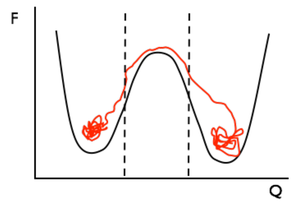
Path Sampling for Unbinding Kinetics

David W.H. Swenson, van 't Hoff Institute for Molecular Sciences, University of Amsterdam

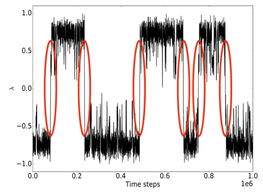
This work has received funding from the European Union's Horizon 2020 research and innovation program under the grant agreement No. 676531 (project E-CAM)

Rare Events



When metastable states are separated by a large free energy barrier, the transition between them is a **rare event**. Properties such as the mechanism and rate of the transition require significant sampling of the transition region. However, the Boltzmann distribution means that the vast majority of time is spent in the metastable states, not the transition region.

Path Sampling



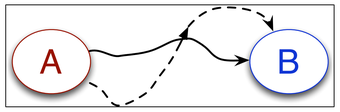
Direct molecular dynamics is, at best, an inefficient way to study a rare event; at worst, intractable. **Path sampling** methods perform a Monte Carlo simulation is the space of paths (trajectories), enabling efficient simulation of rare events by focusing the simulation effort on the transition region and reducing the simulation inside the metastable states.

OpenPathSampling

OpenPathSampling is a recently-developed software package for studying rare events. It includes many tools for performing and analyzing path sampling simulations, as well as tools from other trajectory-based approaches to rare events.

http://openpathsampling.org
http://github.com/openpathsampling/openpathsampling
Twitter: @pathsampling

Transition Path Sampling



TPS is a path sampling approach that uses trajectories that connect the metastable states. While some TPS-based approaches can extract kinetic information such as rates, TPS is best-suited for studying the mechanisms of complex systems.

Many tools for analyzing mechanisms are based on initial sampling using a variant of TPS.

OpenPathSampling supports several variants of TPS:

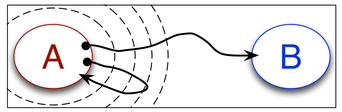
- fixed length
- flexible length
- multiple state

OPS supports several variants of the shooting move (for both TPS and TIS):

- one-way shooting
- two-way shooting
- uniform or biased shooting point selection

Dellago et al. J. Chem. Phys. **108** 1964 (1998).

Transition Interface Sampling



TIS is a path sampling method that is specifically designed to calculate rates. In TIS, the space is foliated by several "interfaces" along an approximate reaction coordinate. The rate is determined based on the successive crossing probabilities of ensembles defined by the interfaces:

$$k_{AB} = \phi_{A_0} \prod_{i=0}^{m-1} P_A(\lambda_{i+1}|\lambda_i)\ P_A(B|\lambda_m)$$

flux — crossing prob. — conditional transition prob.

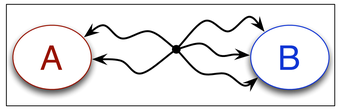
TIS is forgiving with respect to the proposed order parameter. In addition, path from TIS can be reweighted to obtain free energy landscapes in any collective variables of interest, without re-running the sampling.

OpenPathSampling supports several variants of TIS:

- replica exchange
- multiple state
- multiple interface set
- single replica

van Erp, Moroni, Bolhuis. J. Chem. Phys. **118** 7762 (2003).

Committor Analysis

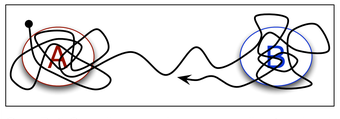


The committor for a given configuration is the probability that a trajectory launched from there (with Boltzmann-distributed velocities) will land in the "product" state. It is an excellent approximation for the reaction coordinate.

OpenPathSampling includes tools to efficiently calculate the committor. Using OPS's monitoring function, the OPS committor simulation ensures that the trajectory stops when a trajectory enters a state, thus avoiding wasted computing effort.

External contributors are building modules based on the committor simulation to add calculations such as reactive flux (Bennet-Chandler) calculation of the rate.

Direct Simulation



OpenPathSampling can also manage direct molecular dynamics simulations. This is particularly useful for using OPS calculate the flux (or, in some cases, the rate) for a process, since the OPS state definitions can be re-used across simulations and are uniquely identified (tracking provenance). The direct simulation module has the option of not saving snapshots, which can be useful for very long simulations.

Analysis tools

In addition to sampling rare events, OPS includes many tools for analyzing rare events, including path density plots, analysis of fluxes, crossing probabilities, and rates from TIS, analysis of replica exchange behavior (trip times, replica flow, mixing matrix) for replica exchange TIS.

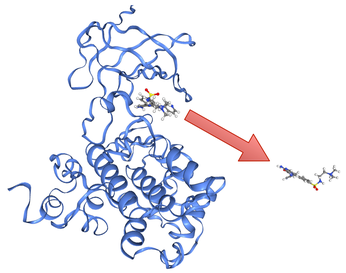
It also includes tools to facilitate creation of new state definitions based on user-annotated trajectories, and tools to identify trajectories that follow proposed mechanisms or classify trajectories according to a specified reaction channel. It is easy to add more analysis tools, and external contributors have already created some, such as reaction coordinate analysis using maximum likelihood analysis.

OPS interfaces with other codes

OpenPathSampling is designed in interface with existing software packages and libraries, in order to provide familiar functionality to users. This includes general scientific software as well as tools for specifically for molecular dynamics.

jupyter · pandas · matplotlib · OpenMM · MDTRAJ · MSM BUILDER · PyEMMA · GROMACS FAST, FLEXIBLE · PLUMED · COMING SOON

Pilot Project: Unbinding of a selective inhibitor from GSK3B



- Potential drug for Alzheimer's Disease
- Selective, reversible inhibitors are excellent drug molecules

Residence times:

- Predictor of in vivo efficacy
- Difficult to capture experimentally
- Long timescale require advanced simulation techniques

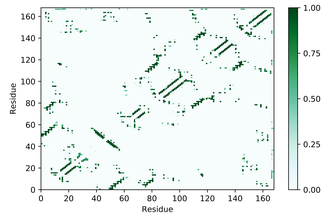
Inhibitor: Berg et al. J. Med. Chem. **55** 9107 (2012).

Steps in Path Sampling:

1. Identify stable states
2. Obtain initial trajectory
3. Run the simulation
4. Analyze

Ligand can attach to decoy (non-productive binding) sites for moderately long timescales (~50-100ns). Can we enumerate the states?

Contact maps



One reasonable approach to identify potential decoy states in this system is by considering the contacts of the ligand with with the protein. This is a general way to characterize the metastable states, and is an approach that can be useful in the analysis of other systems as well: for example, the metastable states of protein folding could be characterized by the contacts that are present.

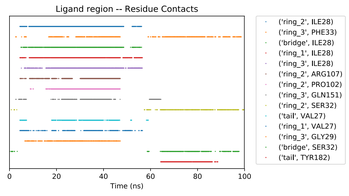
The figure shows a residue-residue contact map (for a different system); the darker the color, the larger the fraction of the trajectory when that contact was present.

The **Contact Maps** module includes code for:

- residue-residue and atom-atom contact matrices
- contact matrices restricted to certain atoms/residues
- finding most common contacts in the trajectory (overall) for for a specific residue
- comparing two contact matrices
- plotting contact matrices

Source: https://github.com/dwhswenson/contact_map
Documentation: http://contact-map.readthedocs.io/
Status: Released (v0.2.0)

Binding Event Analysis



Contact maps show time-averaged behavior. In order to identify metastable states, we need to explicitly look at how things evolve over time. The example here shows why: one set of contacts is present in the first half of the trajectory, another in the second. This indicates a transition between two stable states has occurred.

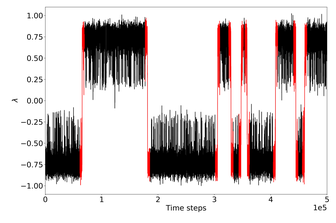
Analysis such as the concurrence plot at the left focuses on the time dependence of the contacts, and highlights which contacts are simultaneously present. This is essential to identify stable states based on MD simulations of a ligand binding process.

The **Binding Event Analysis** module includes:

- concurrence plots and associated data structures
- finding trajectory segments when a ligand is bound
- specific needs for ligand binding contacts that aren't relevant for general contact maps

Source: https://github.com/dwhswenson/binding_md
https://github.com/dwhswenson/contact_map
Status: In development/Testing

Binding Event Sampling



In practice, we found that the number of potential decoy sites was too large to characterize them all. In order to ensure that we could capture all the relevant states, we are developing a new approach which incorporates time into the stable state definition. This approach will be based on specific contacts being stable over some period of time.

This can be seen as increasing the range of time that path sampling focuses on, by adding an extra window to the trajectory when it hits an unknown, potentially stable, state.

The **Binding Event Sampling** module includes:

- OPS-compatible state definition for the stable contact state
- OPS Ensemble and Network objects using this definition, enabling use of OPS's path sampling tools
- Improvements to other code to make the method efficient

Source: https://github.com/dwhswenson/binding_md
https://github.com/dwhswenson/contact_map
Status: In development

E-CAM WP1 modules for path sampling and binding kinetics:

Path Sampling Simulation Methods

- Two-way shooting (D1.2)
- Improved input for OPS networks (D1.2)
- *Binding event path sampling tools*
- *Aimless shooting*
- Shifting mover
- Spring shooting
- *Shooting range algorithm*
- *Web throwing move*
- *PPTIS*

Interfacing OPS with Other Software

- OPSPiggybacker (D1.2)
- *Gromacs support*
- *PLUMED integration*

Analysis Tools for Path Sampling

- Path density (D1.2)
- New WHAM code for OPS (D1.2)
- Annotated trajectories (D1.2)
- Channel analysis (D1.3)
- Resampling statistics (D1.3)
- New TIS analysis framework (D1.3)
- *Binding event analysis tools*
- Maximum likelihood for reaction coordinates (D1.3)

Miscellaneous Modules

- Flux/rate from existing trajectories (D1.2)
- OPS snapshot features (D1.2)
- Contact maps
- Interface optimization (D1.3)

Other Rare Event Simulation Methods

- Direct (on-the-fly) flux and rate calculation (D1.2)
- Committor analysis (D1.3)
- Reactive flux (D1.3)
- Transition state ensemble (D1.3)
- *S-Shooting* (D1.3)

Legend

Developed by: PDRA, ESDW Traunkirchen, ESDW Leiden

Italics indicate modules that are still in development
(Parentheses indicate deliverable for the module)