# Supporting Information

## Investigation of ECD conformational transition mechanism of GLP-1R by molecular dynamics simulations and Markov State Model

Jintu Zhang[1,2,#], Qifeng Bai[1,#,*], Horacio Pérez-Sánchez[3], Shuxia Shang[2], Xiaoli An[1], Xiaojun Yao[2,*]

[1]Key Lab of Preclinical Study for New Drugs of Gansu Province, Institute of Biochemistry and Molecular Biology, School of Basic Medical Sciences, Lanzhou University, Lanzhou, Gansu 730000, P. R. China

[2]College of Chemistry and Chemical Engineering, Lanzhou University, Lanzhou, Gansu 730000, P. R. China

[3]Computer Science Department, Universidad Católica San Antonio de Murcia (UCAM), Murcia, Spain

[#]These authors contributed equally to this work.

* Corresponding author

Tel.: +86-931-8912011
E-mail: baiqf@lzu.edu.cn or xjyao@lzu.edu.cn

**Method S1**

In AIM theory (Bader, 1990), $\rho(r)$ at Bond Critical Point (BCP) represents the interaction strength. Term $sign(\lambda_2)$ is the sign of the secondary large eigenvalue of Hessian matrix. When $sign(\lambda_2) > 0$, it means the interested region exists repulsive interactions, otherwise, attractive interactions should be observed.

Bader, R.F. (1990). Atoms in molecules (Wiley Online Library).

**Table S1.** MD simulated results

| ID | Time | System | Target state* |
|---|---|---|---|
| 1 | 200 ns | Bound-state GLP-1R | -- |
| 2 | 200 ns | Bound-state GLP-1R | -- |
| 3 | 200 ns | Bound-state GLP-1R | -- |
| 4 | 200 ns | Apo-state GLP-1R | appeared |
| 5 | 200 ns | Apo-state GLP-1R | appeared |
| 6 | 200 ns | Apo-state GLP-1R | appeared |
| 7 | 200 ns | Apo-state GLP-1R | Didn't appear |
| 8 | 200 ns | Apo-state GLP-1R | Didn't appear |

* Target state means conformation transition of GLP-1R ECD.

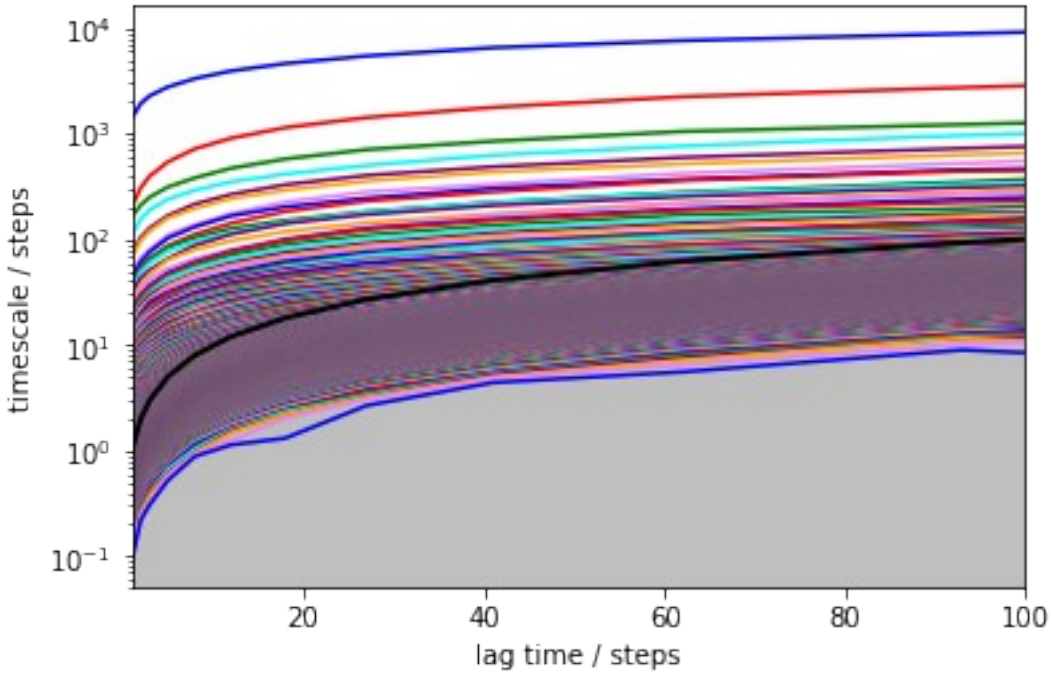


**Figure S1.** Implied timescales *vs.* lag time. The implied timescales of a dynamic system will tend to constants with the increasing of lag time, if the system satisfies the Markov State Model. When estimating the Markov State Model, we use the smallest lag time to obtain model with highest time resolution. Thus, we choose 25 steps (0.25 ns) as the lag time.

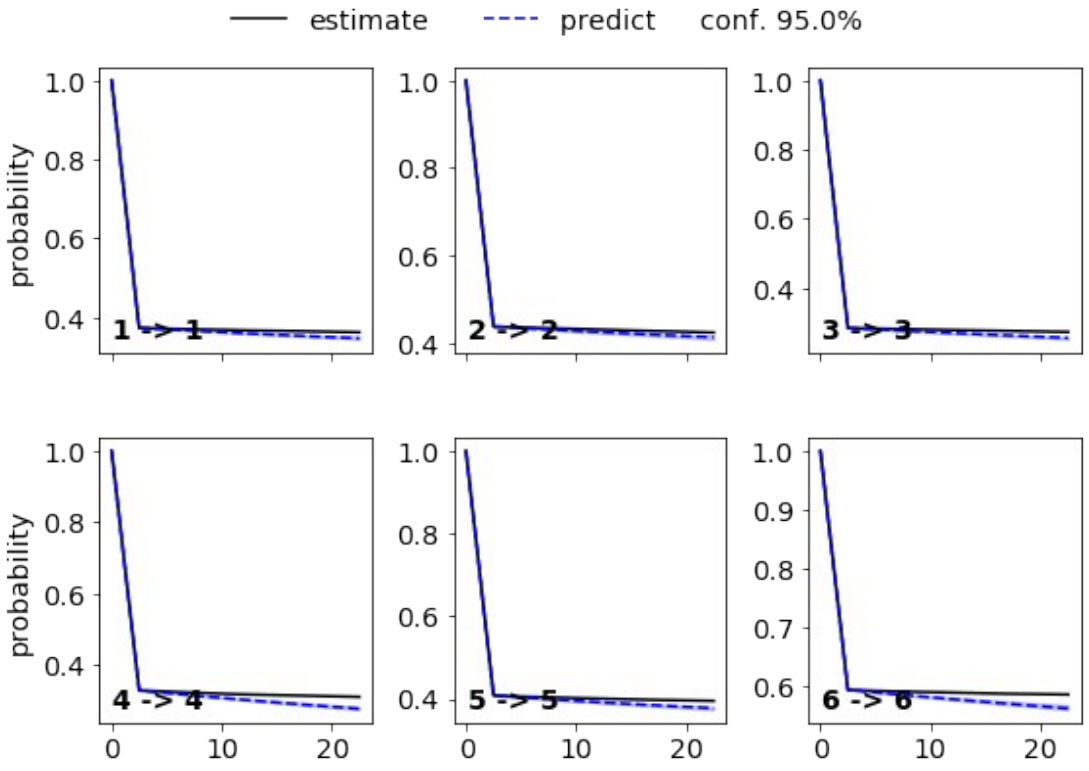


**Figure S2.** Results of the Chapman-Kolmogorov test (cktest). The standard cktest computes the transition probability between metastable states for different lag times. If the estimate probability is similar with the predict probability, this MSM is credible.

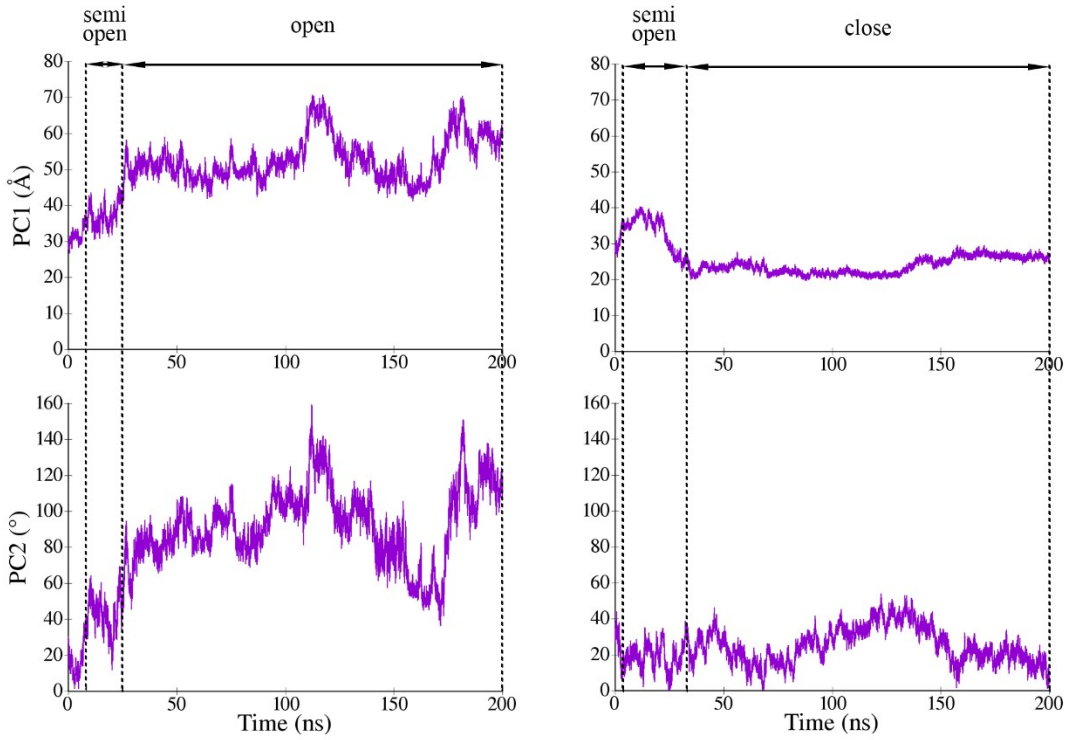


**Figure S3.** The fluctuation of two principal components (PCs) that we chose. We selected them to build the MSM since they have the ability to reflect the conformational conversion process of apo-state GLP-1R. Here we draw the conversion process of these two PCs during two 200 ns trajectories. The trajectories are divided into different stages by the different values of PCs, so we can identify each state and perform following analysis based on these two PCs.

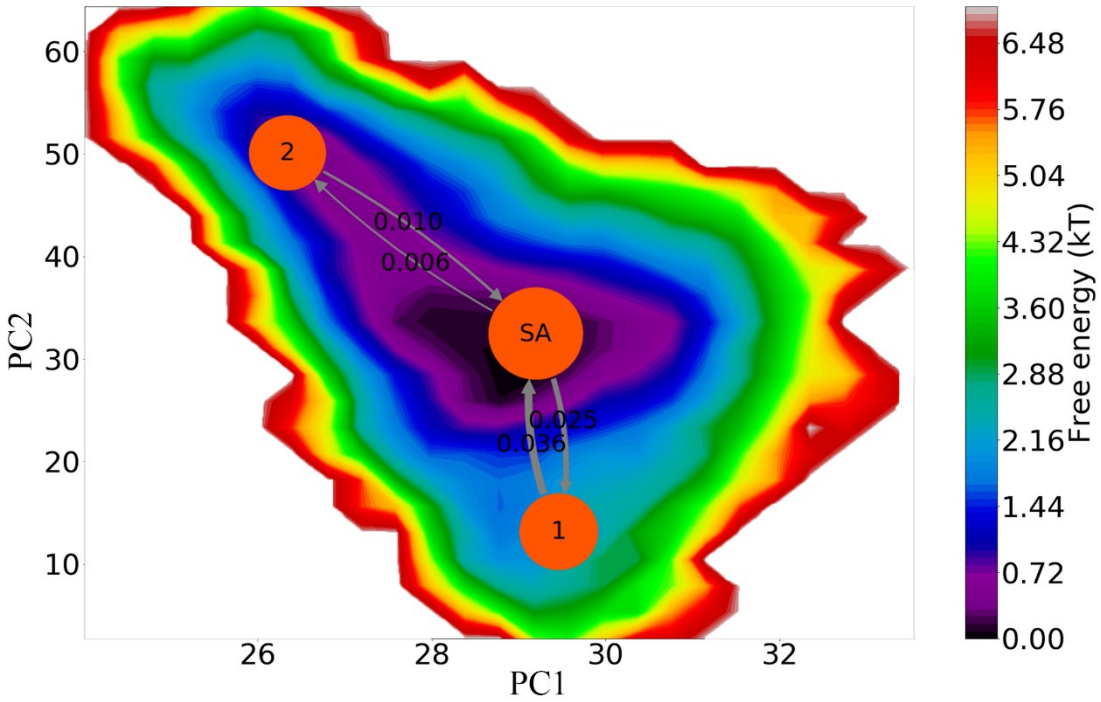


**Figure S4.** The MSM of bound state GLP-1R system. We chose 0.5 ns as the lag time and used 1000 microstates to build this model, and performed PCCA++ to generate the coarse-grained MSM. Finally, MSM with three states was given out. The labels upon the arrows denote the entries in the transition matrix.

**Move S1.** This is one of three 200 ns bound-state GLP-1R MD simulation trajectories. In order to exhibit the movement of protein, we did not draw lipid or water environment here. And we drew protein in periodic cell alone x and z axis each one.

**Move S2.** This is one of three 200 ns apo-state GLP-1R MD simulation trajectories with our target phenomena. In order to exhibit the movement of protein, we did not draw lipid or water environment here. And we drew protein in periodic cell alone x and z axis each one.