Chemical Science

Electronic supplementary information

Investigating allosteric effects on the functional dynamics of β2-adrenergic ternary complexes with enhanced-sampling simulations

Noureldin Saleh, Giorgio Saladino, Francesco Luigi Gervasio, Timothy Clark*

Supplementary Information

Contents

General set up of MD simulations

Topologies for the receptors were generated using the AMBER ff99SBildn force field[1] and inserted to a pre-prepared hydrated pre-equilibrated dioleoylphosphatidylcholine (DOPC) bilayer[2] according to the orientation in the OPM database[3] using g_membed.[4] Ligands were assigned AM1-BCC partial charges and the generalized AMBER force field (GAFF)[5, 6] is used for topology generation. The appropriate number of sodium and chloride ions was added to the systems to simulate a physiological salt concentration of 100mM. Particle mesh Ewald (PME) was used to treat electrostatic interactions, using the cut-off distance of 1.0 nm. The resulting system was geometry-optimized and then equilibrated for 10 ns followed by a production run. All simulations used the SPC/E water model.[7] All simulations were performed using GROMACS[8] with the PLUMED plug-in[9] for the metadynamics simulations. The ternary simulations included box size of 9.5×9.5×16 nm^3 with 37000 water molecules, and 239 DOPC molecules, whereas the binary systems compromised a box of 9.5×9.5×11.5 nm^3 with 21000 water molecules, and 239 DOPC molecules.

ADRB2 Models

An inactive ADRB2 was modeled based on the high-resolution crystal structure (PDB access code 2RH1[10]). The model was then equilibrated and simulated for 500 ns of production MD simulation. The crystal structures of the ADRB2 bound to ICI 118,551 (PDB access code 3NY8[11]), ADRB2 bound to Alprenolol (PDB access code 3NYA[11]), ADRB1 bound to Isoprenaline (PDB access code 2Y03[12]) and ADRB1 bound to Carvedilol (PDB access code 4AMJ[13]), were aligned to the inactive model of ADRB2 after 500ns MD simulations based on the C_α atoms of the residues within 0.5 nm of the binding pocket of each ligand. The ligand-coordinates (structures and protonation states used shown in Scheme S1) were transferred to the inactive models, which were then equilibrated and simulated for 500 ns of MD simulation.

A ternary complex model for the ADRB2 with BI-167107 and $G_{\alpha s}$ after 500ns MD simulation, based on the ternary complex structure of ADRB2 (PDB accession code 3SN6[14]) was used.[15] The inactive binary models after 500 ns of MD simulation were aligned to the ternary complex based on the C_α atoms of the residues within 0.5 nm of the binding pocket of each ligand. The ligand-coordinates were transferred to our ternary models to substitute BI-167107 and were then simulated for 500 ns MD simulation.

An Apo-complex model for the ADRB2 and $G_{\alpha s}$ was based on the simulated model mentioned above for the ternary ADRB2-BI167107-$G_{\alpha s}$,[15] the Apo–ADRB2-$G_{\alpha s}$ complex was then equilibrated and refined using 500 ns of productive MD run.

The active structure of β-arrestin was modelled based on the crystal structure of β-arrestin bound to the C-terminal peptide of the vasopressin-2 receptor.[16] The structural changes upon activation of the finger-loop were modelled, using the Modeller software,[17] based on the changes in the crystal structure of S-arrestin bound rhodopsin (PDB accession code 4ZWJ[18]). The active state of ADRB2 from the ternary complex crystal structure (PDB accession code 3SN6[14]), was aligned to the crystal structure of rhodopsin-bound arrestin and the coordinates for β-arrestin2 were transferred to ADRB2 to match the orientation of the co-crystalized arrestin.

To mimic the experimental procedures reported by Wisler *et al.,*[19] we modelled a chimeric cytoplasmic ADRB2-vasopressin2 receptor tail for our ADRB2 models, which was reported to increase the arrestin affinity. That chimeric tail included residues co-crystalized in β-arrestin-phosphopeptide. These residues were used as an anchor for the interaction of arrestin with the modelled cytoplasmic tail. The cytoplasmic tail was fully phosphorylated and the topology for the phosphorylated residues were generated based on the parameters described by Sticht *et al.*[20] The model was aligned to the ternary ADRB2-G_s models based on the C_α atoms of the residues within 0.5 nm from each ligand. Five models were generated, one for each ligand, in addition to an *apo*-ADRB2-arrestin model (based on the ADRB2-Isoprenaline-Arrestin model after removing the agonist). Each model was then equilibrated and simulated for 500 ns of MD simulation for refinement.

Metadynamics simulation of ligand binding to ADRB2

Three simulations were started for each ligand, one using the ternary-arrestin, ternary-$G_{\alpha s}$ complex and binary complex of the corresponding ligand with ADBR2. Metadynamics simulations were performed in order to obtain the binding

free-energy profiles. We used a combination of the well-tempered variant (WT)[21, 22] of metadynamics and funnel metadynamics (FM).[23] A metadynamics history-dependent bias was applied along the projection. The *z*-component of the distance between the relatively immobile C_α of Trp6.48 deep in the binding region and the center quaternary amine-nitrogen of each ligand was used as a simple collective variable. The funnel restraint was then applied to the relative position on the *xy*-plane to ensure better sampling for the relevant region of the free energy as the ligand moves far into the extracellular solvent. Gaussian hills with initial height of 1.2 kcal mol^{-1} applied every 1 ps were used. The hill width was chosen to be 0.1 nm. The Gaussian functions were rescaled in the WT scheme using a bias factor of 20. We performed an initial metadynamics simulation with a higher bias factor. Nine starting geometries, spanning the bins of the unbinding process, were extracted. These nine starting geometries for each model were then simulated with the multiple walker technique[24] as staring geometries. This ensured faster convergence of the free-energy surface and enhanced the parallelization up to 5,040 CPUs. Each of the four simulations converged within 2 μs (collective over the replicas) in a single run on the Haswell nodes of SuperMUC.

Metadynamics simulation of coupling of $G_{\alpha s}$ and activation of ADBR2

BI-167107 was removed from its ternary complex model and the resulting complex equilibrated for 20 ns. WT-metadynamics was used to map the FES of the activation and coupling of both the *apo*- (ADRB2-$G_{\alpha s}$) and the ternary (ligand-ADRB2-$G_{\alpha s}$) complexes for each ligand. We used Arg3.50 and Glu6.30 to define an initial reaction coordinate for the activation. However, this did not result in full activation of ADRB2. Using the distance between C_α of Arg3.50 and Leu6.34 resulted in a successful full transition (see Figure S2). This distance was used as the first reaction coordinate (β2-activation/ TM3-TM6 distance). The *z*-component of the distance between C_α of Glu392 in the α5-helix of $G_{\alpha s}$ and C_α of Arg3.50 was used as a reaction coordinate to describe the coupling of $G_{\alpha s}$ to ADRB2 (Coupling depth). A harmonic restraint was applied to mimic the palmitoylation site at the N-terminus of the G-protein to the membrane's phospholipid. Gaussian hills with initial height of 1.2 kcal mol^{-1} applied every 1 ps were used. The hill width was chosen to be 0.1 nm. The Gaussian functions were rescaled in the WT scheme using a bias factor of 20. We performed an initial metadynamics simulation with a higher bias factor and a single CV (between Glu392 of $G_{\alpha s}$ and Arg3.50 of ADBR2/ coupling depth). Fifty starting geometries, spanning the bins of the uncoupling process, were extracted. These 50 starting geometries for each model were then simulated with the multiple walker technique as starting geometries. This ensured faster convergence of the free-energy surface and enhanced the parallelization up to 14,000 CPUs. Each of the simulations converged within 7 μs (collective over the replicas) through two runs on the Haswell nodes of SuperMUC.
As discussed in our work,[25] recent GPCRs simulations have shown that the vestibule of the receptor can pre-orient the ligand[26-30] and thus provide a well-defined extracellular end-point for docking pathways, often by a form of electrostatic focusing[31] but also by a simple mechanical effect in which part of the ligand is anchored, decreasing the number of degrees of freedom to be sampled. This property of the extracellular region makes our single CV quite effective; only a few pathways for GPCR-ligand binding are possible, the ligands find the right orientation during the sampling and binding sites along the path are identified and characterized reliably.

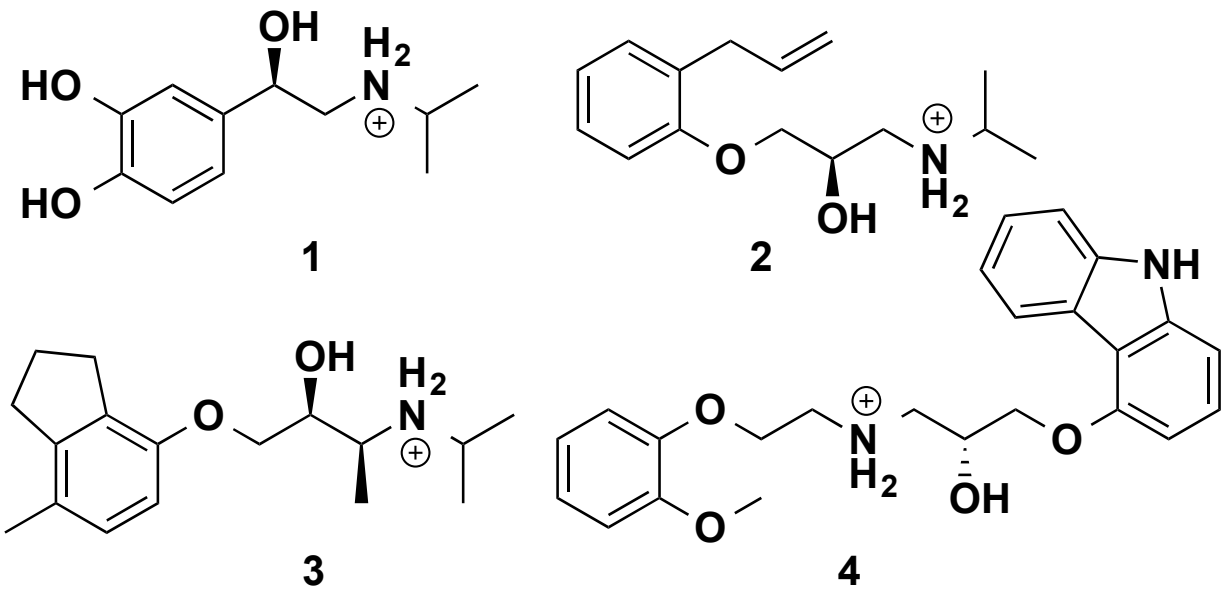
Metadynamics simulation of coupling of the β-arrestin and activation of the ADBR2

The well-tempered variant (WT) of metadynamics was used to map the FES of the activation of ADRB2 and coupling to β-arrestin2 in the presence of each of the four ligands. The *z*-component of the distance between the C_α of the Val71 of the β-arrestin finger loop and the C_α of Arg3.50 was used as a reaction coordinate to describe the coupling of β-arrestin2 to ADRB2 (Coupling depth). The distance between the C_α of the Arg3.50 and Leu6.34 was used as a coordinate for the activation of ADRB2. Gaussian hills with initial height of 1.2 kcal.mol^{-1} applied every 1 ps were used. The hill width was chosen to be 0.1 nm. The Gaussian functions were rescaled in the WT scheme using a bias factor of 20. We performed an initial metadynamics simulation with higher bias factor and a single CV (between the Val71 of the β-arrestin's finger loop and Arg3.50 of ADBR2/ coupling depth). 50 starting geometries, spanning the bins of the uncoupling process, were extracted. These 50 starting geometries for each model were then simulated with the multiple walker technique as staring geometries. This ensured faster convergence of the free-energy surface and enhanced the parallelization up to 14,000 CPUs. Each of the two simulations converged within 7 microseconds (collective over the replicas) through two runs on the Haswell nodes of the SuperMUC.
Detaching the ADRB2's cytoplasmic tail from arrestin required the box size to be doubled in the z-direction. The large increase in the number of bins to converge (points/states on the FES) led to a major increase in the estimated

computational time needed to converge the metadynamics simulation. Thus, our approach allows us to reduce this intractably large landscape to just loss of interaction interface between the receptor and arrestin. This approach allows us to determine the global minima for arrestin-ADRB2 ternary complexes. However, an accurate estimate of the arrestin affinity to ADRB2 would require complete detachment of the arrestin from the C-terminus.

For G-protein and arrestin coupled states, the major conformation change in the two coupling partners, were in the orientation and helical conformation of the α5-helix and finger-loop of Gs and arrestin respectively, which extend/protrude to the coupling interface. In the case of the G-protein/arrestin, our simulations were intended to simulate the uncoupling event. The starting geometries were carefully chosen to start from coupled state with replicas using starting geometries along the uncoupling as described above. Our ability to predict the partially engaged arrestin complexes, recently confirmed by the recently published work of Kumari *et al.*[32] that lies far from the starting geometries for the arrestin complexes that were based on the Rhodopsin-bound arrestin and thus provide an evidence for the success of our approach to model this process.

Schemes



Scheme S1. Structures for the ligands used in this study. The full Gs-protein (Gs)/arrestin native agonist isoprenaline, 1, the Gs/arrestin unselective antagonist alprenolol, 2, the Gs inverse agonist/arrestin antagonist ICI-118,551, 3, and the Gs inverse agonist/arrestin partial agonist, carvedilol, 4.

Figures

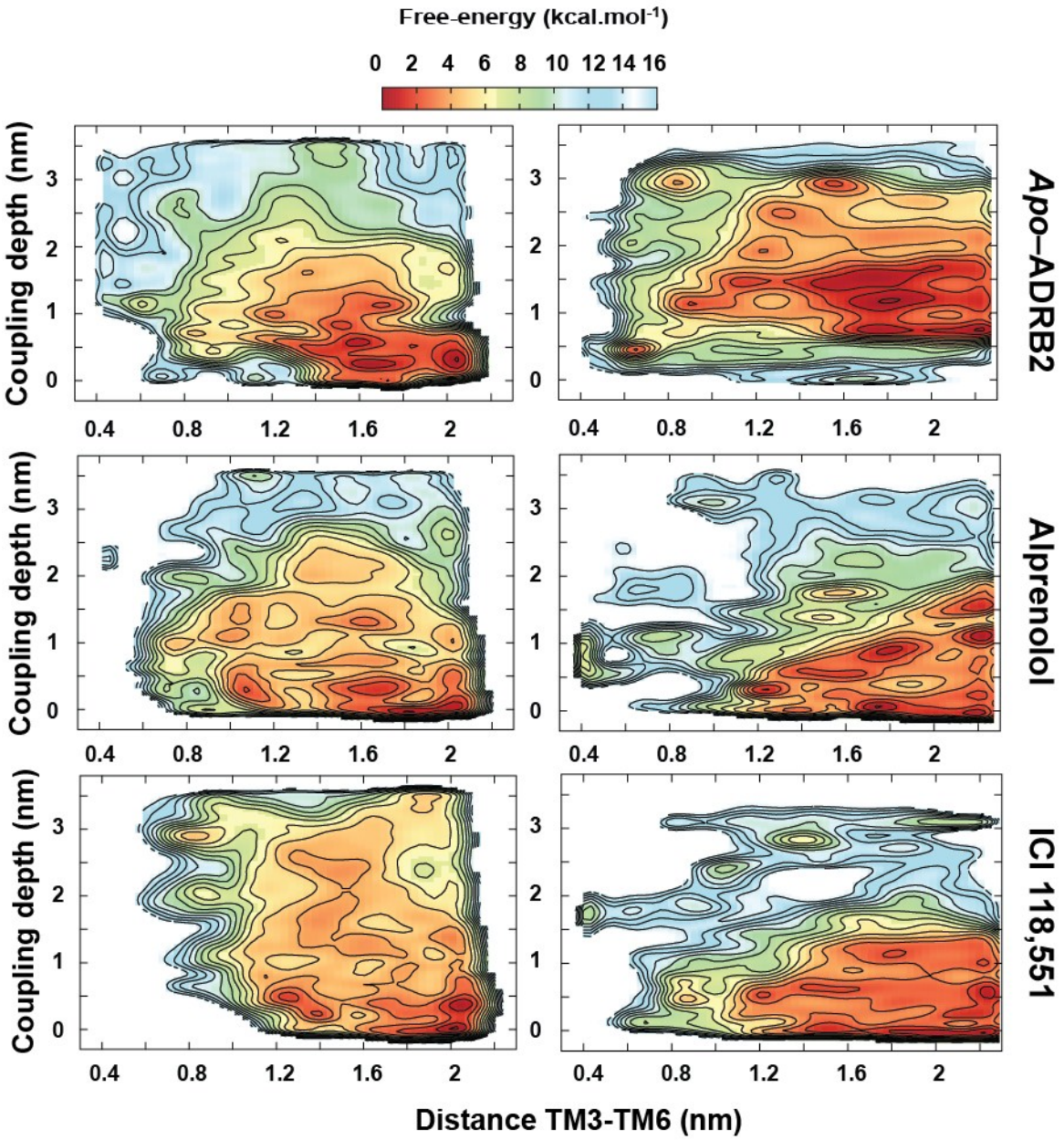


Figure S1. Free-energy contour maps for conformational changes in IBP-ADRB2 coupling. Left: coupling to Gαs and Right: coupling to arrestin.

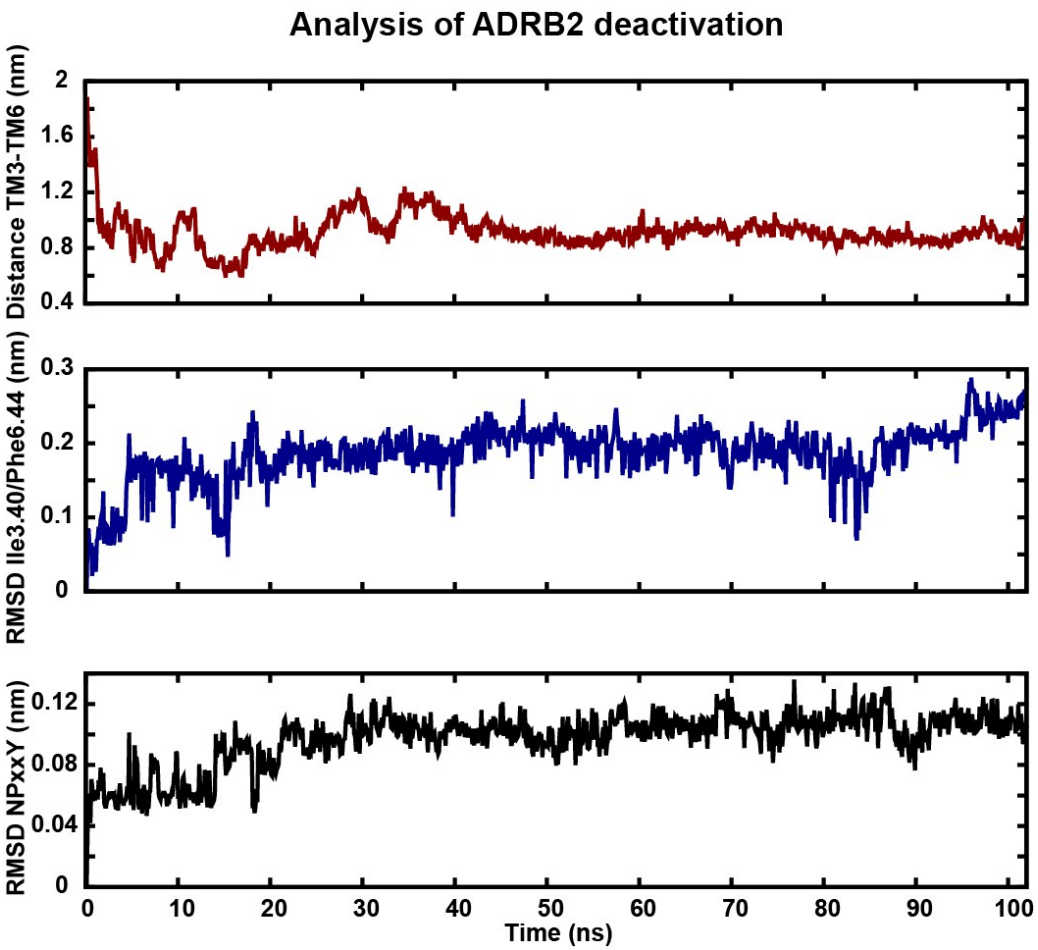


Figure S2. Comparative analysis for the deactivation of the ADRB2 as determined by unbiased simulation.[33] Top: distance between the TM3 and TM6 as determined by the distance from the Cα of Arg3.50 and Leu6.34. Center: RMSD of the connector region Ile3.40 and Phe6.44. Bottom: RMSD of the NPxxY motif at TM7.

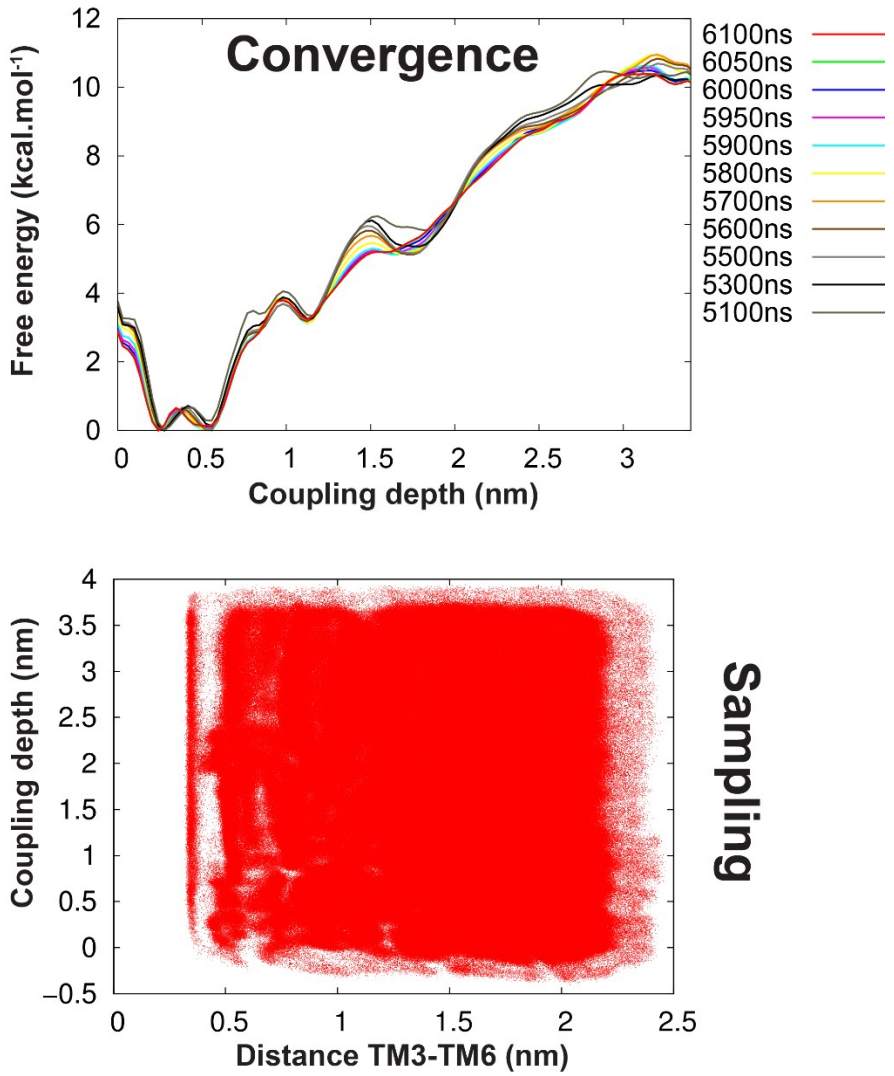


Figure S3. Convergence criteria: Top: convergence of the metadynamics simulation for the coupling and activation of ADRB2 demonstrated by the change in FES as different intervals of sampling. Bottom: Sampling of the metadynamics simulation for the coupling and activation of ADRB2.

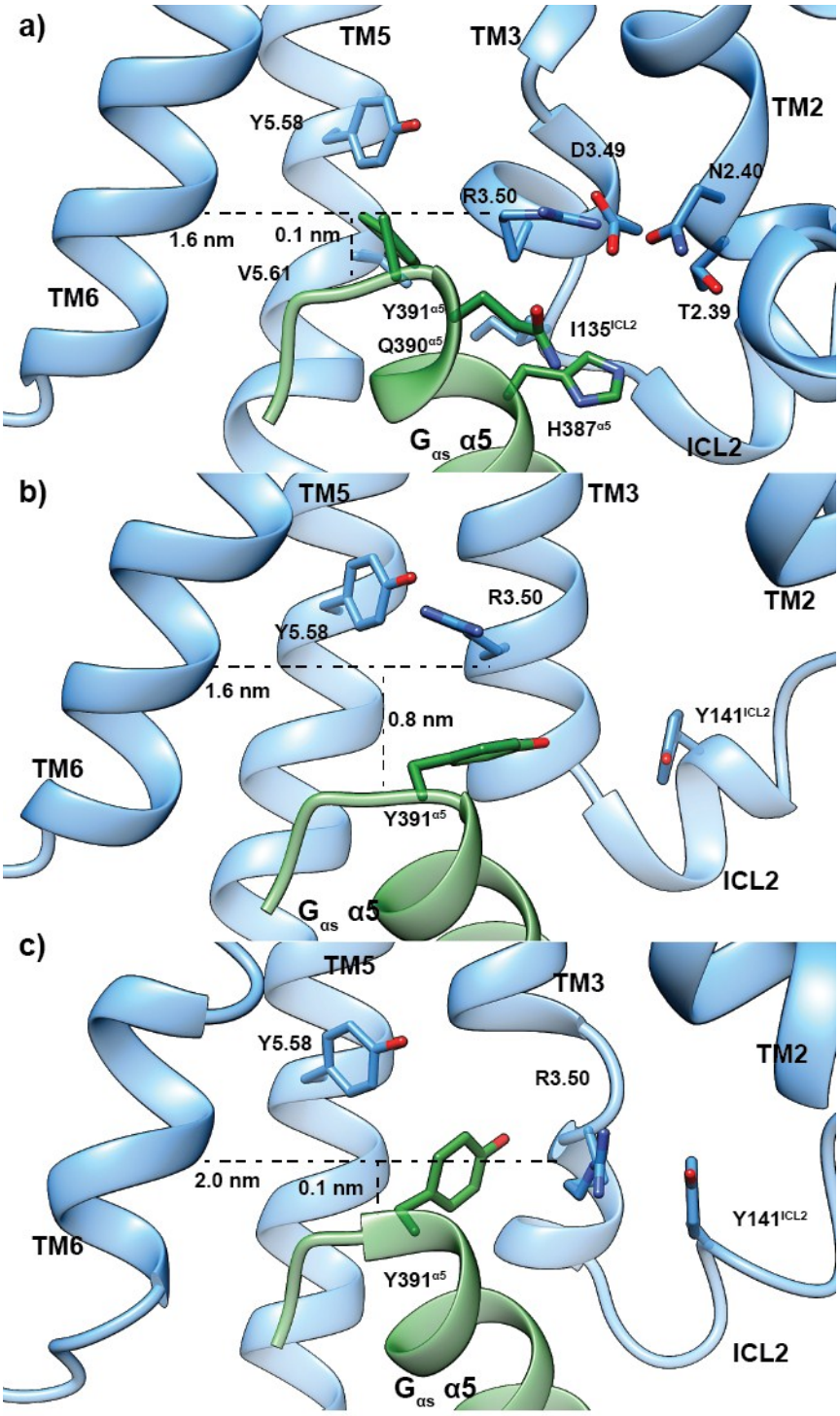


Figure S4. Structural comparison between the three minima for G-protein coupling to ADRB2.

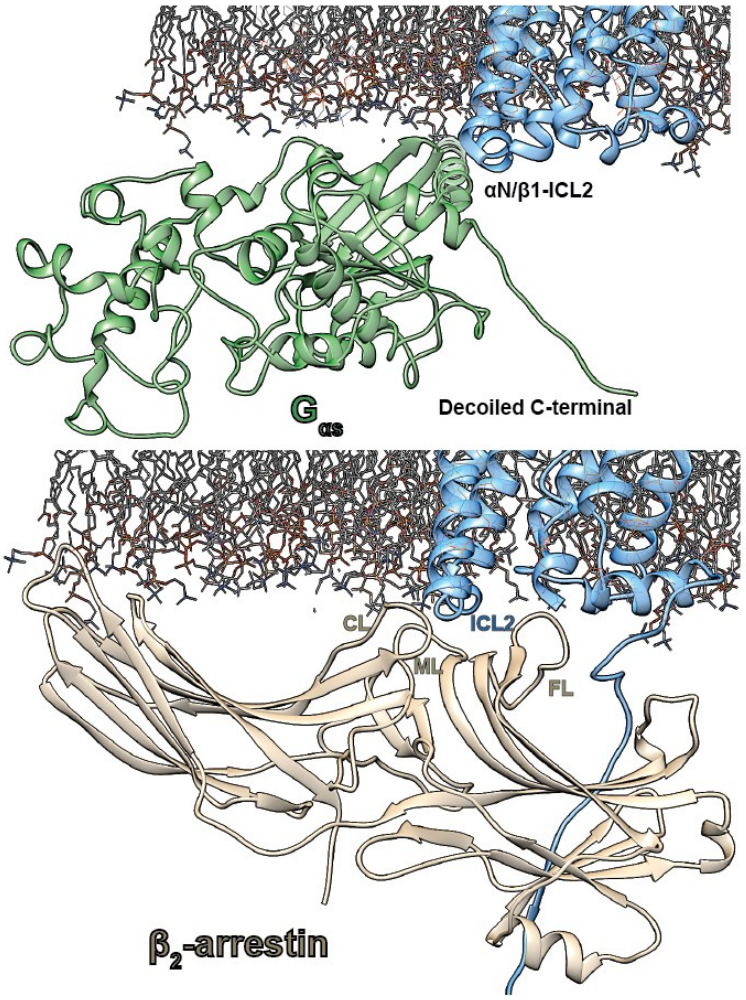


Figure S5. Initial Contact between ADRB2 and IBPs: Top: Gαs bottom: β-arrestin.

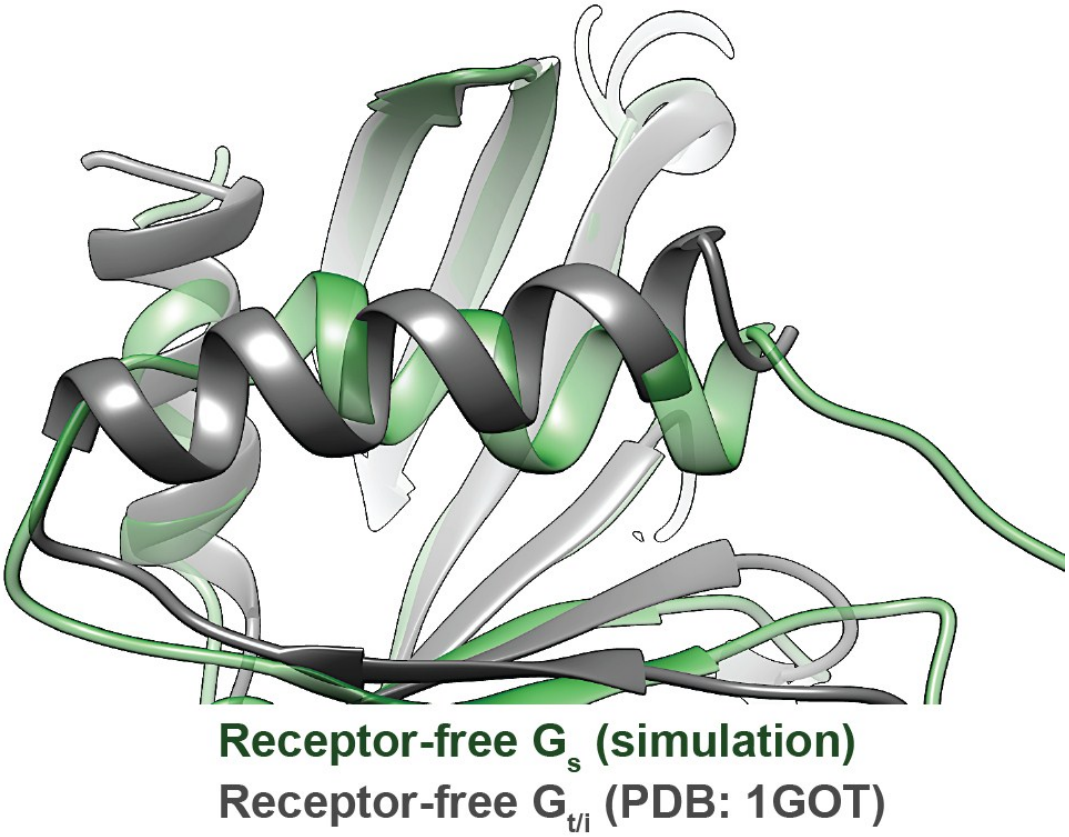


Figure S6. Comparison between the uncoupled Gαs from the metadynamics simulation (green) and the crystallized Gt/i (grey).

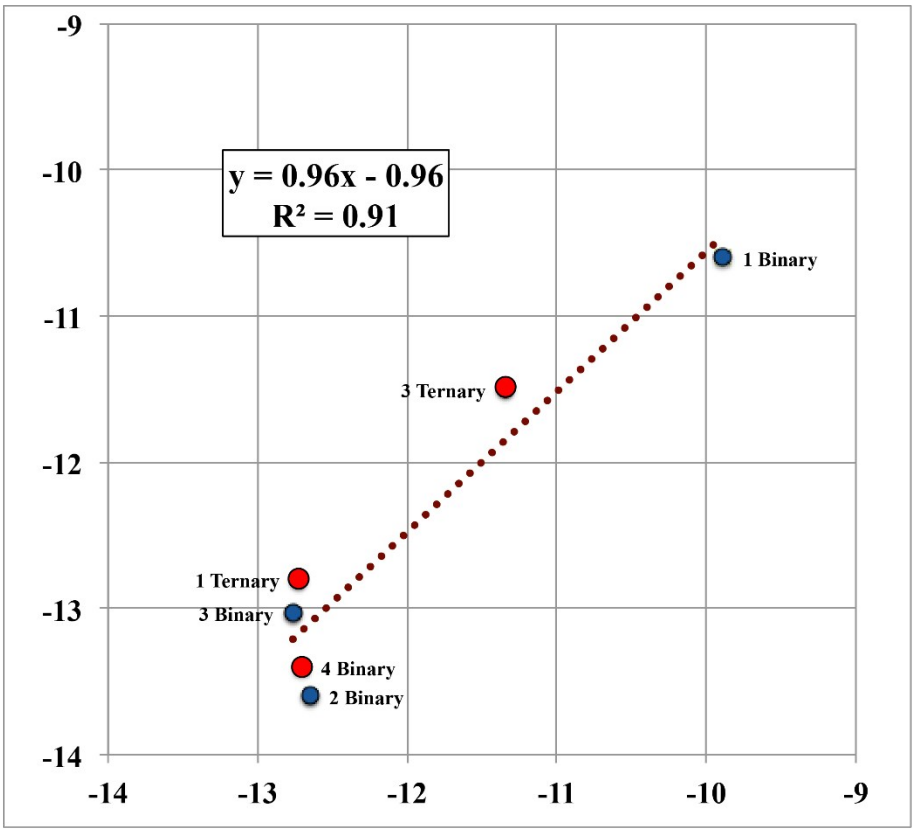


Figure S7. Correlation between the calculated and experimental binding free energies for the four ligands in binary complexes and ternary ones with Gs.

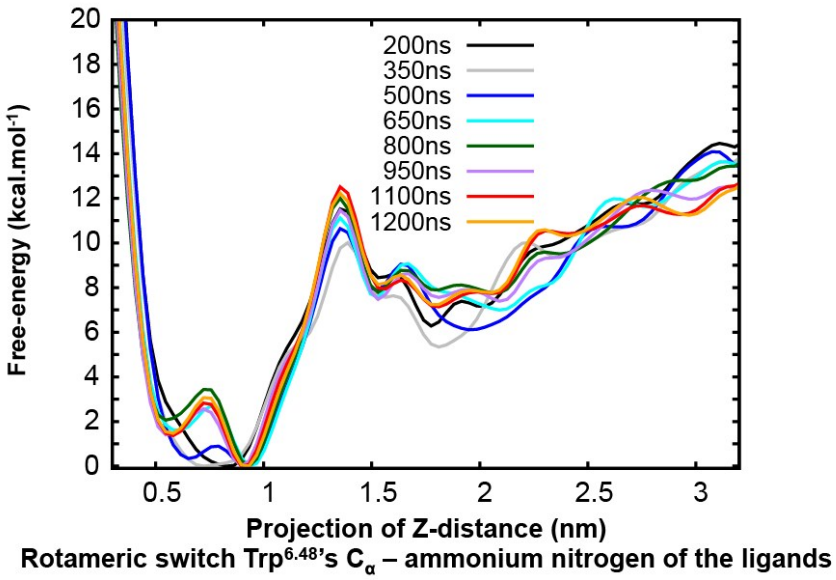


Figure S8. Convergence of the metadynamics simulations for the ligand binding at different sampling exhaustiveness (in nanoseconds collectively over all replica). The free-energy profile doesn't show any significant change (binding free energy changes < 0.1 kcal mol-1) after 950 ns of sampling.

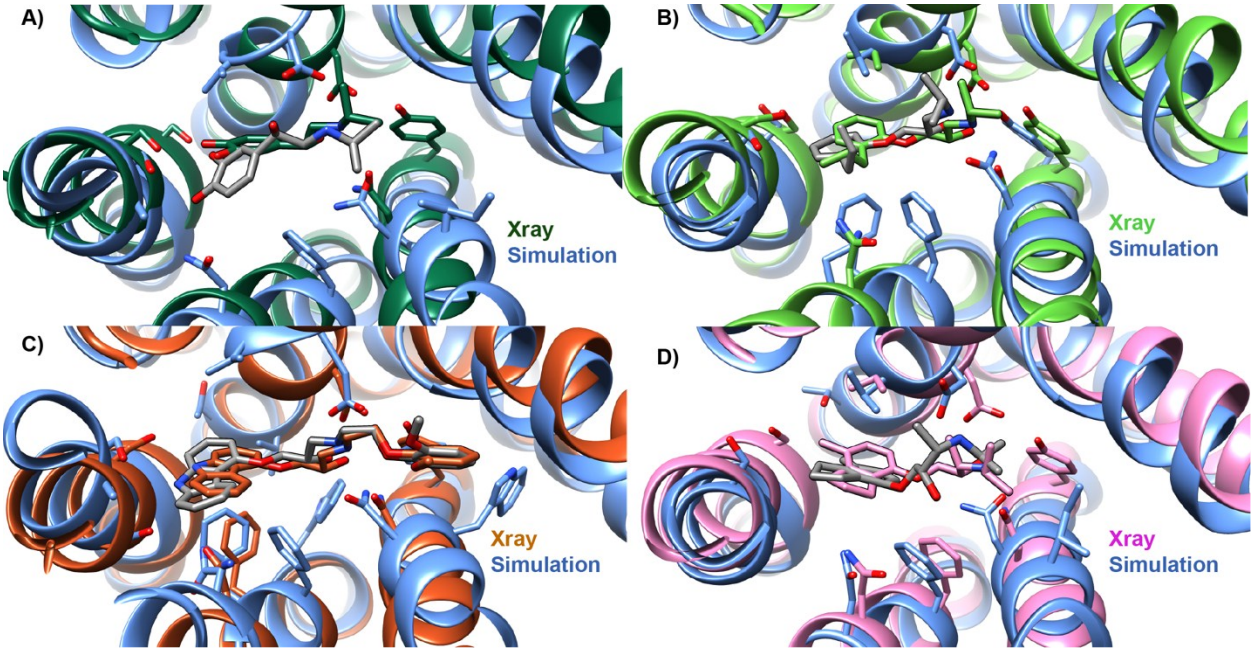


Figure S9. Comparison between binding modes of the global minima for the binary ADRB2 ligands with the crystals structures. A) Isoprenaline with ADRB1 X-ray structure PDB 2Y03. B) Alprenolol with the ADRB2 X-ray structure PDB 3NYA. C) Carvedilol with ADRB1 X-ray structure PDB 4AMJ. C) ICI118551 with ADRB2 X-ray structure PDB 3NY8.

Table S.1. Systems used for metadynamics simulation and their experimental and calculated free energies (in kcal mol-1). Experimental free energies are obtained from the relation ΔG=–RT ln(Ki) at T=298k.

Ligand	Notes	K_i(nM)	$\Delta G_{exp.}$	$\Delta G_{calc.}$	Error
ICI-188551	Binary-inactive	1[34]	-12.8	-13.0	0.3
Alprenolol	Binary-inactive	1.2[34]	-12.7	-13.6	0.9
Carvedilol	Binary-inactive	1.1[35]	-12.7	-13.4	0.7
Isoprenaline	Binary-inactive	107[14]	-9.9	-10.6	0.7
Isoprenaline	Ternary-Gs	1.07[14]	-12.7	-12.8	0.1
ICI-188551	Ternary-Gs	10[34]	-11.3	-11.5	0.2

References

1. K. Lindorff-Larsen, S. Piana, K. Palmo, P. Maragakis, J. L. Klepeis, R. O. Dror and D. E. Shaw, *Proteins*, 2010, **78**, 1950-1958.
2. S. W. Siu, R. Vacha, P. Jungwirth and R. A. Bockmann, *The Journal of chemical physics*, 2008, **128**, 125103.
3. M. A. Lomize, A. L. Lomize, I. D. Pogozheva and H. I. Mosberg, *Bioinformatics*, 2006, **22**, 623-625.
4. M. G. Wolf, M. Hoefling, C. Aponte-Santamaria, H. Grubmuller and G. Groenhof, *J Comput Chem*, 2010, **31**, 2169-2174.
5. J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman and D. A. Case, *J Comput Chem*, 2004, **25**, 1157-1174.
6. A. Jakalian, D. B. Jack and C. I. Bayly, *J Comput Chem*, 2002, **23**, 1623-1641.
7. H. J. C. Berendsen, J. R. Grigera and T. P. Straatsma, *J Phys Chem-Us*, 1987, **91**, 6269-6271.
8. S. Pronk, S. Pall, R. Schulz, P. Larsson, P. Bjelkmar, R. Apostolov, M. R. Shirts, J. C. Smith, P. M. Kasson, D. van der Spoel, B. Hess and E. Lindahl, *Bioinformatics*, 2013, **29**, 845-854.
9. G. A. Tribello, M. Bonomi, D. Branduardi, C. Camilloni and G. Bussi, *Comput Phys Commun*, 2014, **185**, 604-613.
10. V. Cherezov, D. M. Rosenbaum, M. A. Hanson, S. G. Rasmussen, F. S. Thian, T. S. Kobilka, H. J. Choi, P. Kuhn, W. I. Weis, B. K. Kobilka and R. C. Stevens, *Science*, 2007, **318**, 1258-1265.
11. D. Wacker, G. Fenalti, M. A. Brown, V. Katritch, R. Abagyan, V. Cherezov and R. C. Stevens, *Journal of the American Chemical Society*, 2010, **132**, 11443-11445.
12. T. Warne, R. Moukhametzianov, J. G. Baker, R. Nehme, P. C. Edwards, A. G. Leslie, G. F. Schertler and C. G. Tate, *Nature*, 2011, **469**, 241-244.
13. T. Warne, P. C. Edwards, A. G. Leslie and C. G. Tate, *Structure*, 2012, **20**, 841-849.
14. S. G. Rasmussen, B. T. DeVree, Y. Zou, A. C. Kruse, K. Y. Chung, T. S. Kobilka, F. S. Thian, P. S. Chae, E. Pardon, D. Calinski, J. M. Mathiesen, S. T. Shah, J. A. Lyons, M. Caffrey, S. H. Gellman, J. Steyaert, G. Skiniotis, W. I. Weis, R. K. Sunahara and B. K. Kobilka, *Nature*, 2011, **477**, 549-555.
15. R. C. Kling, H. Lanig, T. Clark and P. Gmeiner, *PloS one*, 2013, **8**, e67244.
16. A. K. Shukla, A. Manglik, A. C. Kruse, K. H. Xiao, R. I. Reis, W. C. Tseng, D. P. Staus, D. Hilger, S. Uysal, L. Y. Huang, M. Paduch, P. Tripathi-Shukla, A. Koide, S. Koide, W. I. Weis, A. A. Kossiakoff, B. K. Kobilka and R. J. Lefkowitz, *Nature*, 2013, **497**, 137-+.
17. N. Eswar, D. Eramian, B. Webb, M. Y. Shen and A. Sali, *Methods in molecular biology*, 2008, **426**, 145-159.
18. Y. Kang, X. E. Zhou, X. Gao, Y. He, W. Liu, A. Ishchenko, A. Barty, T. A. White, O. Yefanov, G. W. Han, Q. Xu, P. W. de Waal, J. Ke, M. H. Tan, C. Zhang, A. Moeller, G. M. West, B. D. Pascal, N. Van Eps, L. N. Caro, S. A. Vishnivetskiy, R. J. Lee, K. M. Suino-Powell, X. Gu, K. Pal, J. Ma, X. Zhi, S. Boutet, G. J. Williams, M. Messerschmidt, C. Gati, N. A. Zatsepin, D. Wang, D. James, S. Basu, S. Roy-Chowdhury, C. E. Conrad, J. Coe, H. Liu, S. Lisova, C. Kupitz, I. Grotjohann, R. Fromme, Y. Jiang, M. Tan, H. Yang, J. Li, M. Wang, Z. Zheng, D. Li, N. Howe, Y. Zhao, J. Standfuss, K. Diederichs, Y. Dong, C. S. Potter, B. Carragher, M. Caffrey, H. Jiang, H. N. Chapman, J. C. Spence, P. Fromme, U. Weierstall, O. P. Ernst, V. Katritch, V. V. Gurevich, P. R. Griffin, W. L. Hubbell, R. C. Stevens, V. Cherezov, K. Melcher and H. E. Xu, *Nature*, 2015, **523**, 561-567.
19. J. W. Wisler, S. M. DeWire, E. J. Whalen, J. D. Violin, M. T. Drake, S. Ahn, S. K. Shenoy and R. J. Lefkowitz, *Proceedings of the National Academy of Sciences of the United States of America*, 2007, **104**, 16657-16662.
20. N. Homeyer, A. H. Horn, H. Lanig and H. Sticht, *Journal of molecular modeling*, 2006, **12**, 281-289.
21. A. Laio and M. Parrinello, *Proceedings of the National Academy of Sciences of the United States of America*, 2002, **99**, 12562-12566.
22. A. Barducci, G. Bussi and M. Parrinello, *Physical review letters*, 2008, **100**, 020603.

23. V. Limongelli, M. Bonomi and M. Parrinello, *Proceedings of the National Academy of Sciences of the United States of America*, 2013, **110**, 6358-6363.
24. P. Raiteri, A. Laio, F. L. Gervasio, C. Micheletti and M. Parrinello, *Journal of Physical Chemistry B*, 2006, **110**, 3533-3539.
25. N. Saleh, P. Ibrahim, G. Saladino, F. L. Gervasio and T. Clark, *J Chem Inf Model*, 2017, **56**.
26. R. O. Dror, A. C. Pan, D. H. Arlow, D. W. Borhani, P. Maragakis, Y. Shan, H. Xu and D. E. Shaw, *Proceedings of the National Academy of Sciences of the United States of America*, 2011, **108**, 13118-13123.
27. R. O. Dror, H. F. Green, C. Valant, D. W. Borhani, J. R. Valcourt, A. C. Pan, D. H. Arlow, M. Canals, J. R. Lane, R. Rahmani, J. B. Baell, P. M. Sexton, A. Christopoulos and D. E. Shaw, *Nature*, 2013, **503**, 295-299.
28. A. C. Kruse, A. M. Ring, A. Manglik, J. Hu, K. Hu, K. Eitel, H. Hubner, E. Pardon, C. Valant, P. M. Sexton, A. Christopoulos, C. C. Felder, P. Gmeiner, J. Steyaert, W. I. Weis, K. C. Garcia, J. Wess and B. K. Kobilka, *Nature*, 2013, **504**, 101-106.
29. K. Kappel, Y. Miao and J. A. McCammon, *Quarterly reviews of biophysics*, 2015, **48**, 479-487.
30. Y. Miao and J. A. McCammon, *Proceedings of the National Academy of Sciences of the United States of America*, 2016, **113**, 12162-12167.
31. S. Blumenthal, Y. Tang, W. Yang and B. Y. Chen, *IEEE/ACM Trans Comput Biol Bioinform*, 2013, **10**, 1188-1198.
32. P. Kumari, A. Srivastava, R. Banerjee, E. Ghosh, P. Gupta, R. Ranjan, X. Chen, B. Gupta, C. Gupta, D. Jaiman and A. K. Shukla, *Nature communications*, 2016, **7**, 13416.
33. R. O. Dror, D. H. Arlow, P. Maragakis, T. J. Mildorf, A. C. Pan, H. F. Xu, D. W. Borhani and D. E. Shaw, *Proceedings of the National Academy of Sciences of the United States of America*, 2011, **108**, 18684-18689.
34. S. G. Rasmussen, H. J. Choi, J. J. Fung, E. Pardon, P. Casarosa, P. S. Chae, B. T. Devree, D. M. Rosenbaum, F. S. Thian, T. S. Kobilka, A. Schnapp, I. Konetzki, R. K. Sunahara, S. H. Gellman, A. Pautsch, J. Steyaert, W. I. Weis and B. K. Kobilka, *Nature*, 2011, **469**, 175-180.
35. C. Hoffmann, M. R. Leitz, S. Oberdorf-Maass, M. J. Lohse and K. N. Klotz, *Naunyn-Schmiedeberg's archives of pharmacology*, 2004, **369**, 151-159.