PNAS

Supplementary Information for

The full activation mechanism of the adenosine A_1 receptor revealed by GaMD and supervised GaMD (Su-GaMD) simulations

Yang Li,[#, 1, 2] Jixue Sun,[#, 1] Dongmei Li,[*, 1] and Jianping Lin[*, 1, 3, 4]

[1]State Key Laboratory of Medicinal Chemical Biology, College of Pharmacy and Tianjin Key Laboratory of Molecular Drug Research, Nankai University, Haihe Education Park, 38 Tongyan Road, Tianjin 300350, China
[2]College of Life Sciences, Nankai University, Tianjin 300350, China
[3]Biodesign Center, Tianjin Institute of Industrial Biotechnology, Chinese Academy of Sciences, 32 West 7th Avenue, Tianjin Airport Economic Area, Tianjin 300308, China
[4]Platform of Pharmaceutical Intelligence, Tianjin International Joint Academy of Biomedicine, Tianjin 300457, China

[#]These authors contributed equally to this work.

*Corresponding authors: Dongmei Li and Jianping Lin

Email: dongmeili@nankai.edu.cn and jianpinglin@nankai.edu.cn

This PDF file includes:

Supplementary Materials and Methods
Supplementary Text
Figs. S1 to S14
Tables S1 to S5
Legends for Movies S1 to S5
SI References

Other supplementary materials for this manuscript include the following:

Movies S1 to S5

Supplementary Materials and Methods

General. All MD simulations were carried out using Amber 18 (1) with the PMEMD engine on a GPU cluster equipped with eight NVIDIA GTX 2080. Before running Su-GaMD simulations, the following preliminary phases were carried out: (i) system setup, (ii) system equilibration, (iii) GaMD (2) simulation and parameter calculation. The AMBER FF14SB force field (3) was used for proteins, the general AMBER force field (GAFF) (4) was used for ligands, and the AMBER lipid force field LIPID14 (5) was used for POPCs. A 12 Å cut-off was set for the non-bonded interaction. The SHAKE algorithm (6) integration was used to constrain the covalent bonds involving hydrogen atoms and the Particle Mesh Ewald (PME) algorithm (7) was applied to treat long-range electrostatic interactions. The time step was set to 2 fs. The frames were saved every 5000 steps for analysis. In the present work, three replicates were performed for each production run in all the simulations, except for the long-time 1000-ns GaMD simulation. The trajectories were analyzed with the VMD and CPPTRAJ tools in AMBER18 (1).

System setup. To simulate the association of G_i protein to A_1R in cytoplasm, we built two system groups of Ado-bound A_1R immersed in POPC bilayers. We first built a simplified ternary complex (Ado−A_1R−$G\alpha_i$) with the α subunit of G_{i2} protein ($G\alpha_i$) to present the heterotrimeric G_{i2} protein (Fig. S1A). The Ado, A_1R and $G\alpha_i$ were extracted from the 6D9H structure, the Gβγ and other unnecessary atoms were removed. The missing loop ICL3 of A_1R was built with homology modelling using the I-TASSER online server program (8). The protonation state for titratable residues in A_1R was determined using the H++ program (9) and the Tleap module of AMBER 18 (1). Then, the A_1R structure was inserted into 100 Å × 100 Å POPC bilayers. We placed the $G\alpha_i$ > 20 Å away from A_1R with Ado in the orthosteric site. After that, the system was solvated in a TIP3P water box and neutralized with 0.15 M NaCl (denoted as system A thereafter). The dimension of system A was 100 Å× 100 Å × 165 Å (Fig. S2A).

To investigate the molecular recognition pathways of Ado to A_1R and the whole heterotrimeric G_{i2} protein to A_1R, we built a ternary complex of Ado, G_{i2} and A_1R (Ado−A_1R−G_{i2}) from both active and inactive A_1R (systems B and C). The Ado, active A_1R and G_{i2} were extracted from the 6D9H structure. The inactive state of A_1R was extracted from the 5N2S structure (10). The A_1R structure was inserted into 120 Å × 110 Å POPC bilayers. In system B (Fig. S1B), Ado and G_{i2} were placed > 20 Å away from A_1R separately. System C was similar to system B except that the A_1R was replaced with its inactive state (Fig. S1C). The systems were solvated and neutralized in the same way of system A. The dimensions of systems B and C were 120 Å× 110 Å × 165 Å (system B as an example in Fig. S2B).

System equilibration. Firstly, each system was minimized for 10000 steps (5000 steps of steepest descent minimization followed by 5000 steps of conjugate gradient minimization). Secondly, each system was heated from 0 K to 310 K in 500 ps using the Langevin thermostat (11), the proteins, ligand and lipid head groups were constrained with a force constant of 50 kcal·mol^{-1}·Å^{-2}. Thirdly, a series of equilibrations were performed for each system. The POPCs were equilibrated for 30 ns, and the proteins and ligand were constrained with 50 kcal·mol^{-1}·Å^{-2}. Then, the added missing residues and atoms were optimized for 30 ns, and the other residues of proteins and ligand were constrained with 50 kcal·mol^{-1}·Å^{-2}. Finally, the whole system was released and equilibrated for 20 ns with no constrains.

The equilibrated coordinates of system A were marked as system A_1. And we repeat three times for the final 20 ns equilibration of system A to produce three different positions and orientations of $G\alpha_i$ relative to A_1R, which were marked as systems A_2, A_3 and A_4. Therefore, systems A_1, A_2, A_3 and A_4 were proposed to investigate the A_1R−$G\alpha_i$ binding event with different relative positions and orientations of $G\alpha_i$ (see Fig. 1C).

Gaussian accelerated molecular dynamics (GaMD). We performed 10-ns cMD simulation to calculate the GaMD acceleration parameters and 50-ns GaMD equilibration after adding the boost potential for each system. The GaMD acceleration parameters were applied for the following Su-GaMD simulations. We also performed a 1000-ns GaMD simulation for system A to compare with the Su-GaMD results.

Supervised Gaussian accelerated molecular dynamics (Su-GaMD). The Su-GaMD approach is a standard GaMD simulation in which a parameter (*Q*) is supervised by a tabu-like algorithm.

During the production of the GaMD trajectory, points with different *Q* values are collected "on the flight" at regular intervals (Δ*t*) and fitted into a linear function, $f(x) = m\mathrm{x}$. If the slope (*m*) is negative, the parameter *Q* is likely to decrease, and the GaMD simulation is restarted from the last set of coordinates. Otherwise, the simulation is restored from the original set of coordinates and started over. The supervision is repeated during the GaMD simulation until the parameter *Q* is less than the target value Q_0. Only the steps from which the slope (*m*) is negative are saved to analysis. We implemented the similar Su-GaMD workflow (Fig. S3) as described by Moro's group (12) by using our in-house python script. The GaMD part was performed by using Amber 18 (1). Besides, the CPPTRAJ tools were employed to operate the MD trajectories.

The Su-GaMD requires several configuration parameters. The parameters were included in the python script for processing the whole workflow. The python script is comprised of three major sections containing information about (i) the system, (ii) the supervision procedure, and (iii) the simulation settings.

In the system settings section, the following details about the molecular system were provided: (i) the PDB file name containing the starting coordinates of the complex for investigating the binding event, (ii) the PDB file name containing the targeting coordinates (the 6D9H structure) and (iii) the supervised parameter (*Q*) in simulation.

In the supervision settings section, the following values were declared: (i) the slope (*m*) threshold (default value: 0) and (ii) the target value (Q_0, default value: 5 Å) of the parameter to stop the simulation.

In the simulation settings section, the following details were specified: (i) the topology and input coordinate files, (ii) the parameter file to perform GaMD and (iii) GaMD acceleration parameters calculated in former GaMD simulation. In this section, a Boolean operator was used to supervise the *Q* through the non-supervised GaMD simulation of a definite time interval (Δ*t*). The time interval Δ*t* was set in the parameter file of GaMD simulation.

Design of the Su-GaMD simulations. We used the Su-GaMD method to investigate the A_1R−$G\alpha_i$ binding event and reconstruct the A_1R−$G\alpha_i$ complex. The final coordinates after equilibration of system A_1 was set as the starting coordinates, the 6D9H structure (13) was set as the targeting coordinates. In the supervision settings section, the supervised parameter (*Q*) was set to the $G\alpha_i$ RMSD. In the simulation settings section, to explore the appropriate time interval, the time interval was set to 300, 600 and 900 ps, respectively. To investigate the A_1R−$G\alpha_i$ binding event under different initial positions and orientations of $G\alpha_i$ relative to A_1R, we also performed Su-GaMD simulations for systems A_2, A_3 and A_4.

To compare this Su-GaMD method with previous Su-MD method by Moro's group (12), we performed Su-MD simulations for system A_1, with the same starting coordinates, targeting coordinates and *Q* of the Su-GaMD simulations. In the Su-MD simulations, the cMD was employed (without Gaussian acceleration).

Ado−A_1R binding and A_1R−G_{i2} binding. To investigate the full activation mechanism of A_1R from its active and inactive state, we performed Su-GaMD simulations for systems B and C. The final coordinates after equilibration of systems B and C were set as the starting coordinates, respectively, and the 6D9H structure (13) was set as the targeting coordinates. In the Su-GaMD simulation of system B, we supervised the Ado RMSD in the Ado−A_1R recognition process in presence of G_{i2}, and then supervised the $G\alpha_i$ RMSD in the A_1R−G_{i2} recognition process. While for system C, after the Ado−A_1R recognition Su-GaMD simulation in absence of G_{i2} (Su-GaMD-1, with Ado RMSD supervised), the inactive state of the system was proceeded to a 150 ns GaMD simulation to obtain a preactive state of A_1R. Then the G_{i2} was added to the system by placing it > 20 Å away from the preactive A_1R. After the minimization and equilibration phases, the A_1R−G_{i2} recognition Su-GaMD simulation was performed from the preactive Ado−A_1R complex (Su-GaMD-2, with $G\alpha_i$ RMSD supervised). For more comparison, Su-MD simulation of A_1R−G_{i2} recognition from preactivated A_1R and G_{i2} was performed as well (Su-MD-2, with $G\alpha_i$ RMSD supervised).

Binding free energy calculations. The binding free energies between Ado and A_1R and between A_1R and G_{i2} protein were calculated using the molecular mechanics generalized born surface area (MM/GBSA) (14) approach. The 60 ps trajectories prior to the calculated frame were extracted from the Su-GaMD trajectories to calculate the binding free energies. All the parameters were set as default values in the calculations.

Pocket volume calculations. The volumes of the extracellular Ado-binding pocket and the intracellular G_{i2}-binding pocket were calculated using the POVME program (15). The 1.2 ns trajectories prior to the frame of States a^*, b^*, c^* and the Final State were extracted from the Su-GaMD trajectories to calculate the volumes.

Network analysis. The NetworkView plugin in VMD was employed to perform the protein contact network analysis for States a, b, c and d in the A_1R-G_{i2} recognition pathway. In the network analysis, each protein residue was considered as a node. The nodes which had C_α atom within 4.5 Å were defined as "in-contact" nodes. Edges was added to the network by connecting pairs of "in-contact" nodes.

Supplementary Text

Details of the Su-GaMD and GaMD simulations for $A_1R-G\alpha_i$ recognition process. To investigate the $A_1R-G\alpha_i$ binding event, we placed the $G\alpha_i$ > 20 Å away from A_1R with the Ado in the orthosteric site (system A, Fig. S1A), the Su-GaMD simulations were performed to reconstruct the $A_1R-G\alpha_i$ complex with the $G\alpha_i$ RMSD supervised (system A in Table 1).

Time interval of Su-GaMD. At the beginning, to select an appropriate time interval for Su-GaMD, three replicates of Su-GaMD simulation were performed with time intervals of 300, 600 and 900 ps, respectively. $G\alpha_i$ was successfully observed to enter the intracellular binding site of A_1R (Fig. 1A, the $G\alpha_i$ RMSD reached < 5 Å in Fig. 1B) in less than 50 ns Su-GaMD simulations. $G\alpha_i$ contacted with the residues of A_1R in the intracellular end of TM5-TM7 and helix 8 (H8) through its α5-helix, which was similar to that in the 6D9H structure (red ribbons for α5-helix of 6D9H in Fig. 1A). The $G\alpha_i$ RMSD well as the $A_1R-G\alpha_i$ distance in each trajectory were calculated. During the Su-GaMD simulations, the $G\alpha_i$ RMSD fell from 52.9 Å to ~4.7 Å, and the $A_1R-G\alpha_i$ distance reduced from 69.4 Å to ~35.9 Å (Fig. 1B, system A_1 in Fig. 1C). The minimum $G\alpha_i$ RMSD and $A_1R-G\alpha_i$ distance fell to the approximate value of the 6D9H structure during each Su-GaMD simulation (Table S1). These results indicate that the $A_1R-G\alpha_i$ complex close to the 6D9H structure was reconstructed through these Su-GaMD simulations.

When we employed the time intervals of 300, 600 and 900-ps (system A_1 in Table 1), the Su-GaMD simulation time were 17.2, 31.2 and 41.4 ns, respectively. Thus, the larger time interval we used, the longer simulation time and more computing resources we needed to reconstruct the $A_1R-G\alpha_i$ complex. As a reasonable compromise based on both enough sampling number and a shorter simulation time, we chose the 600-ps interval (same as the previous SuMD works of Moro's group (16, 17)) and employed it for the Su-GaMD simulations for the rest systems.

Different initial positions and orientations of $G\alpha_i$ relative to A_1R. We performed Su-GaMD simulations for systems A_2, A_3 and A_4 which had different initial positions and orientations of $G\alpha_i$ relative to A_1R (systems A_2 to A_4 in Fig. 1C and Table 1). Without exception, after 25.0, 18.2 and 30.0-ns Su-GaMD simulations for systems A_2, A_3 and A_4, the $G\alpha_i$ RMSD fell to ~4.6 Å from the starting value ($RMSD_0$) of 40.5, 34.4 and 24.2 Å. The $A_1R-G\alpha_i$ distance reduced from 57.3, 49.1 and 45.3 Å to ~34.2 Å (systems A_2 to A_4 in Fig. 1C, Tables 1 and S2). These suggests that the $G\alpha_i$ can enter its binding site in A_1R and achieve the $A_1R-G\alpha_i$ complex similar to the 6D9H structure in a reasonable Su-GaMD simulation time no matter where we placed it and what orientation of it in the beginning (Fig. 1C).

For comparison, we performed an 1000-ns unsupervised GaMD simulation for system A_1. Ado was observed to be stable in the extracellular ligand binding site, with the average Ado RMSD of 1.1 Å in the last 10-ns GaMD trajectories (Fig. S4A). However, the minimum $G\alpha_i$ RMSD was 25.8 Å (Fig. S4B), and $G\alpha_i$ only formed some loosely contacts with the receptor intracellular surface in the extremely long-time GaMD simulation, suggesting that the $G\alpha_i$ did not entered the intracellular binding site of A_1R in the long-time unsupervised GaMD simulation. In addition, we performed three parallel Su-MD simulations (without Gaussian acceleration) for system A_1 and compared the results with those of Su-GaMD simulations. The $G\alpha_i$ RMSDs and $G\alpha_i-A_1R$ distances in the three replicates of Su-MD simulation are depicted in Fig. S5. The mimimum $G\alpha_i$ RMSDs and the minimum $A_1R-G\alpha_i$ distances of the Su-MD simulations are depicted in Table 3. We found that the Su-MD simulations could reconstruct the $A_1R-G\alpha_i$ complex as well, but the simulation time were 45.0, 54.6 and 75.6 ns (Fig. S5), which were longer than those of the Su-GaMD simulations (30.0, 30.0 and 33.6 ns, see Fig. 1B). The mimimum $G\alpha_i$ RMSDs of the Su-MD simulations were comparable with those of

the Ga-SuMD simulations (4.6, 5.0 and 4.9 Å for Su-MD *vs*. 4.9, 4.9 and 4.9 Å for Su-GaMD, see Tables S3 and S1). The minimum $A_1R-G\alpha_i$ distances of the Su-MD simulations were longer than those of the Ga-SuMD simulations (37.0, 37.2 and 38.8 Å for Su-MD *vs*. 33.6, 33.6 and 35.2 Å for Su-GaMD, see Tables S3 and S1).

In summary, we can reconstruct the $A_1R-G\alpha_i$ complex in the binding mode similar to that of the 6D9H structure and observed the $A_1R-G\alpha_i$ recognition process in less than 50 ns by using the Su-GaMD strategy, while this $A_1R-G\alpha_i$ complex cannot be reached even in long-time (e.g. 1000 ns) unsupervised GaMD simulation.

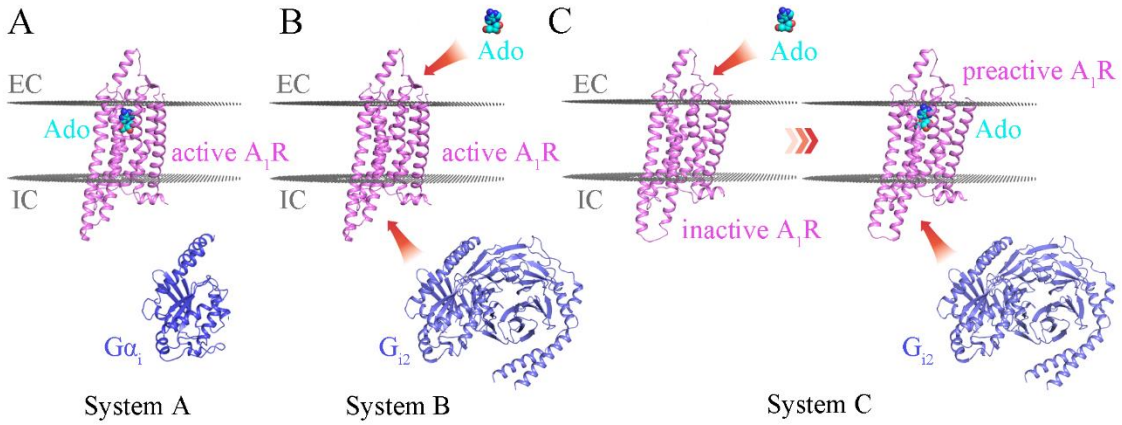

Fig. S1. Overall structures of systems A, B and C.

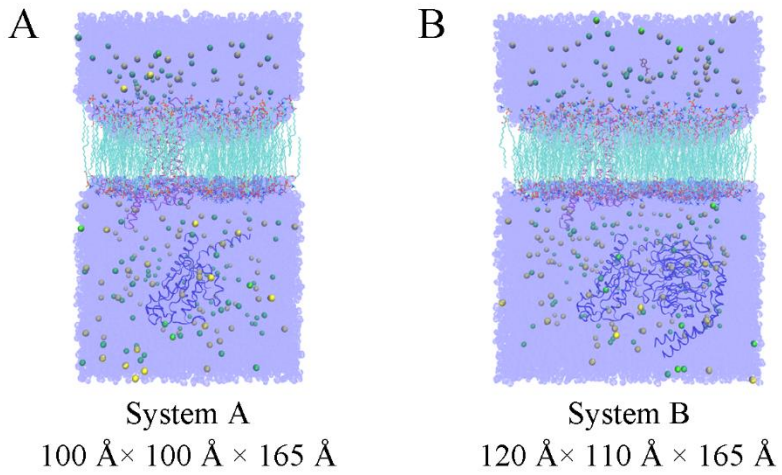

Fig. S2. Schematic representation of (A) system A and (B) system B.

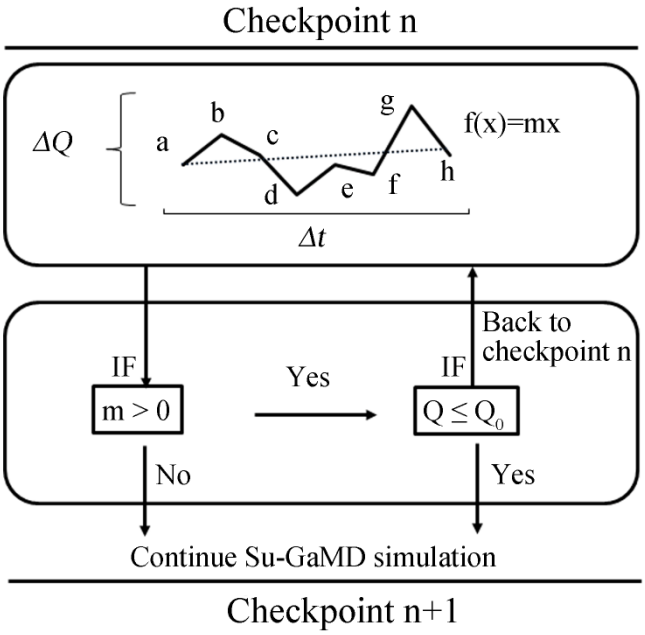

Fig. S3. Workflow of the Su-GaMD simulation.

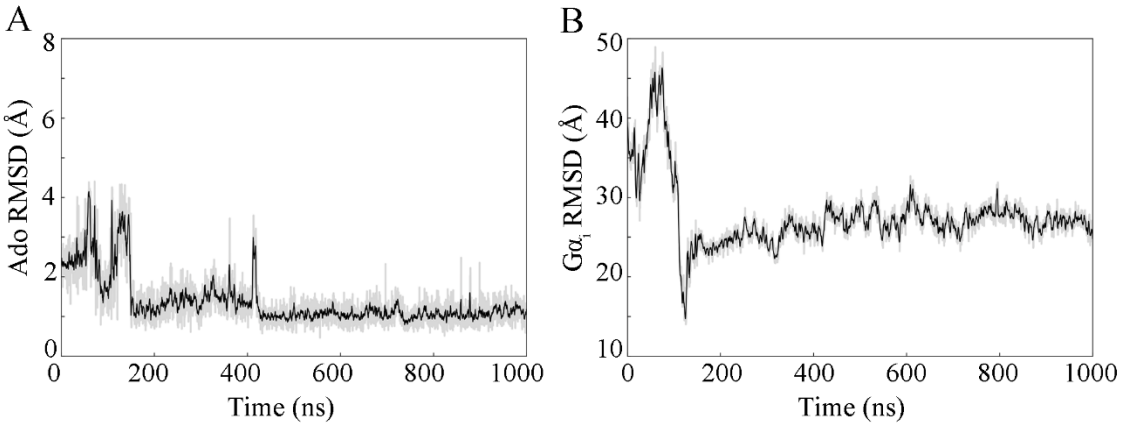

Fig. S4. Time dependences of the (A) Ado RMSD and (B) $G\alpha_i$ RMSD in the 1000-ns GaMD simulation for system A_1.

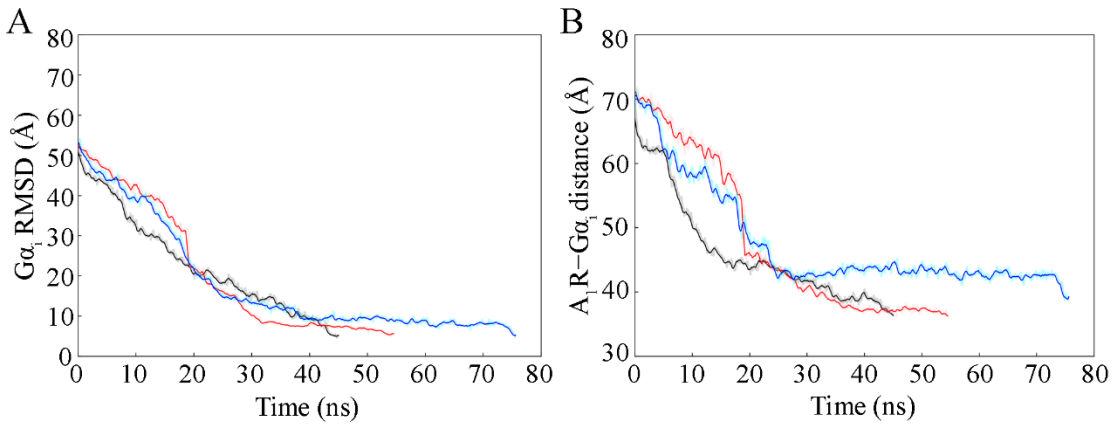

Fig. S5. Time dependences of (A) Gαᵢ RMSDs and (B) A_1R−Gαᵢ distances of three independent Su-MD simulations (without Gaussian acceleration) for system A_1. The three different colored lines represent the results of the three independent Su-MD simulations.

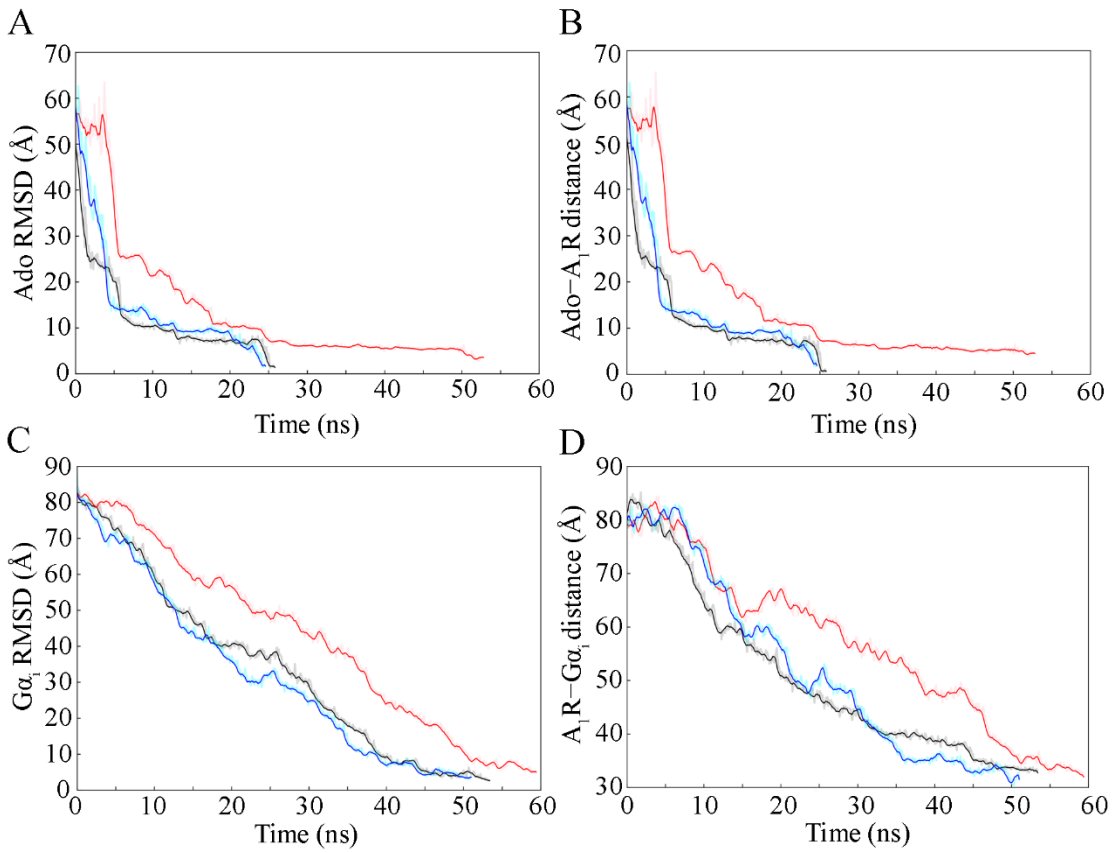

Fig. S6. Time dependences of (A) Ado RMSDs, (B) Ado−A_1R distances, (C) $G\alpha_i$ RMSDs and (D) A_1R−$G\alpha_i$ distances of three independent Su-GaMD simulations for the reconstruction of the Ado−A_1R−G_{i2} complex from the active A_1R structure. The three different colored lines represent the results of the three independent Su-GaMD simulations.

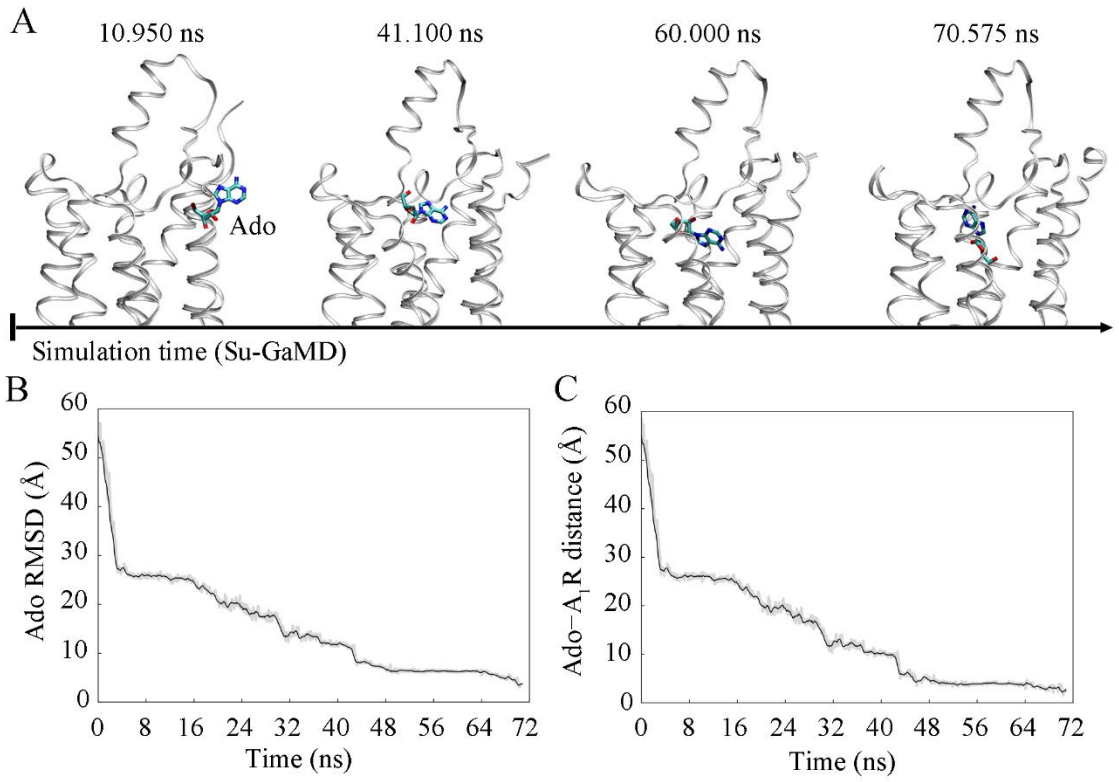

Fig. S7. (A) The landscape of the Ado−A_1R recognition process in the simulation on a model with the five N-terminal residues added. Time dependences of (B) Ado RMSD and (C) Ado−A_1R distance during the simulation.

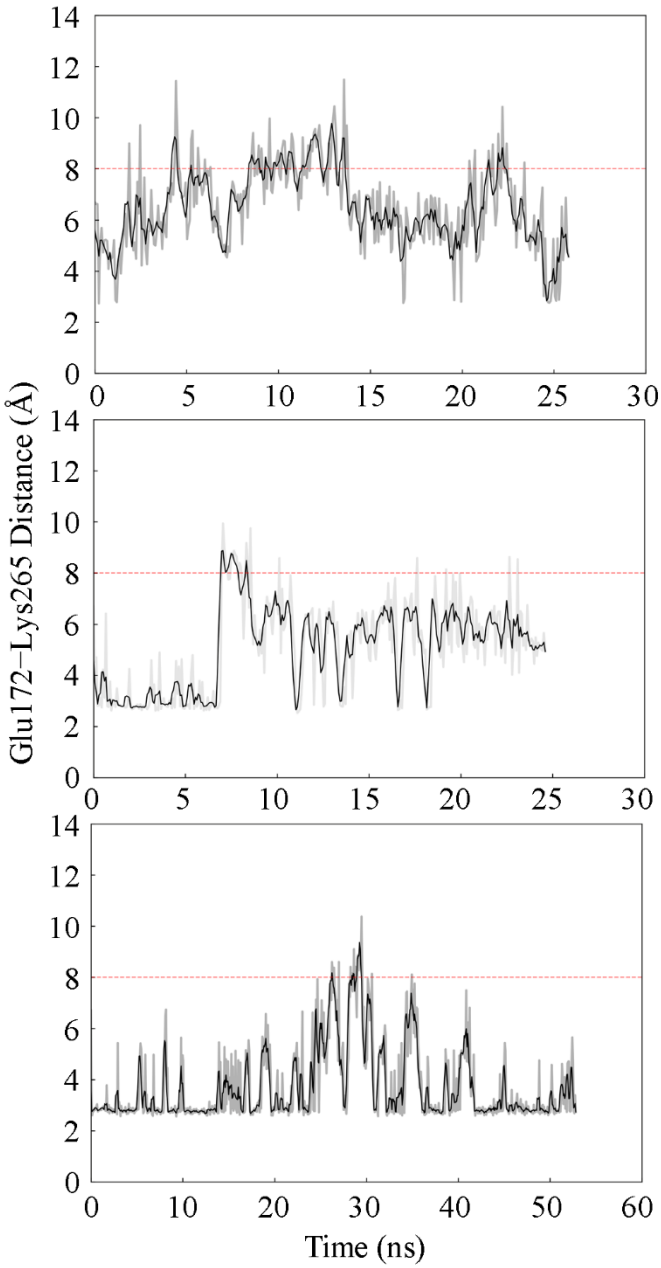

Fig. S8. Time dependences of the Glu172^{ECL2}−Lys265^{ECL3} salt bridge of three independent Su-GaMD simulations for the Ado−A_1R binding process.

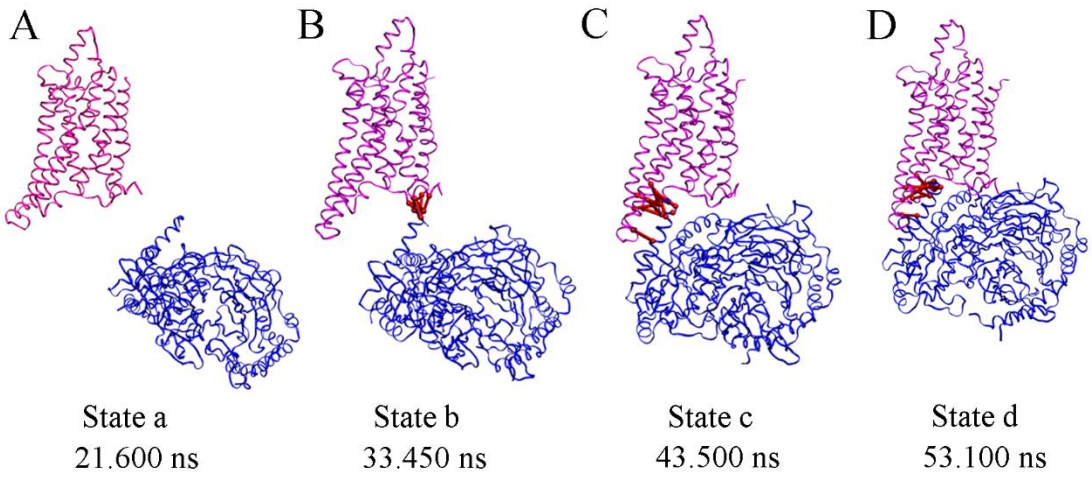

Fig. S9. A_1R-G_{i2} protein contact networks for States a, b, c and d. The "in-contact" between A_1R and G_{i2} were connected with red thick lines.

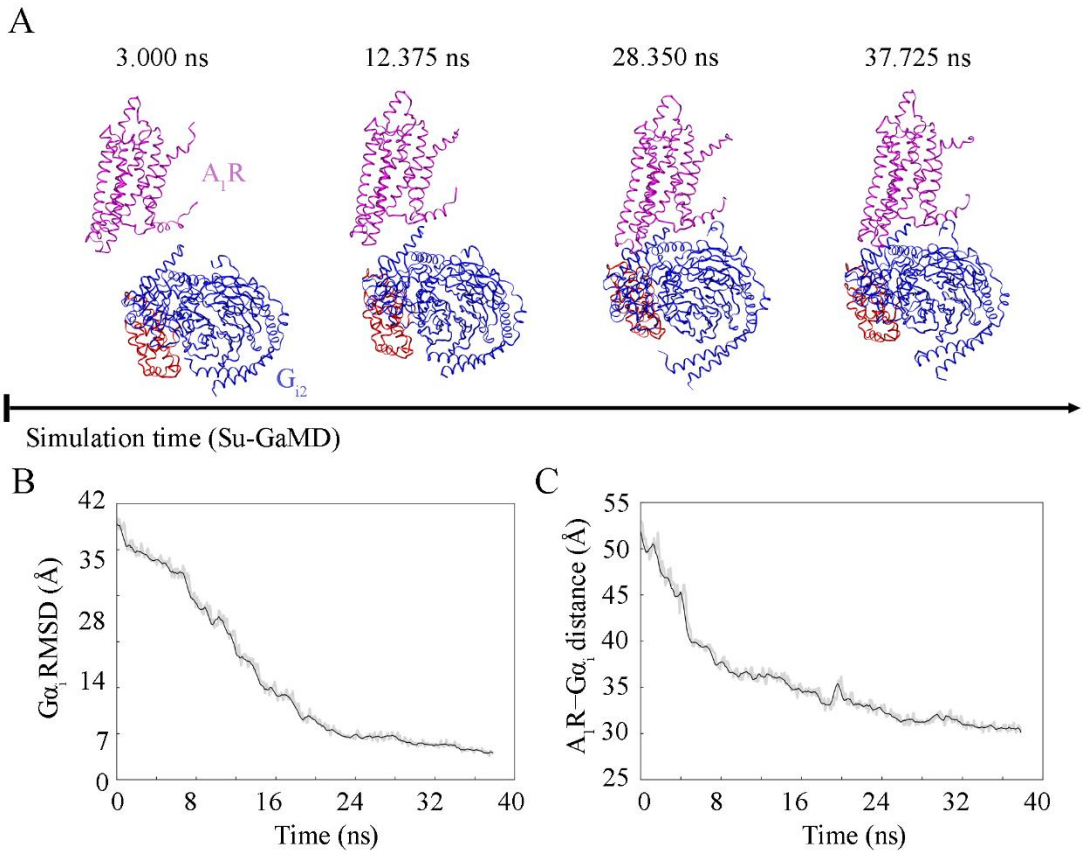

Fig. S10. (A) The landscape of the A_1R-G_{i2} recognition process in the simulation on a model with the helical domain (shown in red) of G_{i2} rebuilt. Time dependences of (B) $G\alpha_i$ RMSD and (C) $A_1R-G\alpha_i$ distance during in the simulation.

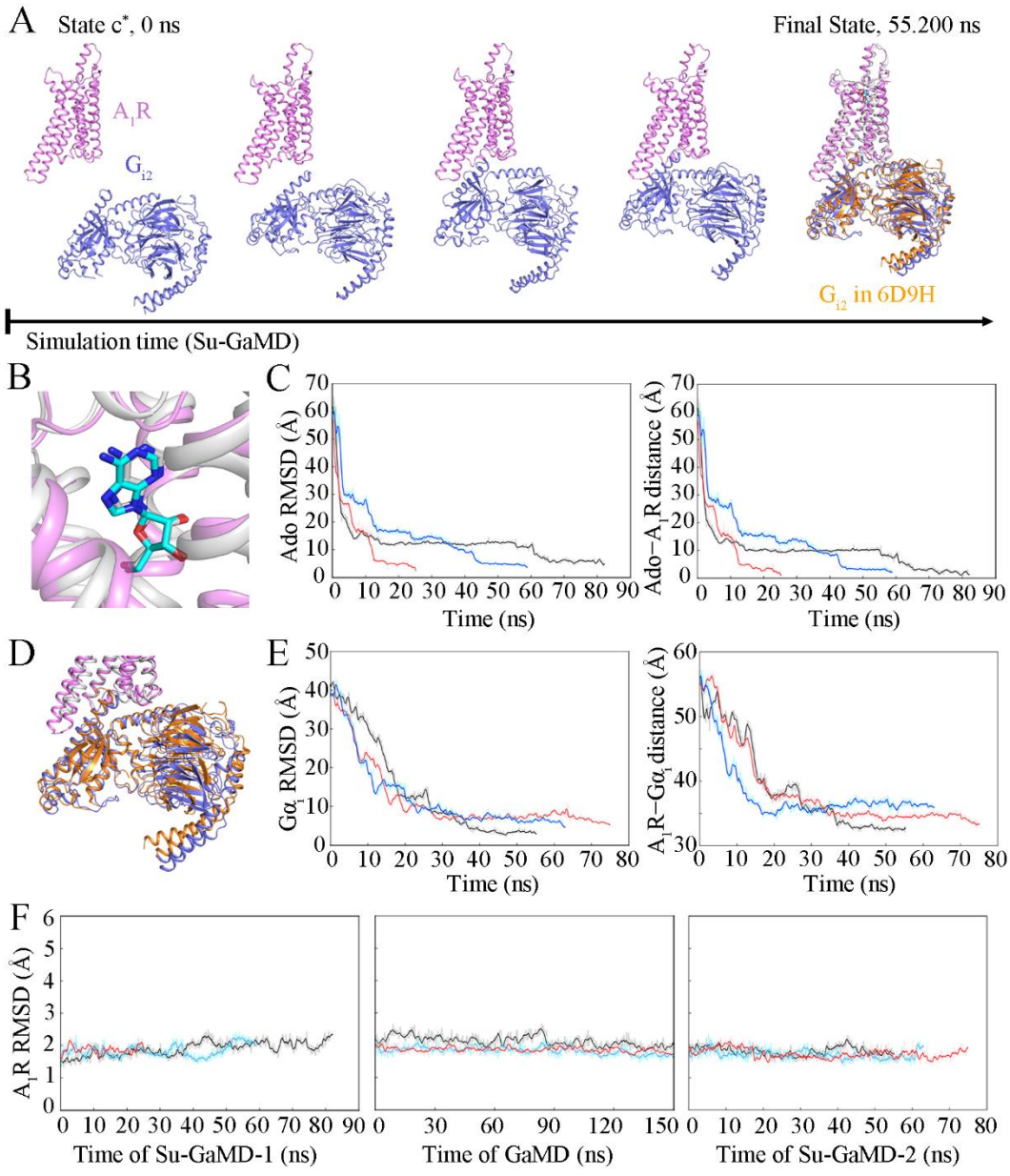

Fig. S11. (A) The landscape of A_1R−G_{i2} recognition pathway from the preactive state of A_1R. Relative position of G_{i2} after global alignment of A_1R in the final snapshot (A_1R is shown in violet, and G_{i2} is shown in blue) of the Su-GaMD simulation to that of the 6D9H structure (G_{i2} is shown in orange) was shown in the Final State. (B) Overlay of the Ado−A_1R complex extracted from the trajectory of Su-GaMD-2 (Ado is shown in cyan) and the 6D9H structure (Ado is shown in silver). (C) Time dependences of the Ado RMSDs and Ado−A_1R distances in the three replicates of Su-GaMD-1. (D) Overlay of the Ado−A_1R−G_{i2} complex extract from the trajectory of Su-GaMD-2 (G_{i2} is shown in blue) and the 6D9H structure (G_{i2} is shown in orange). (E) Time dependences of the $G\alpha_i$ RMSDs and the A_1R−$G\alpha_i$ distances in the three replicates of Su-GaMD-2. (F) Time dependences of the A_1R RMSD during the three stages of simulations.

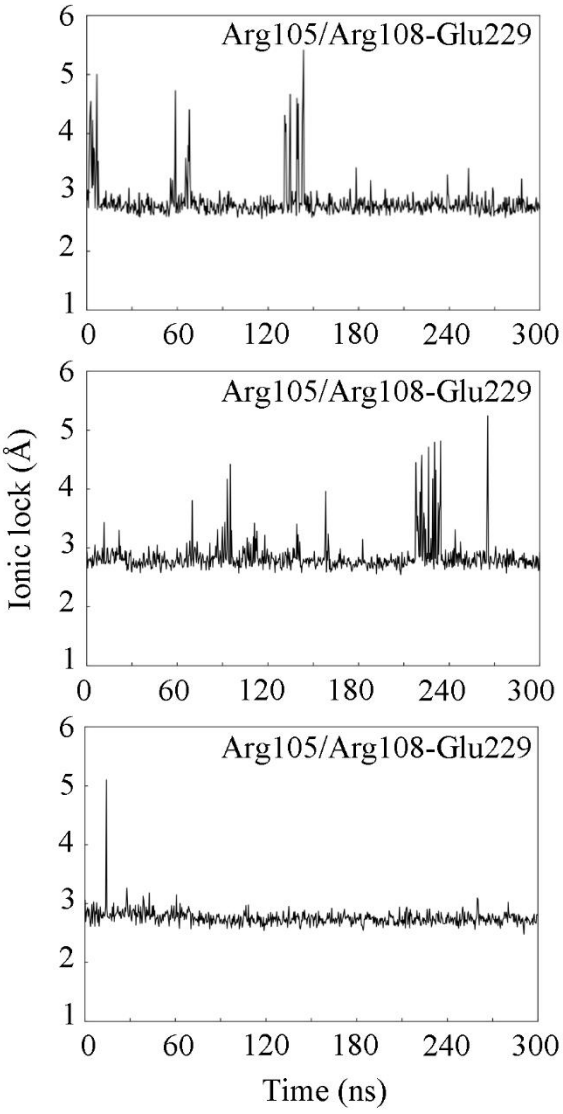

Fig. S12. Time-dependent N--O distances between the guanidinium of Arg105$^{3.50}$/Arg108$^{3.53}$ and the carboxyl of Glu229$^{6.30}$ during the three parallel GaMD simulations of apo-A_1R.

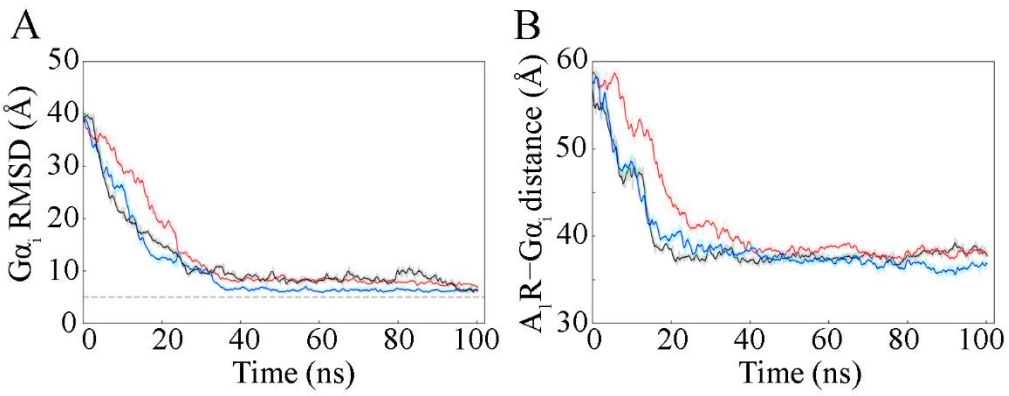

Fig. S13. Time dependences of (A) $G\alpha_i$ RMSDs and (B) $A_1R-G\alpha_i$ distances of three independent Su-MD-2 for the reconstruction of the $Ado-A_1R-G_{i2}$ complex from the preactived $Ado-A_1R$ and G_{i2}. The three different colored lines represent the results of the three independent Su-MD-2 simulations.

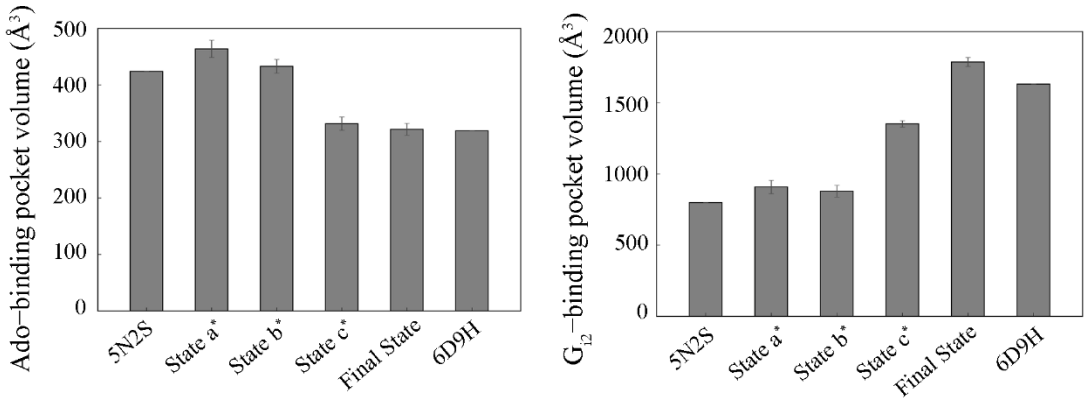

Fig. S14. The volumes of the Ado-binding pocket and the G_{i2}-binding pocket during the full activation process of A_1R.

Table S1. The minimum $G\alpha_i$ RMSDs and minimum $A_1R-G\alpha_i$ distances in the nine replicates of Su-GaMD simulation for system A_1.

Entry	Time interval of Su-GaMD (ps)	Time (ns)	Minimum $G\alpha_i$ RMSD (Å)	Minimum $A_1R-G\alpha_i$ distance (Å)
6D9H structure	–	–	0	32.8
System A_1	300	18.9	4.9	36.4
		15.0	3.5	37.0
		17.7	4.5	37.4
	600	30.0	4.9	33.6
		30.0	4.9	33.6
		33.6	4.9	35.2
	900	41.4	4.9	36.6
		35.1	4.9	35.6
		47.7	4.8	37.3

Table S2. The minimum Gαi RMSDs and minimum A_1R−Gαi distances in the three replicates of Su-GaMD simulation for systems A_2, A_3 and A_4.

Entry	Initial Gαi RMSD/ A_1R−Gαi distance (Å)	Time (ns)	Minimum Gαi RMSD (Å)	Minimum A_1R−Gαi distance (Å)
System A_2	40.5/57.3	26.4	4.9	33.5
		36.0	4.4	33.3
		12.6	3.8	34.6
System A_3	34.4/49.1	16.8	4.9	35.9
		19.2	3.9	34.1
		18.6	4.7	33.5
System A_4	24.2/45.3	22.8	5.0	33.9
		48.0	4.6	34.2
		19.2	4.9	34.7

Table S3. The minimum Gαᵢ RMSDs and minimum A_1R−Gαᵢ distances in the three replicates of Su-MD simulation for system A_1.

Entry	Time interval of Su-MD (ps)	Time (ns)	Minimum Gαᵢ RMSD (Å)	Minimum A_1R−Gαᵢ distance (Å)
System A_1	600	45.0	4.6	37.0
		54.6	5.0	37.2
		75.6	4.9	38.8

Table S4. The minimum $G\alpha_i$ RMSDs and minimum $A_1R-G\alpha_i$ distances in the three replicates of Su-GaMD-2 and Su-MD-2 for the A_1R-G_{i2} recognition process from the preactived A_1R and G_{i2}.

Entry	Time (ns)	Minimum $G\alpha_i$ RMSD (Å)	Minimum $A_1R-G\alpha_i$ distance (Å)
Su-GaMD-2	55.2	2.9	32.1
	63.0	4.4	34.0
	75.0	5.0	32.9
Su-MD-2	100.2	5.7	35.3
	100.2	6.7	37.0
	100.2	5.7	36.2

Table S5. Key residue interactions between A_1R and G_{i2} in the 6D9H structure and State d. Same interactions were colored blue.

	Hydrogen Bond	Salt Bridge
6D9H	Gln210$^{5.68}$-Asp342 Lys228$^{5.68}$-Phe355 Gln210$^{5.68}$-Lys346	Arg108$^{3.53}$-Asp351 Lys294^{H8}-Asp351 Lys213$^{5.71}$-Asp342 Lys224$^{6.25}$-Asp316
State d	Gln210$^{5.68}$-Asp342 Lys228$^{5.68}$-Phe355 Pro112^{ICL2}-Asn348 Tyr115^{ICL2}-Asn348	Arg108$^{3.53}$-Asp351 Lys294^{H8}-Asp351 Lys213$^{5.71}$-Asp342 Lys214$^{5.72}$-Asp342

Movie S1 (separate file). The Ado−A_1R recognition process observed in the Su-GaMD simulation of the reconstruction of the Ado−A_1R−G_{i2} complex from the active A_1R structure.

Movie S2 (separate file). The A_1R−G_{i2} recognition process observed in the Su-GaMD simulation of the reconstruction of the Ado−A_1R−G_{i2} complex from the active A_1R structure.

Movie S3 (separate file). The Ado−A_1R recognition process observed in the Su-GaMD-1 simulation of the reconstruction of the Ado−A_1R−G_{i2} complex from the inactive A_1R structure.

Movie S4 (separate file). The preactivation process of A_1R observed in the GaMD simulation.

Movie S5 (separate file). The A_1R−G_{i2} recognition pathway process observed in the Su-GaMD-2 simulation of the reconstruction of the Ado−A_1R−G_{i2} complex from the inactive A_1R structure.

SI References

1. D. A. Case *et al.* AMBER 2018 (University of California, San Francisco), (2018).
2. Y. Miao, J. A. McCammon, Gaussian accelerated molecular dynamics: Theory, implementation, and applications. *Annu. Rep. Comput. Chem.* **13**, 231-278 (2017).
3. J. A. Maier *et al.*, ff14SB: Improving the accuracy of protein side chain and backbone parameters from ff99SB. *J. Chem. Theory Comput.* **11**, 3696-3713 (2015).
4. J. Wang *et al.*, Development and testing of a general Amber force field. *J. Comput. Chem.* **25**, 1157-1174 (2004).
5. C. J. Dickson *et al.*, Lipid14: The Amber lipid force field. *J. Chem. Theory Comput.* **10**, 865-879 (2014).
6. J.-P. Ryckaert, G. Ciccotti, H. J. C. Berendsen, Numerical integration of the cartesian equations of motion of a system with constraints: Molecular dynamics of N-alkanes. *J. Comput. Phys.* **23**, 327-341 (1977).
7. T. Darden, D. York, L. Pedersen, Particle mesh Ewald: An Nlog(N) method for Ewald sums in large systems. *J. Chem. Phys.* **98**, 10089-10092 (1993).
8. J. Yang *et al.*, The I-TASSER suite: Protein structure and function prediction. *Nat. Methods* **12**, 7-8 (2015).
9. R. Anandakrishnan, B. Aguilar, A. V. Onufriev, H++ 3.0: Automating pK prediction and the preparation of biomolecular structures for atomistic molecular modeling and simulations. *Nucleic Acids Res.* **40**, W537-W541 (2012).
10. R. K. Y. Cheng *et al.*, Structures of human A_1 and A_{2A} adenosine receptors with xanthines reveal determinants of selectivity. *Structure* **25**, 1275-1285 (2017).
11. R. W. Pastor, B. R. Brooks, A. Szabo, An analysis of the accuracy of Langevin and molecular dynamics algorithms. *Mol. Phys.* **65**, 1409-1419 (1988).
12. D. Sabbadin, S. Moro, Supervised molecular dynamics (SuMD) as a helpful tool to depict GPCR–ligand recognition pathway in a nanosecond time scale. *J. Chem. Inf. Model.* **54**, 372-376 (2014).
13. C. J. Draper-Joyce *et al.*, Structure of the adenosine-bound human adenosine A_1 receptor–G_i complex. *Nature* **558**, 559-563 (2018).
14. E. Wang *et al.*, End-point binding free energy calculation with MM/PBSA and MM/GBSA: Strategies and applications in drug design. *Chem. Rev.* **119**, 9478-9508 (2019).
15. J. D. Durrant, C. A. F. de Oliveira, J. A. McCammon, POVME: An algorithm for measuring binding-pocket volumes. *J. Mol. Graph. Model.* **29**, 773-776 (2011).
16. V. Salmaso, M. Sturlese, A. Cuzzolin, S. Moro, Exploring protein-peptide recognition pathways using a supervised molecular dynamics approach. *Structure* **25**, 655-662 (2017).
17. A. Cuzzolin *et al.*, Deciphering the complexity of ligand-protein recognition pathways using supervised molecular dynamics (SuMD) simulations. *J. Chem. Inf. Model.* **56**, 687-705 (2016).