SI: A streamlined, general approach for computing ligand binding free energies and its application to GPCR-bound cholesterol

Reza Salari,[†,‡] Thomas Joseph,[¶,†] Ruchi Lohia,[†] Jérôme Hénin,[§,‖] and Grace Brannigan[*,†,‡,‖]

†*Center for Computational and Integrative Biology, Rutgers University, Camden, NJ, USA*

‡*Department of Physics, Rutgers University, Camden, NJ, USA*

¶*Department of Anesthesia and Critical Care, University of Pennsylvania Perelman School of Medicine, Philadelphia, PA, USA*

§*Laboratoire de Biochimie Théorique, Institut de Biologie Physico-Chimique, CNRS, Paris, France*

‖*These authors contributed equally to supervising this work.*

E-mail: grace.brannigan@rutgers.edu

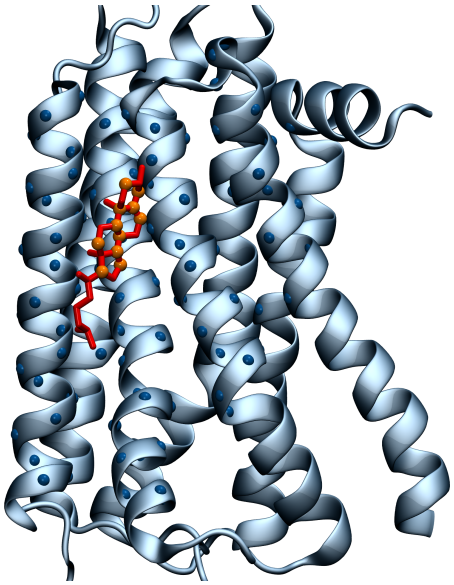

Figure S1. Definition of a distance-to-bound-configuration coordinate. The β2-adrenergic receptor is rendered in blue as ribbons, bound cholesterol in red. Atoms of cholesterol and the receptor involved in defining the DBC coordinate are shown as orange and blue spheres, respectively.

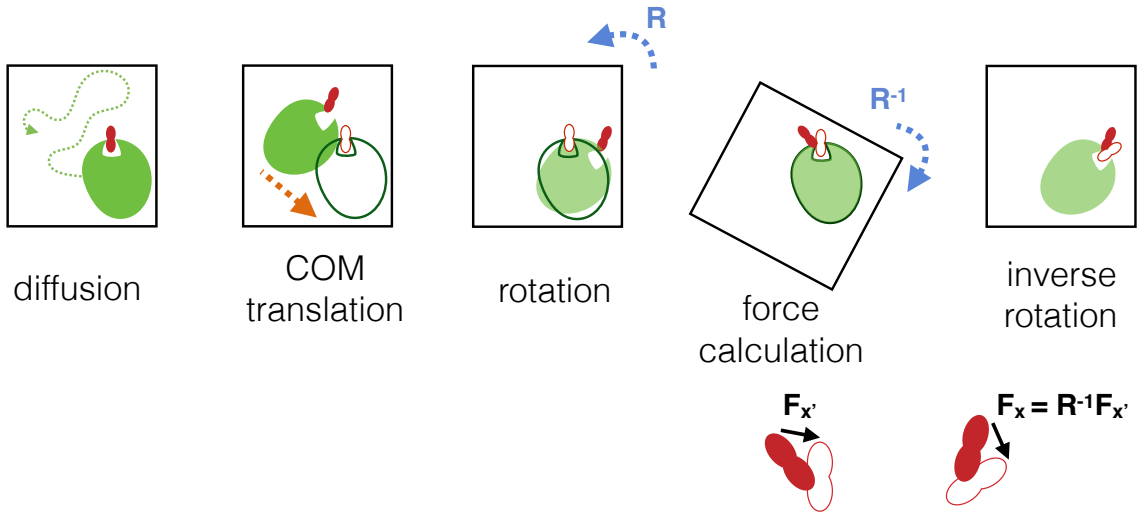

Figure S2. Rotations and translations in calculation of a distance-to-bound-configuration coordinate. A) The green receptor with red ligand bound diffuses away from its initial (reference) configuration. B) The center of mass (COM) is translated back to the reference position. C) The optimal rotation operation $\mathbf{R}$ for aligning the protein onto the reference coordinates is determined. D) A force is derived from the restraint potential, using the RMSD between the ligand and its reference coordinates. E) The direction of the force vector is inversely rotated with $\mathbf{R}^{-1}$

		β2 adrenergic	5HT-2B	μ Opioid
$\Delta G_r = k_B T \ln \frac{Z_r^{\triangle}}{Z_g^{\triangle} Z_b^{-}}$	Replica 1(kcal/mol)	34.4	25.6	23.1
	Replica 2(kcal/mol)	33.9	25.2	25.7
	Replica 3(kcal/mol)	34.5	27.0	25.7
	Arithmetic Mean(kcal/mol)	34.3	25.9	24.8
	Standard Error(kcal/mol)	0.2	0.5	0.9

$\frac{Z_r^{\triangle}}{Z_g^{\triangle} Z_b^{-}}$	Replica 1	2E+25	7E+18	1E+17
	Replica 2	9E+24	4E+18	8E+18
	Replica 3	2E+25	7E+19	8E+18
	Arithmetic Mean	2E+25	3E+19	6E+18
	Standard Error	5E+24	2E+19	3E+18
	Geometric Mean	2E+25	1E+19	2E+18

x_{50}	Replica 1	5E-10	2E-03	
	Replica 2	1E-09	3E-03	1E-03
	Replica 3	4E-09	1E-04	1E-03
	Arithmetic Mean	2E-09	2E-03	1E-03
	Standard Error	1E-09	8E-04	
	Geometric Mean	1E-09	8E-04	1E-03
	Determined using single shaded Geometric Mean	6E-10	1E-03	6E-03

Table S1. Dispersion in protein decoupling $\left(\frac{Z_r^{\triangle}}{Z_g^{\triangle} Z_b^{-}}\right)$ calculations across replicas, and resulting uncertainty in x_{50}. Shaded row corresponds to values in Table 6. For values of $\left(\frac{Z_r^{\triangle}}{Z_g^{\triangle} Z_b^{-}}\right)$ weaker than 10^{18}, no x_{50} could be predicted, because occupancy was expected to plateau below 50%.