Supporting Information

GLOW: A workflow integrating Gaussian accelerated molecular dynamics and Deep Learning for free energy profiling

Hung N. Do[1], Jinan Wang[1], Apurba Bhattarai[1], Yinglong Miao[1,*]

[1]The Center for Computational Biology and Department of Molecular Biosciences, The University of Kansas, Lawrence, Kansas 66047

*Corresponding author: miao@ku.edu

Figure S1. Learning curves of the training and validation datasets for the A_1AR activation (A and B) and allosteric modulation (C and D). The metrics used here are loss (A and C) and accuracy (B and D).

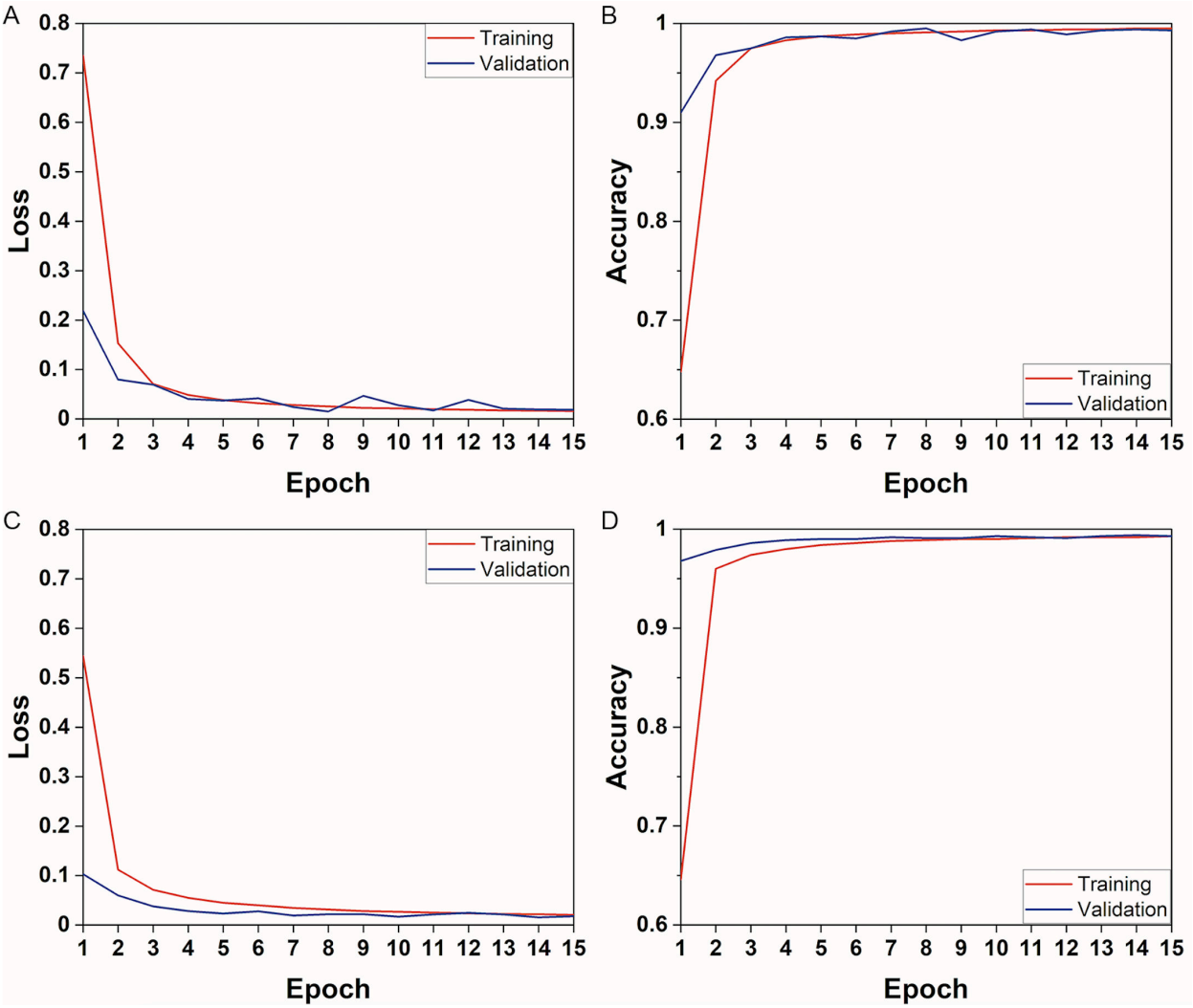

Figure S2. Saliency (attention) maps of residue contact gradients of the A_1AR bound by (A) "Antagonist", (B) "Agonist" and (C) "Agonist – Gi". The seven transmembrane helices are labeled from I to VII. The gradient of each residue contact is shown in a 0 (blue) to 1 (red) color scale.

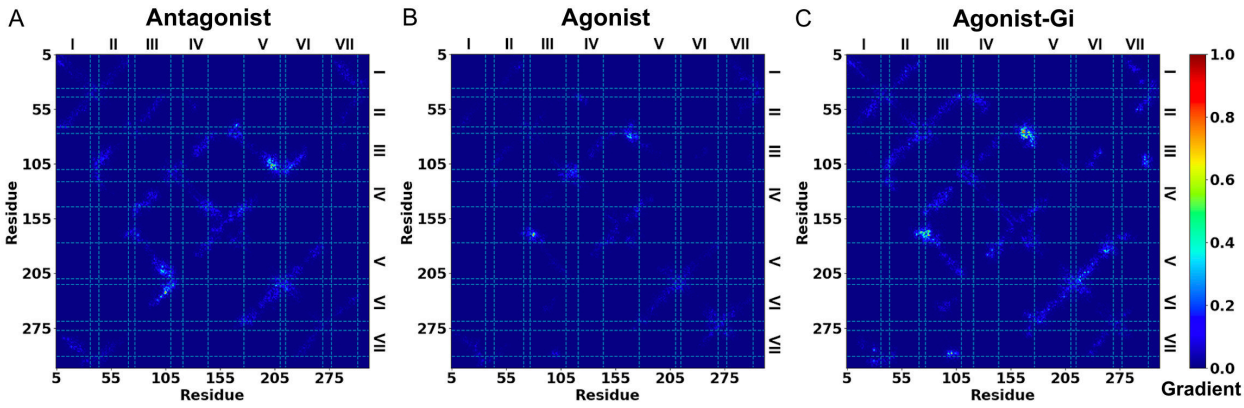

Figure S3. Locations of characteristic residue contacts in the activation of the A_1AR, which are shown in sticks and spheres and colored in red.

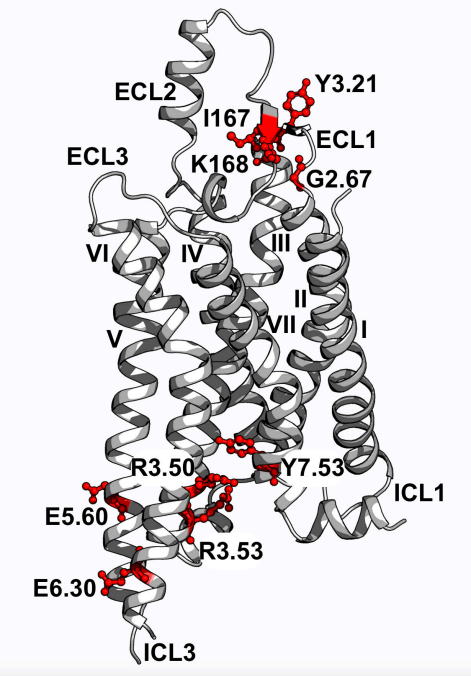

Figure S4. Time courses of selected reaction coordinates for free energy profiling of activation of A_1AR.

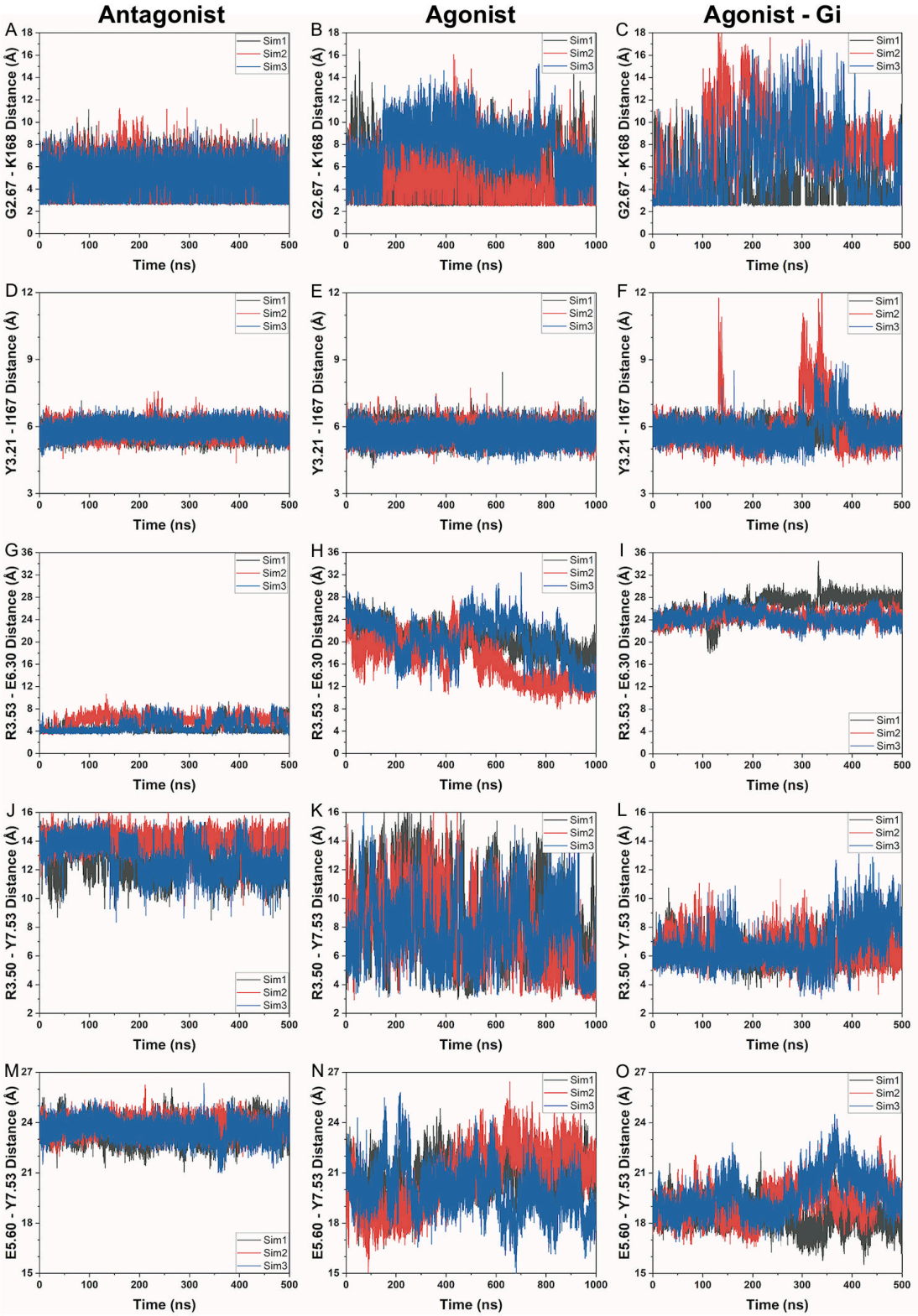

Figure S5. Saliency (attention) maps of residue contact gradients of the A_1AR bound by (A) "Agonist – Gi" and (B) "Agonist – Gi – PAM". The seven transmembrane helices are labeled from I to VII. The gradient of each residue contact is shown in a 0 (blue) to 1 (red) color scale.

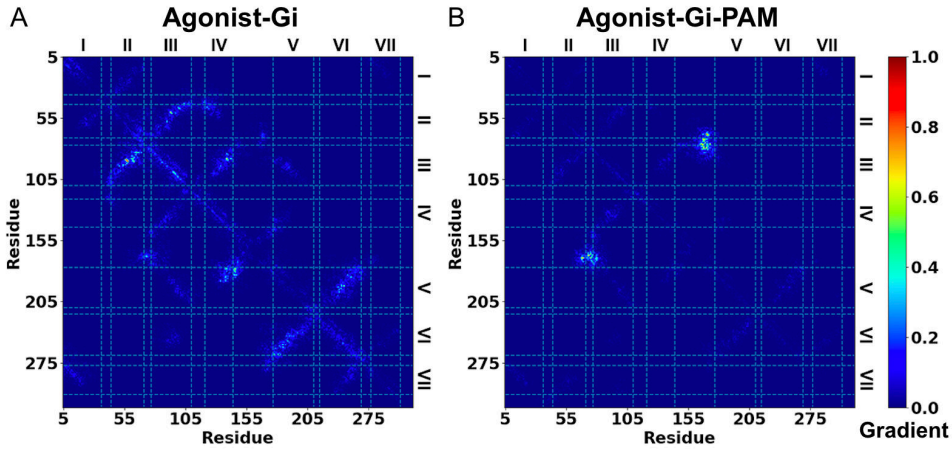

Figure S6. Locations of characteristic residue contacts in the allosteric modulation of the A_1AR, which are shown in sticks and spheres and colored in red.

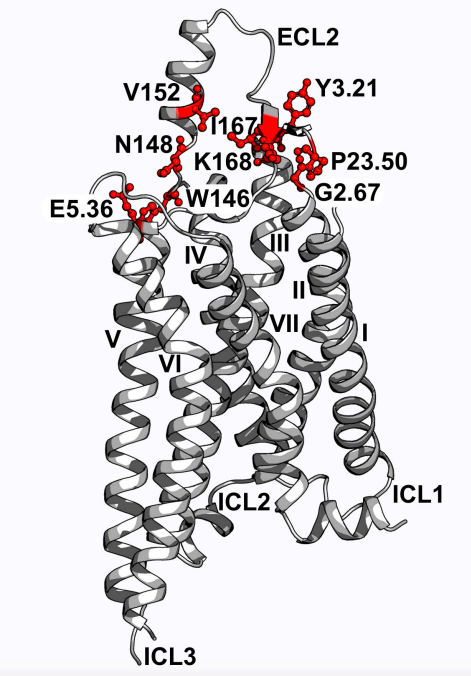

Figure S7. Time courses of selected reaction coordinates for free energy profiling of allosteric modulation of A_1AR.

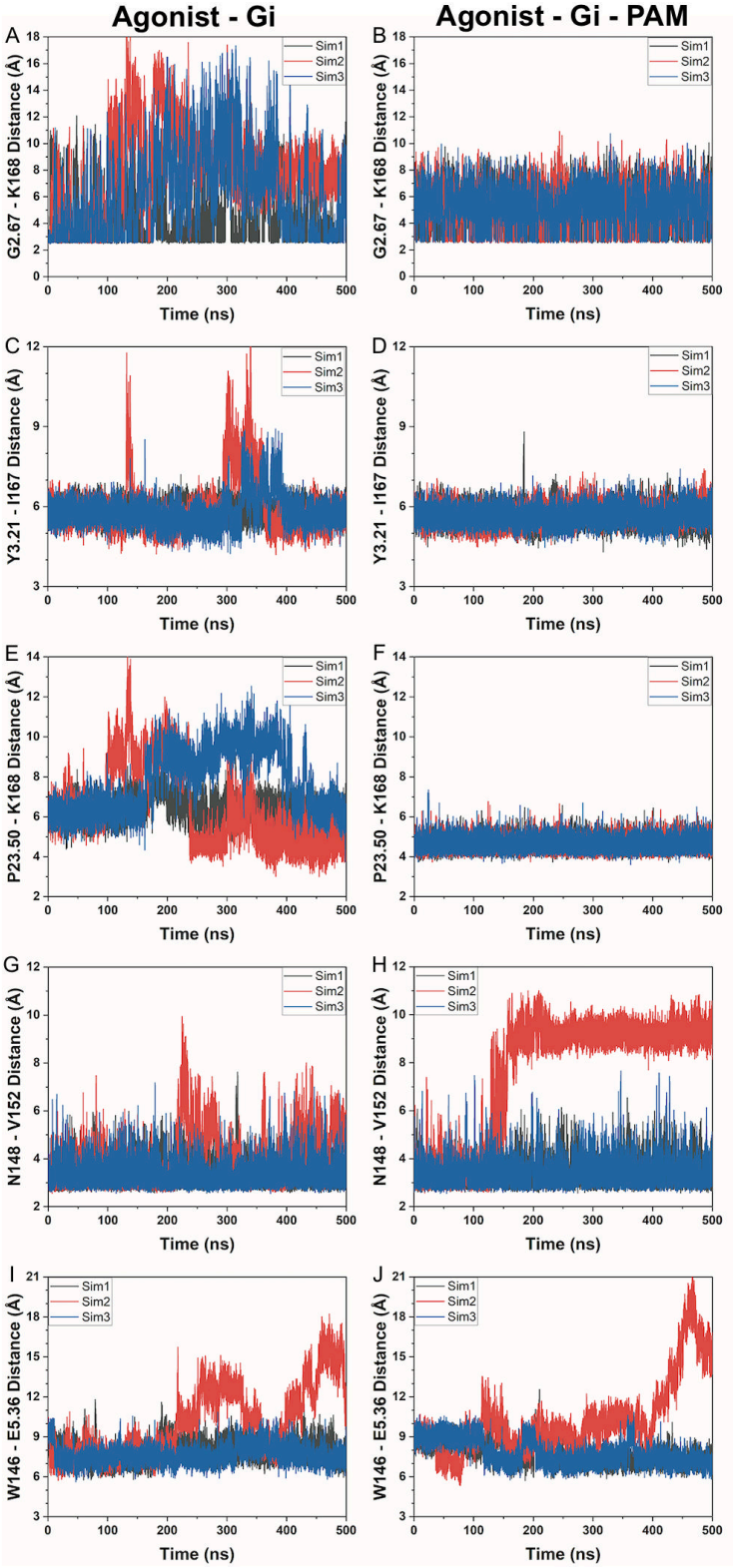

Figure S8. Independent Deep Learning analyses of the A_1AR activation and allosteric modulation with random assignments of training and validation datasets. (**A-G**) Confusion matrices and saliency (attention) maps of residue contact gradients calculated from the first DL model (also shown in **Figures 2** and **5**). (**H-N**) Confusion matrices and saliency (attention) maps of residue contact gradients from the second DL model. (**O-U**) Confusion matrices and saliency (attention) maps of residue contact gradients from the third DL model.

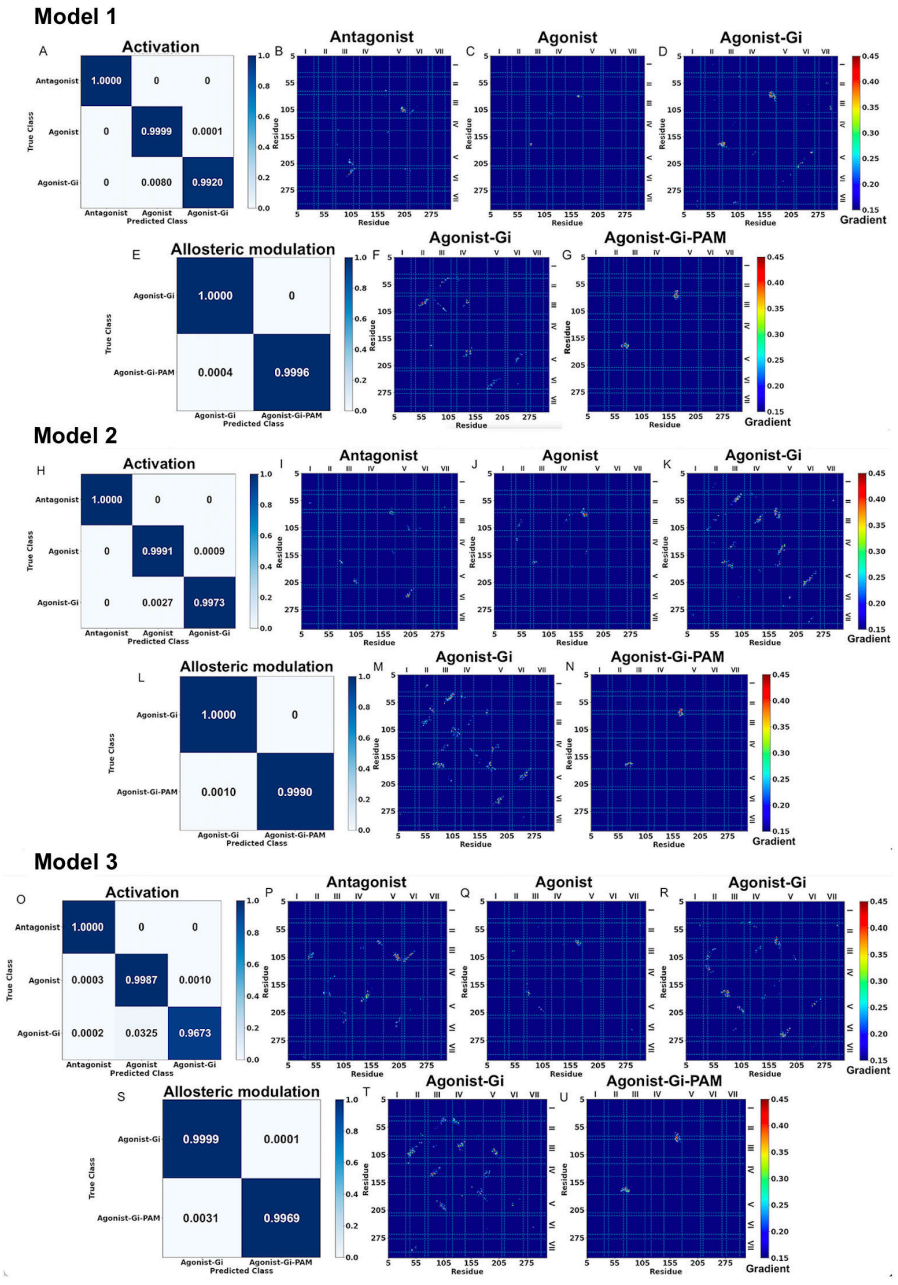

Figure S9. Flexibility of the adenosine receptor domain in the A_1AR bound by "Antagonist", "Agonist", "Agonist – Gi" and "Agonist – Gi – PAM" observed in GaMD simulations. The root-mean-square fluctuations (RMSF) are shown in a blue (0Å) – white – red (4Å) color scale.

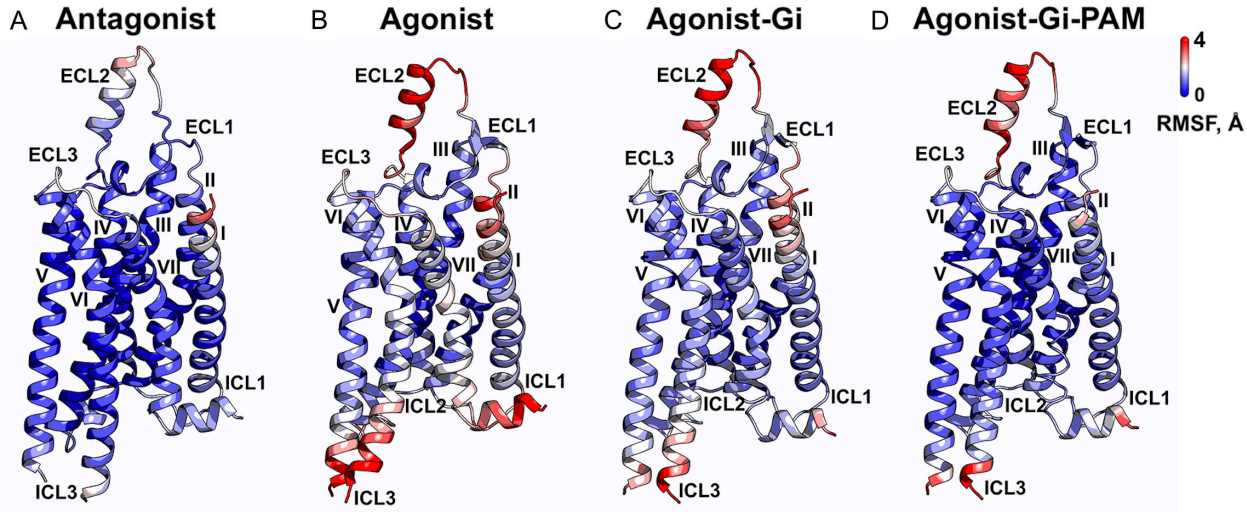

Table S1. Summary of GaMD simulations performed on the model A_1AR systems, including the A_1AR bound by "Antagonist", "Agonist", "Agonist – Gi" and "Agonist – Gi – PAM".

System	Method	Simulation Length	Boost Potential (kcal/mol)
"Antagonist"	GaMD_Dual	3 x 500 ns	15.07 ± 4.53
"Agonist"	GaMD_Dual	3 x 1000 ns	17.07 ± 5.01
"Agonist - Gi"	GaMD_Dual	3 x 500 ns	18.32 ± 6.29
"Agonist - Gi - PAM"	GaMD_Dual	3 x 500 ns	20.53 ± 6.54

Table S2. Characteristic residue contacts (whose gradients from pixel attribution ≥0.4) of the A_1AR activation.

Gradient	"Antagonist"		"Agonist"	"Agonist - Gi"	
0.9	R3.50-I6.33	Y3.51-E5.60	C3.25-E45.51^{ECL2}	Q23.51^{ECL1}-K168^{ECL2}	T23.52^{ECL1}-V166^{ECL2}
				C3.25-C45.50^{ECL2}	
0.8	R3.53-E6.30			Q23.51^{ECL1}-I167^{ECL2}	Y3.21-I167^{ECL2}
				I3.46-Y7.53	
0.7	R3.50-A6.34		C3.25-C45.50^{ECL2}	R3.50-Y7.53	Y3.21-V166^{ECL2}
0.6	V3.54-L5.64			P23.50^{ECL1}-I167^{ECL2}	F3.22-I167^{ECL2}
0.5	R3.50-V5.61	Y3.51-V5.61	C3.25-I167^{ECL2}	G2.67-K168^{ECL2}	Q23.51^{ECL1}-I167^{ECL2}
	P34.50^{ICL2}-Y6.27	I3.46-L6.37		F3.22-P165^{ECL2}	C3.25-I167^{ECL2}
				C3.25-V174^{ECL2}	L3.43-Y7.53
				N5.42-H6.52	F5.62-A6.34
0.4	I2.64-E45.51^{ECL2}	A3.47-V5.61		V1.53-P7.50	I2.64-C45.50^{ECL2}
	Y3.51-I5.57	V3.54-I5.65		I2.66-K168^{ECL2}	P23.50^{ECL1}-K168^{ECL2}
	P34.50^{ICL2}-Y6.26			T23.52^{ECL1}-K168^{ECL2}	C3.25-K168^{ECL2}
				Y3.21-P165^{ECL2}	L3.33-I175^{ECL2}
				Y5.40-F6.60	F5.43-I6.53
				R5.66-E6.30	

Table S3. Characteristic residue contacts (whose gradients from pixel attribution ≥0.4) of the A_1AR allosteric modulation.

Gradient	"Agonist -Gi"		"Agonist - Gi - PAM"	
0.9	V2.53-L3.35			
0.8			I2.64-C45.50^{ECL2}	
0.7	S2.45-S3.38	L3.33-G4.57	Q23.51^{ECL1}-I167^{ECL2}	Y3.21-V166^{ECL2}
			F3.22-I167^{ECL2}	
0.6	A2.52-L3.35	W146^{ECL2}-E5.36	L2.63-C45.50^{ECL2}	I2.64-F45.52^{ECL2}
	M5.35-L6.59		G2.67-K168^{ECL2}	P23.50^{ECL1}-I167^{ECL2}
			P23.50^{ECL1}-K168^{ECL2}	T23.52^{ECL1}-V166^{ECL2}
			T23.52^{ECL1}-C45.50^{ECL2}	Y3.21-I167^{ECL2}
0.5	L2.46-S3.39	L2.60-V3.28	G2.67-C45.50^{ECL2}	C3.25-I167^{ECL2}
	I3.34-L4.58		C3.25-C45.50^{ECL2}	
0.4	T2.39-D3.49	F2.42-A3.45	Q23.51^{ECL1}-K168^{ECL2}	Y3.21-P165^{ECL2}
	A2.49-S3.38	P23.50^{ECL1}-K168^{ECL2}	F3.22-K173^{ECL2}	
	L3.33-P4.60	N148^{ECL2}-V152^{ECL2}		
	I3.40-L3.44	L3.41-A3.45		
	L3.44-V3.48	E5.36-L6.59		
	V5.39-L6.59	V5.47-W6.48		
	P5.50-F6.44			

Table S4. Distances between selected characteristic residue pairs for 2D free energy profiling of the A_1AR activation in the inactive 5N2S and 5UEN and active 6D9H and 7LD3 PDB structures.

Residues	5N2S	5UEN	6D9H	7LD3
G2.67 - K168^{ECL2}	4.8 Å	6.9 Å	5.0 Å	2.5 Å
Y3.21 - I167^{ECL2}	5.1 Å	5.5 Å	5.7 Å	5.6 Å
R3.53 - E6.30	5.4 Å	4.3 Å	21.5 Å	21.4 Å
R3.50 - Y7.53	8.8 Å	11.6 Å	6.3 Å	5.6 Å
E5.60 - Y7.53	22.6 Å	22.6 Å	19.0 Å	18.4 Å

Table S5. Distances between selected characteristic residue pairs for 2D free energy profiling of the A_1AR allosteric modulation in the inactive 5N2S and 5UEN and active 6D9H and 7LD3 PDB structures.

Residues	5N2S	5UEN	6D9H	7LD3
G2.67 - K168^{ECL2}	4.8 Å	6.9 Å	5.0 Å	2.5 Å
Y3.21 - I167^{ECL2}	5.1 Å	5.5 Å	5.7 Å	5.6 Å
P23.50^{ECL1} - K168^{ECL2}	5.0 Å	6.3 Å	5.4 Å	4.7 Å
N148^{ECL2} - V152^{ECL2}	3.6 Å	2.8 Å	3.0 Å	3.1 Å
W146^{ECL2} - E5.36	9.3 Å	9.1 Å	9.5 Å	8.9 Å