-coupled receptors (GPCRs), which are among the most popular drug targets. We applied long-time scale molecular dynamics simulations, coupled with mutagenesis studies, to two prototypical GPCRs, the M3 and M4 muscarinic acetylcholine receptors. Our results indicate that unlike synthetic antagonists, which bind to the classic orthosteric site, the endogenous agonist acetylcholine is able to diffuse into a much deeper binding pocket. We also discovered that the most recently resolved crystal structure of the LTB4 receptor comprised a bound inverse agonist, which extended its benzamidine moiety to the same binding pocket discovered in this work. Analysis on all resolved GPCR crystal structures indicated that this new pocket could exist in most receptors. Our findings provide new opportunities for GPCR drug discovery.

Received 12th April 2018
Accepted 12th July 2018

DOI: 10.1039/c8sc01680a

rsc.li/chemical-science

Open Access Article. Published on 13 July 2018. Downloaded on 2/9/2026 10:52:24 AM.
This article is licensed under a Creative Commons Attribution-NonCommercial 3.0 Unported Licence.

Introduction

Many essential cellular processes, including cell regulation, signal transduction, and the immune response, are mediated by specific protein–ligand interactions. The identification and characterization of specific protein binding sites for particular ligands are crucial for the understanding of the functions of both endogenous ligands and synthetic drug molecules.[1–3] Thus, the detection and characterization of ligand binding sites are important steps towards protein function identification and drug discovery.[4–7] A newly revealed ligand binding site will provide a new opportunity for a drug target, to design new classes of compounds based on new chemical environments.[8] Besides traditional biochemical methods such as systematic mutagenesis experiments, structural biology or NMR,[9] computational methods offer alternative powerful and efficient approaches for ligand binding site detection.[7,10]

In this work, we applied molecular dynamics simulations and computer modeling, combined with functional mutagenesis experiments, to explore new ligand binding sites of GPCRs. GPCRs typically detect extracellular signals (photons, odorants, hormones, and neurotransmitters)[11,12] on the cell surface and undergo multiple conformational changes that enable the binding and activation of intracellular proteins (*e.g.* G-proteins, arrestins and kinases).[13–16] More than a third of modern therapeutic compounds are estimated to target GPCRs.[7] Thus, understanding the ligand binding process and detecting a new ligand binding site for GPCRs in molecular details are of great importance in revealing GPCR-mediated signaling and improving GPCR-targeted drug design. Recent crystal structures of GPCRs show that different small molecules can bind to different regions of receptors including (1) the traditional orthosteric ligand binding site in the vicinity of the highly conserved ($W^{6.48}$),[17] (2) the allosteric ligand binding site next to extracellular loop 2 (ECL2),[18,19] (3) allosteric ligand binding sites between transmembrane (TM) helices TM2 and TM3,[18,19] (4) allosteric ligand binding sites between TM3 and TM4,[20] (5) the intracellular G protein binding region[21] and (6) the outer surface of the receptor in the middle of the TM area[22] (Fig. S1†).

Computational methods have become important tools for understanding the structural and dynamical function of GPCRs.[7,23] Molecular dynamics (MD) simulations and computational modeling can be used to address specific questions about the dynamic properties of the modelled system, which are difficult to illuminate in experiments. For example, with MD simulations, it is possible to sample the process of ligand binding to a specific GPCR.[24–26] Many atomic details, including molecular switches, binding site expansions or domain movements, can be efficiently revealed through MD simulations.[27–30]

[a]*Faculty of Life Sciences, University of Bradford Bradford, UK*

[b]*iHuman Institute, ShanghaiTech University, China. E-mail: liuzhj@shanghaitech.edu.cn*

[c]*School of Medicine, Case Western Reserve University, USA*

[d]*Faculty of Chemistry, Biological and Chemical Research Centre, University of Warsaw, Poland*

[e]*Institute of Chemical Sciences and Engineering, Ecole Polytechnique Fédérale de Lausanne (EPFL), Switzerland. E-mail: shuguang.yuan@gmail.com; horst.vogel@epfl.ch*

† Electronic supplementary information (ESI) available. See DOI: 10.1039/c8sc01680a

‡ Authors contributed equally to this work.

This journal is © The Royal Society of Chemistry 2018

Open Access Article. Published on 13 July 2018. Downloaded on 2/9/2026 10:52:24 AM.
This article is licensed under a Creative Commons Attribution-NonCommercial 3.0 Unported Licence.

Table 1 Receptor–ligand (R–L) complexes investigated by MD simulations

R–L	Receptor PDB	Ligand	Ligand type	Simulation time (μs)
M3–ACh	4DAJ	ACh	agonist	2 × 3
M3–TTP	4DAJ	TTP	antagonist	2 × 3
M4–ACh	5DSG	ACh	agonist	2 × 3
M4–TTP	5DSG	TTP	antagonist	2 × 3

Here we explore a new ligand binding site for two GPCRs, muscarinic M3 and M4 receptors, which specifically activate Gαq and Gαi proteins, respectively (Table 1).[31,32] Crystal structures have been solved for M3 and M4 receptors containing antagonists in orthosteric binding sites, whereas no agonist-bound structures are available. A new binding site for muscarinic acetylcholine receptors could provide new opportunities for the understanding of ligand binding and activation processes. We therefore simulated the entire binding processes of the endogenous agonist acetylcholine (ACh) to both M3 and M4 receptors. Surprisingly, the resulting position of ACh bound to the particular receptors revealed an additional new binding site next to the highly conserved Asp ($D^{2.50}$).[33] We also observed that ligand binding leads to the expansion of the binding site.

We further inspected over 200 ligand bound GPCR crystal structures and discovered that most ligands were located in traditional orthosteric sites. However, the recently resolved crystal structure of the leukotriene B4 (LTB4) receptor comprised a bound bitopic ligand which extended from this typical orthosteric site to the new ligand binding site found in the present work. This observation strongly suggests that our newly discovered binding site might also exist in other GPCRs. Among these >200 structures, we systematically inspected 39 unique family A GPCRs and found that most of them indeed contain this potential binding site for binding small ligands or extending larger ones from the main orthosteric site. Our findings provide a new opportunity for GPCR drug discovery.

Experimental section

Biological testing

Split luciferase biosensor cAMP assay for measuring activation of Gi protein. Promega's split luciferase based GloSensor cAMP biosensor technology was used to determine GPCR mediated cAMP production. M4 receptor DNA (4 μg) and GloSensor cAMP DNA (4 μg, Promega) were co-transfected into HEK293 T cells using Lipofectamine 2000 (Life technologies). After at least 24 h, the cells were seeded in 384 well white clear bottom plates (Greiner) with DMEM (Life technologies) supplemented with 1% dialyzed FBS at a density of 15–20 000 cells in 20 μl medium per well and incubated at 37 °C in 5% CO_2 for at least 6 h before analysis. Wells were loaded for 20 min at 37 °C with 20 μl of 2 mg ml^{-1} Luciferin D sodium salt prepared in HBSS at pH 7.4. All the following steps were carried out at room temperature. To measure agonist activity of the M4 receptor, 10 μl of acetylcholine solutions ranging from 0 to 30 000 μM was added to the cells 15 min before addition of 10 μl of isoproterenol at a final concentration of 200 nM, followed by counting of the plate for chemiluminescence using EnVision (Perkin Elmer) after 15 min. Chemiluminescence intensity was plotted as a function of ACh concentration and normalized to percent acetylcholine with 100% for E_{max} acetylcholine cAMP inhibition and 0% for the isoproterenol-stimulated cAMP baseline. Data were analyzed using log (ACh) *vs.* response in GraphPad Prism.

Ca^{2+} Mobilization assay for measuring activation of Gq protein. To measure ACh-induced G protein coupling to the M3 receptor and the subsequent increase of intracellular calcium ion concentration, HEK293T cells were seeded in a 10 cm dish and incubated overnight. After 24 h, the cells were transfected with 4 μg M3 receptor plasmid and 20 μl Lipofectamine 2000 (Life technologies). 6 h later, the cells were seeded in 384 well plates at a density of 15 000 cells per well in DMEM containing 1% dialyzed FBS and incubated overnight. On the assay day, the cells were incubated (20 μl per well) for 1 h at 37 °C with Fluo-4 Direct dye (Invitrogen) reconstituted in FLIPR buffer (1 × HBSS, 2.5 mM probenecid, and 20 mM HEPES, pH 7.4). After the dye was loaded, the cells were placed in a $FLIPR^{TETRA}$ fluorescence imaging plate reader (Molecular Devices); acetylcholine dilutions were prepared at 3× final concentration in FLIPR buffer and aliquoted into 384 well plates. The fluidics module and plate reader of the $FLIPR^{TETRA}$ were programmed to read baseline fluorescence for 10 s (1 read per s), and then 10 μl of drug/well was added to read for 6 min (1 read per s). Fluorescence in each well was normalized to the average of the first 10 reads (*i.e.*, baseline fluorescence). The maximal fluorescence intensity increase was measured 60 s after acetylcholine addition. Data were analyzed using GraphPad Prism.

Loop modelling and structural preparations

Loop filling and refinements. The resolved crystal structures of M3 and M4 receptors comprised the engineered receptors and inserted proteins in the intracellular loop ICL2 to facilitate crystallisation. Before starting MD simulations, we removed the corresponding inserted proteins from the M3 and M4 crystal structures and used the loop refinement protocol in Modeller[34] V9.10 to reconstruct and refine the ICL2 region. A total of 20 000 loops were generated for each receptor, and the conformation with the lowest DOPE (Discrete Optimized Protein Energy) score was chosen for receptor construction. Repaired models were submitted to Rosetta V3.4 for loop refinement with kinematic loop modeling methods.[35] Kinematic closure (KIC) is an analytic calculation inspired by robotics techniques for rapidly determining possible conformations of linked objects subject to constraints. In the Rosetta KIC implementation, 2N – 6 backbone torsions of an N-residue peptide segment (called non-pivot torsions) were set to values randomly drawn from the Ramachandran space of each residue type, and the remaining 6 phi/psi torsions (called pivot torsions) were solved analytically by KIC.

Protein structure preparations. All protein models were prepared using Schrodinger suite software under the

Open Access Article. Published on 13 July 2018. Downloaded on 2/9/2026 10:52:24 AM.
This article is licensed under a Creative Commons Attribution-NonCommercial 3.0 Unported Licence.

OPLS_2005 force field.[36] Hydrogen atoms were added to the repaired crystal structures at physiological pH (7.4) with the PROPKA[37] tool to optimize the hydrogen bond network provided by the Protein preparation tool in Schrodinger. Constrained energy minimizations were carried out on the full-atomic models, allowing the maxium RMSD for heavy atoms of 0.4 Å.

Ligand structure preparations. All ligand structures were obtained from the PubChem[38] online database. The LigPrep module in Schrodinger 2015 suite software was introduced for geometric optimization by using the OPLS_2005 force field. The ionization state of ligands was calculated with the Epik[39] tool employing Hammett and Taft methods in combination with ionization and tautomerization tools.[39]

Molecular simulations

Molecular dynamics simulations. Membrane systems were built with the membrane building tool g_membed[40] in Gromacs with the receptor crystal structure pre-aligned in the OPM (Orientations of Proteins in Membranes) database.[41] Pre-equilibrated 120 POPC lipids coupled with 92 000 TIP3P water molecules in a box of ~ 68 Å $\times$ 68 Å $\times$ 96 Å were used to build the protein/membrane/water system. We modeled the protein, lipids, water and ions using the CHARMM36 force field.[42,43] The ionization states of both protein and the ligand were assigned properly according to the results from Schrodinger software. Ligands were assigned to the CHARMM CgenFF force field.[44] The ligand geometry was submitted to the GAUSSIAN 09 program[45] for optimization at the Hartree-Fock 6-31G* level when generating force field parameters. The system was gradually heated from 0 K to 310 K followed by a 1 ns initial equilibration at constant volume with the temperature set at 310 K. Both the ligand molecule and protein backbone were restrained by a force constant of 10 kcal mol^{-1} Å^{-2} during this step. Next, an additional 40 ns restrained equilibration was performed at constant pressure and temperature (NPT ensemble; 310 K, 1 bar), and the force constant was tapped off gradually from 10 to 0 kcal mol^{-1}. The backbones of the proteins and the heavy atoms of the ligands were restrained during the equilibration steps. All bond lengths to hydrogen atoms were constrained with M-SHAKE. The van der Waals interactions were included using the switching function in the range of 10–12 Å.

Long-range electrostatic interactions were computed using the Particle Mesh Ewald (PME) summation scheme. All MD simulations were done using Gromacs.[46] The simulation parameter files were obtained from the CHARMM-GUI website.[47] The MD simulation results were analyzed using Gromacs[46] and VMD.[48] The solvent accessible surface area was calculated using Gromacs. Figures were prepared using PyMOL and Inkscape.[49]

Metadynamics simulations

Free-energy profiles of the systems were calculated using well-tempered metadynamics in Gromacs[46] V5.1.4 with Plumed[50] V2.2.1 patches. Metadynamics adds a history-dependent potential $V(s, t)$ to accelerate sampling of the specific collective variables (CVs) s (s_1, s_2, …, s_m).[51] $V(s,t)$ is usually constructed as the sum of multiple Gaussians centered along the trajectory of the collective variables (eqn (1)).

$$V(s,t) = \sum_{j=1}^{n} G(s,t_j) = \sum_{j=1}^{n} w_j \prod_{i=1}^{m} \exp\left(-\frac{\left[s_i(t) - s_i(t_j)\right]^2}{2\sigma^2} \right) \quad (1)$$

During the simulation, another Gaussian potential, whose location is dictated by the current values of the collective variables, is periodically added to $V(s,t)$.[51] In our simulations, distances between the quaternary N atom of ACh and a side chain oxygen of D$^{2.50}$ were assigned as the collective variables s_1, while the width of Gaussians, σ, was set as 0.05. The time interval, τ, was 0.09 ps. Well-tempered metadynamics involves adjusting the height, w_j, in a manner that depended on $V(s,t)$ where the initial height of Gaussians w was 0.05 kcal mol^{-1}, the simulation temperature was 310 K, and the sampling temperature ΔT was 298 K. The convergence of our simulations was judged by using the free energy difference between states A and B at 10 ns intervals. Once the results stopped changing over time, the simulation was considered as converged.[51]

Each metadynamics simulation was performed for 100 ns, and the results were analyzed upon convergence.

Analysis

Interaction fingerprint calculations. The IFP was performed using PLIP software.[52] PLIP detects frequent non-covalent protein–ligand interactions including hydrogen bonds, hydrophobic contacts, pi-stacking, pi-cation interactions, salt bridges, water bridges and halogen bonds.[52] We used 200 frames from the final 20 ns MD simulations for the IFP analysis. A plot was generated for the combination of all kinds of interactions. Parameters used for IFP calculations were kept as default.

New ligand binding site predictions and calculations. Inspections for the 39 unique crystal structures of the family A GPCRs were performed using ConCavity,[53] a widely used program for ligand site prediction. Both the grid size and resolutions were set to 0.1 Å. The volume of the newly discovered binding site for each receptor was calculated using the POVME[54] program. All the other analyses have been done using Gromacs and VMD. Figures were prepared using PyMOL and Inkscape.

Results and discussion

The entrance pathway of M3–ACh

We first present the analysis of eight 3 μs-long MD trajectories, each of which starts with either the antagonist tiotropium (TTP)-bound receptor or the ACh bound at the extracellular vestibule of the receptor (Fig. 1). The crystal structures of both inactive M3 (PDB: 4DAJ)[31] and M4 (PDB: 5DSG)[32] receptors were used directly for long time scale MD simulations.

In order to probe the binding mode of ACh in the crystal structure of the M3 receptor,[32] we first placed ACh at the entrance of the orthosteric binding pocket. Next, we submitted this M3–ACh structure to all-atom MD

 This journal is © The Royal Society of Chemistry 2018

Open Access Article. Published on 13 July 2018. Downloaded on 2/9/2026 10:52:24 AM.
This article is licensed under a Creative Commons Attribution-NonCommercial 3.0 Unported Licence.

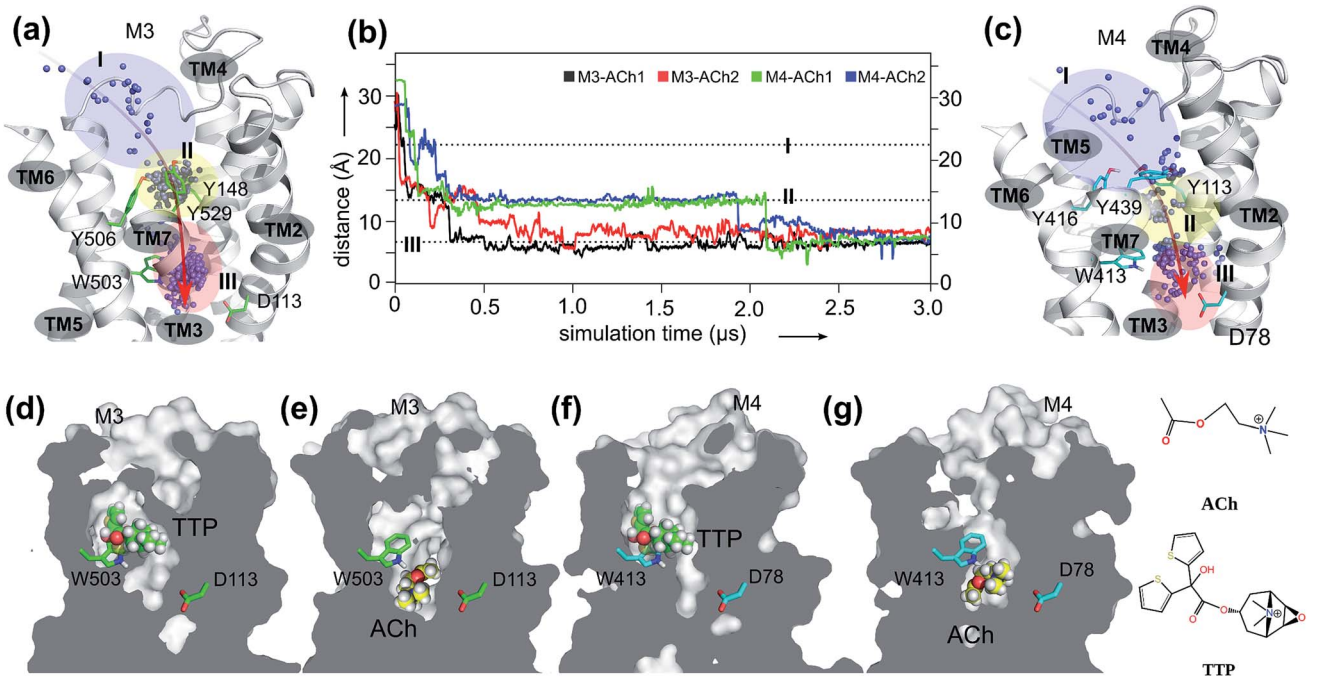

Fig. 1 The orthosteric and novel ligand binding sites of the muscarinic M3 and M4 receptors: (a) view of the M3 receptor showing the overall movement (red arrow) of the ACh molecule from the extracellular vestibule towards the orthosteric site of the receptor during the 3.0 μs MD simulations. Blue dots: mass centers of ACh at different simulation times showing its positional fluctuations within and between the extracellular vestibule, and orthosteric and novel ligand binding sites. Blue-shaded area: extracellular vestibule (zone I). Yellow-shaded area: orthosteric site (zone II). Red-shaded area: novel ligand site (zone III). (b) Time dependence of distance between ACh and oxygen at the sidechain of $D^{2.50}$ during two independent MD simulations for M3–ACh (black and red) and for M4–ACh (green and blue). The simulations reveal the atomistic details of ACh movements from the extracellular vestibule to the novel ligand binding site of both the M3 and the M4 receptors. (c) View of the M4 receptor showing the overall movement (red arrow) of the ACh molecule from the extracellular vestibule towards the orthosteric site of the receptor during the 3.0 μs MD simulations. Details as in A. (d–g) Cross-sections through the centers of M3 and M4 receptors revealing the final positions of the indicated ligands in different receptor binding sites.

simulations. At the beginning of the simulations (Fig. 1a and b, Movie S1†), ACh quickly enters (during 0.1–0.2 μs) the orthosteric site (zone II). Then, ACh is first confined to an aromatic cage characterized by residues $Y148^{3.33}$–$Y506^{6.51}$–$W503^{6.48}$–$Y529^{7.39}$. The volume of this pocket increases substantially from the initial 86 ± 2 Å^3 to 121 ± 3 Å^3 at the end of MD simulations. We superimposed a typical conformation of ACh (at ~0.1 μs snapshot) with the crystal structure of M3–TTP (4DAJ)[31] and found that the conformation of ACh is identical to that of TTP (Fig. S2†). This supports the hypothesis that the binding site of ACh in the two mAChRs might be the same as observed in the corresponding orthosteric structures (zone II in Fig. 1).[32]

However, during the 0.5–0.6 μs simulation period, ACh drifts into a much deeper new binding site (zone III) next to the highly conserved residue $D113^{2.50}$ (Fig. 1a and b). Consequently, the volume of zone III expands from the initial 57 ± 1 Å^3 to 148 ± 2 Å^3 at the end of the MD simulations. Interestingly, the side chain of the highly conserved $W503^{6.48}$ undergoes a corresponding flip, leaving space for the entry of the ACh molecule (Movie S1 and Fig. S3†). Such molecular switches have been observed both computationally[55] and experimentally[56] for adenosine receptors. In contrast, in the antagonist bound M3–TTP system, the antagonist TTP remains stably bound in the orthosteric site (zone II) (Fig. S4†) throughout the entire simulations. The volume of the orthosteric site (zone II) and that of the new binding site (zone III) is 162 ± 2 Å^3 and 56 ± 1 Å^3, respectively, in this M3–TTP system.

We also investigated the crystal structures of M2 receptors, bound to both an agonist (PDB: 4MQS)[57] and to an antagonist (PDB: 3UON).[58] The highly conserved $W400^{6.48}$ also underwent molecular switches (Fig. S5†) comparable with that found for the M3 receptor in the present work.

The binding mode of M3–ACh

To validate the above-mentioned findings, we first employed a protein–ligand interaction fingerprint (IFP) analysis for M3–ACh (Fig. 2a) followed by mutagenesis experiments (Fig. 2c and S6†) for M3 receptors. IFP was performed based on the snapshots from the final 20 ns MD simulations. IFP identified that ACh frequently interacts with five residues including $A112^{2.57}$, $I116^{2.53}$, $A150^{3.35}$, $S154^{3.39}$ and $W503^{6.48}$. Thus, we next mutated $I116^{2.53}$, $S154^{3.39}$ and $W503^{6.48}$ individually into Ala residues. The signaling assays of the different M3 receptor mutants showed noticeably decreased activities (Table 2). Specifically, the ACh induced activation of Gq by the wild-type (WT) M3 shows an EC_{50} of 0.5 μM, whereas receptor mutants $I116^{2.53}$A, $S154^{3.39}$A and $W503^{6.48}$A show an EC_{50} of 1.8 μM, 1.7 μM and 4.0 μM, respectively.

Open Access Article. Published on 13 July 2018. Downloaded on 2/9/2026 10:52:24 AM.
This article is licensed under a Creative Commons Attribution-NonCommercial 3.0 Unported Licence.

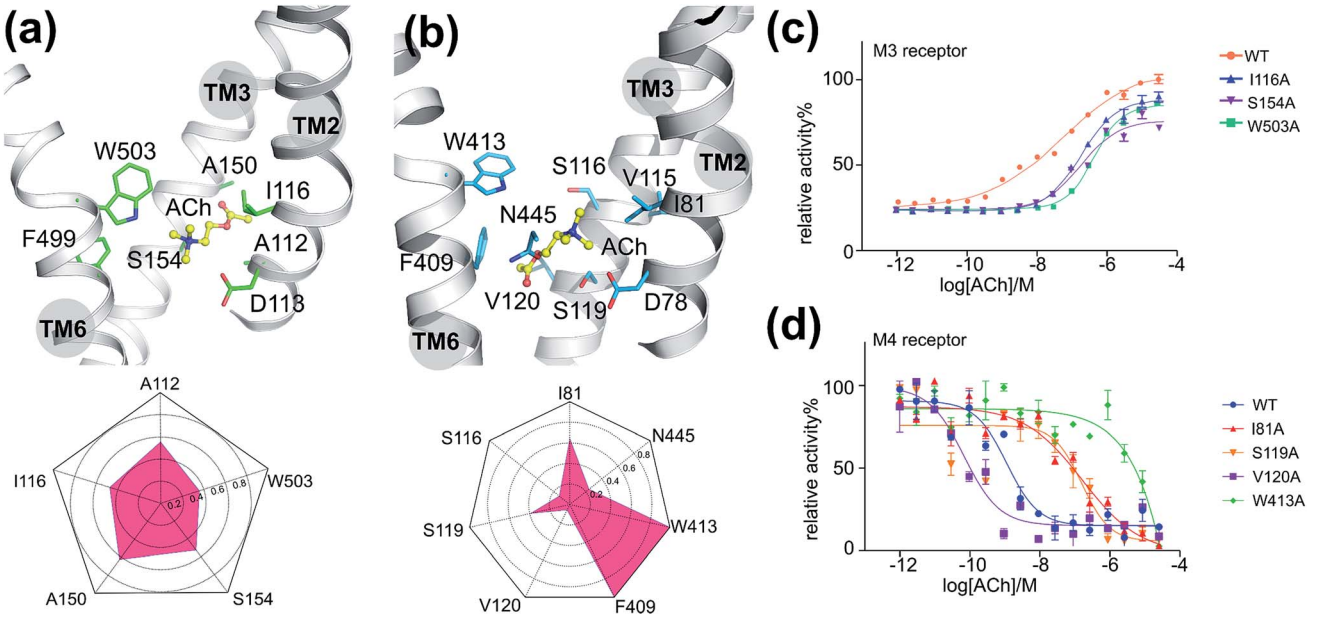

Fig. 2 The interactions between ACh and mAChRs: (a) the final binding position of ACh in the MD simulation of M3–ACh (top) and the protein–ligand interaction fingerprint (IFP) of M3–ACh calculated based on the final 20 ns MD simulations (bottom). Yellow balls-and-sticks: ACh molecule. Green sticks: side chains of M3. (b) The final binding position of ACh in the MD simulation of M4–ACh (top) and the protein–ligand interaction fingerprint (IFP) of M4–ACh calculated based on the final 20 ns MD simulations (bottom). Yellow balls-and-sticks: ACh molecule. Green sticks: side chains of M3. (c) Gq protein activation assay of M3 receptor mutants as a function of ACh concentration. (d) Gi protein activation assay of M4 receptor mutants as a function of ACh concentration.

Table 2 ACh induced activation of G proteins in different mutants of M3 (Gq protein) and M4 receptors (Gi protein)

M3	WT	I116A	S154A	W503A	M4	WT	I81A	S116A	S119A	V120A	F409A	W413A	N445A
EC_{50} (μM)	0.5	1.8	1.7	4.0	EC_{50} (nM)	1.6	174.0	192.4	23.0	0.1	6.0	NA	12.9

The entrance pathway of ACh into the M4 receptor

Furthermore, we extended our investigations to M4 receptors and performed similar MD simulations as before for the M3 receptor. We found that ACh also diffuses into a much deeper binding site of the M4 receptor next to the highly conserved $D78^{2.50}$, however at a different time scale. It first spends a longer period (up to 0.25 μs) in the extracellular vestibule (zone I) before reaching the orthosteric site (zone II) at 0.3–0.4 μs (Movie S2,† Fig. 1b, c, f and g). After stabilizing for a considerable time in this region, ACh continues moving from the orthosteric site (zone II) to the new binding site (zone III) at 1.9–2.1 μs, where it remains firmly bound until the end of the MD simulations. The volumes of zone II and zone III in M4–ACh also change significantly during the MD simulations. The initial volume of zone II of 92 ± 1 $Å^3$ expands to 140 ± 2 $Å^3$ in M4–ACh at the end of MD simulations. The corresponding volume changes in zone III are from 46 ± 1 $Å^3$ to 139 ± 2 $Å^3$. These results are in complete agreement with our previous studies that agonist binding leads to an expansion of the extracellular orthosteric regions.[19]

The binding mode of M4–ACh

As described for the M3 receptor, we also applied both IFP analysis and mutagenesis experiments to validate our findings for M4–ACh. IFP indicated that there are more residues involved in M4–ACh than in M3–ACh interactions, including $I81^{2.53}$, $S116^{3.36}$, $S119^{3.39}$, $V120^{3.40}$, $F409^{6.44}$, $W413^{6.48}$ and $N445^{7.45}$. We then employed mutagenesis to all these residues individually which led to dramatic effects on the M4 activity in each case (Fig. 2d and Table 2). Specifically, $W413^{6.48}$A led to a completely inactive M4 mutant towards ACh; the activities of mutants $I81^{2.53}$A and $S116^{3.36}$A also decreased 100 times compared to that of the WT M4 receptor. In contrast, the mutations of $F409^{6.44}$A, $S119^{3.39}$A and $N445^{7.45}$A show smaller effects with 4–14 fold decreased activities (Fig. S6† and Table 2), while the mutation $V120^{3.40}$A increased the EC_{50} by 16 fold.

The binding mode differences between M3–ACh and M4–ACh

Moreover, we found that the final binding positions of ACh in M3 and M4 are different (Fig. 1b and 2a) at the end of MD simulations. In M3–ACh, ACh forms (i) hydrophobic interactions with $A112^{2.57}$, $I116^{2.53}$, $A150^{3.35}$, $S154^{3.39}$ and $W503^{6.48}$, and (ii) ionic interactions between its quaternary nitrogen and highly conserved residue $D113^{2.50}$. The flexible acetyl tail sits in the void space between TM2 and TM3, next to $A150^{3.35}$. In contrast, ACh flipped by 180 degrees in M4–ACh compared to the binding configuration in M3–ACh: the ester head group comes into

This journal is © The Royal Society of Chemistry 2018

Open Access Article. Published on 13 July 2018. Downloaded on 2/9/2026 10:52:24 AM.
This article is licensed under a Creative Commons Attribution-NonCommercial 3.0 Unported Licence.

contact with V120$^{3.40}$, F409$^{6.44}$, W413$^{6.48}$ and N445$^{7.45}$, whereas the quaternary nitrogen interacts with I81$^{2.53}$, V115$^{3.35}$, S116$^{3.36}$ and S119$^{3.39}$. Such variations probably stem from the differences in position 3.35: A150$^{3.35}$ in M3 and V115$^{3.35}$ in M4. All other residues in the ligand binding sites of M3 and M4 are exactly the same. The smaller A150$^{3.35}$ in M3 creates more space which can accommodate the ester group but it is still too small to fit the bigger quaternary nitrogen (Fig. 2a). When ACh approaches this region of M3, the ester group can be stabilized between TM2 and TM3. The V115$^{3.35}$ in M4 takes up more space than A150$^{3.35}$ in M3. The hydrophobic sidechain of V115$^{3.35}$ favors the hydrophobic interactions with $-CH_3$ groups from the quaternary nitrogen. The positively charged quaternary nitrogen could be further stabilized in this area by the highly conserved D$^{2.50}$ through ionic interactions (Fig. 2b) in both M3 and M4.

The kinetic differences in binding M3–ACh and M4–ACh

Noticeably, the time scales for ACh entering the new ligand sites of M3 (0.5 μs) and M4 (2.0 μs) are quite different. To understand these observations, we introduced well-tempered metadynamics MD.

MD simulations are employed to explore the free energy profile of the entire binding processes of ACh for each receptor (Fig. S7†). We defined the distance between the quaternary nitrogen of ACh and the carboxyl group of D$^{2.50}$ as the collective variable (CV). The results of our metadynamics MD simulations are in complete agreement with our unbiased long-time scale MD simulations, revealing three distinct energy states: the transient locations of ACh in the extracellular vestibule zone I, the orthosteric site zone II, and the final stable new binding site zone III (Fig. 1a and c). Interestingly, we found that the free energy barrier for the movement between zone I and zone II is very low (~0.1–0.5 kcal mol^{-1}) facilitating the diffusion of ACh from the extracellular vestibule towards the orthosteric site. However, the free energy barrier between zone II and zone III is around 1.6 ± 0.1 kcal mol^{-1} for M3–ACh and 3.8 ± 0.2 kcal mol^{-1} for M4–ACh. These large differences in the energy barriers explain why the diffusion of ACh from zone II to zone III takes longer in the M4 than in the M3 receptor. This is probably due to the fact that in position 3.35 in TM3, M4 has a larger residue (Val) than M3 (Ala) which creates a higher steric barrier for ACh to enter the deep binding pocket.

New putative binding sites in crystal structures

To validate whether our newly discovered ligand binding pocket is also applicable to other GPCRs, we analyzed more than 200 crystal structures of family A GPCRs comprising bound ligands (Fig. 3a and S8†) from GPCRDB.[59] In the recently deposited crystal structure of the leukotriene B4 (LTB4) receptor (PDB:

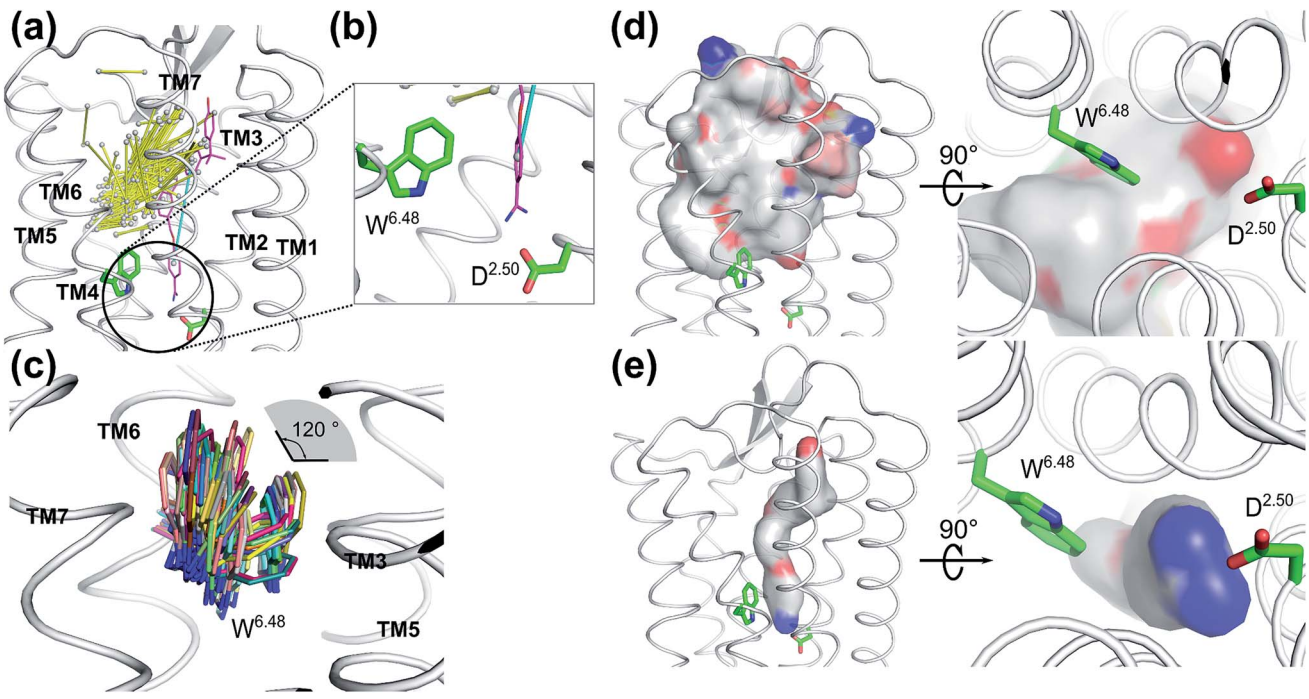

Fig. 3 Ligands in the vicinity of the classic orthosteric binding site of all family A GPCR crystal structures available in the protein structure databank: (a) ligands observed in the vicinity of the classic orthosteric binding sites of family A GPCR crystal structures after superimposing the receptor coordinates. Yellow stick: each ligand is represented by its mass center of the two heaviest moieties. Purple stick: inverse agonist BIIL260 observed in the leukotriene B4 (LTB4) receptor. Cyan stick: the mass center vector of BIIL260. (b) Extended view of the encircled region in A. An inverse agonist (purple stick) bound to the leukotriene B4 receptor in the crystal structure (PDB: 5X33) is extending to the region which was newly discovered in the present work. (c) Superimposed positions of the highly conserved W$^{6.48}$ in the crystal structures of family A GPCRs. The χ1 angle of W$^{6.48}$ can undergo 120° flipping which has a great impact on the volume changes induced by ligand binding to the adjacent orthosteric site. (d) Superposition of the normalized electrostatic potential surfaces of all ligands in the GPCR crystal structures deposited in the structure data bank, excluding ligand BIIL260. Left: side view and right: enlarged bottom view. White: hydrophobic areas. Red: negatively charged areas. Blue: positively charged areas. (e) The normalized electrostatic potential surface for BIIL260 in the leukotriene B4 (LTB4) receptor. White: hydrophobic area. Red: negatively charged area. Blue: positively charged area.

Open Access Article. Published on 13 July 2018. Downloaded on 2/9/2026 10:52:24 AM.
This article is licensed under a Creative Commons Attribution-NonCommercial 3.0 Unported Licence.

5X33),[60] the benzamidine moiety of the inverse agonist BIIL260 occupies the same new ligand binding site which we have discovered for ACh in the M3 and M4 receptors in this work, contacting $D66^{2.50}$ through an ion-lock interaction (Fig. 3b). Moreover, the highly conserved $W236^{6.48}$ in the LTB4 receptor was also found undergoing molecular switches induced by the binding of BIIL260 (ref. 60) (Fig. S5b†). Interestingly, the mass center vector of BIIL260 was found at a unique position compared to those of all other ligands resolved in the crystal structures: it directly pointed to the highly conserved $D^{2.50}$, whereas the mass center vectors of all other ligands are located far away from $D^{2.50}$. The normalized electrostatic potential surface area (Fig. 3d and e) reveals positively charged surfaces for BIL260 which facilitate the ion lock interaction with $D^{2.50}$, very similar to that observed for ACh (Fig. 2a and b). In contrast, all other ligands resolved in GPCR crystal structures have either hydrophobic or negatively charged surfaces in the same region, which is very unfavorable to interact with the negatively charged $D^{2.50}$. This is probably the reason why all the other ligands do not bind to this newly discovered pocket. We propose that a well-designed ligand with both a proper mass center vector and a positively charged head group might be able to move into this new binding pocket of the receptors addressed in Fig. 3.

New putative ligand binding sites in other GPCRs

To further investigate whether our newly discovered binding sites of the M3 and M4 receptors are also present in other GPCRs, we inspected 39 unique crystal structures of family A GPCRs (Fig. S9† and 4). Here, we used ConCavity[53] for ligand site prediction and found that all investigated GPCRs also potentially have an additional ligand binding site between positions 6.48 and 2.50. To further characterize this binding site for each receptor, we calculated the volume of this region using the POVME tool.[54] The calculation indicates that most receptors have a reasonable volume of 30–85 $Å^3$ for this newly discovered binding site. Specifically, $5\text{-}HT_{2B}$, M1, AT1R, APJ, H1R, κOR, OX1, $P2Y_{12}$, PAR2 and LPA6 receptors all have a relatively small binding site, with volumes in the range of 30–40 $Å^3$. In contrast, the volumes of the putative additional binding pockets in $5\text{-}HT_{1B}$, $5\text{-}HT_{2C}$, AT2R, M3, M4, C5aR, CCR9, CXCR4, D2R, D3R, ETB, LPA1, δOR, μOR, RHO, OX2, $P2Y_1$ and $S1P_1$ are in the range of 40–60 $Å^3$. Finally, receptors $A_{1A}R$, M2, β_1AR, β_2AR, CCR2, CCR5, D4R, NOP, LTB4 and CX3CL1 have distinct larger putative ligand binding pockets, in the range of 60–85 $Å^3$ which is the size of two water molecules or a small ligand. However, the volume of this pocket in each receptor might be substantially enlarged upon agonist binding as indicated by our MD simulation in this work.

Interestingly, the final position of the sidechain of $W236^{6.48}$ observed in LTB4 is identical to that of the corresponding $W400^{6.48}$ in the agonist bound activated M2 receptor (Fig. S5†): both sidechains underwent rotamer changes. The superimposed structures of all family A GPCRs in the data bank indicate that the $\chi1$ angles of $W^{6.48}$ can undergo changes as large as 120° (Fig. 3C).

Although these new potential orthosteric sites are small in size, their volumes can be increased accordingly after the rotation of the highly conserved $W^{6.48}$ upon proper ligand binding. This is further confirmed by our previously performed long time scale MD simulations on FPR1,[61] $S1P_1R$[62] and $A_{2A}R$,[55] of which the volumes for newly discovered binding sites increased from 46 ± 1 $Å^3$ to 132 ± 2 $Å^3$, 58 ± 3 $Å^3$ to 163 ± 3 $Å^3$ and 51 ± 2 $Å^3$ to 156 ± 2 $Å^3$, respectively, after the flipping of the corresponding $W^{6.48}$ sidechains.

Finally, we inspected the conservations of residues in this newly discovered cavity, based on all resolved crystal structures of family A GPCRs (Fig. 5). We found that most residues of this

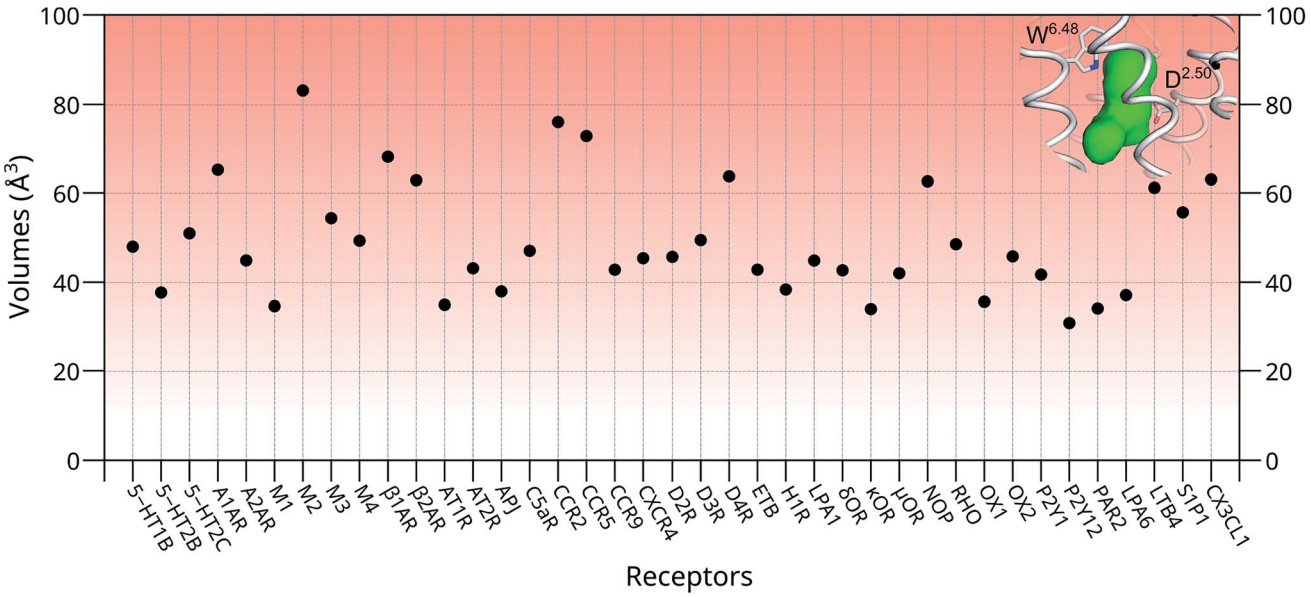

Fig. 4 The volumes of the newly discovered putative binding sites in the absence of ligands in 39 unique structures of family A GPCRs. Green region: the volume of the newly discovered binding site located between the highly conserved $W^{6.48}$ and $D^{2.50}$ of the M4 receptor. Most of the proposed empty binding pockets have volumes >40 $Å^3$.

 This journal is © The Royal Society of Chemistry 2018

Open Access Article. Published on 13 July 2018. Downloaded on 2/9/2026 10:52:24 AM.
This article is licensed under a Creative Commons Attribution-NonCommercial 3.0 Unported Licence.

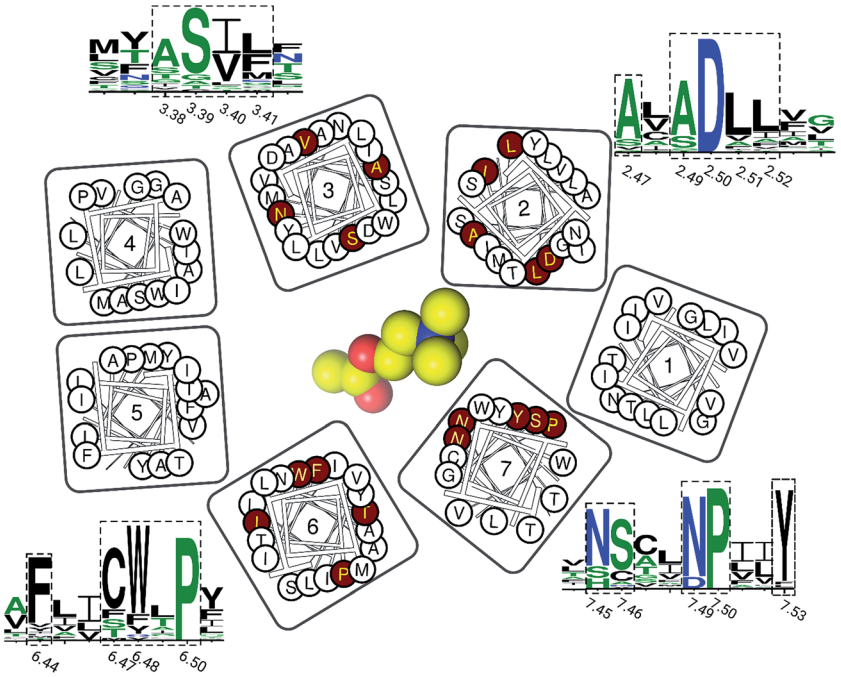

Fig. 5 The residue conservations of the newly discovered binding pocket. The logos of conserved residues (highlighted in dashed rectangles) are based on the multiple sequence alignment on all resolved crystal structures of family A GPCRs. The height of each single letter in the depicted sequence segments scales with its conservation at the corresponding position. Their positions on TM helices are highlighted in red circles in the residue diagrams. An ACh molecule is represented as a yellow sphere to show its relative position.

binding pocket are highly conserved, including residues in positions 2.47, 2.49, 2.50, 2.51, 2.52, 3.38, 3.39, 3.40, 3.41, 6.44, 6.47, 6.48, 4.50, 7.45, 7.46, 7.49 and 7.50. This observation indicates that this newly discovered binding pocket indeed exists and is highly conserved in most resolved structures of family A GPCRs.

Conclusions

In summary, discovering a new ligand binding site for a particular protein is an essential first step for structure-based drug discovery. Computer modelling is a very useful tool for modern structural biology. By using MD simulations coupled with mutagenesis experiments, we identified a potential new binding pocket for the endogenous agonist ACh in M3 and M4 muscarinic acetylcholine receptors (Fig. 6), whereas the antagonist molecule TTP is trapped in the traditional orthosteric binding site. The small and flexible ACh diffuses deeper inside the receptor towards a new binding site next to the highly conserved residue $D^{2.50}$. Molecular switches take place at the highly conserved $W^{6.48}$ induced by ligand binding. As the volumes of both orthosteric and new sites expand upon the binding of ACh in the case of the M3 and M4 receptors, we predict similar volume changes in the other putative binding sites upon ligand binding.

In our present work, we systematically inspected the binding sites of all published family A GPCR crystal structures. We discovered that most residues, forming this newly discovered pocket, are highly conserved. We also found that most GPCRs potentially comprise these newly discovered binding sites, which have been confirmed by both a recently resolved crystal structure of LTB4 and computer modeling. Our newly uncovered ligand binding sites open new opportunities for GPCR drug discovery and design.

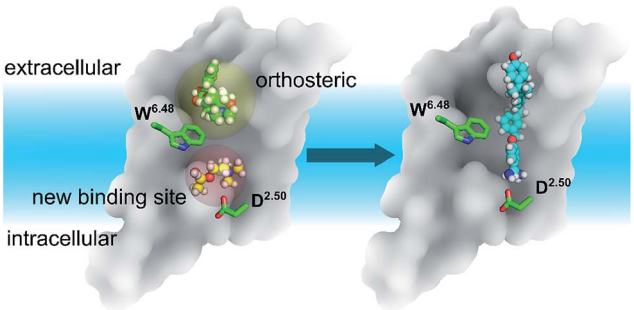

Fig. 6 The relative positions of the classic orthosteric site (left panel, yellow shaded sphere) and the newly discovered ligand binding site (left panel, red shaded sphere) located between $W^{6.48}$ and $D^{2.50}$. A new drug molecule (right panel) can be designed, based on the newly discovered site in this work. Green balls-and-sticks: the TTP molecule observed in the crystal structures of M3 & M4 receptors. Yellow balls-and-sticks: ACh observed in this work. Cyan stick: an inverse agonist observed in the crystal structure of the LTB4 receptor.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

Support from the Interdisciplinary Centre for Mathematical and Computational Modelling in Warsaw (grants GB70-3 & GB71-3

Open Access Article. Published on 13 July 2018. Downloaded on 2/9/2026 10:52:24 AM. This article is licensed under a Creative Commons Attribution-NonCommercial 3.0 Unported Licence.

to SY and G07-13 & GA65-23 to SF), the National Center of Science, Poland (grant 2013/08/M/ST6/00788) to SF), the European Community (project SynSignal, grant no. FP7-KBBE-2013-613879) and the EPFL to HV is acknowledged. Additional funding came from the Arnold and Mabel Beckman Foundation to KP. HV and SF participated in the European COST Action CM1207 (GLISTEN). KP is the John H. Hord Professor of Pharmacology. Dedicated to Prof. Horst Vogel on the occasion of his 70th birthday.

References

1 J. Skolnick, M. Gao, A. Roy, B. Srinivasan and H. Zhou, *Bioorg. Med. Chem. Lett.*, 2015, **25**, 1163–1170.
2 M. Marti-Solano, D. Schmidt, P. Kolb and J. Selent, *Drug discovery today*, 2016, **21**, 625–631.
3 Z. Guo, B. Li, L. T. Cheng, S. Zhou, J. A. McCammon and J. Che, *J. Chem. Theory Comput.*, 2015, **11**, 753–765.
4 A. Stank, D. B. Kokh, J. C. Fuller and R. C. Wade, *Acc. Chem. Res.*, 2016, **49**, 809–815.
5 S. J. Teague, *Nat. Rev. Drug Discovery*, 2003, **2**, 527–541.
6 D. Guo, A. C. Pan, R. O. Dror, T. Mocking, R. Liu, L. H. Heitman, D. E. Shaw and A. P. IJzerman, *Mol. Pharmacol.*, 2016, **89**, 485–491.
7 A. S. Hauser, M. M. Attwood, M. Rask-Andersen, H. B. Schioth and D. E. Gloriam, *Nat. Rev. Drug Discovery*, 2017, **16**, 829–842.
8 S. C. Erlandson, C. McMahon and A. C. Kruse, *Annu. Rev. Biophys.*, 2018, **47**, 1–18.
9 J. J. Ziarek, F. C. Peterson, B. L. Lytle and B. F. Volkman, *Methods Enzymol.*, 2011, **493**, 241–275.
10 M. McCarthy, P. Prakash and A. A. Gorfe, *Acta Biochim. Biophys. Sin.*, 2016, **48**, 3–10.
11 J. Huang, S. K. Lakkaraju, A. Coop and A. D. MacKerell Jr, *J. Phys. Chem. B*, 2016, **120**, 11897–11904.
12 S. K. Lakkaraju, J. A. Lemkul, J. Huang and A. D. MacKerell Jr, *J. Comput. Chem.*, 2016, **37**, 416–425.
13 N. Van Eps, L. N. Caro, T. Morizumi, A. K. Kusnetzow, M. Szczepek, K. P. Hofmann, T. H. Bayburt, S. G. Sligar, O. P. Ernst and W. L. Hubbell, *Proc. Natl. Acad. Sci. U. S. A.*, 2017, **114**, E3268–E3275.
14 S. Wolf, N. Jovancevic, L. Gelis, S. Pietsch, H. Hatt and K. Gerwert, *Sci. Rep.*, 2017, **7**, 16007.
15 D. Hilger, M. Masureel and B. K. Kobilka, *Nat. Struct. Mol. Biol.*, 2018, **25**, 4–12.
16 L. Ye, N. Van Eps, M. Zimmer, O. P. Ernst and R. S. Prosser, *Nature*, 2016, **533**, 265–268.
17 W. Liu, E. Chun, A. A. Thompson, P. Chubukov, F. Xu, V. Katritch, G. W. Han, C. B. Roth, L. H. Heitman, A. P. IJzerman, V. Cherezov and R. C. Stevens, *Science*, 2012, **337**, 232–236.
18 D. Zhang, Z. G. Gao, K. Zhang, E. Kiselev, S. Crane, J. Wang, S. Paoletta, C. Yi, L. Ma, W. Zhang, G. W. Han, H. Liu, V. Cherezov, V. Katritch, H. Jiang, R. C. Stevens, K. A. Jacobson, Q. Zhao and B. Wu, *Nature*, 2015, **520**, 317–321.
19 S. Yuan, H. C. Chan, H. Vogel, S. Filipek, R. C. Stevens and K. Palczewski, *Angew. Chem.*, 2016, **55**, 10331–10335.
20 A. Srivastava, J. Yano, Y. Hirozane, G. Kefala, F. Gruswitz, G. Snell, W. Lane, A. Ivetac, K. Aertgeerts, J. Nguyen, A. Jennings and K. Okada, *Nature*, 2014, **513**, 124–127.
21 C. Oswald, M. Rappas, J. Kean, A. S. Dore, J. C. Errey, K. Bennett, F. Deflorian, J. A. Christopher, A. Jazayeri, J. S. Mason, M. Congreve, R. M. Cooke and F. H. Marshall, *Nature*, 2016, **540**, 462–465.
22 N. Robertson, M. Rappas, A. S. Dore, J. Brown, G. Bottegoni, M. Koglin, J. Cansfield, A. Jazayeri, R. M. Cooke and F. H. Marshall, *Nature*, 2018, **553**, 111–114.
23 M. De Vivo, M. Masetti, G. Bottegoni and A. Cavalli, *J. Med. Chem.*, 2016, **59**, 4035–4061.
24 D. Sabbadin and S. Moro, *J. Chem. Inf. Model.*, 2014, **54**, 372–376.
25 K. A. Marino and M. Filizola, *Methods Mol. Biol.*, 2018, **1705**, 351–364.
26 R. O. Dror, A. C. Pan, D. H. Arlow, D. W. Borhani, P. Maragakis, Y. Shan, H. Xu and D. E. Shaw, *Proc. Natl. Acad. Sci. U. S. A.*, 2011, **108**, 13118–13123.
27 A. Strasser, H. J. Wittmann and R. Seifert, *Trends Pharmacol. Sci.*, 2017, **38**, 717–732.
28 Y. Miao and J. A. McCammon, *Curr. Opin. Struct. Biol.*, 2016, **41**, 83–89.
29 A. S. Rose, M. Elgeti, U. Zachariae, H. Grubmuller, K. P. Hofmann, P. Scheerer and P. W. Hildebrand, *J. Am. Chem. Soc.*, 2014, **136**, 11244–11247.
30 A. S. Rose, U. Zachariae, H. Grubmuller, K. P. Hofmann, P. Scheerer and P. W. Hildebrand, *PLoS One*, 2015, **10**, e0143399.
31 A. C. Kruse, J. Hu, A. C. Pan, D. H. Arlow, D. M. Rosenbaum, E. Rosemond, H. F. Green, T. Liu, P. S. Chae, R. O. Dror, D. E. Shaw, W. I. Weis, J. Wess and B. K. Kobilka, *Nature*, 2012, **482**, 552–556.
32 D. M. Thal, B. Sun, D. Feng, V. Nawaratne, K. Leach, C. C. Felder, M. G. Bures, D. A. Evans, W. I. Weis, P. Bachhawat, T. S. Kobilka, P. M. Sexton, B. K. Kobilka and A. Christopoulos, *Nature*, 2016, **531**, 335–340.
33 V. Isberg, C. de Graaf, A. Bortolato, V. Cherezov, V. Katritch, F. H. Marshall, S. Mordalski, J. P. Pin, R. C. Stevens, G. Vriend and D. E. Gloriam, *Trends Pharmacol. Sci.*, 2015, **36**, 22–31.
34 N. Eswar, B. Webb, M. A. Marti-Renom, M. S. Madhusudhan, D. Eramian, M. Y. Shen, U. Pieper and A. Sali, *Current protocols in protein science*, ed. John E. Coligan, 2007, ch. 2, Unit 2 9.
35 D. J. Mandell, E. A. Coutsias and T. Kortemme, *Nat. Methods*, 2009, **6**, 551–552.
36 D. Shivakumar, J. Williams, Y. J. Wu, W. Damm, J. Shelley and W. Sherman, *J. Chem. Theory Comput.*, 2010, **6**, 1509–1519.
37 C. R. Sondergaard, M. H. M. Olsson, M. Rostkowski and J. H. Jensen, *J. Chem. Theory Comput.*, 2011, **7**, 2284–2295.
38 Y. L. Wang, J. W. Xiao, T. O. Suzek, J. Zhang, J. Y. Wang, Z. G. Zhou, L. Y. Han, K. Karapetyan, S. Dracheva,

 This journal is © The Royal Society of Chemistry 2018

Open Access Article. Published on 13 July 2018. Downloaded on 2/9/2026 10:52:24 AM.
This article is licensed under a Creative Commons Attribution-NonCommercial 3.0 Unported Licence.

B. A. Shoemaker, E. Bolton, A. Gindulyte and S. H. Bryant, *Nucleic Acids Res.*, 2012, **40**, D400–D412.

39 J. R. Greenwood, D. Calkins, A. P. Sullivan and J. C. Shelley, *J. Comput.-Aided Mol. Des.*, 2010, **24**, 591–604.

40 M. G. Wolf, M. Hoefling, C. Aponte-Santamaria, H. Grubmuller and G. Groenhof, *J. Comput. Chem.*, 2010, **31**, 2169–2174.

41 A. L. Lomize, I. D. Pogozheva and H. I. Mosberg, *J. Chem. Inf. Model.*, 2011, **51**, 918–929.

42 J. B. Klauda, R. M. Venable, J. A. Freites, J. W. O'Connor, D. J. Tobias, C. Mondragon-Ramirez, I. Vorobyov, A. D. MacKerell and R. W. Pastor, *J. Phys. Chem. B*, 2010, **114**, 7830–7843.

43 J. Huang, S. Rauscher, G. Nawrocki, T. Ran, M. Feig, B. L. de Groot, H. Grubmuller and A. D. MacKerell Jr, *Nat. Methods*, 2017, **14**, 71–73.

44 K. Vanommeslaeghe, E. P. Raman and A. D. MacKerell Jr, *J. Chem. Inf. Model.*, 2012, **52**, 3155–3168.

45 M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery Jr, J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, O. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski and D. J. Fox, *Gaussian 09, Revision A.1*, Gaussian, Inc., Wallingford CT, 2009.

46 S. Pronk, S. Pall, R. Schulz, P. Larsson, P. Bjelkmar, R. Apostolov, M. R. Shirts, J. C. Smith, P. M. Kasson, D. van der Spoel, B. Hess and E. Lindahl, *Bioinformatics*, 2013, **29**, 845–854.

47 S. Jo, T. Kim, V. G. Iyer and W. Im, *J. Comput. Chem.*, 2008, **29**, 1859–1865.

48 W. Humphrey, A. Dalke and K. Schulten, *J. Mol. Graphics Modell.*, 1996, **14**, 33–38.

49 S. Yuan, H. C. Chan, S. Filipek and H. Vogel, *Structure*, 2016, **24**, 2041–2042.

50 M. Bonomi, D. Branduardi, G. Bussi, C. Camilloni, D. Provasi, P. Raiteri, D. Donadio, F. Marinelli, F. Pietrucci, R. A. Broglia and M. Parrinello, *Comput. Phys. Commun.*, 2009, **180**, 1961–1972.

51 A. Barducci, G. Bussi and M. Parrinello, *Phys. Rev. Lett.*, 2008, **100**.

52 S. Salentin, S. Schreiber, V. J. Haupt, M. F. Adasme and M. Schroeder, *Nucleic Acids Res.*, 2015, **43**, W443–W447.

53 J. A. Capra, R. A. Laskowski, J. M. Thornton, M. Singh and T. A. Funkhouser, *PLoS Comput. Biol.*, 2009, **5**, e1000585.

54 J. D. Durrant, L. Votapka, J. Sorensen and R. E. Amaro, *J. Chem. Theory Comput.*, 2014, **10**, 5047–5056.

55 S. Yuan, Z. Hu, S. Filipek and H. Vogel, *Angew. Chem.*, 2014, **54**, 556–559.

56 M. T. Eddy, M. Y. Lee, Z. G. Gao, K. L. White, T. Didenko, R. Horst, M. Audet, P. Stanczak, K. M. McClary, G. W. Han, K. A. Jacobson, R. C. Stevens and K. Wuthrich, *Cell*, 2018, **172**, 68–80.

57 A. C. Kruse, A. M. Ring, A. Manglik, J. Hu, K. Hu, K. Eitel, H. Hubner, E. Pardon, C. Valant, P. M. Sexton, A. Christopoulos, C. C. Felder, P. Gmeiner, J. Steyaert, W. I. Weis, K. C. Garcia, J. Wess and B. K. Kobilka, *Nature*, 2013, **504**, 101–106.

58 K. Haga, A. C. Kruse, H. Asada, T. Yurugi-Kobayashi, M. Shiroishi, C. Zhang, W. I. Weis, T. Okada, B. K. Kobilka, T. Haga and T. Kobayashi, *Nature*, 2012, **482**, 547–551.

59 V. Isberg, S. Mordalski, C. Munk, K. Rataj, K. Harpsoe, A. S. Hauser, B. Vroling, A. J. Bojarski, G. Vriend and D. E. Gloriam, *Nucleic Acids Res.*, 2016, **44**, D356–D364.

60 T. Hori, T. Okuno, K. Hirata, K. Yamashita, Y. Kawano, M. Yamamoto, M. Hato, M. Nakamura, T. Shimizu, T. Yokomizo, M. Miyano and S. Yokoyama, *Nat. Chem. Biol.*, 2018, **14**, 262–269.

61 S. Yuan, U. Ghoshdastider, D. Latek, B. Trzaskowski, A. Debinski and S. Filipek, *PLoS One*, 2012.

62 S. Yuan, R. Wu, D. Latek, B. Trzaskowski and S. Filipek, *PLoS Comput. Biol.*, 2013, **9**, e1003261.