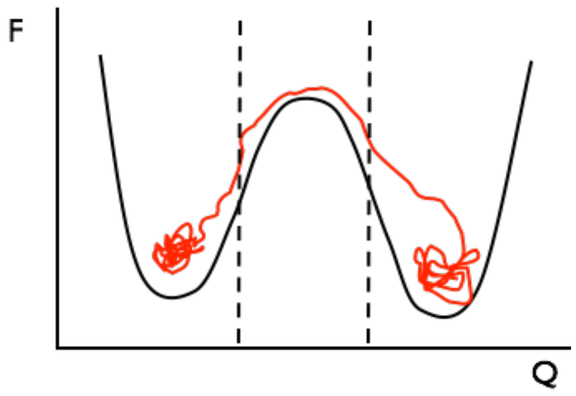
Path Sampling: Software and Practical Applications

David W.H. Swenson

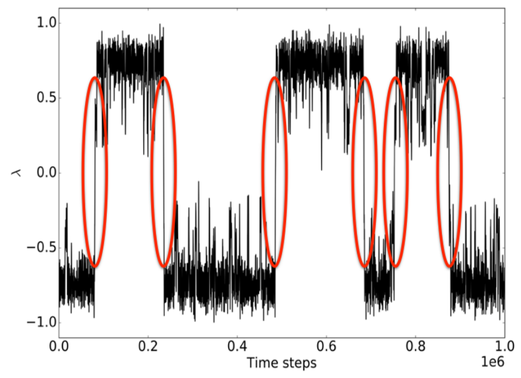
This work has received funding from the European Union's Horizon 2020 research and innovation program under the grant agreement No. 676531 (project E-CAM)

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.

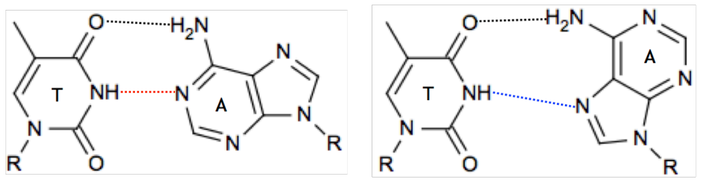
OpenPathSampling

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

Watson-Crick to Hoogsteen Conversion in DNA

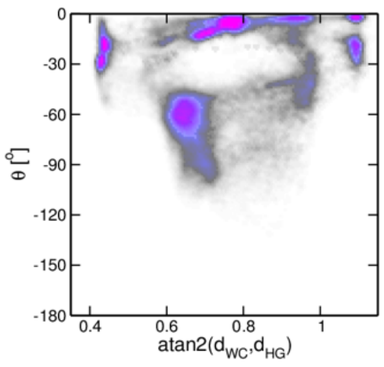
In addition to the well-known Watson-Crick (WC) motif, DNA at physiological conditions can be found in the Hoogsteen (HG) motif, where the purine is flipped 180° relative to Watson-Crick. This motif plays a role in some DNA replication processes, and may be important in many other biological processes.



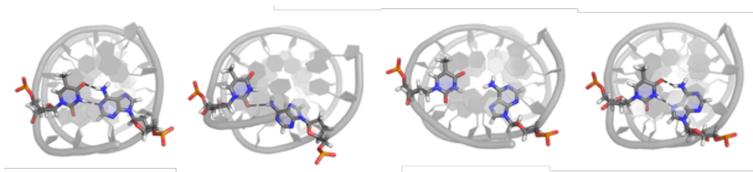
Watson-Crick Hoogsteen

In TPS simulations, we observe multiple mechanisms (channels) for this transition. In particular, we distinguish between an "inside" mechanism, where the purine rotates while inside the double helix structure, and an "outside" mechanism, where the purine exits the helical structure (base flipping) before rotating and re-entering the helix. Previous work had emphasized the importance of the inside transition, but we see evidence that the outside transition is preferred.

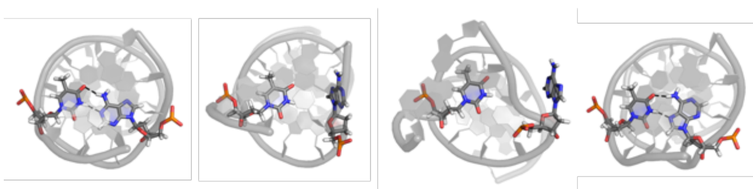
The characteristic hydrogen bond distances d_{WC} and d_{HG} are illustrated in the figure. We use atan2(d_{WC}, d_{HG}) as an approximate reaction coordinate (order parameter) for TIS.



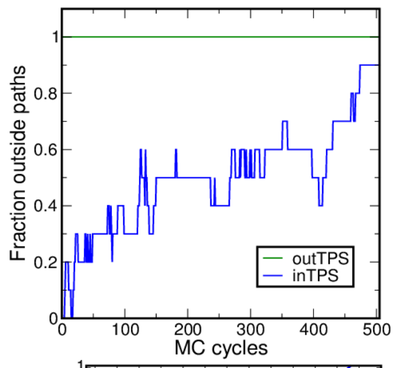
Path density plot with density in both the inside transition (base opening angle (where $\theta > -32°$) and the outside transition ($\theta < -32°$). WC state is at the left; HG at the right.



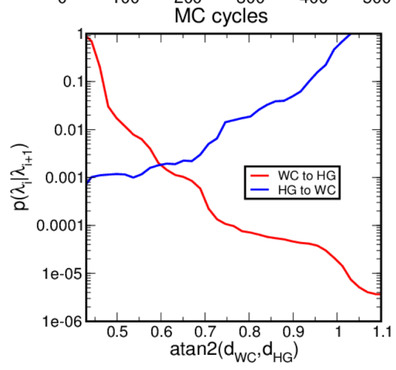
Representative snapshots for inside transition



Representative snapshots for outside transition



Path sampling naturally switches to the outside path: 10 simulations starting in outside (outTPS) all stay in outside; 9/10 simulations starting in inside (inTPS) have switched to outside after 500 MC steps.



Total crossing probability for the outside transition, in both directions, along the atan2(d_{WC}, d_{HG}) order parameter, as calculated by TIS and used to determine the rates.

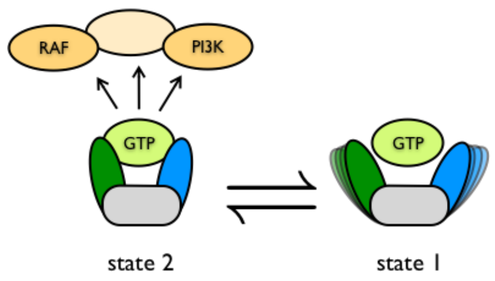
	expt	TIS
$k_{WC \to HG}$ (s^{-1})	14.2 ± 1.03	742
$k_{HG \to WC}$ (s^{-1})	3670 ± 200	$1.6 \cdot 10^5$
ΔG ($k_B T$)	5.5	5.4

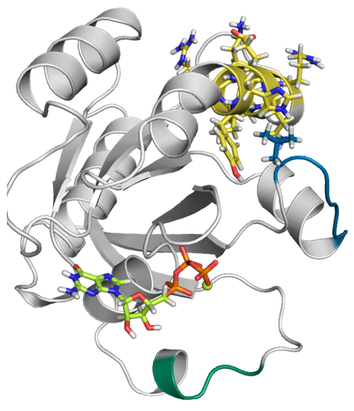
Comparing rates/free energy differences from TIS and from experiment. Rates are overestimated, but free energies are very good.

Vreede, Bolhuis, DWHS. In prep.

Differences in the dynamics of the oncogenic K-Ras mutant Q61L

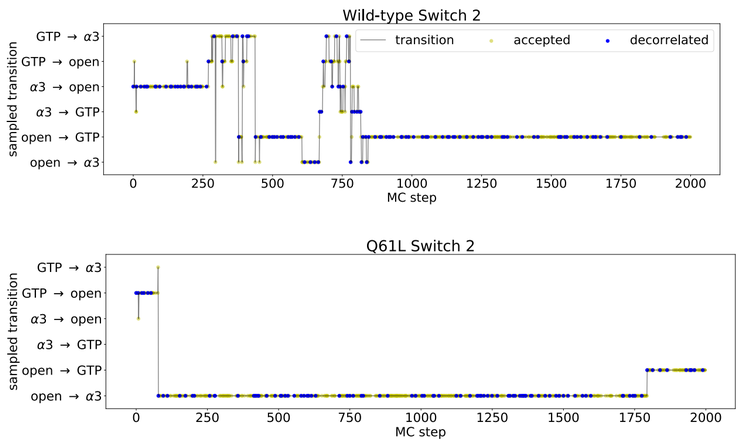
Mutations in the Ras protein family play a role in 15% of all cancers. The GTPase K-Ras is a member of this family, and 85% of oncogenic Ras mutations are found in K-Ras. Experiments suggest that GTP-bound K-Ras exists in two substates: an inactive state 1, in which two loop regions are flexible, and an active (and less flexible) state 2. One of the oncogenic mutations, Q61L, involves a mutation in the switch 2 region. This work asks the question: **How does a mutation (such as Q61L in K-Ras) change the fundamental dynamics of a protein in a way that can lead to cancer?**



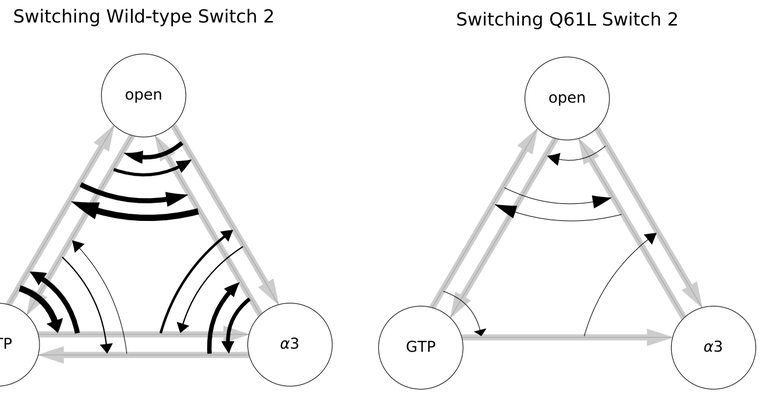


The K-Ras mutant Q61L, with the switch 1, switch 2, and alpha helix 3 regions highlighted. The residue 61L is drawn in licorice, as is α3 and the GTP.

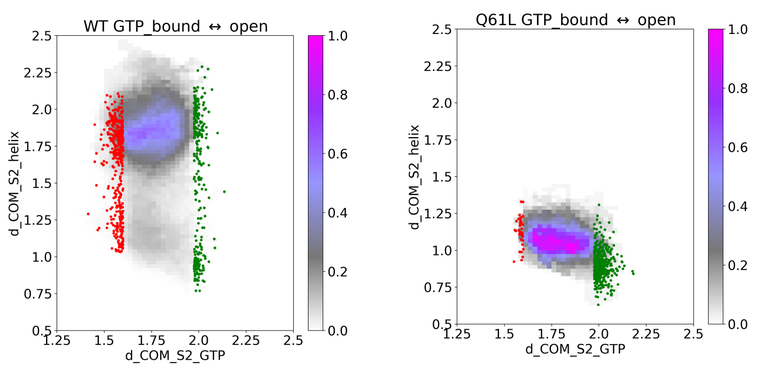
Path sampling identified 3 metastable states for switch 2 in GTP-bound K-Ras: with the switch bound to the GTP, bound to the α3, or unbound.



Transition switching in the TPS simulations of wild type and Q61L K-Ras. The three substates lead to six possible transitions. In a multiple state TPS simulation, the transition being sampled can change as a function of the Monte Carlo step. Yellow dots mark accepted trajectories; blue dots mark successive decorrelated trajectories (no frames in common.)



Another view of transition switching in the TPS simulations. Circles indicate the metastable substates, grey arrows indicate the sampled transitions, and curved black arrows indicate the number of switches between transitions. The width of the curved black arrows is proportional to the number of switches. There is much more switching in the wild type than in the mutant.



Path density plots for the wild type and Q61L mutant, for paths making the transition from GTP-bound to unbound (open) or vice versa. The horizontal axis measures the distance between switch 2 and GTP, which separates the states (bound shown in red dots to the left; unbound in green dots to the right). The vertical axis measures the distance from switch 2 to the α3 helix. In the wild type, two channels are sampled: one with a low value of switch 2-α3 distance, and one with a high value. Only the low value is sampled in the mutant. The low distance to α3 corresponds to sliding along the helix to unbind; the high distance corresponds to direct solvation of the switch region.

The Q61L mutation makes direct solvation unlikely, leading to more time in the closed, active "state 2," causing abnormal cell growth and proliferation.

Roet, Hooft, Bolhuis, DWHS, Vreede. In prep.