Electronic Supplementary Material (ESI) for Physical Chemistry Chemical Physics.

Supporting Information

The effect of dimerization on the activation of adenosine A_1 receptor

Yang Li,[a] Mukuo Wang,[a] Na Gao,[a,b] Dongmei Li,[,a] and Jianping Lin*[,a,b]*

Contents:

RMSD of NECA inside the orthosteric binding site was measured with respect to its docking pose coordinates for each protomer over the three parallel cMD trajectories.

- **Fig. S5.** 2D Structure of NECA.
- **Table S2.** Averaged distances and standard deviations between the hydroxyl oxygen of the Thr91$^{3.36}$ side chain and the N$^{5'}$ of NECA, the centers of the phenyl group of Phe171^{ECL2} and the purine group of NECA, the carbonyl oxygen of Glu172^{ECL2} and the N^{6} of NECA, the imidazole ε-nitrogen of His251$^{6.52}$ and the O$^{5'}$ of NECA, hydrogen bond Hbond1 between the acylamino δ1-oxygen of Asn254$^{6.55}$ and the N^{6} of NECA (Hbond1), the acylamino δ2-nitrogen of Asn254$^{6.55}$ and the N^{7} of NECA (Hbond2), the distance between the hydroxyl oxygen of Thr277$^{7.42}$ and the O$^{3'}$ of NECA and the distance between the imidazole ε-nitrogen of His278$^{7.43}$ and O$^{2'}$ of NECA in the three parallel GaMD simulations.
- **Fig. S6.** Time dependences of RMSD of NECA respect to the relative to the docking pose for the NECA-A_1R-M system in the cMD simulation.
- **Table S3.** Average number of water molecules in the G-protein binding crevice of the inactive and active states of A_1R.
- **Fig. S6.** Time dependences of (A) TM3-TM6 distance, (B) RMSD of the DRY motif relative to the inactive starting structure, (C) the N--O distance between the guanidinium of Arg105$^{3.50}$/Arg108$^{3.53}$ and the carboxyl of Glu229$^{6.30}$ and (D) RMSD of the KxxK motif relative to the inactive starting structure in the GaMD simulation of the second NECA-A_1R trajectory.
- **Fig. S7.** Time dependences of (A) TM3-TM6 distance, (B) RMSD of the DRY motif relative to the inactive starting structure, (C) the N--O distance between the guanidinium of

Arg105$^{3.50}$/Arg108$^{3.53}$ and the carboxyl of Glu229$^{6.30}$ and (D) RMSD of the KxxK motif relative to the inactive starting structure in the GaMD simulation of the third NECA-A_1R trajectory.

- **Fig. S8.** Time dependences of (A) TM3-TM6 distance, (B) RMSD of the DRY motif relative to the inactive starting structure, (C) the N--O distance between the guanidinium of Arg105$^{3.50}$/Arg108$^{3.53}$ and the carboxyl of Glu229$^{6.30}$ and (D) RMSD of the KxxK motif relative to the inactive starting structure in the GaMD simulation of the second dNECA-A_1R trajectory.
- **Fig. S9.** Time dependences of (A) TM3-TM6 distance, (B) RMSD of the DRY motif relative to the inactive starting structure, (C) the N--O distance between the guanidinium of Arg105$^{3.50}$/Arg108$^{3.53}$ and the carboxyl of Glu229$^{6.30}$ and (D) RMSD of the KxxK motif relative to the inactive starting structure in the GaMD simulation of the third dNECA-A_1R trajectory.
- **Fig. S10.** Potential of mean force (PMF) calculated for the TM3-TM6 distance vs. the RMSD of the DRY motif relative to the inactive starting structure for (A) P_1 and (B) P_2 of the apo-A_1R system. Time dependences of the (C) TM3-TM6 distance, (D) RMSD of the DRY motif relative to the inactive starting structure, (E) N--O distance between the guanidinium of Arg105$^{3.50}$/Arg108$^{3.53}$ and the carboxyl of Glu229$^{6.30}$ and (F) RMSD of the KxxK motif relative to the inactive starting structure in the GaMD simulation of the apo-A_1R system.
- **Fig. S11.** Dynamic cross-correlations map of the apo-A_1R system in the GaMD simulations.
- **Fig. S12.** Dynamic cross-correlations map of the (A) NECA-A_1R based on the snapshots in 110-140 ns of the second NECA-A_1R trajectory, (B) NECA-A_1R based on the snapshots in

200-230 ns of the third NECA-A_1R trajectory, (C) dNECA-A_1R based on the snapshots in 130-160 ns of the second dNECA-A_1R trajectory, and (D) dNECA-A_1R based on the snapshots in 230-260 ns of the second NECA-A_1R trajectory in the GaMD simulations.

- **Table S4.** The highly correlated residues with the absolute values of correlation coefficients larger than 0.6 for the receptor in the NECA-A_1R and dNECA-A_1R systems.
- **Table S5.** The networks in the intracellular sides of the P_1 and P_2 in the GaMD simulation of the second NECA-A_1R and dNECA-A_1R trajectory.
- **Table S6.** The networks in the intracellular sides of the P_1 and P_2 in the GaMD simulation of the third NECA-A_1R and dNECA-A_1R trajectory.
- **Table S7.** Averaged values and standard deviations of the distances between COMs of P_1 and P_2 in the apo-A_1R, NECA-A_1R, and dNECA-A_1R systems over the three parallel GaMD trajectories.

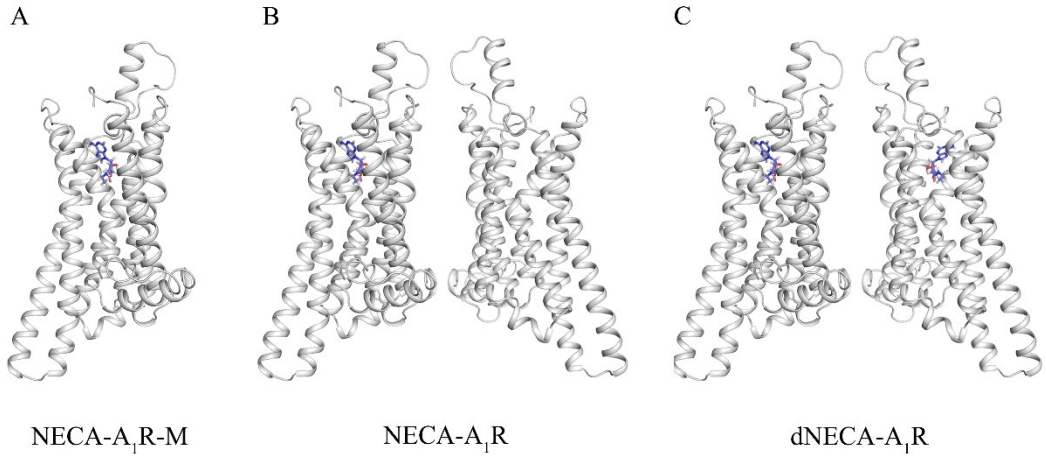

Fig. S1. Schematic representation of the NECA-A_1R-M, NECA-A_1R and dNECA-A_1R systems used in the MD simulations.

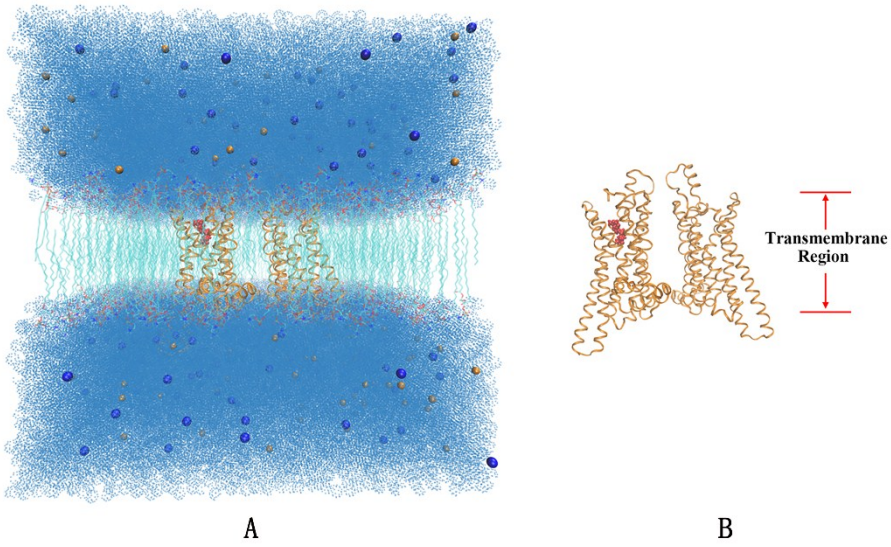

Fig. S2. Structure of A_1R and NECA in the NECA-A_1R system. A_1R is shown in a cartoon model, and NECA in the binding site is shown in a stick model.

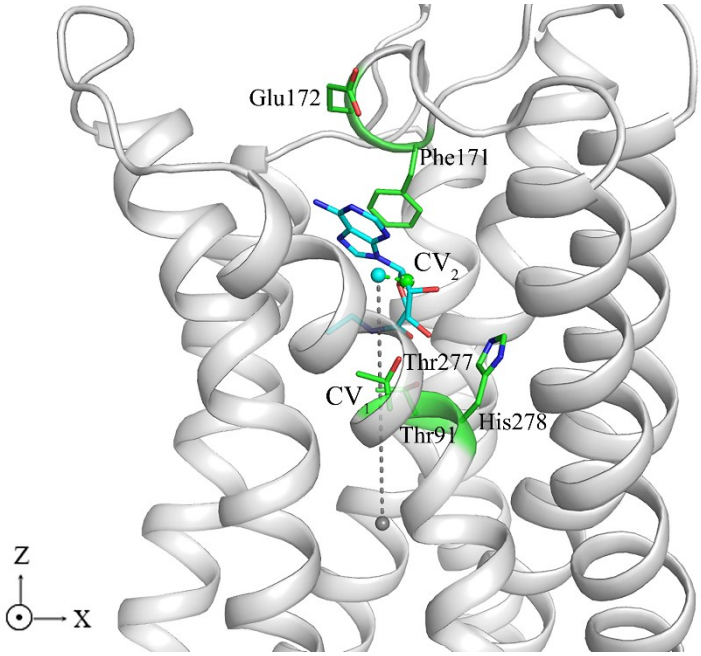

Fig. S3. Definitions of the collective variables CV_1 and CV_2 for metaMD. CV_1 (shown with gray dash line) is the distance between the centers of mass (COMs) of NECA and the seven transmembrane helical bundles of A_1R along the direction perpendicular to membrane (Z-direction), and CV_2 (shown with green dash line) is the distance between the COMs of NECA and residues Thr91$^{3.36}$, Phe171^{ECL2}, Glu172^{ECL2}, Thr277$^{7.42}$ and His278$^{7.43}$.

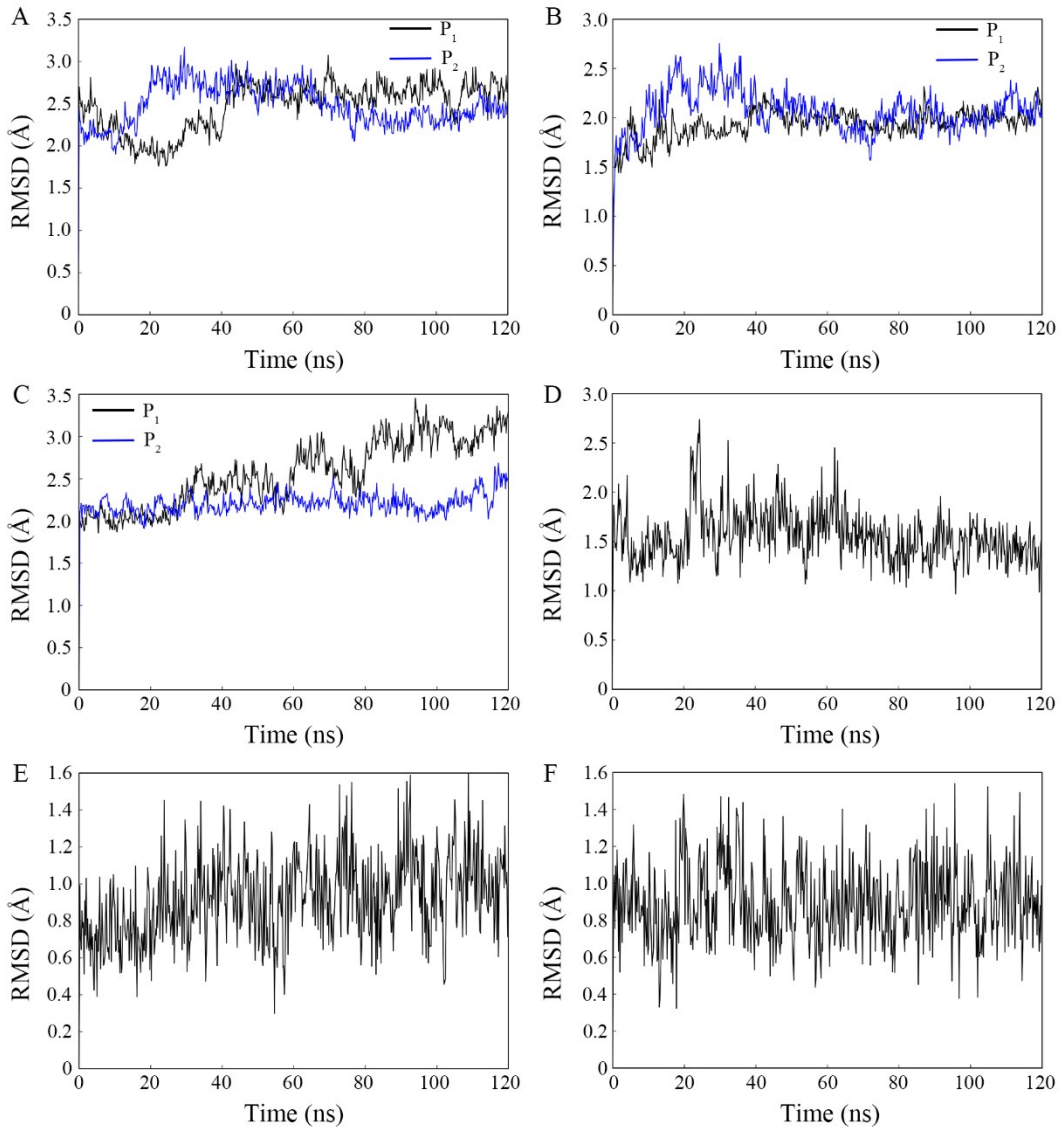

Fig. S4. Time dependences of RMSD of the protein of each protomers with respect to the initial X-ray crystallographic structure for the (A) apo-A_1R, (B) NECA-A_1R and (C) dNECA-A_1R system. Time dependences of RMSD of NECA inside the orthosteric binding site was measured with respect to its docking pose coordinates for the (E) apo-A_1R, (F) NECA-A_1R and (G) dNECA-A_1R system in the cMD simulation.

Table S1. Averaged values and standard deviations of the RMSD of the protein of each protomers with respect to the initial X-ray crystallographic structure, and the RMSD of NECA inside the orthosteric binding site was measured with respect to its docking pose coordinates for each protomer over the three parallel cMD trajectories.

System	RMSD of Protein (Å)		RMSD of NECA (Å)	
	Average	SD	Average	SD
P_1 of apo-A_1R (cMD)	2.4	0.3	-	
P_2 of apo-A_1R (cMD)	2.4	0.1	-	
P_1 of NECA-A_1R (cMD)	1.9	0.1	1.4	0.1
P_2 of NECA-A_1R (cMD)	2.3	0.2	-	
P_1 of dNECA-A_1R (cMD)	2.7	0.6	1.0	0.3
P_2 of dNECA-A_1R (cMD)	2.3	0.01	1.2	0.7

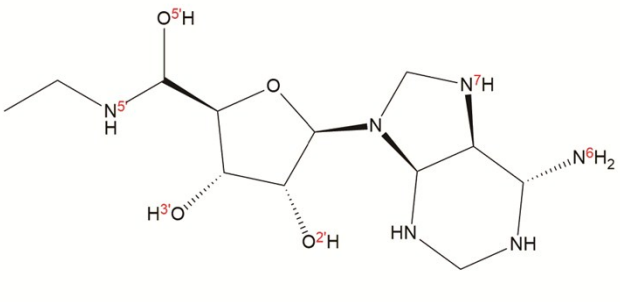

Fig. S5. 2D Structure of NECA.

Table S2. Averaged distances and standard deviations between the hydroxyl oxygen of the Thr91$^{3.36}$ side chain and the $N^{5'}$ of NECA, the centers of the phenyl group of Phe171^{ECL2} and the purine group of NECA, the carbonyl oxygen of Glu172^{ECL2} and the N^6 of NECA, the imidazole ε-nitrogen of His251$^{6.52}$ and the $O^{5'}$ of NECA, hydrogen bond Hbond1 between the acylamino δ1-oxygen of Asn254$^{6.55}$ and the N^6 of NECA (Hbond1), the acylamino δ2-nitrogen of Asn254$^{6.55}$ and the N^7 of NECA (Hbond2), the distance between the hydroxyl oxygen of Thr277$^{7.42}$ and the $O^{3'}$ of NECA and the distance between the imidazole ε-nitrogen of His278$^{7.43}$ and $O^{2'}$ of NECA in the three parallel GaMD simulations.

Residue	P_1 of NECA-A_1R		P_1 of dNECA-A_1R		P_2 of dNECA-A_1R	
	Average	SD	Average	SD	Average	SD
Thr91$^{3.36}$	3.5	0.7	3.0	0.02	3.0	0.01
Phe171^{ECL2}	4.0	0.1	3.8	0.1	4.0	0.03
Glu172^{ECL2}	4.5	1.1	3.2	0.01	3.3	0.1
His251$^{6.52}$	2.9	0.03	3.0	0.04	2.9	0.01
Asn254$^{6.55}$ (Hbond1)	2.9	0.01	2.9	0.01	2.9	0.004
Asn254$^{6.55}$ (Hbond2)	3.1	0.002	3.1	0.01	3.1	0.01
Thr277$^{7.42}$	3.1	0.1	3.7	0.6	2.8	0.01
His278$^{7.43}$	3.1	0.4	4.3	0.8	3.2	0.2

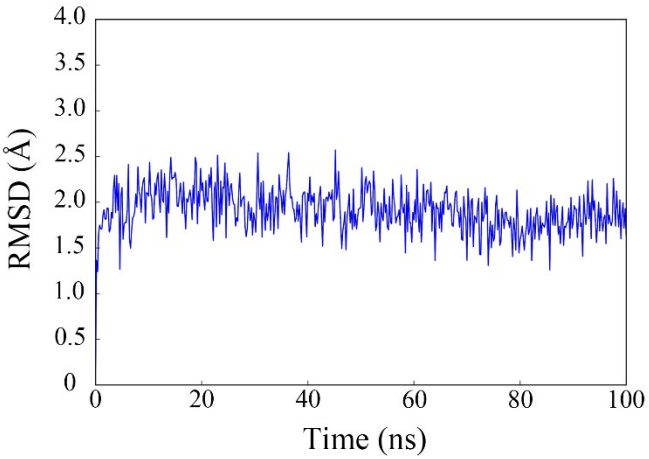

Fig. S6. Time dependences of RMSD of NECA respect to the relative to the docking pose for the NECA-A_1R-M system in the cMD simulation.

Table S3. Average number of water molecules in the G-protein binding crevice of the inactive and active states of A_1R.

System	P_1		P_2	
	Average	**SD**	**Average**	**SD**
NECA-A_1R	15	1.5	[a]20	0.6
dNECA-A_1R	[b]22	1.7	18	0.6

[a]The snapshots of the three parallel GaMD simulations in 260-290 ns, 110-140 ns, and 200-230 ns which correspond to the active state for the NECA-A_1R system are extracted from the trajectories to calculate the average number of water molecules here.

[b]The snapshots of the three parallel GaMD simulations in 235-265 ns, 130-160 ns and 230-260 ns which correspond to the active state for the dNECA-A_1R system are extracted from the trajectories to calculate the average number of water molecules here.

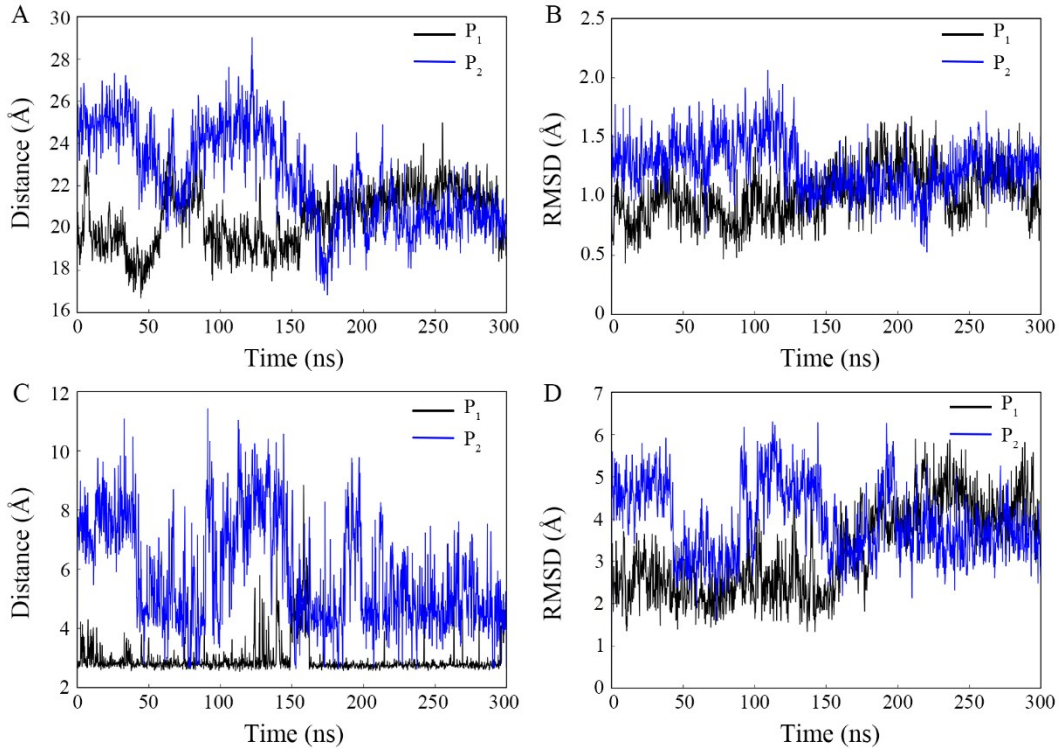

Fig. S7. Time dependences of (A) TM3-TM6 distance, (B) RMSD of the DRY motif relative to the inactive starting structure, (C) the N--O distance between the guanidinium of Arg105$^{3.50}$/Arg108$^{3.53}$ and the carboxyl of Glu229$^{6.30}$ and (D) RMSD of the KxxK motif relative to the inactive starting structure in the GaMD simulation of the second NECA-A_1R trajectory.

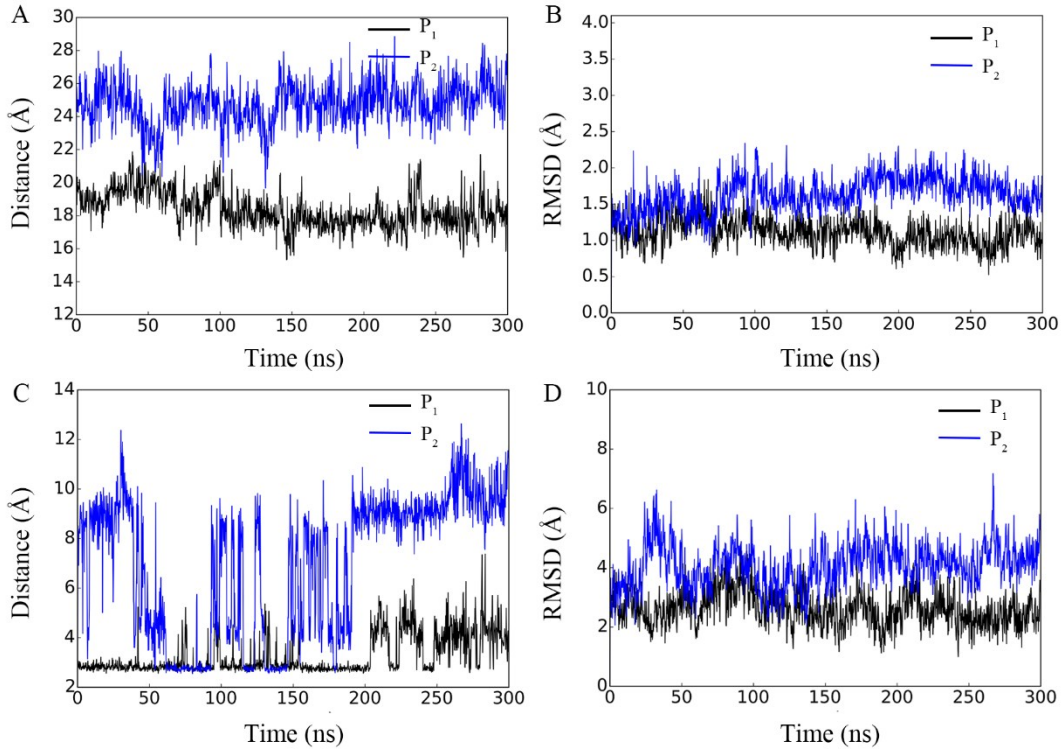

Fig. S8. Time dependences of (A) TM3-TM6 distance, (B) RMSD of the DRY motif relative to the inactive starting structure, (C) the N--O distance between the guanidinium of Arg105$^{3.50}$/Arg108$^{3.53}$ and the carboxyl of Glu229$^{6.30}$ and (D) RMSD of the KxxK motif relative to the inactive starting structure in the GaMD simulation of the third NECA-A_1R trajectory.

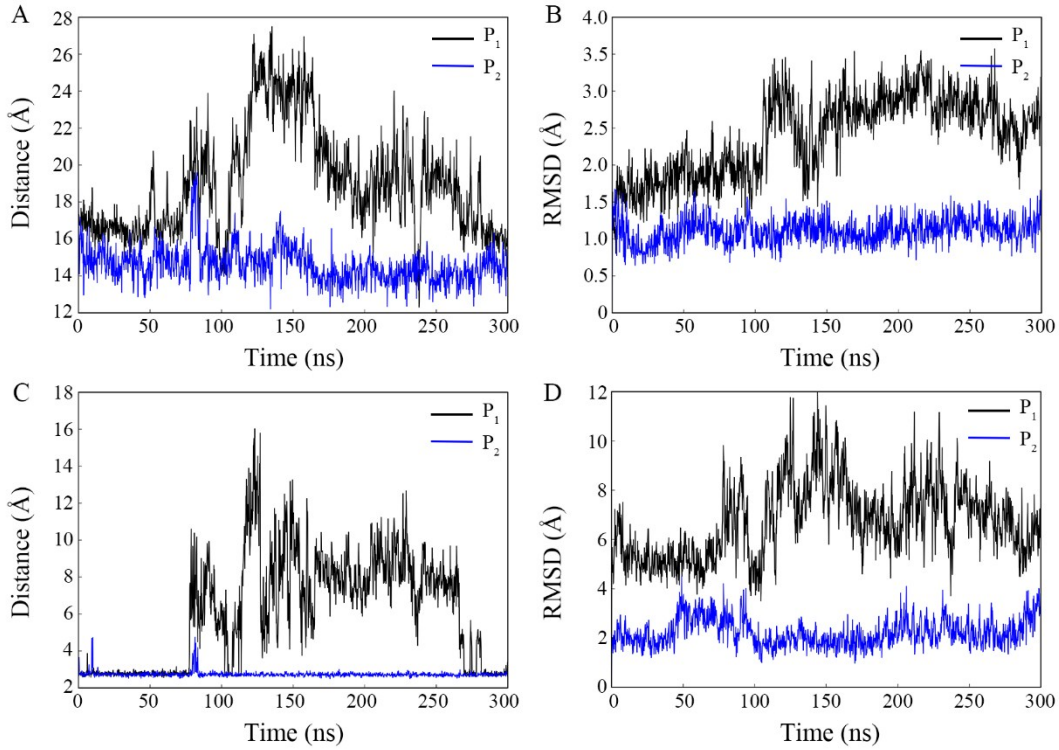

Fig. S9. Time dependences of (A) TM3-TM6 distance, (B) RMSD of the DRY motif relative to the inactive starting structure, (C) the N--O distance between the guanidinium of Arg105$^{3.50}$/Arg108$^{3.53}$ and the carboxyl of Glu229$^{6.30}$ and (D) RMSD of the KxxK motif relative to the inactive starting structure in the GaMD simulation of the second dNECA-A_1R trajectory.

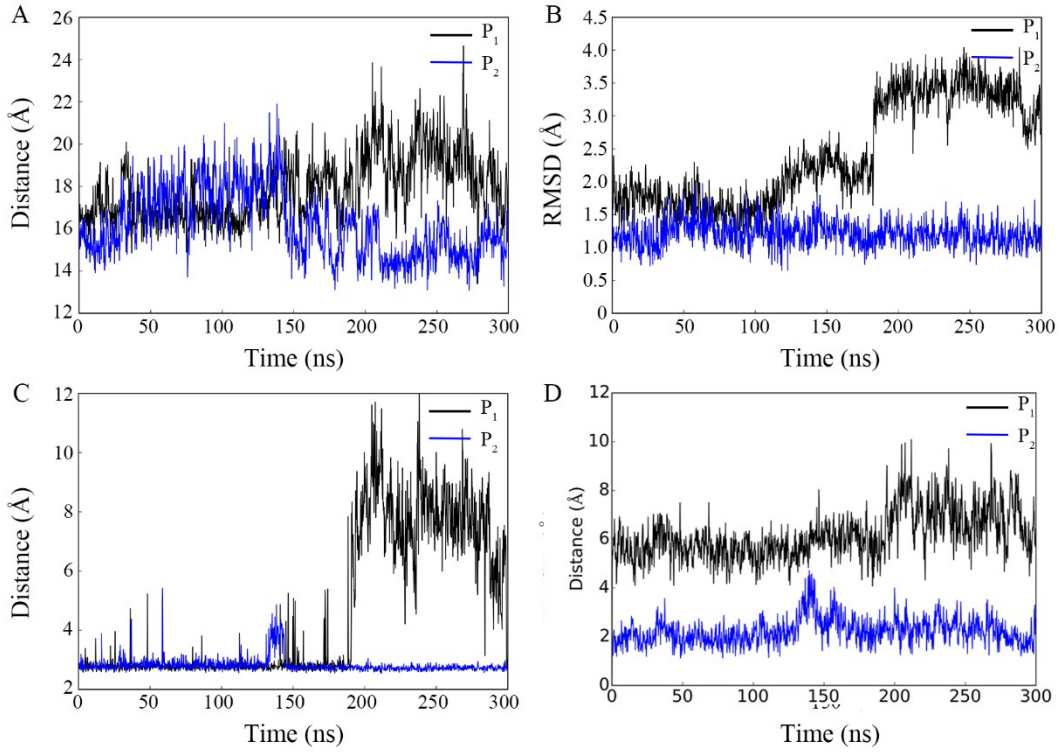

Fig. S10. Time dependences of (A) TM3-TM6 distance, (B) RMSD of the DRY motif relative to the inactive starting structure, (C) the N--O distance between the guanidinium of Arg105$^{3.50}$/Arg108$^{3.53}$ and the carboxyl of Glu229$^{6.30}$ and (D) RMSD of the KxxK motif relative to the inactive starting structure in the GaMD simulation of the third dNECA-A_1R trajectory.

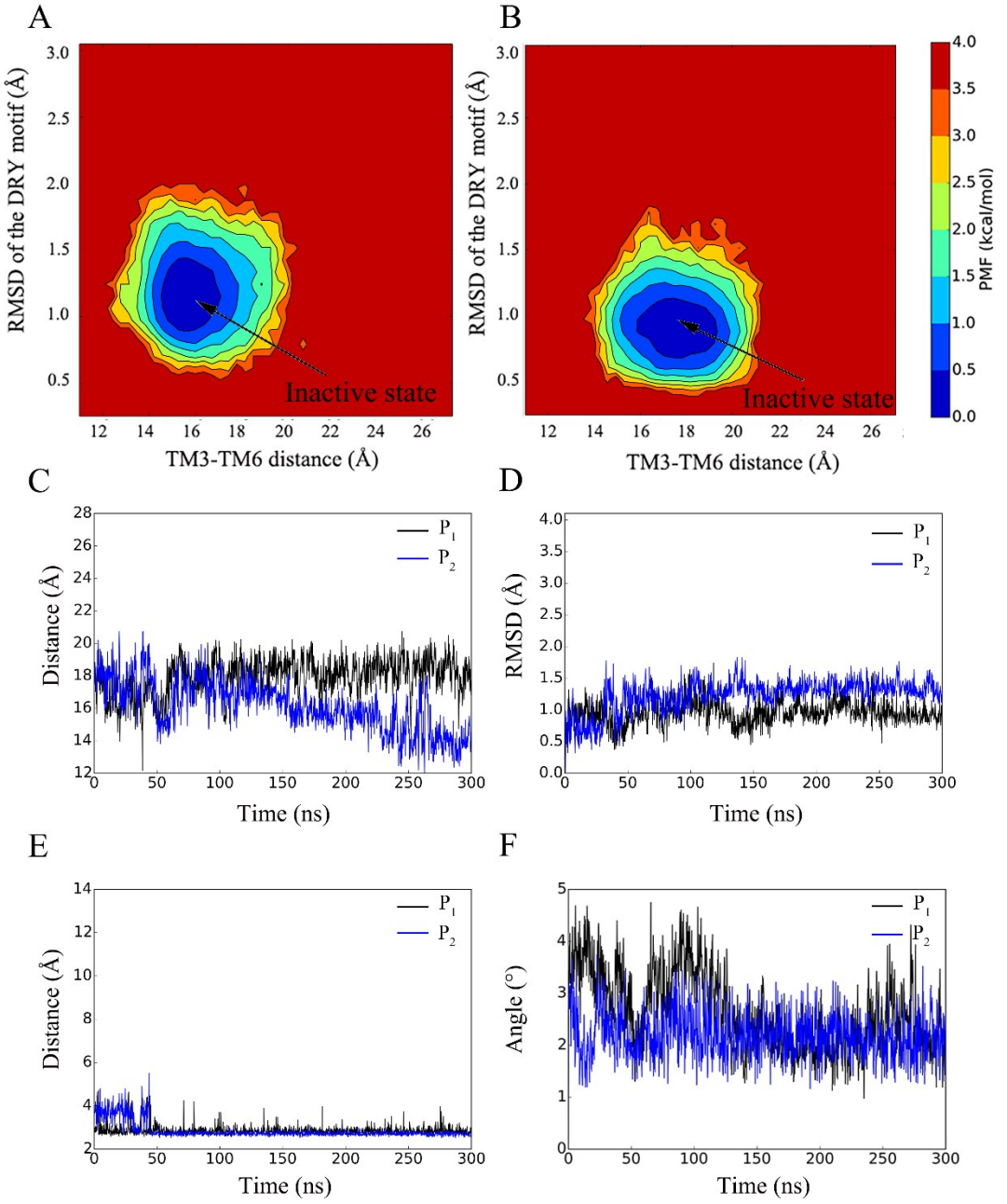

Fig. S11. Potential of mean force (PMF) calculated for the TM3-TM6 distance vs. the RMSD of the DRY motif relative to the inactive starting structure for (A) P_1 and (B) P_2 of the apo-A_1R system. Time dependences of the (C) TM3-TM6 distance, (D) RMSD of the DRY motif relative to the inactive starting structure, (E) N--O distance between the guanidinium of Arg105$^{3.50}$/Arg108$^{3.53}$ and the carboxyl of Glu229$^{6.30}$ and (F) RMSD of the KxxK motif relative to the inactive starting structure in the GaMD simulation of the apo-A_1R system.

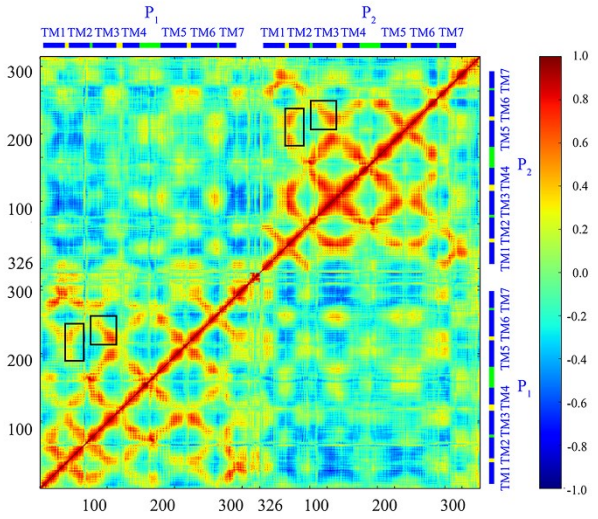

Fig. S12. Dynamic cross-correlations map of the apo-A_1R system in the GaMD simulations.

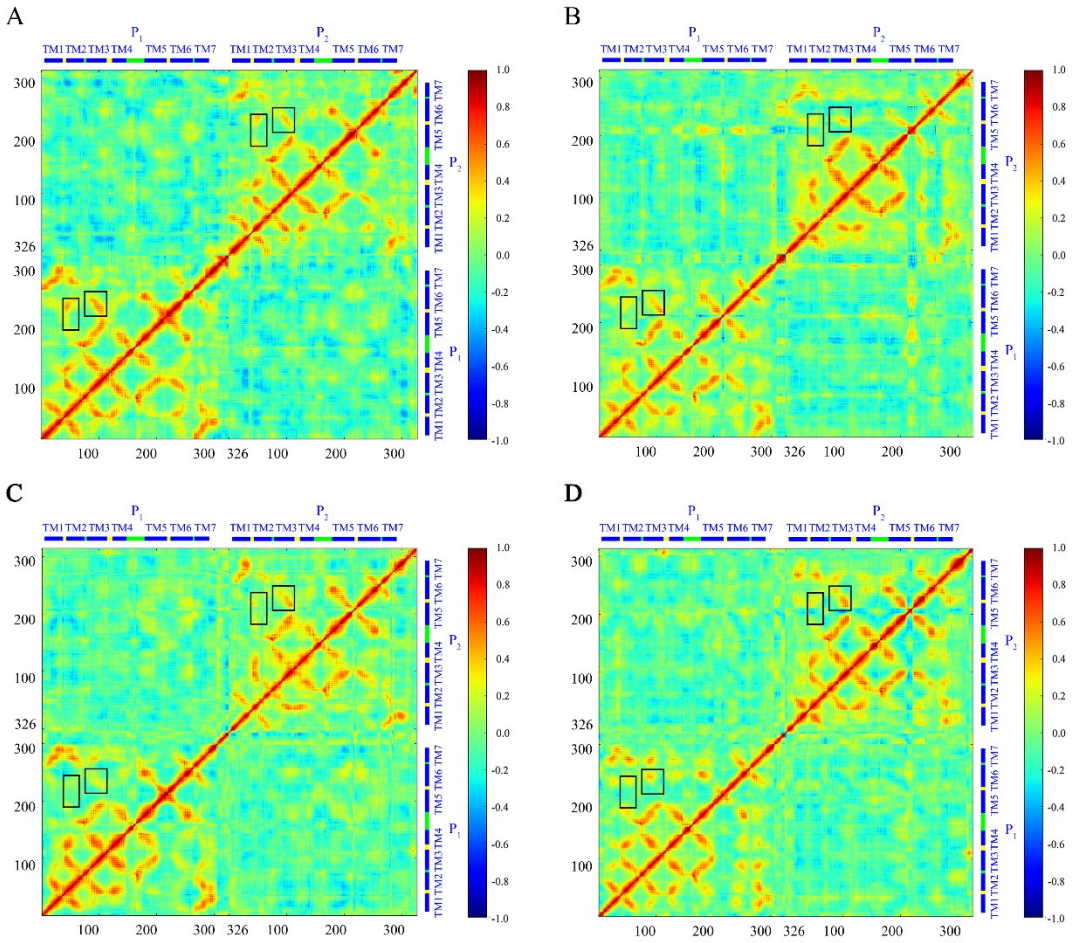

Fig. S13. Dynamic cross-correlations map of the (A) NECA-A_1R based on the snapshots in 110-140 ns of the second NECA-A_1R trajectory, (B) NECA-A_1R based on the snapshots in 200-230 ns of the third NECA-A_1R trajectory, (C) dNECA-A_1R based on the snapshots in 130-160 ns of the second dNECA-A_1R trajectory, and (D) dNECA-A_1R based on the snapshots in 230-260 ns of the second NECA-A_1R trajectory in the GaMD simulations.

Table S4. The highly correlated residues with the absolute values of correlation coefficients larger than 0.6 for the receptor in the NECA-A_1R and dNECA-A_1R systems.

System	P_1	P_2
NECA-A_1R	*15-28 in TM1 and 282-287 in TM7* *20-35 in TM1 and 49-64 in TM2* *42-52 in TM2 and 94-109 in TM3* *42-52 in TM2 and 229-238 in TM6* *44-54 in TM2 and 125-133 in TM4* *47-51 in TM2 and 200-205 in TM5* *52-57 in TM2 and 281-286 in TM7* *76-82 in TM3 and 167-176 in ECL2* *88-101 in TM3 and 132-143 in TM4* *93-109 in TM3 and 124-130 in TM4* *96-110 in TM3 and 193-208 in TM5* *98-110 in TM3 and 226-242 in TM6* *135-155 in TM4 and 174-189 in TM5* *177-188 in TM5 and 251-261 in TM6* *204-209 in TM5 and 225-238 in TM6* *247-254 in TM6 and 274-277 in TM7*	*15-19 in TM1 and 51-62 in TM2* *16-29 in TM1 and 278-284 in TM7* *66-72 in TM2 and 168-170 in ECL2* *76-82 in TM3 and 166-174 in ECL2* *52-57 in TM2 and 281-285 in TM7* *86-94 in TM3 and 132-143 in TM4* *143-150 in TM4 and 173-183 in TM5* *178-183 in TM5 and 253-258 in TM6* *205-212 in TM5 and 225-230 in TM5*
dNECA-A_1R	*15-35 in TM1 and 49-64 in TM2* *19-35 in TM1 and 279-289 in TM7* *46-55 in TM2 and 92-108 in TM3* *50-62 in TM2 and 87-90 in TM3* *45-54 in TM2 and 123-135 in TM4* *65-88 in TM2 and 164-176 in ECL2* *50-63 in TM2 and 278-285 in TM7* *81-104 in TM3 and 124-145 in TM4* *132-152 in TM4 and 171-189 in TM5* *176-190 in TM5 and 247-258 in TM6* *205-215 in TM5 and 219-230 in TM6*	*17-35 in TM1 and 44-64 in TM2* *18-35 in TM1 and 275-291 in TM7* *44-62 in TM2 and 90-108 in TM3* *44-52 in TM2 and 200-204 in TM5* *44-55 in TM2 and 230-240 in TM6* *46-50 in TM2 and 124-128 in TM4* *48-63 in TM2 and 278-290 in TM7* *66-72 in TM2 and 167-173 in ECL2* *77-88 in TM3 and 165-180 in ECL2* *77-100 in TM3 and 127-145 in TM4* *92-111 in TM3 and 192-207 in TM5* *95-110 in TM3 and 226-244 in TM6* *134-153 in TM4 and 171-189 in TM5* *146-149 in ECL2 and 254-259 in TM6* *172-191 in TM5 and 248-260 in TM6* *247-254 in TM6 and 269-277 in TM7*

Table S5. The networks in the intracellular sides of the P_1 and P_2 in the GaMD simulation of the second NECA-A_1R and dNECA-A_1R trajectory.

NECA-A_1R			**dNECA-A_1R**		
P_1		**P_2**	**P_1**	**P_2**	
Leu99 – Ile199	Tyr106 – Ile199	Ala102 – Ile199	Tyr106 –Ile199	Leu99 - Met196	Arg108 – Glu229
Ala102 – Ile199	Tyr106 – Glu202	Arg105 – Val203		Ala102 – Met196	
Ala102 – Tyr200	Tyr106 – Val203	Tyr106 –Ile199		Ala102 – Ile199	
Val103 – Ile199	Arg108 – Glu229	Tyr106 – Glu202		Ala102 – Tyr200	
Arg105 – Val203	Val109 – Val203	Tyr106 – Val203		Val103 – Ile199	
Arg105 – Ile232	Val109 – Leu206	Val109 – Val203		Arg105 – Val203	
Arg105 – Ala233	Lys110 – Leu206	Val109 – Leu206		Tyr106 –Ile199	
Arg105 – Leu236				Tyr106 – Val203	

Table S6. The networks in the intracellular sides of the P_1 and P_2 in the GaMD simulation of the third NECA-A_1R and dNECA-A_1R trajectory.

NECA-A_1R			dNECA-A_1R		
P_1	P_2		P_1	P_2	
Ala102 – Ile199	Arg108 – Glu229	Ala102 – Ile199	Ala102 – Ile199	Leu99 – Met196	Tyr106 – Ile199
Ala102 – Tyr200	Val109 – Leu206	Val103 – Ile199	Val103 – Ile199	Ala102 – Met196	Tyr106 – Glu202
Val103 – Ile199	Val109 – Ile207	Ala102 – Met196	Tyr106 –Ile199	Ala102 – Ile199	Tyr106 – Val203
Arg105 – Val203	Lys110 – Leu206			Ala102 – Tyr200	Arg108 – Glu225
Tyr106 –Ile199				Val103 – Ile199	Val109 – Ile207
Tyr106 – Glu202				Arg105 – Tyr200	Arg108 – Glu229
Tyr106 – Val203				Arg105 – Val203	Val109 – Glu229
Tyr106– Leu206				Arg105 – Glu229	Lys110 – Leu206

Table S7. Averaged values and standard deviations of the distances between COMs of P_1 and P_2 in the apo-A_1R, NECA-A_1R, and dNECA-A_1R systems over the three parallel GaMD trajectories.

	apo-A_1R (Å)		**NECA-A_1R (Å)**		**dNECA-A_1R (Å)**	
	Average	**SD**	**Average**	**SD**	**Average**	**SD**
P_1-P_2 distance	33.3	0.2	35.2	0.7	35.5	0.3