



Delineating the conformational landscape and intrinsic properties of the angiotensin II type 2 receptor using a computational study



Xiaoliang Cong^{a,1}, Xiaogang Zhang^{b,1}, Xin Liang^{a,1}, Xinheng He^{c,1}, Yehua Tang^a, Xing Zheng^d, Shaoyong Lu^{c,*}, Jiayou Zhang^{a,*}, Ting Chen^{a,*}

^a Department of Cardiology, Shanghai Changzheng Hospital, the Second Affiliated Hospital of Naval Medical University, Shanghai 200003, China

^b Department of Cardiology, Shanghai University of Medicine & Health Sciences Affiliated Zhoupu Hospital, Shanghai 201318, China

^c Medicinal Chemistry and Bioinformatics Centre, Shanghai Jiao Tong University, School of Medicine, Shanghai 200025, China

^d Department of Cardiology, Changhai Hospital, Naval Medical University, Shanghai 200433, China

ARTICLE INFO

Article history:

Received 9 March 2022

Received in revised form 4 May 2022

Accepted 6 May 2022

Available online 10 May 2022

Keywords:

G protein-coupled receptors

AT₂R

MD simulations

Conformational dynamics

Cryptic pocket

ABSTRACT

As a key regulator for the renin-angiotensin system, a class A G protein-coupled receptor (GPCR), AngII type 2 receptor (AT₂R), plays a pivotal role in the homeostasis of the cardiovascular system. Compared with other GPCRs, AT₂R has a unique antagonist-bound conformation and its mechanism is still an enigma. Here, we applied combined dynamic and evolutionary approaches to investigate the conformational space and intrinsic properties of AT₂R. With molecular dynamic simulations, Markov State Models, and statistics coupled analysis, we captured the conformational landscape of AT₂R and identified its unicity from both dynamical and evolutionary viewpoints. A cryptic pocket was also discovered in the intermediate state during conformation transitions. These findings offer a deeper understanding of the AT₂R mechanism at an atomic level and provide hints for the design of novel AT₂R modulators.

© 2022 The Author(s). Published by Elsevier B.V. on behalf of Research Network of Computational and Structural Biotechnology. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The renin-angiotensin system (RAS) plays a critical role in maintaining the homeostasis of the cardiovascular system and its dysfunction leads to hypertension, heart disease and nephropathy [1–4]. The main function of RAS is the direct regulation of the blood pressure via the octapeptide angiotensin II (AngII), whose receptors are two subtypes of class A G protein-coupled receptors (GPCRs), AngII type 1 receptor (AT₁R) and AngII type 2 receptor (AT₂R) [1,5,6]. Although sharing 34% sequence similarity, the two receptors show contrast function in cardiovascular system regulation, while the activation of AT₁R increases the blood pressure, the activation of AT₂R decreases it [7–9]. Moreover, AT₂R has abundant functions such as inhibiting the cardiomyocytes autophagy [10], promoting vascular growth [11] causing anti-inflammatory effects [12,13] and improving insulin resistance and metabolism [14,15]. Thus, the regulation of AT₂R is a promising field in new drug development and is capable of treating a number of pathological processes.

Normally, class A GPCRs have seven transmembrane (TM) helices linked by three extra- and three intracellular loops (ECLs, ICLs) with the helix 8 (H8) in the intracellular C-terminus [16,17]. During its activation process, the endogenous AngII enters the orthosteric site located in the extracellular region of the TM bundle center. Then, the induced signal dynamically transfers through the TM domain and the cytoplasmic side of the activated receptor engages G proteins or β -arrestins [18–20]. The classical property for class A GPCR activation is the outward movement of TM6 helix [17]. However, AT₂R is an outlier in the subfamily of GPCRs despite its similar sequence to AT₁R. It influences the cellular activity via possible G protein, protein phosphatases and phospholipase pathways, but the affirmatory downstream protein has not been defined [8,21–25].

Recently, its structures in both inactive-like, antagonist-bound and active-like, agonist-bound have become available [6,21,26]. In the antagonist-bound structure (PDB ID: 5UNG), its TM6 helix resembles the active conformation of other class A GPCRs but its helix 8 (H8) shows abnormal movement towards the center of the TM bundle [6]. Thus, it is regarded as an inactive-like state. The agonist-bound structure (PDB ID: 6JOD) shows an indistinct TM6 outward movement, while its H8 moves out from the TM bundle and become parallel with the membrane, like other class A

* Corresponding authors.

E-mail addresses: lushaoyong@yeah.net (S. Lu), zhangjiayou@medmail.com.cn (J. Zhang), pipidan@126.com (T. Chen).

¹ These authors contributed equally to this work.

GPCRs [26]. Because its ligand is the endogenous agonist angiotensin II, the structure is considered an active-like one. Although the static active-like, and inactive-like states of AT₂R exhibit valuable structural divergences, it is still challenging to completely describe the conformational space of AT₂R experimentally. Hence, it remains unclear how a dynamic pathway connects the two states of AT₂R, thereby hindering the elucidation of how AT₂R reaches the unique inactive state and the understanding of its activation and downstream signaling.

To dynamically unravel the GPCR conformational space, molecular dynamics (MD) simulations have been a well-established technique. With the application of MD simulations, the classical activation pathway for class A GPCR has been observed in β₂ adrenergic receptor [27] and a cryptic pocket hidden in the metastable state of AT₁R was discovered [28]. MD simulations have also captured precise molecule interplay with GPCRs (e.g. AT₁R and GPR120) [29,30]. To infer downstream signal protein in a dynamic way, MD simulations have also been applied to the A_{2A} adenosine receptor and μ-opioid receptor [31,32]. Using statistical algorithms, such as Markov State Model (MSM), the detailed transition processes between states provide increasing insights into the activation [33–37].

Here, we used extensive all-atom MD simulations (25 μs) to depict the conformational landscape of AT₂R and provided hints for its unique activation mechanism. Evolutional approaches were also referred to during our analysis. A hidden intermediate state was discovered, and it has a novel pocket for potential drug design. Our study not only sheds light on the research of AT₂R mechanism but also provides an opportunity for the design of novel AT₂R regulators.

2. Materials and methods

2.1. System setup

The AT₂R structure in complex with an antagonist-like ligand, quinazolinone-biphenyltetrazole derivative 1 (compound 1 in [6], PDB ID: 5UNG), and structure complexed with the endogenous ligand, AngII (PDB ID: 6JOD), were downloaded from the Protein Data Bank (PDB). In the AngII bound structure, the ligand and co-crystallized antibody were removed. Meanwhile, the mutation on the structures was back-mutated referring to the wild-type sequence. The hydrogens were then added, while the termini were capped with acetyl and methylamine groups. The antagonist-bound complex was set as the inactive-like structure for simulations.

Compared with the antagonist-bound complex, the extra residues in AngII-bound structure were deleted to unify the atom connecting information for the following process. Then, we docked compound 1 to the apo AngII-bound structure using Molecular Operation Environment (MOE). After receptor preparation and minimization, ligands were initially placed in the pocket with the triangle matcher and London dG scoring. Next, rigid receptor with GBVI/WSA dG scoring was used for refinement. The output ligand pose with the smallest root mean square deviation (RMSD) with compound 1 in the antagonist-bound structure was picked for ligand coordinates for the active-like structure.

2.2. NEB sampling

To explore the conformational space of the AT₂R, nudged elastic band (NEB) algorithm was first induced to generate initial structures on the free energy landscape. NEB determines the transition between different conformations by inserting a series of replicas between the initial and final states. The replicas have the same

atom connection information as the inactive-like and active-like AT₂R structures, but the coordinates are different from each other. The initial coordinates for each replica are evenly distributed along the pathway connecting the initial and final states. Then, elastic bands, a virtual spring force, were used to connect the replicas to their neighbors. With the elastic bands, structures can sample reasonable conformations in a simulated annealing process meanwhile avoiding sliding down to the energy basins [38]. The minimal energy pathway between the inactive-like and active-like AT₂R structures was then described by the conformations of replicas.

The elastic bands were represented as $3 \times N_{\text{atoms}}$ dimensional vectors $[\mathbf{R}_0, \mathbf{R}_1, \dots, \mathbf{R}_n]$, which act on each atom to provide the restraint. \mathbf{R}_x means the replicas between the start state \mathbf{R}_0 and end state \mathbf{R}_n . During calculation, a tangent vector $\boldsymbol{\tau}_i$ was induced to prevent the interference of elastic band forces and forces in MD force fields. Defined in Eq. (1), $\boldsymbol{\tau}_i$ controls the coordinates for each replica referring to the energy of this replica i (V_i) and its neighbors.

$$\boldsymbol{\tau}_i = \begin{cases} \mathbf{R}_{i+1} - \mathbf{R}_i, V_{i+1} > V_i > V_{i-1} \\ \mathbf{R}_i - \mathbf{R}_{i-1}, V_{i+1} < V_i < V_{i-1} \\ (\mathbf{R}_{i+1} - \mathbf{R}_i)\Delta V_i^{\max} + (\mathbf{R}_i - \mathbf{R}_{i-1})\Delta V_i^{\min}, V_i > V_{i+1} > V_{i-1} \text{ or } V_{i+1} > V_{i-1} > V_i \\ (\mathbf{R}_{i+1} - \mathbf{R}_i)\Delta V_i^{\min} + (\mathbf{R}_i - \mathbf{R}_{i-1})\Delta V_i^{\max}, V_i > V_{i-1} > V_{i+1} \text{ or } V_{i-1} > V_{i+1} > V_i \end{cases} \quad (1)$$

where $\Delta V_i^{\max} = \max(|V_{i+1} - V_i|, |V_{i-1} - V_i|)$, $\Delta V_i^{\min} = \min(|V_{i+1} - V_i|, |V_{i-1} - V_i|)$.

With $\boldsymbol{\tau}_i$, Eq. (2) and (3) adjust the total force in its perpendicular (\mathbf{F}_i^\perp) and parallel (\mathbf{F}_i^\parallel) components, respectively.

$$\mathbf{F}_i^\perp = -\nabla V(P_i) + (\nabla V(P_i)\hat{\mathbf{A}} \cdot \boldsymbol{\tau}_i)\hat{\mathbf{A}} \cdot \boldsymbol{\tau}_i \quad (2)$$

$$\mathbf{F}_i^\parallel = (\mathbf{F}_i \cdot \hat{\mathbf{A}}) \hat{\mathbf{A}} \cdot \boldsymbol{\tau}_i \quad (3)$$

$$\mathbf{F}_{\text{total}} = \mathbf{F}_i^\perp + \mathbf{F}_i^\parallel \quad (4)$$

where $\nabla V(P_i)$ is the gradient of the potential energy according to the coordinates in replica i , namely the opposite of forces provided by MD force field. \mathbf{F}^s is the force from the elastic bands. Consequently, the force field and elastic bonds contribute to the perpendicular and parallel part of the total force and avoid interference between them [39,40].

After the preparation, the active-like (PDB ID: 6JOD) and inactive-like (PDB ID: 5UNG) structures with antagonist-like ligand compound 1 were set as the initial and end states, respectively. The Amber ff19SB force field was employed for the description of atom interactions [41]. We firstly conducted 10,000 minimization cycles for our two systems. Then, 52 replicas were created between the initial and end structures.

During the NEB process, replicas were aligned to the center-of-mass and rotated by an optimal rotation matrix. The matrix minimizes the RMSD between structures in order to exclude the translational and rotational differences. In the simulated annealing process, the systems were firstly gradually heated to 300 K in 500 ps, with a spring force of $10 \text{ kcal}\cdot\text{mol}^{-1}\cdot\text{\AA}^{-2}$. Then, replicas were equilibrated for 600 ps with a spring force of $50 \text{ kcal}\cdot\text{mol}^{-1}\cdot\text{\AA}^{-2}$, which is kept in the coming processes. In the next simulated annealing runs, the system was generally heated to 500 K and cooled to 0 K in 1.5 ns. At last, the replicas were completely cooled at 0 K for 2 ns. The NEB workflow has been described and confirmed in the previous studies [28,42].

2.3. MD simulations and analysis

To obtain different initial conformation in production MD simulations, we calculated the RMSD between adjacent replicas and

picked the most different 10 replicas among 52 NEB outputs, including the active-like and inactive-like crystal structures. The following MD simulations are based on these structures.

The initial structures were inserted into a POPC (palmitoyl-2-oleoyl-*sn*-glycero-3-phosphocholine) membrane in the CHARMM-GUI server [43]. TIP3P waters with a length of 15 Å with 0.15 mol·L⁻¹ KCl were added to the top and bottom of the system. FF19SB, LIPID17, and GAFF2 force field were applied for the parameter of amino acids, lipids, and ligand, respectively [41,44,45]. The components of bilayers have been commonly used in other simulations [27,46–48].

The systems were firstly minimized for 15,000 cycles with a restraint of 500 kcal·mol⁻¹·Å⁻² on the protein and lipids. Then, all atoms encountered 30,000 cycles of minimization. Next, the systems were heated from 0 to 300 K in 300 ps and equilibrated for 700 ps with 10 kcal·mol⁻¹·Å⁻² position restraint on non-solvent atoms. At last, the 10 systems encountered 5 rounds of 500 ns production MD simulations, leading to 50 independent repeat trajectories of 25 μs in total. During simulations, the temperature (300 K) and pressure (1 atm) were controlled by the Langevin thermostat and Berendsen barostat, respectively. Long-range electrostatic interactions were treated by the Particle mesh Ewald algorithm and a cutoff of 10 Å was employed for short-range electrostatic and van der Waals interactions. The SHAKE algorithm was applied to restrain the bond with hydrogens. for covalent bonds containing hydrogen. All simulations were finished on Amber20, pmemd.cuda on NVIDIA Tesla V100 PCIe 16 GB.

The analyses were accomplished by Amber20 CPPTRAJ [49]. In particular, RMSF was calculated by the “atomicfluct” command, PCA was calculated by “projection” commands, distance and dihedral were calculated by “distance” and “dihedral” commands, and DCCM was calculated by the “matrix” command. Figures were drawn by MATLAB and PyMOL.

2.4. Statistics coupled analysis (SCA)

Multiple sequence alignment (MSA) of AT₂R was based on the alignment from the GPCRdb database [50]. All source and all species available AT₂R sequences were used. The sequences were then adjusted to fit the sequences of 6JOD. In total, 121 sequences and a row of gaps to prevent the zero-frequency problem were set as input for SCA [51].

In the SCA process, the conservation score for position *i* with amino acid *a* (D_i^a) was estimated by the cross-entropy loss between actual a frequency at the position (f_i^a) and background frequency for this residue (q^a). The loss was calculated in Eq. (5).

$$D_i^a = f_i^a \times \ln(f_i^a/q^a) + (1 - f_i^a) \times \ln((1 - f_i^a)/(1 - q^a)) \quad (5)$$

In each position, the most dominant amino acid was picked to calculate the frequency f_i^a . Then, the correlations between positions *i* and *j* with the dominant residue *a* and *b* (C_{ij}^{ab}), namely the co-evolution score, was estimated by Eq. (6).

$$C_{ij}^{ab} = \frac{\partial D_i^a}{\partial f_i^a} \hat{A} \cdot \frac{\partial D_j^b}{\partial f_j^b} \hat{A} \cdot |f_{ij}^{ab} - f_i^a \hat{A} \cdot f_j^b| \quad (6)$$

where f_{ij}^{ab} represents the frequency that position *i* has residue *a* as the dominant one, meanwhile position *j* has residue *b* as the dominant one. C_{ij}^{ab} values constitute the SCA matrix.

The eigenvectors (principal components, PC) of the SCA matrix were next calculated to identify the evolutionary related positions (sectors). As the first mode reflects the global fluctuations during the evolution process, it was dropped in the definition of sectors. In the sector definition, different sectors are divided clearly on the PC surface and have limited evolutionary relationships intra-

sectors. Thus, red sectors are defined as the positions whose weight of the 2nd eigenvector is larger than the weight of the 4th eigenvector and larger than ϵ . Blue sectors are the positions whose weight of the 2nd eigenvector is smaller than the weight of the 4th eigenvector and $-\epsilon$, or whose weight of the 4th eigenvector is larger than the weight of the 2nd eigenvector and larger than ϵ . In the application of AT₂R, the threshold value ϵ was 0.05. The sector selection and SCA calculation referred to previous publications [51,52].

2.5. Markov state Model (MSM) construction

According to the activation parameters, an MSM was built using the Python Emma's Markov Model Algorithms (PyEMMA) package [53]. Firstly, the points in the free energy landscape were clustered into 500 microstates by the k-means algorithm. Then, multiple transition probability matrixes (TPMs) were calculated according to the transitions among microstates. Referring to Eq. (7), the implied timescale test was performed to confirm the Markovian of microstates.

$$\tau_i = -\tau / \ln \lambda_i \quad (7)$$

where τ represents the lag time for the TPMs, λ_i is the *i*th eigenvalue of the TPM and τ_i is the implied timescale for the *i*th relaxation of the MSM. As a function of the lag time τ , τ_i (especially τ^1 for the slowest transition) is a constant when the transition between microstates is Markovian [33,54]. As shown in Fig. 5C, the lagtime for MSM construction was 4 ns. From Markovian microstates, macrostates were clustered via the PCCA + algorithm. Using transition path theory (TPT), the properties for transition, such as transition time and direction, were calculated [55]. To obtain the representative structures, the snapshots around the microstate centers of each macrostate were extracted to a trajectory. Then, the representative conformation of each macrostate was picked according to the similarity score S_{ij} estimated via Eq. (8).

$$S_{ij} = e^{-d_{ij}/d_{scale}} \quad (8)$$

where the d_{ij} is the RMSD between the snapshots *i* and *j* and d_{scale} is the standard deviation of *d*.

2.6. Pocket identification and molecule docking

Using Fpocket, we identified the potential pockets of different macrostates [56]. Fpocket defines a sphere that contacts four atoms on its boundary and contains no internal atom as an alpha sphere. It then selects alpha spheres defined by zones of tight atom packing. In the following cluster step, it excludes large spheres at the protein surface which solely composes a sphere cluster, then aggregates clusters with a close center of mass to a large cluster. Next, multiple linkage clustering approaches are used to further merge clusters. At last, the ability to bind a small molecule was evaluated by Partial Least Squares fitting to pocket descriptors, and top-scored pockets are shown.

From the pocket, we also conducted molecular screening using the allosteric GPCR sublibrary of Enamine. During the screening, all compounds were firstly prepared at pH 7.0 with the OPLS3 force field. Then, docking in standard precision was applied for all compounds. Then, 1,000 top-scored compounds encountered extra precision docking to produce the final output. The docking procedure was performed by glide in Maestro, Schrödinger suites.

3. Results

3.1. Large-scale unbiased MD simulations unravel the conformational space of AT₂R.

To sample the conformational landscape of AT₂R and understand its transition pathway towards the inactive-like conformation, we first generated a series of replicas connecting the inactive-like (PDB ID: 5UNG) and active-like (PDB ID: 6JOD) crystal structures with agonist bound. Then, nudged elastic band algorithm was applied to sample 10 initial structures distributed on the transition pathway. The structures were embedded in the POPC membrane and encountered 500 ns*5 independent MD simulation runs, leading to a total simulation time of 25 μ s. With the trajectories, we first calculated the root-mean square fluctuation (RMSF) of each residue (Fig. 1A and B) and decomposed the movements shown in the principal component analysis (PCA) to residues (Fig. 1C), to determine the flexible domains and their intrinsic movement patterns.

As shown in Fig. 1A, H8 is the most flexible domain during simulations. Besides H8, the intracellular loops are all flexible and ECL2 also shows obvious movement. According to the RMSF values shown in Fig. 1B, ICL1 and ICL2 are only flexible in the loop domain but ICL3 leads the movement for close intracellular TMs. As for the decomposition of movement shown in PCA (Fig. 1C), the dominant movement comes from H8 while the movement of intracellular TM5-TM6 is also evident in PC3. However, other domains with high RMSF show no obvious movement tendency in the decomposition of PCA. Thus, the major movements in AT₂R conformation transition were the movement of H8 and the intracellular TM5-TM6. Since H8 conformation is distinctive in the start and end structures of NEB, it is expected to show highly dynamic but the in-detail transition process is still worth investigation.

Given that the common transducer pocket is composed of TM5-TM6 and H8 [17,28] and their aforementioned dominant movement, we defined the conformation-describing parameter for AT₂R according to the movement of intracellular TM5-TM6 and H8 (Fig. 2A). Since TM2 is relatively stable, the sum of the distance from K^{5.63} to S^{2.40} and K^{6.25} to S^{2.40} reflects the movement of intracellular TM5-TM6 (superscripts indicate the Ballesteros-Weinstein numbering system [57]). In addition, the fluctuation of H8 in AT₂R covers the space from the center of the TM bundle to the position in parallel with the membrane, which is described by the dihedral among L^{8.54}, V^{7.56}, C^{7.54}, and C^{7.47}. As shown in Fig. 2A, the distance and dihedral values are distinctive between the inactive-like and active-like states, thereby confirming their power to depict the difference in different AT₂R states.

With the conformation-describing parameters, the conformational space of AT₂R during MD simulations was plotted as a free energy landscape in Fig. 2B. The active-like structure (50.7 Å, 101.2°) locates at an energy basin with a distance of ~52–56 Å and a dihedral of ~100–110°. The fixed dihedral indicates that H8 paralleled with the membrane is stable, maybe attributing to the interaction with membrane parts. This H8 conformation is also widely adopted by other class A GPCRs [17,58]. Upon the increasing of dihedral from active-like basin, the sum of distance also increases to 54–57 Å, in order to cross the energy barrier between a dihedral of ~130–150°. It infers that the movement of H8 couples with the open degree of intracellular TM5 and TM6. From a dihedral of around 150°, the relative energy of conformations becomes lower and two different energy basins (~51–56 Å and 180–197°, ~53–58 Å and 203v230°) locates there. Situated at (57.7 Å, 194.7°), the inactive-like structure locates at the low energy area and the specific basin for the inactive-like state will be identified in the coming parts. From a single trajectory, a small number of

points are out of major conformations at a dihedral of 80–250°, reflecting a disordered H8. The isolated phenomenon should be excluded from the main conformation change. Collectively, the free energy landscape shows the conformational space of AT₂R on its intracellular side and illustrates the intrinsic dynamic of AT₂R.

3.2. Evolutional and dynamical analysis reveals the relationship between domains in signal transportation of AT₂R

To unravel the relationship between intracellular movement and orthosteric pocket, we applied evolutional and dynamical algorithms on AT₂R. Statistics coupled analysis (SCA) is an approach to estimating the co-evolution degree between residue pairs, thereby identifying the evolutional- and functional-related residue sectors according to the degree [51,52,59]. We used all sequences of AT₂R from different species (see methods) and calculated the conservation and co-evolution property in each position (Fig. 3). The conservation score and logoplots for each position are shown in Fig. 3A, 3B, respectively. Also, Table 1 summarizes the conservation score for each domain. The co-evolution score between each residue pair forms the SCA matrix in Fig. 3C. Then, the evolutional-related sectors were identified by eigenvector decomposition (Fig. 3D). The red and blue sectors have a clustered co-evolution matrix in Fig. 3E, while their positions are shown in Fig. 3F.

As shown in Fig. 3A and Table 1, the conservation degree is different among domains. TMD (score = 2.12 \pm 0.78) is more conserved than loops (score = 1.53 \pm 1.08), especially for the ICLs (score = 0.59 \pm 0.86). Since ECLs are the gate for the orthosteric pocket [60], its conservation (score = 2.03 \pm 0.82) infers that AT₂R changes its ligand binding process a little during the evolution process. In contrast, ICLs traditionally interact with downstream proteins and the variation of ICLs reflects possible distinct AT₂R-interacting proteins (ATIP) in species [61,62]. Variation in conservation is not obvious in TM2-TM7 but TM1 (score = 1.69 \pm 0.83) and H8 (score = 1.70 \pm 0.73) are not as conserved as other TM helices, inferring that TM1 and H8 may play unique roles in different AT₂Rs.

The logoplots in Fig. 3B correspond with the conservation score in Fig. 3A and directly show the most conserved sequence in AT₂R. Globally, most positions have mutations and some of them are even highly mixed (e.g. 64, 119, 169, and 243), which provides information for the following co-evolution analysis. Also, key activation motifs for class A GPCRs, such as DR142^{3.50}Y, P223^{5.50}-I132^{3.40}-F265^{6.44}, CWxP271^{6.50}, and NP315^{7.50}xxY are still conserved in AT₂R logoplots. It suggests that AT₂R maintains a traditional signal pathway in class A GPCRs.

In Fig. 3C, the spots with the most co-evolution scores to other domains are N156^{4.38} to G210^{5.37}, L239^{5.66} to T250^{5.29}, and G285^{6.64} to L298^{7.33}, which represent the extracellular TM4-TM5, intracellular TM5-TM6, and extracellular TM6-TM7, respectively. We also identified that H8 has higher co-evolution scores with the H8 interaction domain in [6] than other residues, as labeled in Fig. 3C. For instance, R324^{8.49} has a high co-evolution score with directly interacting residues Q253^{6.32} (2.87) and M257^{6.26} (2.84). However, for other residues around, the average co-evolution score is 0.70 \pm 0.64. It confirms that the co-evolution score reflects interactions between domains.

In Fig. 3D, using PCA, we decomposed the SCA matrix to PC2-PC4 and defined evolutional related blue and red sectors (see method). In the clustered SCA matrix (Fig. 3E), residues in blue and red sectors show limited intra-action but evident interaction, confirming the separation of sectors. As shown in Fig. 3F, the blue and red sectors include the residues with the most co-evolution scores. The blue sector is mostly on the membrane side, while the red sector consists of residues around the orthosteric pocket

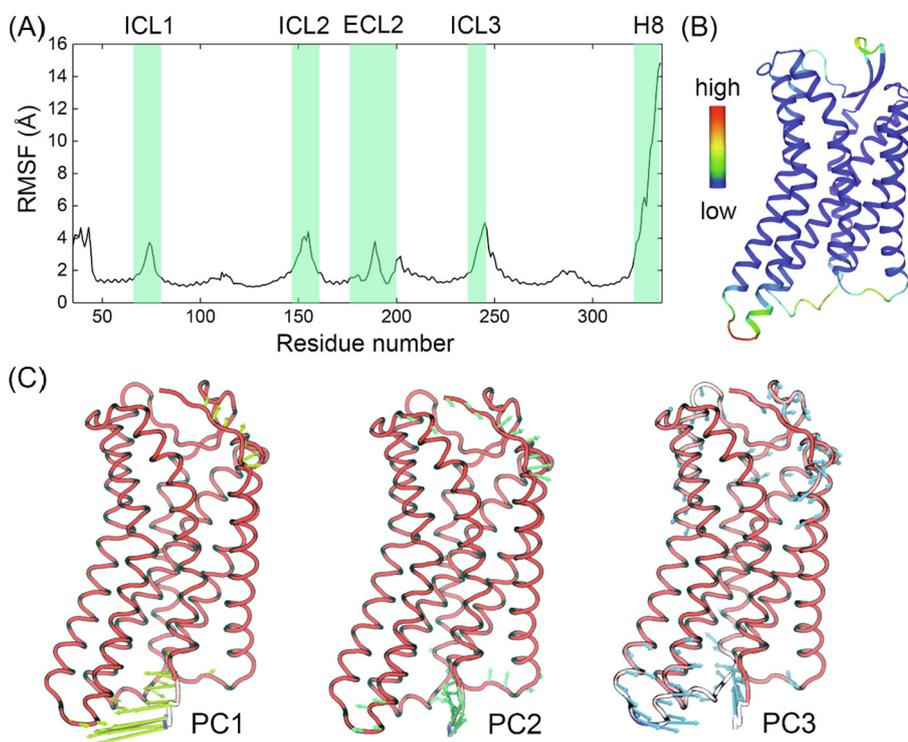


Fig. 1. (A) RMSF value for each residue during MD simulations. Flexible domains are highlighted in green rectangles. (B) The structure colored by the RMSF value. Highly flexible domains such as N-term and H8 were removed for clarity. (C) The movement expressed by major PCs. Yellow, green, and blue arrows represent the movement shown by PC1, PC2, and PC3, respectively. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

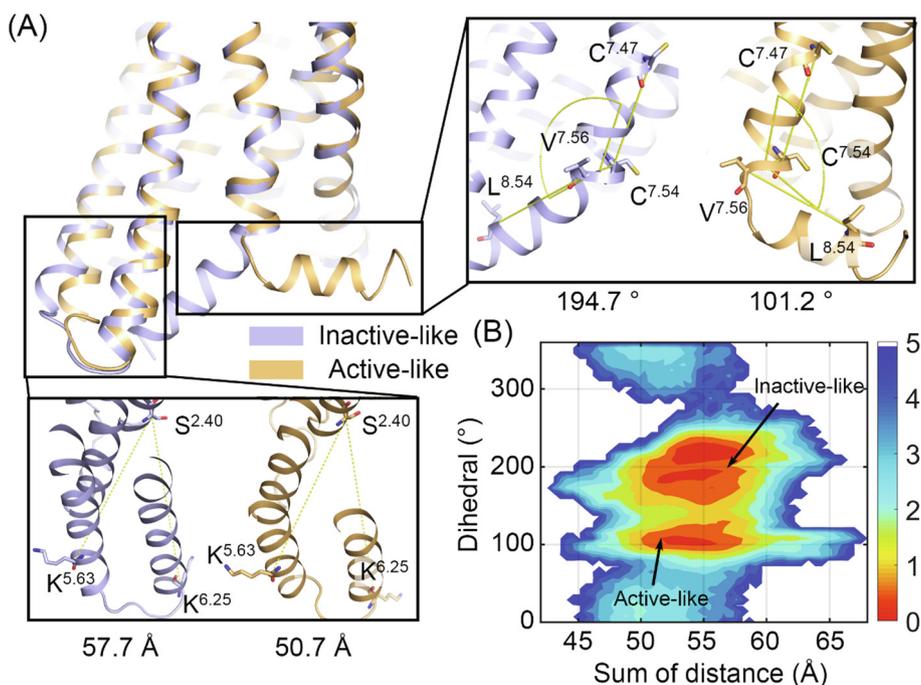


Fig. 2. (A) The distinctions between the AT₂R in the two states. Purple and orange cartoons show inactive-like and active-like AT₂R, respectively. In the zoom-in views, the sum of distance between the C α atoms of S^{2.40} to K^{5.63} and S^{2.40} to K^{6.25} was shown to measure the movement of TM5-TM6. The dihedral among the C α atoms of L^{8.54}, V^{7.56}, C^{7.54}, and C^{7.47} was applied for the movement of H8. (B) The free energy landscape composed of the sum of distance and dihedral. Arrows show the position of inactive-like and active-like AT₂R on the landscape. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

and intracellular TM5-TM6. Evolutional related sectors are spatial and functional closed residues [52]. Thus, considering their position, the blue sector may contribute to the stability of AT₂R in the membrane, while the red sector shows the relationship between the ligand and classical transducer pockets. Since the con-

served motifs are also maintained in AT₂R sequences, the original signal pathway may still exist in AT₂R, though it is controversial to interact with classical transducer proteins.

The dynamic cross correlation matrix (DCCM) was also applied to show the relationship of movement between residue pairs.

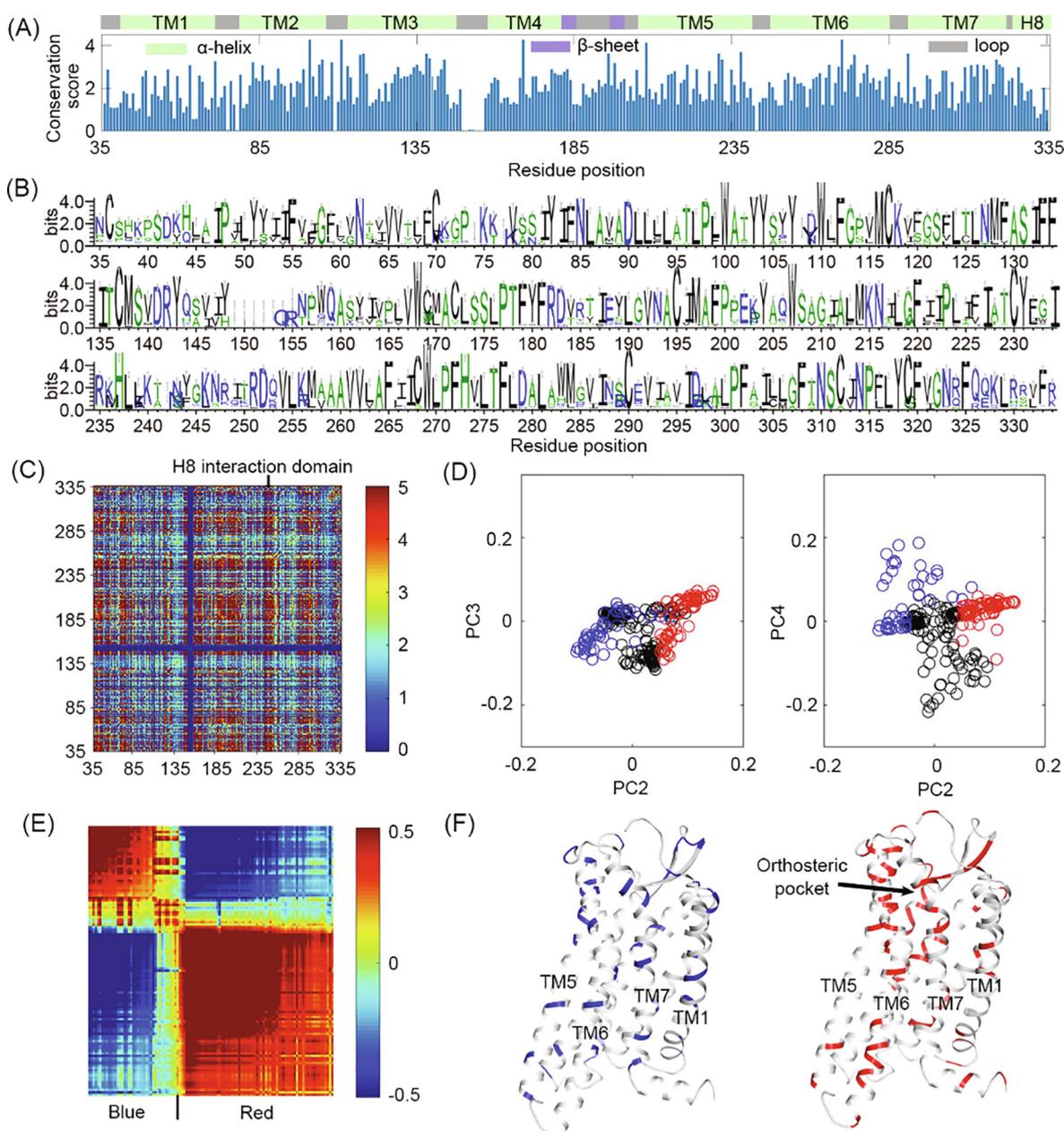


Fig. 3. (A) The conservation score of each position in a multiple sequence alignment for AT₂R sequences. Secondary structures for positions were labelled above. Color scheme: green, α -helix; purple, β -sheet; gray, loop. (B) Logoplots visualizing residue frequency in each position. (C) SCA matrix for residue pairs. From blue to red, the evolutionary correlation increases. H8 interaction domain is labelled by a line. (D) The projection of residues on the PC2-PC3 and PC2-PC4 decomposition for SCA matrix. Blue and red circles identify corresponding sector residues. (E) The clustered co-evolution matrix for blue and red sectors. (F) The location of blue and red evolutionary sectors on AT₂R. Sector residues are colored correspondingly and the other residues are shown in gray cartoons. Key helices and position of the orthosteric pocket are labelled. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

DCCM algorithm estimates the correlation between residue movement, which is shown as a DCCM matrix shown in Fig. 4C [63]. Red and blue in the matrix reflect correlated and anti-correlated movements, respectively. To explore the correlation between the orthosteric pocket and intracellular part for conformation-describing parameters, we first identified residues highly contacting with ligand in the pocket (Fig. 4B). Then, the DCCM matrix for these residues with other residues was visualized in Fig. 4C.

From the global DCCM matrix, the intracellular TM5 (A228^{5,55} to T241^{5,68}) mostly correlates with TM3 (D1, A130^{3,38} to Q144^{3,52}), while the intracellular TM6 (G245^{6,24} to K256^{6,35}) has no obvious interaction with other residues. As for the H8 (G322^{8,47} to R334^{8,59}), its movement is mostly related to TM1, ICL1, and TM2 (D2, D45^{1,33} to V88^{2,48}), which may be attributed

to their close position. In particular, the highly-contacting residues in Fig. 4B show a weak movement relationship with classical activation helices (W100^{2,60} with TM6 and TM7, I304^{7,39} with TM6 and H8, and F308^{7,43} with H8). Since the transferring of a signal from orthosteric pocket to TM6-H8 is the base of class A GPCR activation [17], the relationship between highly contacting residues and TM6-H8 infers that a typical signal pathway still exists in AT₂R. In total, evolutionary and dynamical analysis shows the unicity of AT₂R, but the classical GPCR signal pathway is still maintained.

3.3. Markov state model discovers distinct dynamic properties of AT₂R

To describe the detailed movement of the intracellular domain and further explain the free energy landscape, we built a kinetic

Table 1
Average conservation score for different AT₂R domains.

Domain name	Residue range	Average conservation score
N-terminal	35–44	1.52 ± 0.57
TM1	45–72	1.69 ± 0.83
TM2	78–107	2.28 ± 0.78
TM3	112–148	2.31 ± 0.77
TM4	156–182	2.15 ± 0.69
TM5	203–241	2.12 ± 0.78
TM6	245–285	2.24 ± 0.72
TM7	289–321	2.13 ± 0.81
H8	322–335	1.70 ± 0.73
ICLs	73–77, 149–155, 242–244	0.59 ± 0.86
ECLs	108–112, 183–202, 286–288	2.03 ± 0.82
All loops		1.53 ± 1.08
All TMD		2.12 ± 0.78

network MSM with the conformation-describing parameters. Using the estimation of transition between different states, MSM offers a statistical viewpoint for the conformational ensemble during MD simulations [33]. Three macrostates were clustered from MSM (Fig. 5A) and the transitions between them are shown in Fig. 5B. The activation state of each macrostate is determined via their corresponding representative structure. The result of the implied timescale test in Fig. 5C shows a flattened curve from 4 ns in all timescales. Thus, the choice of 4 ns as lagtime confirms the Markovian for our model. We also show the representative structure for each macrostate in Fig. 6.

Since our systems include an antagonist, the transition in Fig. 5B does not favor active-like structure. The active-like structure is easy to transfer to the intermediate state (5.10 μ s), but it would be difficult to go reversely (25.68 μ s). Although no direct pathway exists between active-like and inactive-like states, the transition time (24.12 μ s to inactive-like and 37.67 μ s to active-like) reflects the effect of the antagonist. The proportion of active-like state is, therefore, the least (25.08%). Notably, the proportion of intermediate (39.64%) and inactive-like (35.28%) states are similar, even the transition between the two states leans towards the intermediate state. It infers that the antagonist also permits the existence of the intermediate state. Considering a 500 ns timescale for each trajectory, the proportions may include many snapshots during the transition, so the specific values are for reference.

As the inactive-like crystal structure locates at (57.7 Å, 194.7°), the inactive-like macrostate shows a larger H8 dihedral (more than 200°), suggesting a conformation closer to the TM bundle center (Fig. 5A and Fig. 6A). The intermediate macrostate shows an H8 vertical to the membrane, alike an extended loop from TM7. The hook shape of H8 with TM3 is possible to form a new binding site (Fig. 6B). The active-like macrostate is overall similar to the crystal structure, but its TM6 shows an increasing open conformation (Fig. 6C).

Besides the distinctions around the conformation-describing parameters, the ligand pocket also shows differences. The different contact profiles in the inactive-like and active-like macrostates show the movement of TM6 in the active-like one (Fig. 6A and

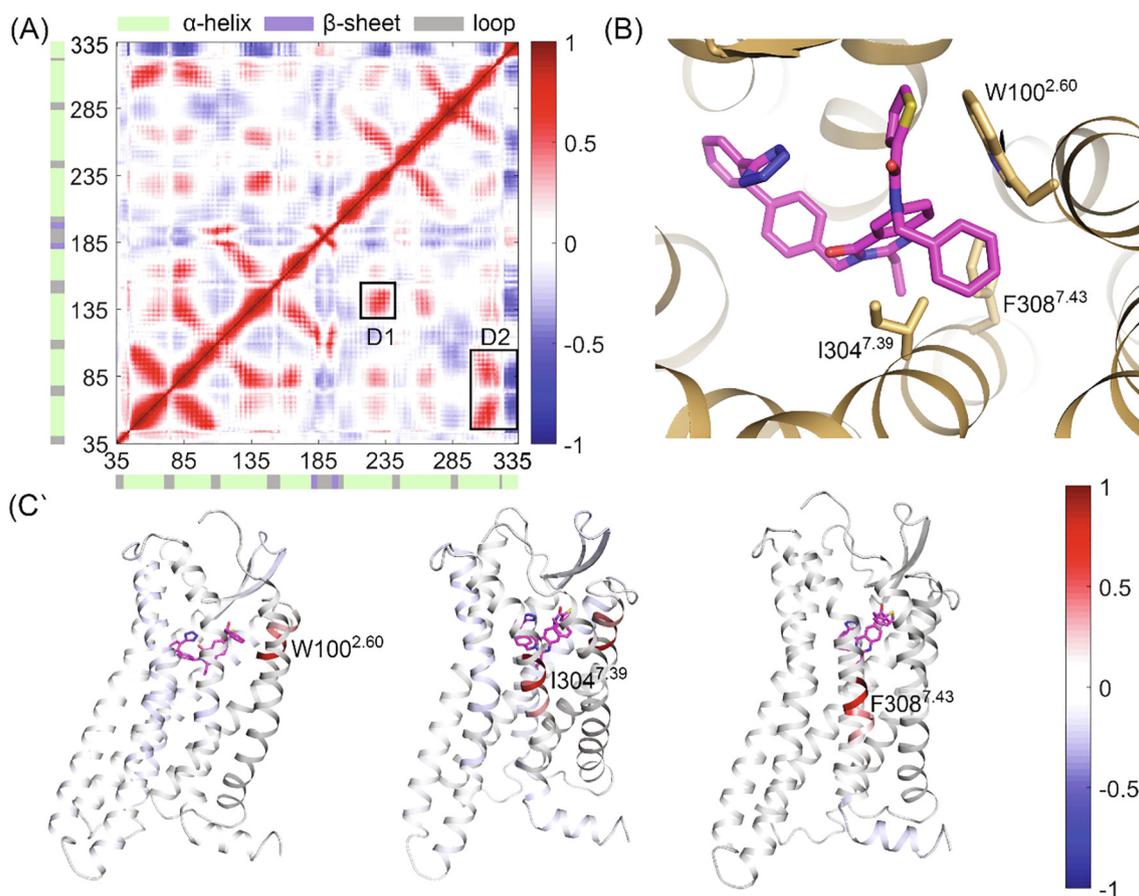


Fig. 4. (A) The DCCM matrix between residue pairs. Red and blue show correlated and anti-correlated movement between the pair, respectively. The correlation with a constant smaller than 0.3 was colored white for clarity. Rectangles show key correlation domains. Secondary structures of residues are shown around the residue index. Color scheme: green, α -helix; purple, β -sheet; gray, loop. (B) The position of highly contacting residues (W100^{2.60}, I304^{7.39}, and F308^{7.43}, more than 60% in contact with ligand at the threshold of 4 Å) and ligand. AT₂R and ligand were colored orange and magentas, respectively. (C) The projection of DCCM score for highly contacting residues on AT₂R structure. The color scheme is the same as (A). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

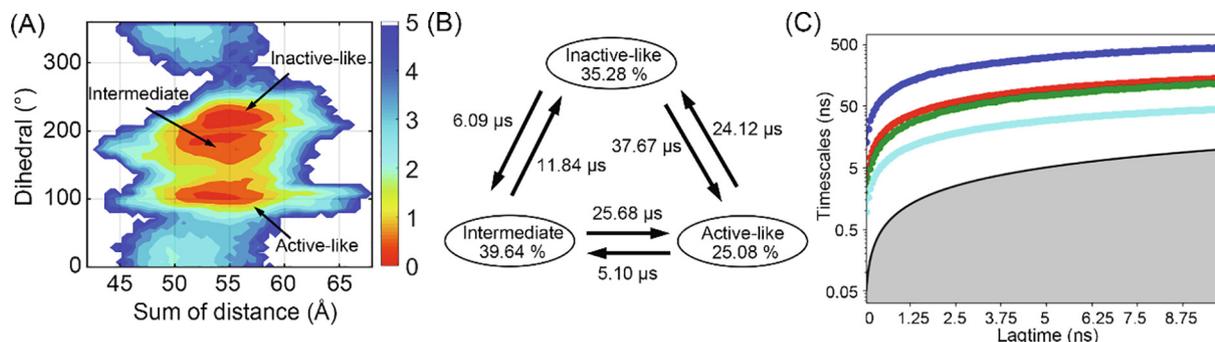


Fig. 5. (A) The position of macrostates on the free energy landscape and their corresponding states in activation. (B) The transition between macrostates and proportions of each state. (C) The implied timescale test for the MSM. Different timescales τ_1 , τ_2 , τ_3 , and τ_4 were represented as blue, red, green, and cyan lines changing with lag time. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

C). It is noteworthy that the intermediate state shows a different ligand binding mode, which shows an extended conformation towards the intracellular side and contacts with the toggle switch residue W269^{6,48} (Fig. 6B). The novel binding pose shows the plasticity of the AT₂R orthosteric pocket and potential multiple signal pathways inside it.

3.4. The discovery of the potential cryptic allosteric site in the intermediate state

We also detected potential pockets on the 3 macrostates to provide possible regulatory sites in AT₂R. Using Fpocket [64], 6 pockets were identified in the intermediate state and P6 was shown as the sole pocket hidden in the state (Fig. 7). From the extracellular to the intracellular side, 6 pockets are distributed on the cavities of AT₂R. P1 locates at the center of ECL2 and TM4, which overlaps with the site in the inactive-like macrostate. It is close to the LY2119620 allosteric pocket in the M2 muscarinic acetylcholine receptor (PDB 4MQT) [65]. Below it, P2 is the binding site of the antagonist. P3 situates at the membrane side among TM3-TM5, while P4 is at a lower position. The two pockets are also predicted in both two other macrostates. P5 locates at the intracellular side, the center of TM1, TM2, and TM7, which coincides with the vercirnon allosteric site in CCR9 (PDB 5LWE) [66]. Since the hook shape of H8 with TM6 is the uniqueness of the intermediate state, P6 is the only pocket hidden in the intermediate state. Considering the large proportion of the intermediate state in the ensemble, the potential of P6 is worth noting.

With the obtained P6, we applied virtual screening on it from the allosteric GPCR sublibrary of Enamine and different binding modes of potential ligands are shown in Fig. 8. The ligand conformations reflected two possible ways for the pocket to interact with regulators. One is the “up” conformation contacting with R142^{3,50} and intracellular TM6 residues (Fig. 8A), which shows a large binding interface and polar contact with R142^{3,50}. Z164965728 is the representative ligand, whose binding is also stabilized with intensively hydrophobic contacts with V254^{6,33}, F333^{8,58}, and F150^{ICL2}. The other is the “down” conformation constrained in the hook shape of H8 and interacting mainly with H8 residues (Fig. 8B). Z1450372712 shows a representative binding pose which is stabilized via polar contacts with S145^{3,53}, S331^{8,56}, and K328^{8,53}. The two binding modes may provide insights into further structure-based drug design on P6.

4. Discussion

Captured by our all-atom unbiased MD simulations, the conformational space of AT₂R shows a transition pathway from active-

like to inactive-like states. With MSM, the macrostate distribution reflects that the stable inactive-like structures show a larger dihedral, indicating an increasing inward movement of H8 towards the TM bundle center. The active-like macrostate also shows a broader distribution than the static crystal structure. Combined with an evolutionary algorithm and structure-based prediction methods [67–70], the stabilized H8 position in MD simulation may provide a potential site for downstream factors, thereby offering insights into the search for the downstream pathway of AT₂R.

Compared with the conformational ensemble in its homolog AT₁R without ligand [28], antagonist-bound AT₂R shows fewer active-like state conformations and a shorter transition time between different states. It infers that the antagonist disfavors the active-like conformation and accelerates the transition between states. Considering the 500 ns length of our simulation, many snapshots during the transition are defined as their original states. Hence, we did not observe a very small number of active-like state despite the antagonist. However, the transition time has been captured to favor non-active states, inferring a biased distribution towards inactive state upon enough simulation time.

Several ATIPs have been discovered to interact with the H8 and C-terminal of active AT₂R [71,72]. From the conformational ensemble of AT₂R, the interface for AT₂R and ATIP is hidden in the TM bundle center in the inactive state, preventing AT₂R from interactions. Upon activation, the interface residues have various conformations as shown by the dihedral in the free energy landscape. Since the active-like structure has half of H8 covered by the membrane, the interface residues may not be able to bind with ATIP in the state. Hence, the interactions with ATIP may involve a metastable intermediate state during the activation process and the state could be stabilized by ATIP.

From our dynamical and evolutionary information, the signal transferring inside AT₂R is overall similar to normal class A GPCRs. For example, the red sector in SCA shows the relationship between the ligand site and intracellular TM5-H8. Moreover, DCCM also infers the correlation of movement from extracellular to intracellular sides. But subtle differences still exist. The weak connection between the orthosteric site and classical downstream protein site reflects that AT₂R is not dependent on the classical signal pathway. Considering that AT₂R is controversial to interact with classical downstream proteins, the remains of the normal class A GPCR pathway may not perform as well as other GPCRs.

The DCCM matrix identified TM1, ICL1, and TM2 as the structures highly related to the movement of H8, inferring the specific conformational ensemble of H8 correlate with TM1-TM2 sequences. Also, experimental evidence has shown that the specific sequence in AT₂R ICL1 leads to the observed atypical conformation of H8 [73]. Additionally, ICL3 has been identified as the key domain

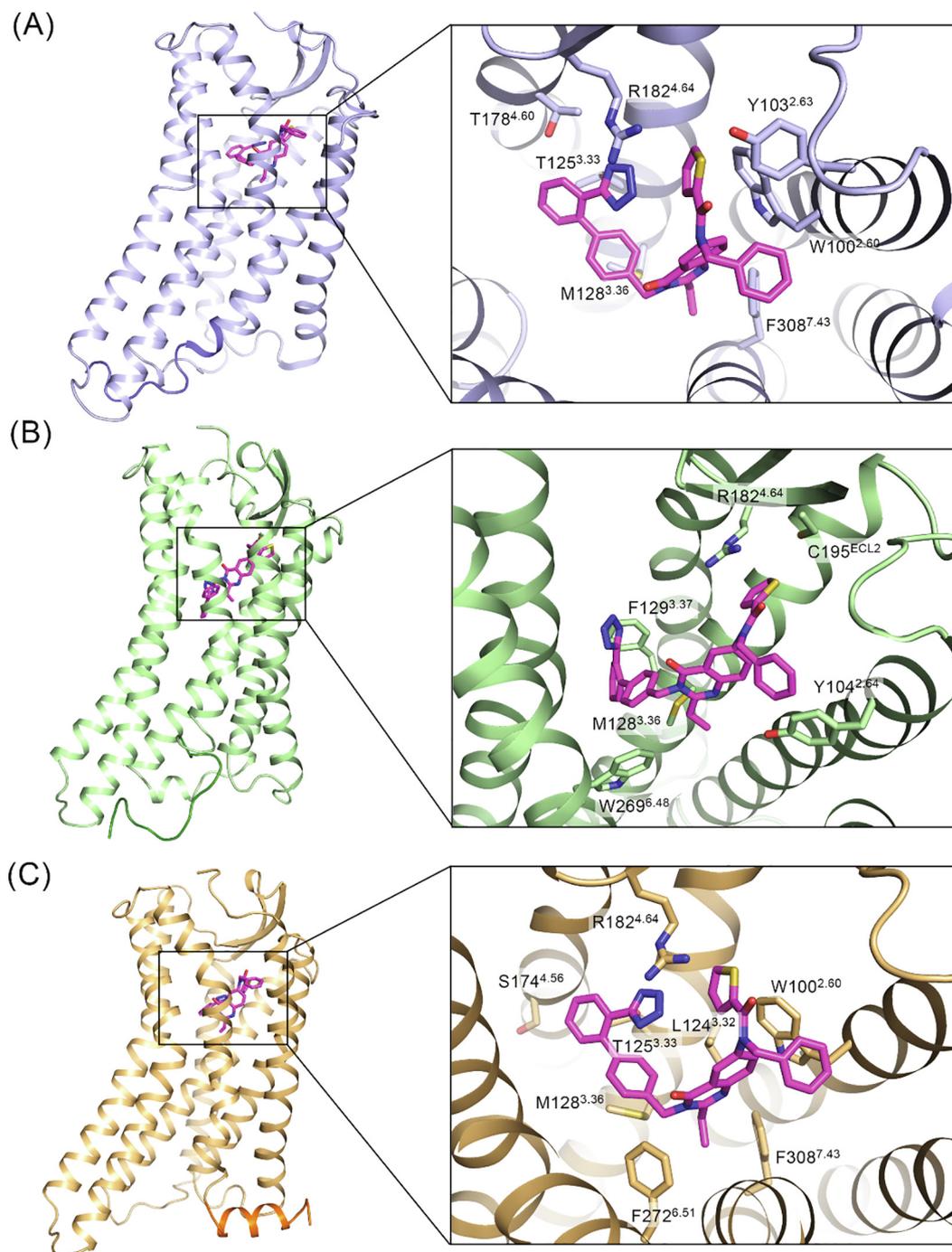


Fig. 6. The representative structure and pocket conformation for inactive-like (A), intermediate (B), and active-like (C) AT₂R, which are shown in purple, green, and orange cartoons. The dark color identifies the position of H8. The interface residues and ligand are shown in sticks. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

for initiating downstream signals of AT₂R [74,75]. In our *in silico* results, it shows evident movement in the decomposition of PCA and the free energy landscape. It reflects that ICL3 is a structural component with obvious dynamics, probably related to its function in signal initiating.

Since a potential pocket P6 has been identified in the hidden intermediate state, the regulation of AT₂R may be benefits from the drug developed targeting P6. The position is always considered as the protein-protein interface in class A GPCRs and the

potential of developing small molecule ligands targeting the interface has been proved [76–78]. In addition, the P6 site tends to be an allosteric one since classical downstream proteins bound here do not interact with AT₂R [8,79,80]. Hence, regulators targeting P6 may show greater selectivity and less toxicity as the common advantage for allosteric modulators [81–86]. Notably, the result of virtual screening are shown for possible ligand binding mode for P6 but their bioactivity was not confirmed by experiments.

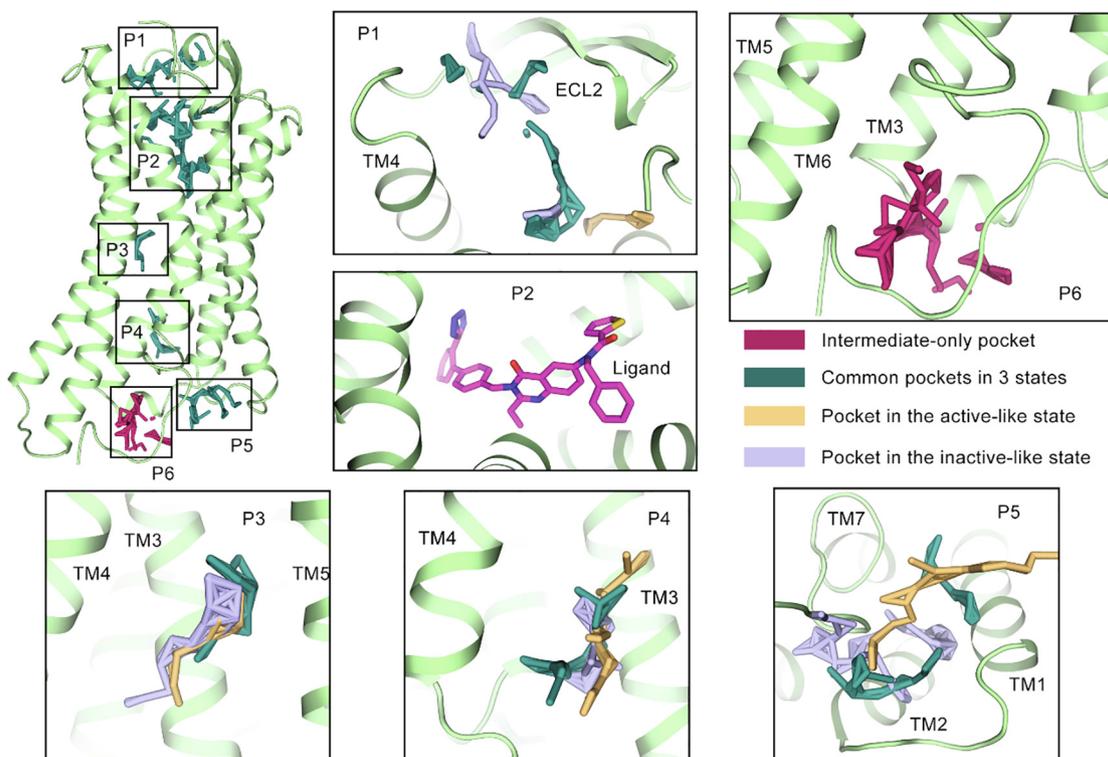


Fig. 7. The pockets predicted by Fpocket in the intermediate state. Pockets are shown in sticks and zoom-in subplots show the specific position of each pocket. The pockets overlapping with those in other macrostates are colored in deep green, while the unique pocket P6 for the intermediate state was colored in red. Purple and orange sticks depict pockets in the representative inactive-like and active-like structures, respectively. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

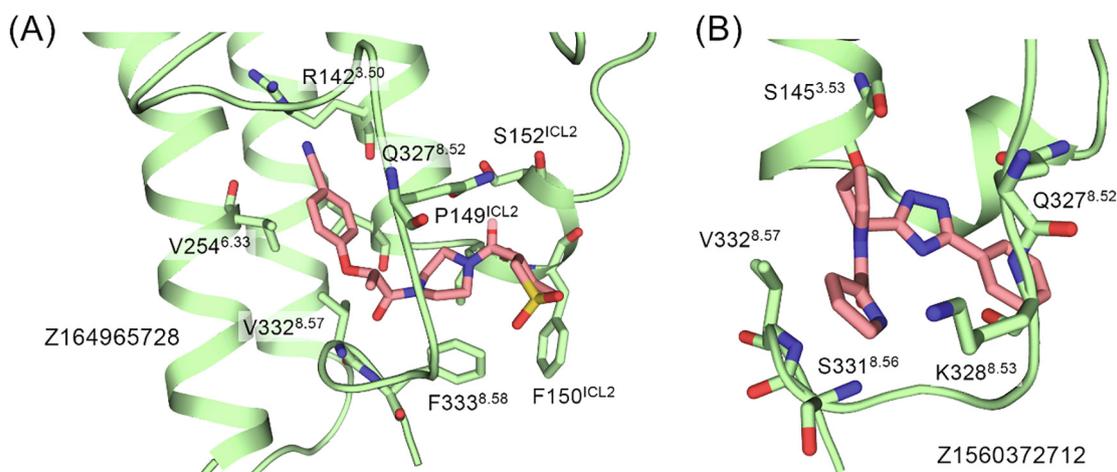


Fig. 8. The two possible ligand binding modes for P6. (A) The docked pose for Z164965728, representing the "up" pose interacting with ICL2, TM3, TM6, and H8. (B) The docked pose for Z1560372712, representing the "down" pose mainly interacting with H8.

5. Conclusion

Here, the all-atom unbiased MD simulations completely show the conformational space of AT₂R and dynamically explained the transitions between different macrostates. With evolutionary insights, the dynamic analysis provides the difference of AT₂R compared with other class A GPCRs, thereby suggesting the attribution of its unique conformational space. The cryptic pocket hidden in the intermediate state also offers an opportunity for the development of novel AT₂R regulators. Overall, our research elucidated the dynamic properties of AT₂R and aids in the explanation of its unique mechanism.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work was supported partly by grants from Postdoctoral Startup Foundation of Naval Military Medical University, the Key Sub Specialty of Pudong New Area Health Committee (No.

PWZy2020-08) and the National Natural Science Foundation of China (No. 81673704)

References

- [1] Karnik SS, Unal H, Kemp JR, Tirupula KC, Eguchi S, Vanderheyden PML, et al. International Union of Basic and Clinical Pharmacology. XCIX. Angiotensin Receptors: Interpreters of Pathophysiological Angiotensinergic Stimuli [corrected]. *Pharmacol Rev* 2015;67:754–819.
- [2] Mehta PK, Griendling KK. Angiotensin II cell signaling: physiological and pathological effects in the cardiovascular system. *Am J Physiol Cell Physiol* 2007;292:82–97.
- [3] Paz Ocaranza M, Riquelme JA, García L, Jalil JE, Chiong M, Santos RAS, et al. Counter-regulatory renin-angiotensin system in cardiovascular disease. *Nat Rev Cardiol* 2020;17:116–29.
- [4] Zaman MA, Oparil S, Calhoun DA. Drugs targeting the renin-angiotensin-aldosterone system. *Nat Rev Drug Discov* 2002;1:621–36.
- [5] de Gasparo M, Catt KJ, Inagami T, Wright JW, Unger T. International union of pharmacology. XXIII. The angiotensin II receptors. *Pharmacol Rev* 2000;52:415–72.
- [6] Zhang H, Han GW, Batyuk A, Ishchenko A, White KL, Patel N, et al. Structural basis for selectivity and diversity in angiotensin II receptors. *Nature* 2017;544:327–32.
- [7] Miura S, Matsuo Y, Kiya Y, Karnik SS, Saku K. Molecular mechanisms of the antagonistic action between AT1 and AT2 receptors. *Biochem Biophys Res Commun* 2010;391:85–90.
- [8] Fatima N, Patel SN, Hussain T. Angiotensin II Type 2 Receptor: A Target for Protection Against Hypertension. *Metabolic Dysfunction, and Organ Remodeling Hypertens* (Dallas, Tex 1979;2021(77):1845–56.
- [9] Ichiki T, Labosky PA, Shiota C, Okuyama S, Imagawa Y, Fogo A, et al. Effects on blood pressure and exploratory behaviour of mice lacking angiotensin II type-2 receptor. *Nature* 1995;377:748–50.
- [10] Porrello ER, D'Amore A, Curl CL, Allen AM, Harrap SB, Thomas WG, et al. Angiotensin II type 2 receptor antagonizes angiotensin II type 1 receptor-mediated cardiomyocyte autophagy. *Hypertens* (Dallas, Tex 1979;2009 (53):1032–40.
- [11] Porrello ER, Delbridge LMD, Thomas WG. The angiotensin II type 2 (AT2) receptor: an enigmatic seven transmembrane receptor. *Front Biosci (Landmark Ed)* 2009;14:958–72.
- [12] Menk M, Graw JA, von Haefen C, Siffringer M, Schwaiberger D, Unger T, et al. Stimulation of the Angiotensin II AT2 Receptor is Anti-inflammatory in Human Lipopolysaccharide-Activated Monocytic Cells. *Inflammation* 2015;38:1690–9.
- [13] Zhu R, Song Y, Liu H, Yang Y, Wang S, Yi C, et al. Allosteric histidine switch for regulation of intracellular zinc(II) fluctuation. *Proc Natl Acad Sci U S A* 2017;114:13661–6.
- [14] Nag S, Patel S, Mani S, Hussain T. Role of angiotensin type 2 receptor in improving lipid metabolism and preventing adiposity. *Mol Cell Biochem* 2019;461:195–204.
- [15] Shum M, Pinaud S, Guimond M-O, Labbé SM, Roberge C, Baillargeon J-P, et al. Angiotensin II type 2 receptor promotes adipocyte differentiation and restores adipocyte size in high-fat/high-fructose diet-induced insulin resistance in rats. *Am J Physiol Endocrinol Metab* 2013;304:197–210.
- [16] Mansoor S, Kayik G, Durdagi S, Sensoy O. Mechanistic insight into the impact of a bivalent ligand on the structure and dynamics of a GPCR oligomer. *Comput Struct. Biotechnol J* 2022.
- [17] Zhou Q, Yang D, Wu M, Guo Y, Guo W, Zhong L, et al. Common activation mechanism of class A GPCRs. *Elife* 2019;8:e50279.
- [18] Baldessari F, Capelli R, Carloni P, Giorgetti A. Coevolutionary data-based interaction networks approach highlighting key residues across protein families: The case of the G-protein coupled receptors. *Comput Struct Biotechnol J* 2020;18:1153–9.
- [19] Pierce KL, Premont RT, Lefkowitz RJ. Seven-transmembrane receptors. *Nat Rev Mol Cell Biol* 2002;3:639–50.
- [20] Che K, Muttenthaler M, Kurzbach D. Conformational selection of vasopressin upon V1a receptor binding. *Comput Struct Biotechnol J* 2021;19:5826–33.
- [21] Asada H, Horita S, Hirata K, Shirosaki M, Shiimura Y, Iwanari H, et al. Crystal structure of the human angiotensin II type 2 receptor bound to an angiotensin II analog. *Nat Struct Mol Biol* 2018;25:570–6.
- [22] Moore SA, Huang N, Hinthong O, Andres RD, Grammatopoulos TN, Weyhenmeyer JA. Human angiotensin II type-2 receptor inhibition of insulin-mediated ERK-2 activity via a G-protein coupled signaling pathway. *Brain Res Mol Brain Res* 2004;124:62–9.
- [23] Hansen JL, Servant G, Baranski TJ, Fujita T, Iiri T, Sheikh SP. Functional reconstitution of the angiotensin II type 2 receptor and G(i) activation. *Circ Res* 2000;87:753–9.
- [24] AbdAlla S, Lother H, Abdel-tawab AM, Quitterer U. The angiotensin II AT2 receptor is an AT1 receptor antagonist. *J Biol Chem* 2001;276:39721–6.
- [25] Juillerat-Jeanneret L. The Other Angiotensin II Receptor: AT2R as a Therapeutic Target. *J Med Chem* 2020;63:1978–95.
- [26] Asada H, Inoue A, Ngako Kadji FM, Hirata K, Shiimura Y, Im D, et al. The Crystal Structure of Angiotensin II Type 2 Receptor with Endogenous Peptide Hormone. *Structure* 2020;28:418–25.
- [27] Dror RO, Arlow DH, Maragakis P, Mildorf TJ, Pan AC, Xu H, et al. Activation mechanism of the β_2 -adrenergic receptor. *Proc Natl Acad Sci U S A* 2011;108:18684–9.
- [28] Lu S, He X, Yang Z, Chai Z, Zhou S, Wang J, et al. Activation pathway of a G protein-coupled receptor uncovers conformational intermediates as targets for allosteric drug design. *Nat Commun* 2021;12:4721.
- [29] Kiriakidi S, Chatzigiannis C, Papaemmanouil C, Tzakos AG, Cournia Z, Mavromoustakos T. Interplay of cholesterol, membrane bilayers and the AT1R: A cholesterol consensus motif on AT1R is revealed. *Comput Struct Biotechnol J* 2021;19:110–20.
- [30] Pal A, Curtin JF, Kinsella GK. Structure based prediction of a novel GPR120 antagonist based on pharmacophore screening and molecular dynamics simulations. *Comput Struct Biotechnol J* 2021;19:6050–63.
- [31] Kelly B, Hollingsworth SA, Blakemore DC, Owen RM, Storer RI, Swain NA, et al. Delineating the Ligand-Receptor Interactions That Lead to Biased Signaling at the μ -Opioid Receptor. *J Chem Inf Model* 2021;61:3696–707.
- [32] Navarro G, Gonzalez A, Campanacci S, Rivas-Santisteban R, Reyes-Resina I, Casajuana-Martin N, et al. Experimental and computational analysis of biased agonism on full-length and a C-terminally truncated adenosine A2A receptor. *Comput Struct Biotechnol J* 2020;18:2723–32.
- [33] Husic BE, Pande VS. Markov State Models: From an Art to a Science. *J Am Chem Soc* 2018;140:2386–96.
- [34] Zhuang H, Fan X, Ji D, Wang Y, Fan J, et al. Elucidation of the conformational dynamics and assembly of Argonaute-RNA complexes by distinct yet coordinated actions of the supplementary microRNA. *Comput Struct Biotechnol J* 2022;20:1352–65.
- [35] Wang Y, Li M, Liang W, Shi X, Fan J, Kong R, et al. Delineating the activation mechanism and conformational landscape of a class B G protein-coupled receptor glucagon receptor. *Comput Struct Biotechnol J* 2022;20:628–39.
- [36] Lan J, He X, Ren Y, Wang Z, Zhou H, Fan S, et al. Structural insights into the SARS-CoV-2 Omicron RBD-ACE2 interaction. *Cell Res* 2022;1–3.
- [37] Lu S, Ni D, Wang C, He X, Lin H, Wang Z, et al. Deactivation Pathway of Ras GTPase Underlies Conformational Substates as Targets for Drug Design. *ACS Catal* 2019;9:7188–96.
- [38] Jónsson H, Mills G, Jacobsen KW. Nudged elastic band method for finding minimum energy paths of transitions. *CI Quantum Dyn Condens Phase Simulations* 1997:385–404.
- [39] Henkelman G, Jónsson H. Improved tangent estimate in the nudged elastic band method for finding minimum energy paths and saddle points. *J Chem Phys* 2000;113:9978–85.
- [40] Ghoreishi D, Cerutti DS, Fallon Z, Simmerling C, Roitberg AE. Fast Implementation of the Nudged Elastic Band Method in AMBER. *J Chem Theory Comput* 2019;15:4699–707.
- [41] Tian C, Kasavajhala K, Belfon KAA, Raguette L, Huang H, Miguez AN, et al. ff19SB: Amino-Acid-Specific Protein Backbone Parameters Trained against Quantum Mechanics Energy Surfaces in Solution. *J Chem Theory Comput* 2020;16:528–52.
- [42] Huang M, Lu S, Shi T, Zhao Y, Chen Y, Li X, et al. Conformational transition pathway in the activation process of allosteric glucokinase. *PLoS ONE* 2013;8:e55857.
- [43] Jo S, Cheng X, Lee J, Kim S, Park S-J, Patel DS, et al. CHARMM-GUI 10 years for biomolecular modeling and simulation. *J Comput Chem* 2017;38:1114–24.
- [44] Lee J, Hitznerberger M, Rieger M, Kern NR, Zacharias M, Im W. CHARMM-GUI supports the Amber force fields. *J Chem Phys* 2020;153:35103.
- [45] He X, Man VH, Yang W, Lee T-S, Wang J. A fast and high-quality charge model for the next generation general AMBER force field. *J Chem Phys* 2020;153:114502.
- [46] Suomivuori C-M, Latorraca NR, Wingler LM, Eismann S, King MC, Kleinhenz ALW, et al. Molecular mechanism of biased signaling in a prototypical G protein-coupled receptor. *Science* 2020;367:881–7.
- [47] Kohlhoff KJ, Shukla D, Lawrenz M, Bowman GR, Konerding DE, Belov D, et al. Cloud-based simulations on Google Exacyle reveal ligand modulation of GPCR activation pathways. *Nat Chem* 2014;6:15–21.
- [48] Hollingsworth SA, Kelly B, Valant K, Michaelis JA, Mastromihalis O, Thompson G, et al. Cryptic pocket formation underlies allosteric modulator selectivity at muscarinic GPCRs. *Nat Commun* 2019;10:3289.
- [49] Roe DR, Cheatham III TE. PTRAJ and CPPTRAJ: software for processing and analysis of molecular dynamics trajectory data. *J Chem Theory Comput* 2013;9:3084–95.
- [50] Munk C, Mutt E, Isberg V, Nikolajsen LF, Bibbe JM, Flock T, et al. An online resource for GPCR structure determination and analysis. *Nat Methods* 2019;16:151–62.
- [51] Seo MJ, Heo J, Kim K, Chung KY, Yu W. Coevolution underlies GPCR-G protein selectivity and functionality. *Sci Rep* 2021;11:7858.
- [52] Halabi N, Rivoire O, Leibler S, Ranganathan R. Protein sectors: evolutionary units of three-dimensional structure. *Cell* 2009;138:774–86.
- [53] Scherer MK, Trendelkamp-Schroer B, Paul F, Pérez-Hernández G, Hoffmann M, Platner N, et al. PyEMMA 2: A Software Package for Estimation, Validation, and Analysis of Markov Models. *J Chem Theory Comput* 2015;11:5525–42.
- [54] Swope WC, Pitera JW, Suits F. Describing Protein Folding Kinetics by Molecular Dynamics Simulations. 1. Theory. *J Phys Chem B* 2004;108:6571–81.
- [55] Chodera JD, Singhal N, Pande VS, Dill KA, Swope WC. Automatic discovery of metastable states for the construction of Markov models of macromolecular conformational dynamics. *J Chem Phys* 2007;126:155101.
- [56] Le Guilloux V, Schmidtke P, Tuffery P. Fpocket: an open source platform for ligand pocket detection. *BMC Bioinf* 2009;10:168.

- [57] Ballesteros JA, Weinstein H. Integrated methods for the construction of three-dimensional models and computational probing of structure-function relations in G protein-coupled receptors. In: Sealfon SCBT-M in N, editor. *Recept. Mol. Biol.*, vol. 25, Academic Press; 1995, p. 366–428.
- [58] Ibrahim P, Wifling D, Clark T. Universal Activation Index for Class A GPCRs. *J Chem Inf Model* 2019;59:3938–45.
- [59] Lockless SW, Ranganathan R. Evolutionarily conserved pathways of energetic connectivity in protein families. *Science* 1999;286:295–9.
- [60] Peeters MC, van Westen GJP, Li Q, IJzerman, AP.. Importance of the extracellular loops in G protein-coupled receptors for ligand recognition and receptor activation. *Trends Pharmacol Sci* 2011;32:35–42.
- [61] Xu Y, Cai H, You C, He X, Yuan Q, Jiang H, et al. Structural insights into ligand binding and activation of the human thyrotropin-releasing hormone receptor. *Cell Res* 2022.
- [62] Duan J, Shen D-D, Zhao T, Guo S, He X, Yin W, et al. Molecular basis for allosteric agonism and G protein subtype selectivity of galanin receptors. *Nat Commun* 2022;13:1364.
- [63] Swaminathan S, Harte Jr WE, Beveridge DL. Investigation of domain structure in proteins via molecular dynamics simulation: application to HIV-1 protease dimer. *J Am Chem Soc* 1991;113:2717–21.
- [64] Palomo V, Soteras I, Perez DI, Perez C, Gil C, Campillo NE, et al. Exploring the binding sites of glycogen synthase kinase 3. Identification and characterization of allosteric modulation cavities. *J Med Chem* 2011;54:8461–70.
- [65] Kruse AC, Ring AM, Manglik A, Hu J, Hu K, Eitel K, et al. Activation and allosteric modulation of a muscarinic acetylcholine receptor. *Nature* 2013;504:101–6.
- [66] Oswald C, Rappas M, Kean J, Doré AS, Errey JC, Bennett K, et al. Intracellular allosteric antagonism of the CCR9 receptor. *Nature* 2016;540:462–5.
- [67] Evans R, O'Neill M, Pritzel A, Antropova N, Senior A, Green T, et al. Protein complex prediction with AlphaFold-Multimer. *BioRxiv* 2021:2021.10.04.463034.
- [68] Mukherjee I, Chakrabarti S. Co-evolutionary landscape at the interface and non-interface regions of protein-protein interaction complexes. *Comput Struct Biotechnol J* 2021;19:3779–95.
- [69] Ma C-Y, Liao C-S. A review of protein-protein interaction network alignment: From pathway comparison to global alignment. *Comput Struct Biotechnol J* 2020;18:2647–56.
- [70] Northey TC, Barešić A, Martin ACR. IntPred: a structure-based predictor of protein-protein interaction sites. *Bioinformatics* 2018;34:223–9.
- [71] Haykal MM, Rodrigues-Ferreira S, Nahmias C. Microtubule-Associated Protein ATP3, an Emerging Target for Personalized Medicine in Breast Cancer. *Cells* 2021;10:1080.
- [72] Nouet S, Amzallag N, Li J-M, Louis S, Seitz I, Cui T-X, et al. Trans-inactivation of Receptor Tyrosine Kinases by Novel Angiotensin II AT2 Receptor-interacting Protein. *ATIP J Biol Chem* 2004;279:28989–97.
- [73] Connolly A, Holleran BJ, Simard É, Baillargeon J-P, Lavigne P, Leduc R. Interplay between intracellular loop 1 and helix VIII of the angiotensin II type 2 receptor controls its activation. *Biochem Pharmacol* 2019;168:330–8.
- [74] Lehtonen JYA, Daviet L, Nahmias C, Horiuchi M, Dzau VJ. Analysis of Functional Domains of Angiotensin II Type 2 Receptor Involved in Apoptosis. *Mol Endocrinol* 1999;13:1051–60.
- [75] Daviet L, Lehtonen JY, Hayashida W, Dzau VJ, Horiuchi M. Intracellular third loops in AT1 and AT2 receptors determine subtype specificity. *Life Sci* 2001;69:509–16.
- [76] Tangella LP, Arooj M, Deplazes E, Gray ES, Mancera RL. Identification and characterisation of putative drug binding sites in human ATP-binding cassette B5 (ABCB5) transporter. *Comput Struct Biotechnol J* 2021;19:691–704.
- [77] Rosell M, Fernández-Recio J. Docking-based identification of small-molecule binding sites at protein-protein interfaces. *Comput Struct Biotechnol J* 2020;18:3750–61.
- [78] Chang C, Lin S-M, Satange R, Lin S-C, Sun S-C, Wu H-Y, et al. Targeting protein-protein interaction interfaces in COVID-19 drug discovery. *Comput Struct Biotechnol J* 2021;19:2246–55.
- [79] Li X, Wang C, Peng T, Chai Z, Ni D, Liu Y, et al. Atomic-scale insights into allosteric inhibition and evolutionary rescue mechanism of *Streptococcus thermophilus* Cas9 by the anti-CRISPR protein AcrIIA6. *Comput Struct Biotechnol J* 2021;19:6108–24.
- [80] Wang Y, Yu Z, Xiao W, Lu S, Zhang J. Allosteric binding sites at the receptor-lipid bilayer interface: novel targets for GPCR drug discovery. *Drug Discov Today* 2021;26:690–703.
- [81] Lu S, Shen Q, Zhang J. Allosteric Methods and Their Applications: Facilitating the Discovery of Allosteric Drugs and the Investigation of Allosteric Mechanisms. *Acc Chem Res* 2019;52:492–500.
- [82] Lu S, He X, Ni D, Zhang J. Allosteric Modulator Discovery: From Serendipity to Structure-Based Design. *J Med Chem* 2019;62:6405–21.
- [83] Lu S, Chen Y, Wei J, Zhao M, Ni D, He X, et al. Mechanism of allosteric activation of SIRT6 revealed by the action of rationally designed activators. *Acta Pharm Sin B* 2021;11:1355–61.
- [84] Ni D, Wei J, He X, Rehman AU, Li X, Qiu Y, et al. Discovery of cryptic allosteric sites using reversed allosteric communication by a combined computational and experimental strategy. *Chem Sci* 2021;12:464–76.
- [85] Lu S, Zhang J. Small molecule allosteric modulators of G-protein-coupled receptors: drug-target interactions. *J Med Chem* 2019;62:24–45.
- [86] Fan J, Liu Y, Kong R, Ni D, Yu Z, Lu S, et al. Harnessing reversed allosteric communication: a novel strategy for allosteric drug discovery. *J Med Chem* 2021;64:17728–43.