Characterization of Rhodopsin's Activation Mechanism Using Multi-Basin Structure-Based Models

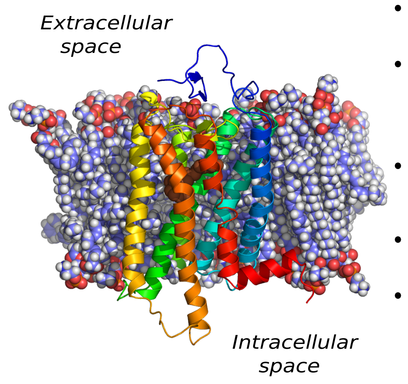
Letty Salas-Estrada[1], Tod D. Romo[1], Nicholas Leioatts[2], Alan Grossfield[1]
[1]University of Rochester Medical School, Rochester, NY, USA
[2]Max Planck Institute for Biophysical Chemistry, Göttingen, Germany

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

Abstract

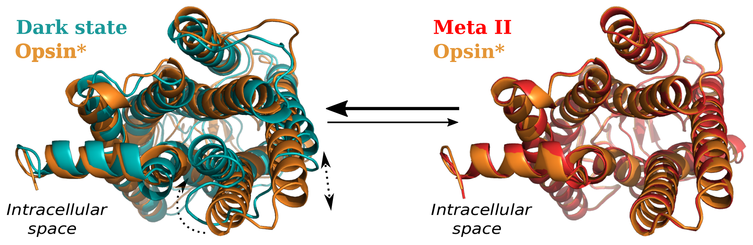
G protein-coupled receptors (GPCRs) are integral membrane proteins that can transduce extrinsic information across lipid bilayers. This allosteric process, which involves large conformational changes upon receptor activation, is not yet well understood. Crystal structures of the inactive and active conformations of several GPCRs showing these activation-induced structural differences have been successfully solved in the last decade. However, to fully understand the molecular basis of GPCR activation and deactivation, it is also necessary to elucidate the nature of the pathway or pathways linking these conformations. To date, the intrinsic timescales of such processes still pose a substantial difficulty for capturing state transitions using unbiased molecular dynamics simulations with standard force fields at high molecular resolutions. Here, we use simpler structure-based (Gō-like) potentials to model the energy landscape of the GPCR rhodopsin. This approach allows for extensive sampling of the receptor's transition pathways with relatively inexpensive all-heavy atom simulations. We also incorporate two potential mixing strategies that facilitate the inclusion of multiple protein state and the study of states interconversion in equilibrium.

G Protein-Coupled Receptors



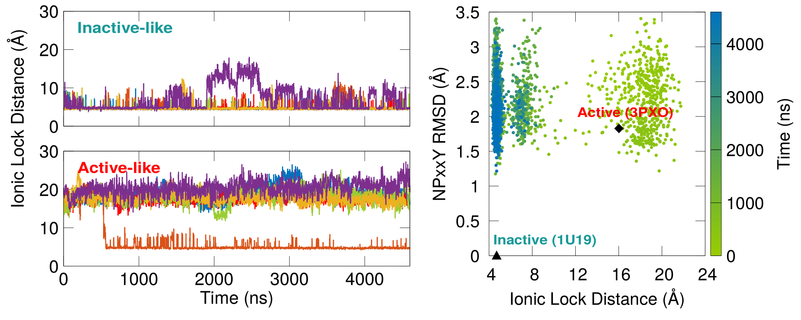
- Ubiquitous 7 transmembrane (7TM) α-helical proteins
- Transduce information across lipid bilayers in response to stimuli
  - Ligand binding/isomerization
  - Environmental cues: light, ions, pH, etc.
- 825 distinct members in humans
  - Highly conserved topology
  - Functional diversity and specificity
- Target of ~40% small-molecule drugs
- Conformational changes at intracellular side involved in activation
  - Allow G protein binding
  - Mediate downstream signal cascade

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

All-Atom Opsin Simulations

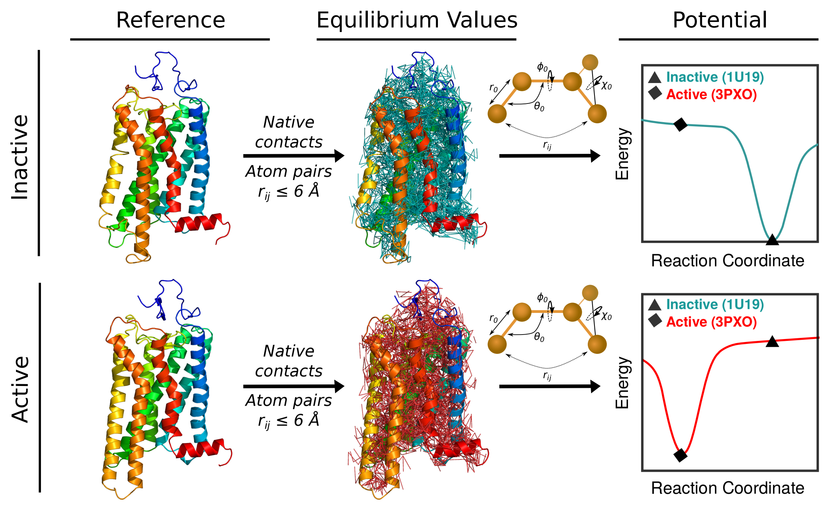


- Inactive-like and active-like **Opsin**
  - Starting structures:
    - Inactive-like: 1U19 (apo)
    - Active-like: 3CAP
  - 6 replicates per ensemble
  - Aggregate simulation time: ~55.2 μs
  - Standard MD is good at capturing fluctuations within a state (Å/ps resolution)
  - Less effective at sampling transitions
- Transitions are rarely observed
  - Activation is in the millisecond regime
  - Inactive-like: 2 partial activations
  - Active-like: 1 deactivation
    - Projected onto ionic lock/NPxxY RMSD
- Ionic lock: R135(N)-E247(O) distance
  - Increases 10-12 Å upon activation
- NPxxY RMSD to inactive:
  - Increases 2-4 Å upon activation

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

Gō-Like Models Can Capture GPCR State Transitions

Gō-Like Models Produce Funnel-Like Potentials



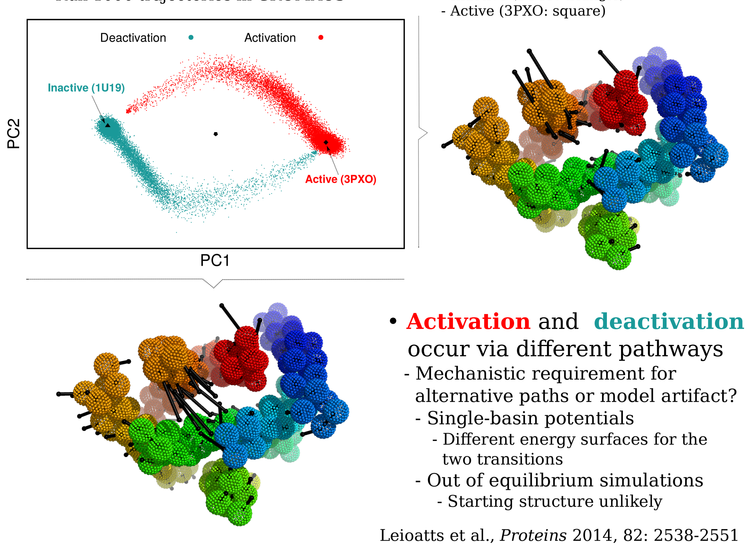
- Transitions unattainable for simulations with standard force fields
  - Requirements for studying state transitions:
    (1) Frequent state interconversion
    (2) Reasonable computational cost for extensive sampling
    (3) High structural resolution
    (4) Allow multiple states and substates
- Structure-based Gō-like models:
  - Reference structure is the global minimum of the potential (native structure)
    - Equilibrium values and native contacts derived from reference structure
  - Conventional bonded interactions
  - Non-bonded contributions only attractive between native contacts
  - SMOG: All-heavy atom implementation

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

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

Opsin Transition Pathways from Single-Basin Gō-Like Potentials

- **Active** and **inactive** potentials generated with SMOG
  - Study inactive-to-active and active-to-inactive transitions separately
  - **Activation** (**deactivation**):
    - Use **active** (**inactive**) potential
    - Start the system from **inactive** (**active**) structure
    - Run 1000 trajectories in GROMACS
- Identify most concerted atomic motions in transitions
  - Principal component analysis on TMC_αs of aggregate data set
  - Each dot is a trajectory frame of a transition in PC-space
  - Crystal structures:
    - Inactive (1U19: triangle)
    - Active (3PXO: square)

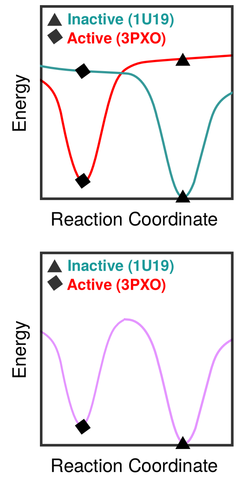


- **Activation** and **deactivation** occur via different pathways
  - Mechanistic requirement for alternative paths or model artifact?
  - Single-basin potentials
    - Different energy surfaces for the two transitions
  - Out of equilibrium simulations
    - Starting structure unlikely

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

Transitions in Equilibrium Using Multi-Basin Potentials

Dual-Basin Potentials



- Mixed potential should preserve location and shape of each basin
- Transition barrier between basins should be continuous and smooth

Boltzmann-Weighting

$$V(\mathbf{R}) = -\frac{1}{\beta_{mix}} \ln \left\{ \exp(-\beta_{mix} V_1(\mathbf{R}) + \varepsilon_1) + \exp(-\beta_{mix} V_2(\mathbf{R}) + \varepsilon_2) \right\}$$

- Exponential average of potentials every energy evaluation
  - β_{mix}: Contribution of each potential
  - ε_i: Relative stability of ith state
  - Each individual potential dominates mixed potential around corresponding basin
- 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

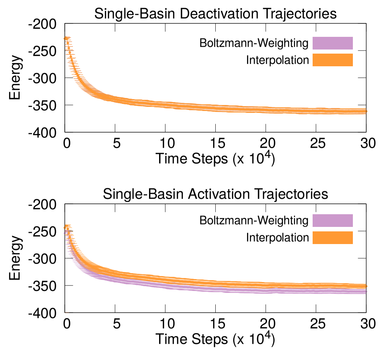
Interpolation

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

- Computationally convenient function couples potentials
  - Δ: Modulates barrier height
  - ΔV_i: Relative stability of ith state
  - Native basins are still preserved
- Barrier heights can be tuned individually
- n-basin generalization requires solving nxn equation

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

Validation

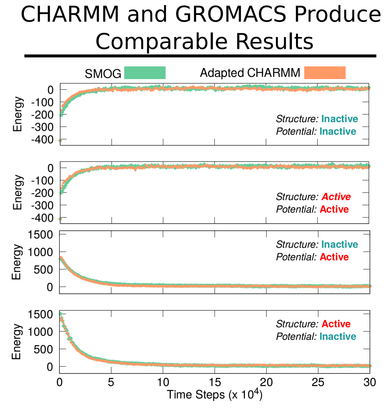


- Energy relaxation curves from single-basin trajectories
  - Test interpolation implementation in CHARMM
  - Energies evaluated every time step using Boltzmann-weighting and interpolation
  - Average of 100 simulations
  - Similar results for two basins

CHARMM Implementation

- SMOG Hamiltonian:
  - CHARMM's functional form adapted using BLOCK and GOPAIR facilities
- Bonds:
  - SHAKE requires matching bond lengths among sets of potentials
    - Averaged each over all potentials
- Repulsive VDW:
  - r_{ij} varies among reference structures
    - Large and unphysical repulsive VDW
  - Make repulsive interactions identical
    - Smallest reference distance used
  - Attractive VDW defines each basin
  - Modified potential still preserves native basins
- Interpolation:
  - Implemented via ENSEMBLE module

CHARMM and GROMACS Produce Comparable Results

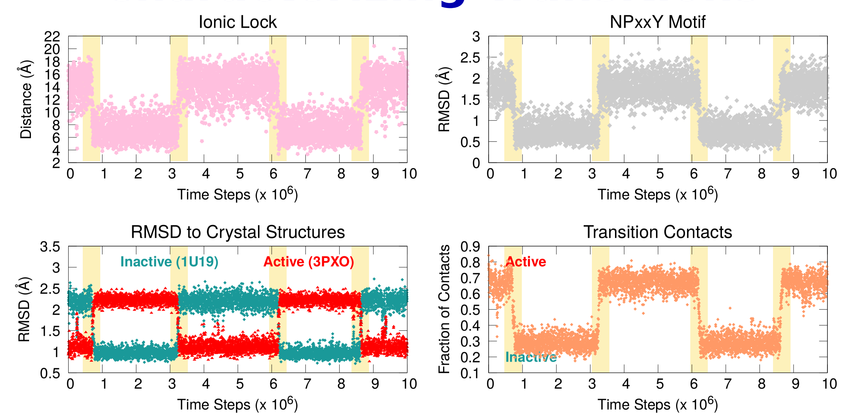


Simulation Details

Starting Structure	PDB Accession Codes	Potential	Simulations (Time Steps)
Inactive	1U19 (apo)	Dual-basin using interpolation	12 runs x 1*10^7
Active	3PXO (apo)		12 runs x 1*10^7
			Total ~ 10^8

- System size: ~2,600 atoms
  - Protein stable in the absence of a lipid bilayer and water
    - Implicitly accounted for
- Force field: SMOG1/2
  - Whitford et al., *Proteins* 2009, 75: 430-441
  - Parameters ported to CHARMM
- Ensemble: NVT
  - T = 50 K in reduced units
- VDW cutoff: 10 Å
- No electrostatics
- Langevin dynamics:
  - Collision frequency = 0.1 ps^{-1}
- Time step: 2 fs
  - SHAKE for all bonds
- Software:
  - CHARMM C39b1 on Linux Cluster

Characterizing Transitions



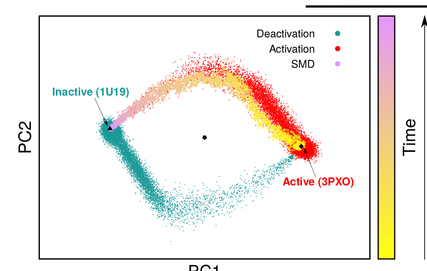
- Double-basin interpolation trajectory
  - Starting from **active** structure
- Biologically motivated reaction coordinates capture states
  - Transition in NPxxY motif still small
- Transitions take the system within 1 Å of crystal structures
- Contact-based rxn coordinate:
  - Set of contacts derived 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)

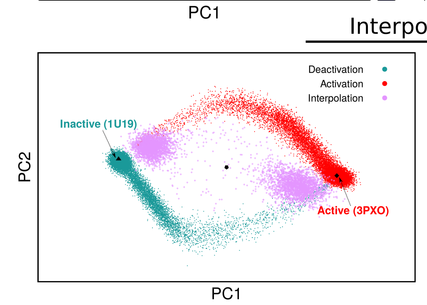
Opsin Transition Pathways from Dual-Basin Potentials

Boltzmann-Weighting



- Steered molecular dynamics used to drive transition
  - System started from **active** structure
    - RMSD constraints on TMC_αs
    - Force constant: 2,000 kcal/mol•$Å^2$
    - Steps: 200,000
  - Trajectory projected onto PC1 and PC2 of single-basin runs
  - Transition follows single-basin paths, but in opposite direction

Interpolation



- Transitions seem to occur through different paths
  - System started from **active** structure
  - Trajectory projected onto PC1 and PC2 of single-basin runs
    - Pathways resemble single-basin results, but do not exactly overlap
  - Basins are systematically off from crystal structures
    - Offset is small: RMSDs to crystals ~1 Å
    - Interpolation parameters chosen to favor transitions might be at fault

Conclusions

- Gō-like models can capture state transitions in macromolecules
  - Study state interconversion at negligible computational cost
  - Retain all-heavy atom resolution
  - Model multiple states in the same energy surface
- Transitions can be characterized using receptor motifs or contact-based reaction coordinates
- Gō-like models can be combined with enhanced sampling methods
  - Replica exchange could improve convergence especially with many states

Future Directions

- Parametrize states from all-atom data using structural clustering:
  - Study substates and state fluctuations
- Guide all-atom simulations:
  - Generate reasonable reaction coordinates quickly using Gō-like potentials
- Characterize effects of ligand and analog molecules on equilibrium
  - SMOG2 supports ligand parametrization
- Extend analysis to other GPCRs
  - Compare activation mechanism and pathway heterogeneity

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

Becas Complemento Becas SEP
AAUW empowering women since 1881