Identification of GPCR Transition Pathways Using Gō Models

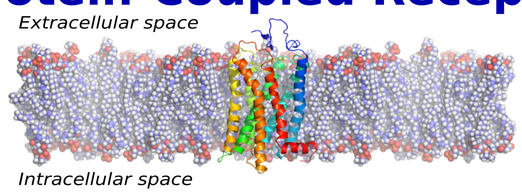
Leslie A. Salas-Estrada[1], Stephen J. Constable[1], Nicholas Leioatts[2], Tod D. Romo[1] and Alan Grossfield[1]
[1]University of Rochester, Rochester, NY, USA
[2]Max Planck Institute for Biophysical Chemistry, Göttingen, Germany

Poster PDF
http://tinyurl.com/gpcr-go

Abstract

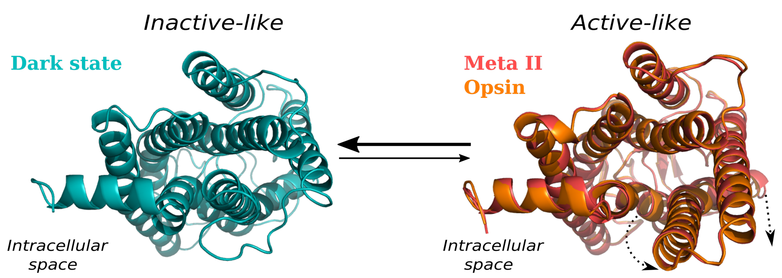
One caveat of standard all-atom molecular dynamics (MD) simulations is that they are resource-intensive, particularly when inspecting state transition pathways. To date, the timescales involved in processes such as G protein-coupled receptor (GPCR) activation make it largely impossible to study them using standard all-atom MD under equilibrium conditions. Our group previously showed that simpler structure-based (Gō-like) potentials could be used to study GPCR state transitions, yielding results in qualitative agreement with those obtained from standard all-atom simulations but with negligible computational cost. These potentials are energetically smooth by design with few non-native traps and a single global energy minimum corresponding to the end-point of the transition of interest. In the case of rhodopsin, a prototypical GPCR, two distinct but reproducible pathways were observed during activation and deactivation. While it is possible that a separation of pathways is mechanistically required for function, it could also be the result of simulating state transitions in two separate energy surfaces out of equilibrium (one for activation and one for deactivation). To study state interconversion in equilibrium, we have implemented a strategy for modeling Gō-like potentials with multiple energy minima at all-heavy-atom resolution. This approach allows for extensive sampling still at a low computational cost and can be extrapolated to almost any biomolecular system in a straightforward manner to complement current standard all-atom MD strategies.

G Protein-Coupled Receptors



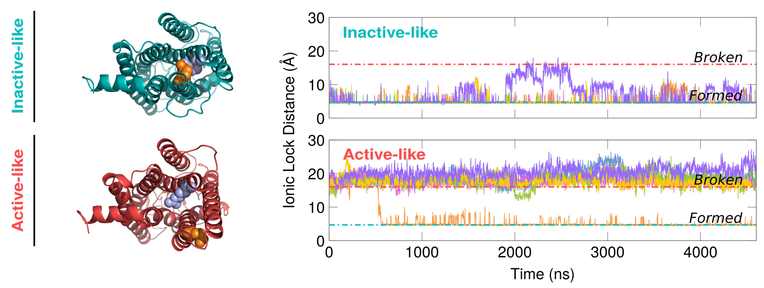
- Ubiquitous 7 transmembrane (7 TM) α-helical proteins
  - Highly conserved topology
  - Functional diversity and specificity
- 826 distinct members in humans
- Target of 40% small-molecule drugs
- Transduce information across membranes in response to stimuli
  - Conformational changes mediate downstream signaling

Opsin Exists in Equilibrium Between Inactive-Like and Active-Like Conformations



- Rhodopsin is a prototypical GPCR
  - Highly efficient photoreceptor
  - Mediates visual signaling cascade
- Opsin: ligand-free (apo) rhodopsin
  - Extremely low signaling
- Crystal looks active-like
  - Opsin/Meta II Cα RMSD is 0.51 Å
- Inactive ensemble more populated at physiological pH
  - FTIR spectra inactive-like

Standard MD Undersamples Functional Transitions

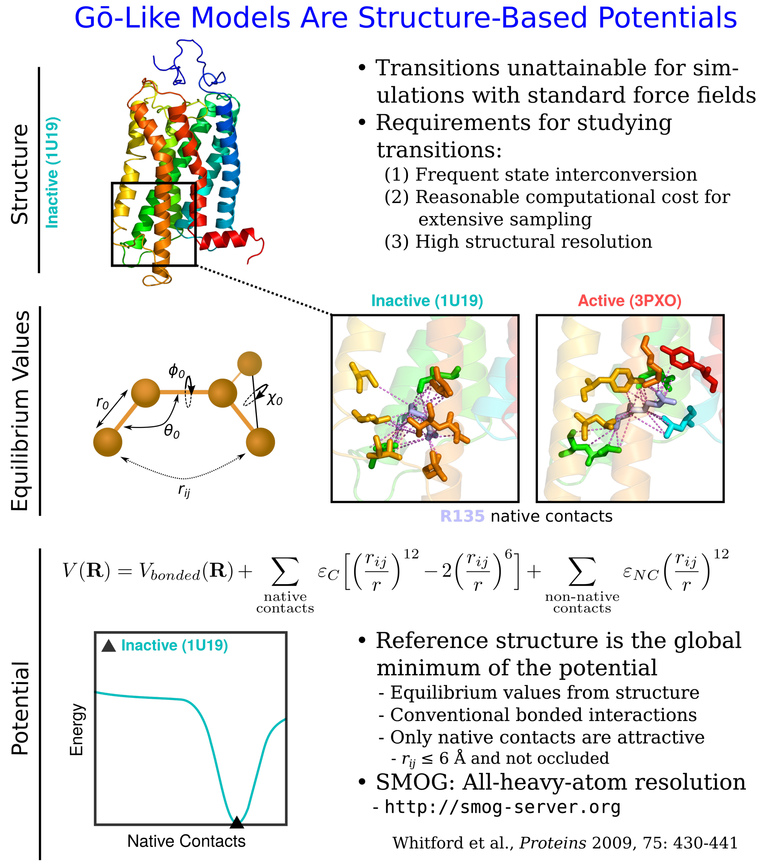


- Inactive-like and active-like opsin
  - Starting structures:
    - Inactive-like: 1U19 (apo)
    - Active-like: 3CAP
- Standard all-atom MD simulations
  - 6 replicates per starting structure
  - Aggregate simulation time: 55.2 μs
  - 1.5 years of supercomputer time
- Ionic lock: R135(N)-E247(O) distance
  - Inactive-like: 4.7 Å
  - Active-like: 15.8 Å
- Standard MD captures fluctuations within a state
  - Transitions are rare

Leioatts et al., *Biophys. J.* 2015, 109: 608-617

Opsin Transition Pathways from Single-Basin Gō-Like Potentials

Gō-Like Models Are Structure-Based Potentials



- Transitions unattainable for simulations with standard force fields
- Requirements for studying transitions:
  (1) Frequent state interconversion
  (2) Reasonable computational cost for extensive sampling
  (3) High structural resolution

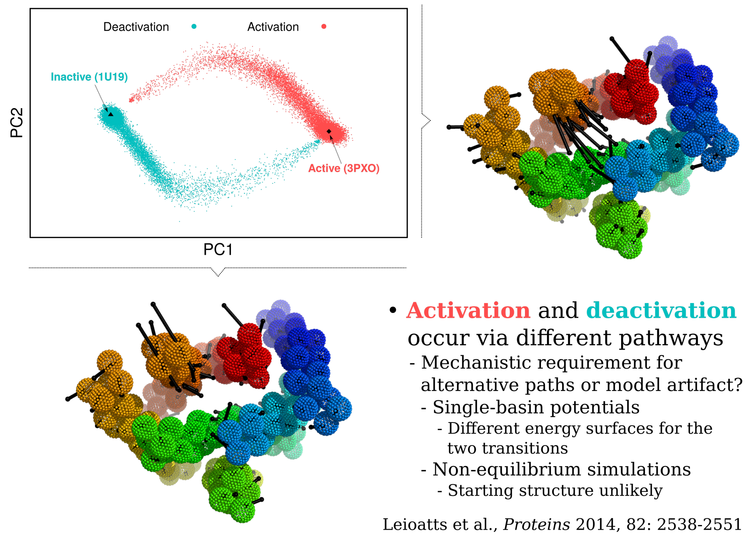
$$V(\mathbf{R}) = V_{bonded}(\mathbf{R}) + \sum_{\substack{\text{native}\\\text{contacts}}} \varepsilon_C \left[\left(\frac{r_{ij}}{r}\right)^{12} - 2\left(\frac{r_{ij}}{r}\right)^{6} \right] + \sum_{\substack{\text{non-native}\\\text{contacts}}} \varepsilon_{NC} \left(\frac{r_{ij}}{r}\right)^{12}$$

- Reference structure is the global minimum of the potential
  - Equilibrium values from structure
  - Conventional bonded interactions
  - Only native contacts are attractive
  - $r_{ij} \leq 6$ Å and not occluded
- SMOG: All-heavy-atom resolution
  - `http://smog-server.org`

Whitford et al., *Proteins* 2009, 75: 430-441

Opsin Populates Two Different Pathways

- **Active** and **inactive** potentials generated with SMOG
  - Study inactive-to-active and active-to-inactive transitions separately
  - **Activation** (**deactivation**):
    - Use **active** (**inactive**) potential
    - Start from **inactive** (**active**) structure
    - Run 1,000 trajectories in GROMACS
- Identify most concerted atomic motions in transitions
  - Principal component analysis on TMC_αs of aggregate data set
  - Data projected onto PC1 & PC2
  - Each dot is a trajectory frame of a transition in PC-space
  - Crystal structures:
    - Inactive (1U19:▲)
    - Active (3PXO: ◆)

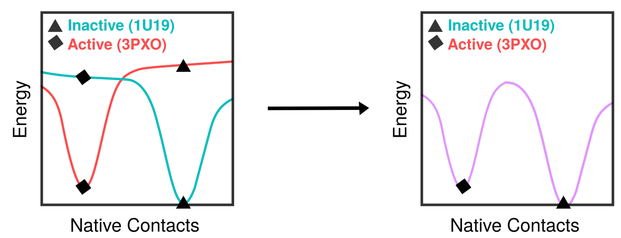


- **Activation** and **deactivation** occur via different pathways
  - Mechanistic requirement for alternative paths or model artifact?
  - Single-basin potentials
    - Different energy surfaces for the two transitions
  - Non-equilibrium simulations
    - Starting structure unlikely

Leioatts et al., *Proteins* 2014, 82: 2538-2551

Equilibrium Transitions Using Multi-Basin Potentials

- Requirements for studying state transitions *in equilibrium*:
  (1) Single energy surface with multiple minima corresponding to (sub)states
  (2) Straightforward automation
    - Designing multi-basin potentials by hand is usually hard and error prone
  (3) Use global mixing schemes
    - Compute energies of each basin locally and then mix them macroscopically
- Mixed potential should preserve location and shape of each basin
  - Local native contacts govern dynamics in the vicinity of minimum
- Transition barrier(s) between basins should be continuous and smooth
  - Dynamics propagation requires differentiable energies



Boltzmann-Weighting

$$V_{mix}(\mathbf{R}) = -\frac{1}{\beta_{mix}} \ln \left\{ \exp(-\beta_{mix}(V_1(\mathbf{R}) + \Delta V_1)) + \exp(-\beta_{mix}(V_2(\mathbf{R}) + \Delta V_2)) \right\}$$

- Exponential average of potentials at every energy evaluation
  - β_{mix}: Contribution of each potential
    - $\uparrow\beta_{mix}$ → Less distortion, higher barrier
    - $\downarrow\beta_{mix}$ → More distortion, smaller barrier
  - ΔV_i: Relative stability of i^{th} state
- Natural extension to multi-basin potentials
  - Each new potential adds an exponential term
  - Relative stability of each basin can be controlled individually

Best et al., *Structure* 2005, 13: 1755-1763

Interpolation

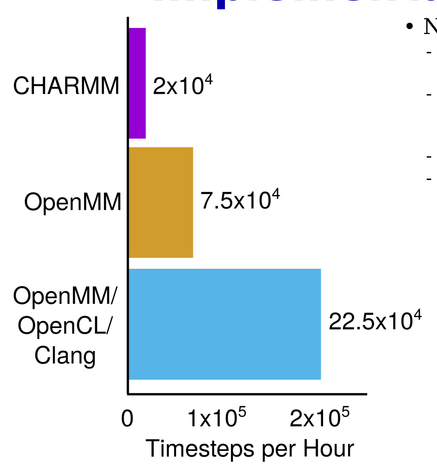
$$V_{mix}(\mathbf{R}) = \frac{V_1(\mathbf{R}) + \Delta V_1 + V_2(\mathbf{R}) + \Delta V_2}{2} - \sqrt{\left(\frac{V_1(\mathbf{R}) + \Delta V_1 - V_2(\mathbf{R}) - \Delta V_2}{2}\right)^2 + \Delta^2}$$

- Computationally convenient function couples potentials
  - Δ: Coupling constant
    - $\uparrow\Delta$ → More distortion, smaller barrier
    - $\downarrow\Delta$ → Less distortion, higher barrier
  - ΔV_i: Relative stability of i^{th} state
- Barrier heights can be tuned individually
- *n*-basin generalization requires solving *nxn* equation

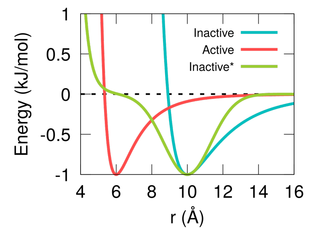
Okazaki et al., *PNAS* 2006, 103: 11844-11849

Implementation

- Initial implementation in CHARMM
  - Dynamics propagation is slow
- Current implementation in OpenMM
  - Patch to support SMOG input files in OpenMM GitHub as of Apr, 2017
- Multi-basin potentials implemented as a Python module
- Interpolation and Boltzmann-weighted formalisms implemented as `CustomIntegrators`
  - Mixed potential energy accessible as a `StateDataReporter`
- SHAKE requires matching bond lengths among sets of potentials
  - Bonds averaged over all potentials



- Non-bonded interactions
  - Critical distance varies among states
    - Large and unphysical energy barrier
  - Make repulsive interactions identical
    - Decouple excluded volume from position of minimum
  - Attractive interactions define basins
  - Modified potential preserves native basins



Local native contacts

$$V^*(r) = \varepsilon_{ij} \left(\left(1 + \left(\frac{\varepsilon_o}{\varepsilon_{ij}}\right)\left(\frac{r_{\text{rep}}}{r}\right)^{12}\right)\left(1 - e^{-\frac{(r-r_{ij})^2}{2\sigma_{ij}^2}}\right) - 1 \right)$$

Local non-native contacts

$$V^*(r) = \varepsilon_o \left(\frac{r_{\text{rep}}}{r}\right)^{12}$$

where,
- r_{ij}: Critical distance
- ε_{ij}: Basin depth
- r_{rep}: Excluded volume
- $\sigma_{ij}^2 = r_{ij}^2 / 50\ ln(2)$
- $\varepsilon_o = 1$ kJ/mol

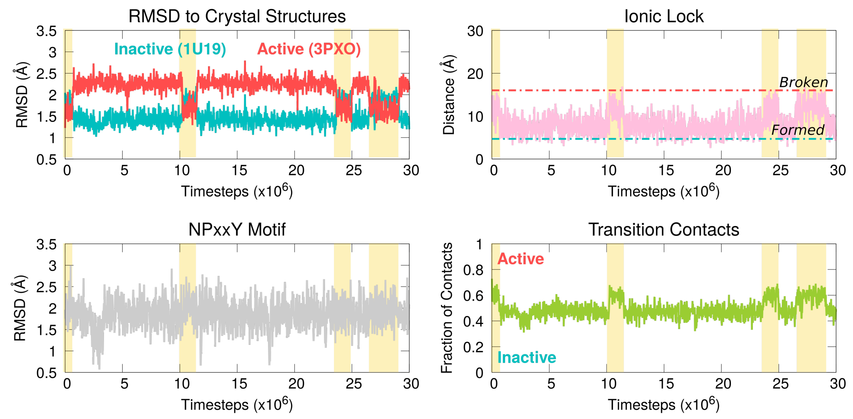
Okazaki et al., *PNAS* 2006, 103: 11844-11849
Noel and Onuchic, *Springer US* 2012, 31-54

Simulation Details

Starting Structure	PDB Accession Codes	Potentials	Resources
Inactive	1U19 (apo)	Dual-basin using Boltzmann-Weighting and Interpolation	10^6 timesteps 1 CPU 5.8 hours
Active	3PXO (apo)		

- System size: ~2,600 atoms
  - Protein stable in the absence of a lipid bilayer and water
- Force field: SMOG1/2
  - Whitford et al., *Proteins* 2009, 75: 430-441
- Ensemble: NVT
  - T = 50 K in reduced units
- Langevin dynamics
  - Collision frequency = 0.1 ps^{-1}
- VDW cutoff: 10 Å
- No electrostatics
- Timestep: 2 fs
- Software: OpenMM compiled with Clang and OpenCL

Opsin Transitions in Equilibrium



- Double-basin interpolation trajectory
  - Starting from **active** structure
  - Δ = 800 kJ/mol, $\Delta V_{inactive}$ = 0 kJ/mol and ΔV_{active} = 0 kJ/mol
- Transitions take the system within 1-1.5 Å of crystal structures
- Biologically motivated reaction coordinates might not work
  - NPxxY motif does not capture transition
- Transition contacts
  - Set of contacts derived from transition end-points: unique active (inactive)
  - Evolution of state-specific contacts
    - Side-chain centroids within 8 Å in TM region

$$\frac{Transition}{contacts} = \frac{Inactive\,broken + Active\,formed}{Inactive + Active}$$

1: All unique active (inactive) formed (broken)
0: No unique active (inactive) formed (broken)

Leioatts et al., *Biophys. J.* 2015, 109: 608-617

Conclusions

- Standard MD undersamples state transitions in complex systems
- Gō-like models can capture state transitions
  - Study state interconversion at negligible computational cost
  - Retain all-heavy-atom resolution
- We implemented a strategy for simulating multi-basin potentials
  - Model multiple states in the same energy surface
  - Study transitions in equilibrium
- Formalism is system-independent
  - Readily extrapolated to biomolecular systems with two or more known states

Future Directions

- Parameterize states from all-atom data using clustering algorithms
  - Study substates and state fluctuations
- Guide all-atom simulations
  - Generate reasonable transition paths quickly using Gō-like potentials
- Study the effects of "mutations" on equilibrium
  - Remove or scale interaction between specific pairs of amino acids
- Extend analysis to other GPCRs
  - Compare activation mechanism and pathway heterogeneity

Grossfield Laboratory
BECAS COMPLEMENTO BECAS SEP
AAUW empowering women since 1881

Data analysis was performed using LOOS (Lightweight Object-Oriented Structure library), an open source C++ and Python library for MD analysis developed by the Grossfield lab:
http://loos.sourceforge.net
https://github.com/GrossfieldLab/loos