

***Milestoning: Use of short trajectories to
compute long time kinetics and
thermodynamics***

Anthony West, Peter Majek, Serdal Kirmizialtin, Alfredo Cardenas,
Steven Kreuzer, and RE
Department of Chemistry and Biochemistry and
Institute of Computational Sciences and Engineering (ICES)
University of Texas at Austin

\$ NIH

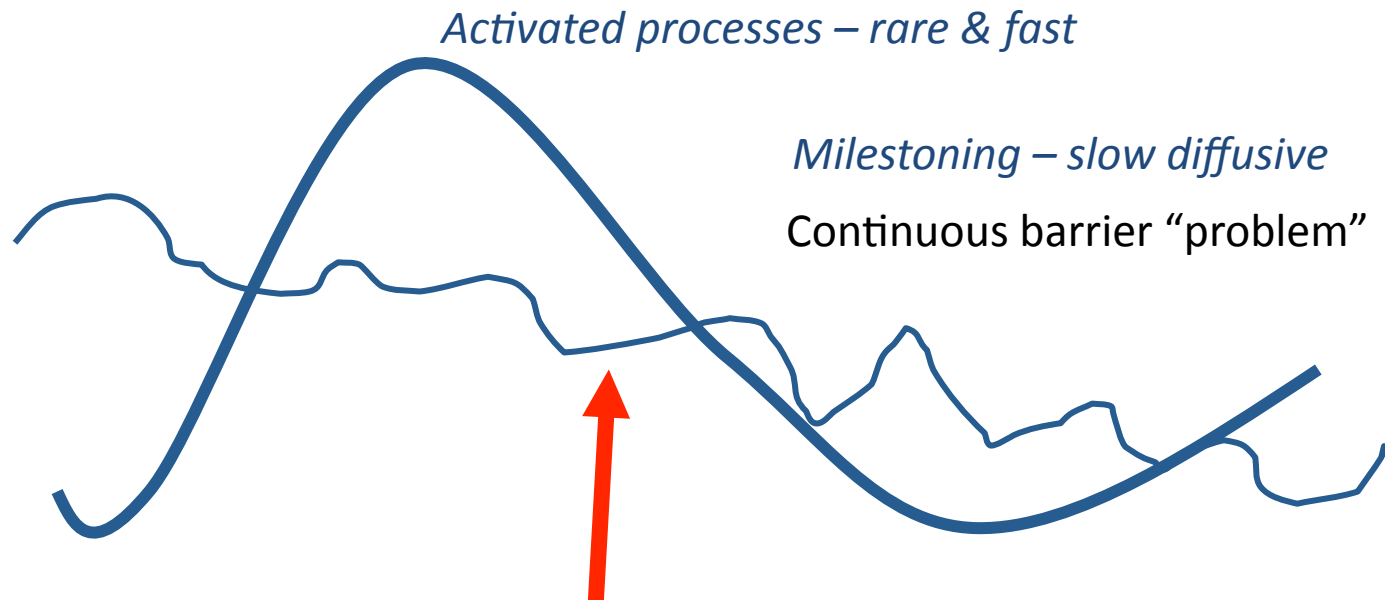
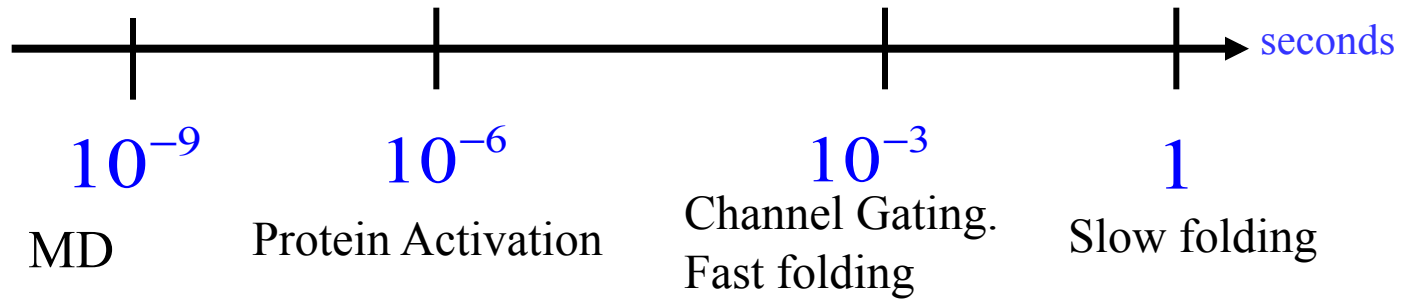
Dresden, March, 2012

Program

- The problem
- A taste of theory

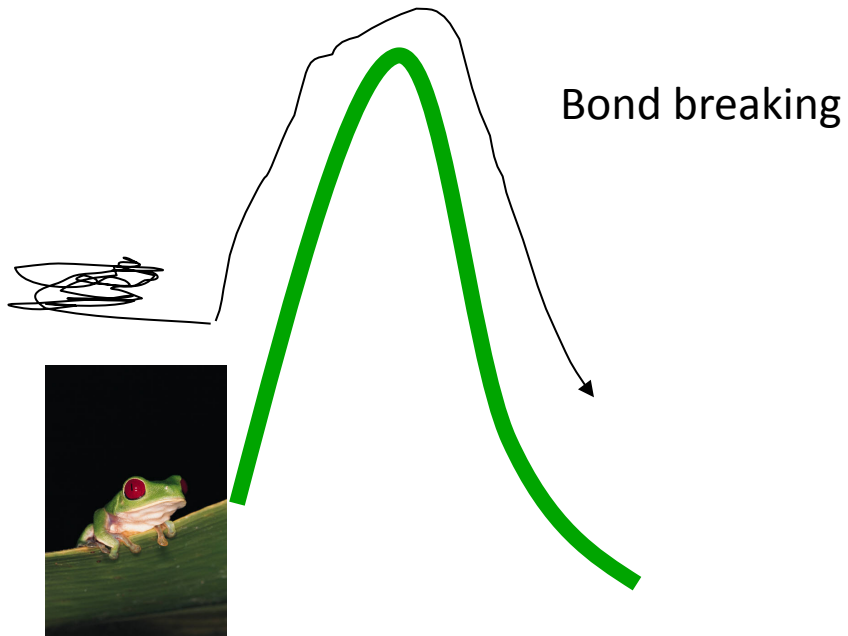
- Helix folding
- Efficiency and selectivity in HIV reverse transcriptase
- Unassisted membrane transport

Long time processes in biophysics: activation or long range diffusion



Long time dynamics

- Rare events (short infrequent trajectories)



TST (Eyring)

Sampling of complete (rare) traj. :

1. TPS (Chandler, Dellago, Bolhuis)
2. TIS (Moroni, Bolhuis, van Erp)
3. FFS (Allen, Frenkel, ten Wolde)
4. WE (Kim, Huber)
5. Hyper-dynamics (Voter)

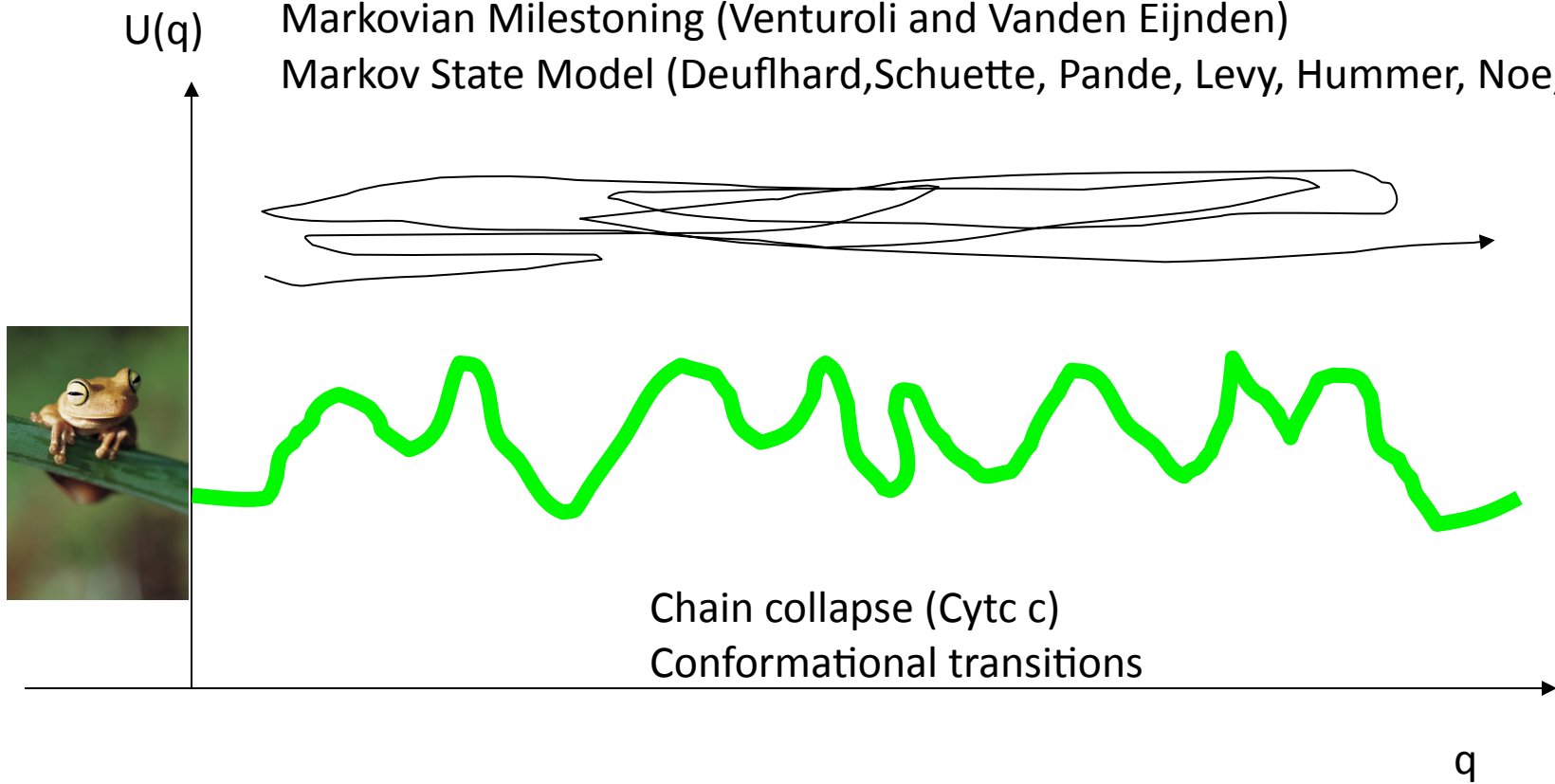
Long time dynamics: Diffusion on rough energy landscape

Milestoning (West, Kirmizialtin, Cardenas, Elber)

PPTIS (Bolhuis, van Erp)

Markovian Milestoning (Venturoli and Vanden Eijnden)

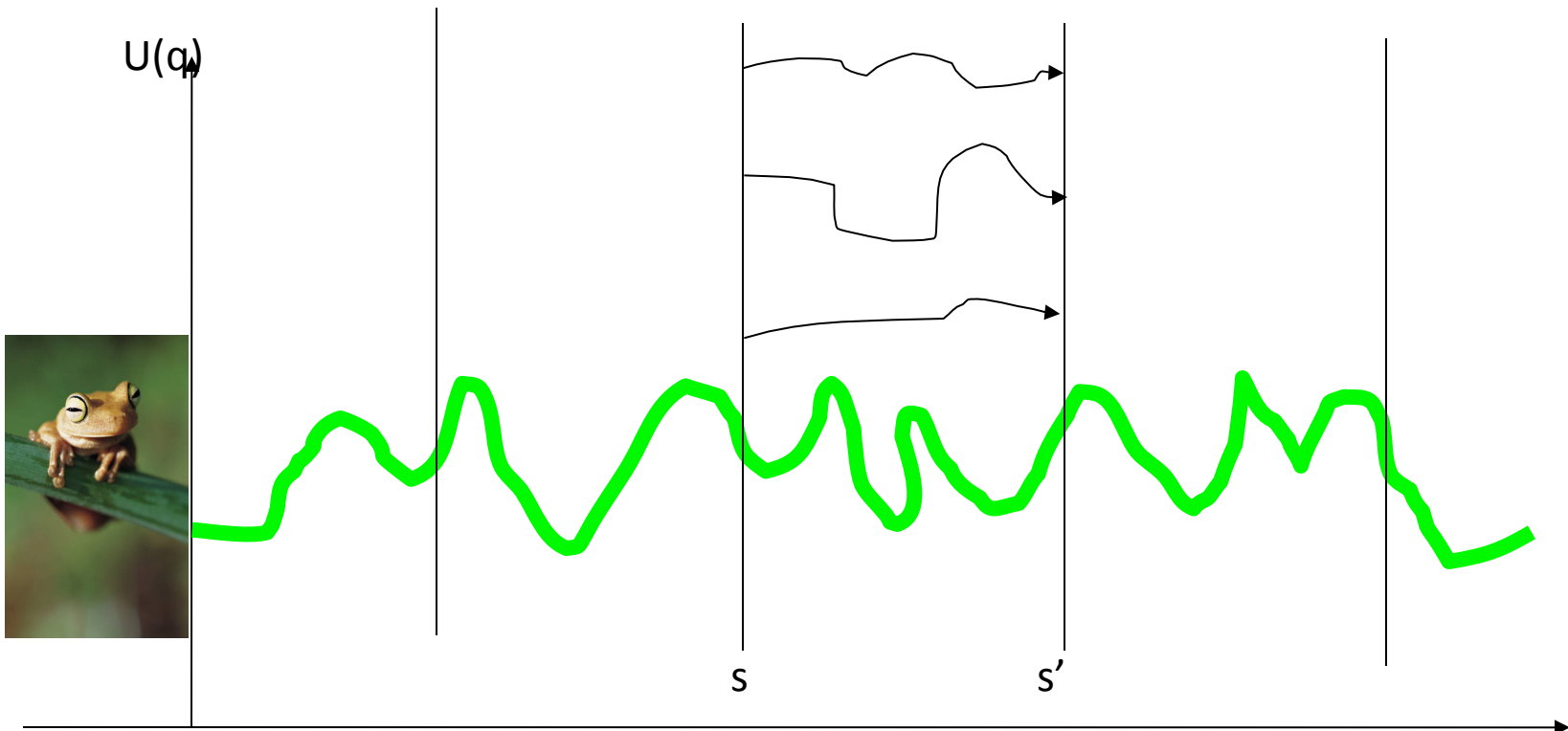
Markov State Model (Deuffhard, Schuette, Pande, Levy, Hummer, Noe,...)



Milestoning:

$$K_{s,s'}(t)$$

The probability density that traj. that hit Milestone s for the first time will reach Milestone s' for the first time exactly at time t



Assume that we know $K_{s,s'}(t)$

How can we calculate the overall time dependence of the system?

With the matrix $K_{s,s'}(\tau)$ determined, compute kinetics

$$Q_s(t) = \eta_s \delta(t - 0^+) + \int_0^t \left[\sum_{s'} Q_{s'}(t') K_{s',s}(t - t') \right] dt'$$

$$P_s(t) = \int_0^t Q_s(t') \left[1 - \int_0^{t-t'} \sum_{s'} [K_{s,s'}(\tau)] d\tau \right] dt'$$

$$\langle t \rangle = \mathbf{P}(0)^t \cdot \left[\mathbf{I} - \int_0^\infty \mathbf{K}(\tau) d\tau \right]^{-1} \cdot \int_0^\infty \tau \mathbf{K}(\tau) d\tau \quad (\mathbf{K}(\tau))_{s,s'} = K_{s,s'}(\tau)$$

$$\mathbf{Q}_{stat} \left(\mathbf{I} - \int_0^\infty \mathbf{K}(\tau) d\tau \right) = 0 \quad \mathbf{P}_{s,stat} = \mathbf{Q}_{s,stat} \bar{t}_s$$

s, s' -- milestones

$Q_s(t)$ -- probability of passing milestone (interface) s at time t

$P_s(t)$ -- probability that the last milestone passed at time t is s

- by direct integration (with West, JCP 2004)
- by Laplace transform and moments of the first passage time (with Shalloway, JCP 2007)
- by trajectory statistics (Vanden Eijnden, JCP 2008)

Equivalent to Generalized Master Equation

- **The generalized Master equation has time dependent rate coefficients**

$$\frac{dP_s(t)}{dt} = \int_0^t \sum_{s'} \left[-R_{s',s}(\tau) P_s(t-\tau) + R_{s,s'}(\tau) P_{s'}(t-\tau) \right] d\tau$$

- **K in the QK formulation is easier to compute than R and the Laplace transforms are related by**

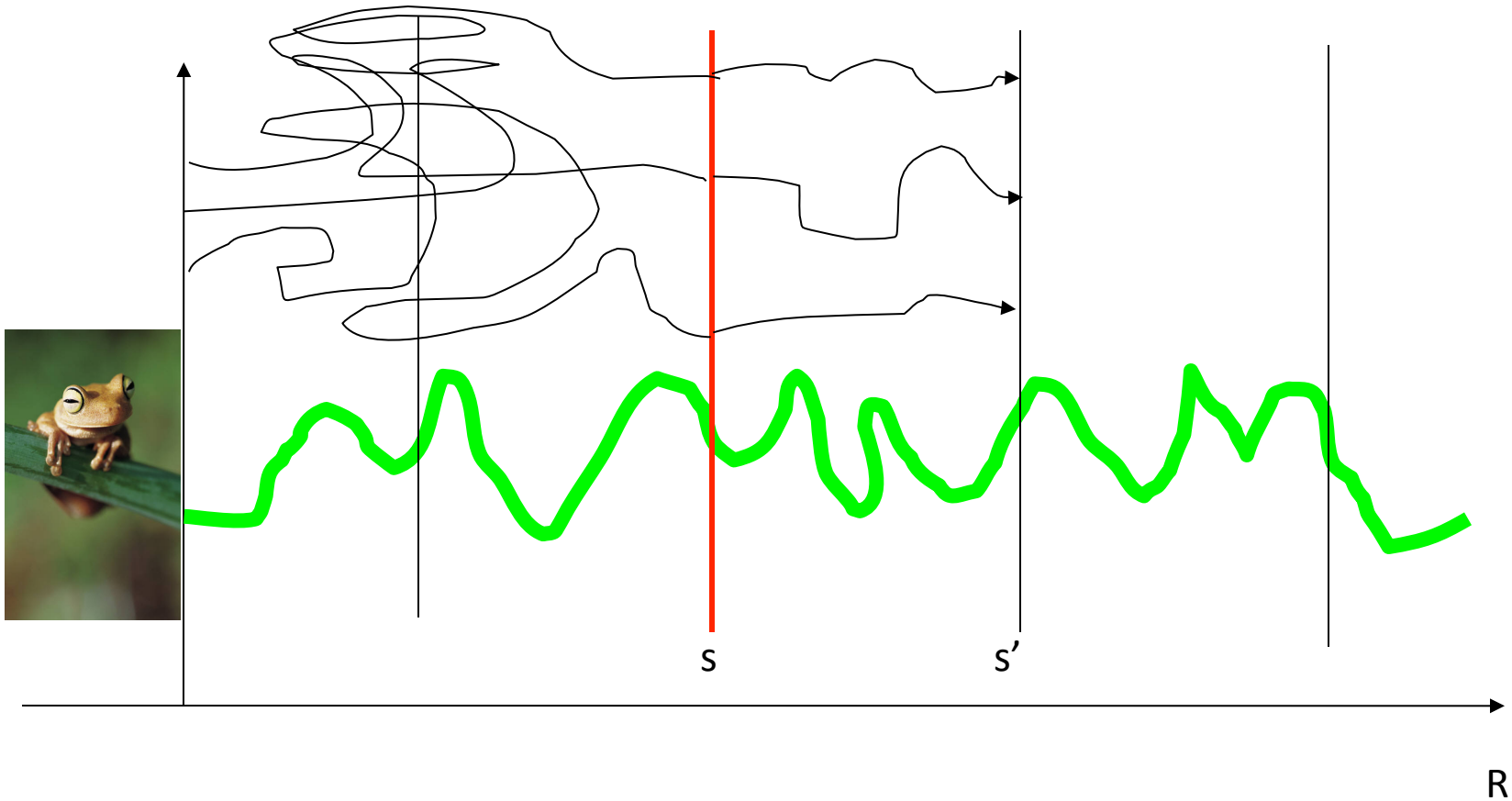
$$\tilde{R}_{s,s'} = \frac{u \tilde{K}_{s,s'}(u)}{\left(1 - \sum_{s'} \tilde{K}_{s',s}(u) \right)}$$

How to compute K?

$$K_{s,s'}(t)$$

How to obtain the “appropriate” initial distribution?

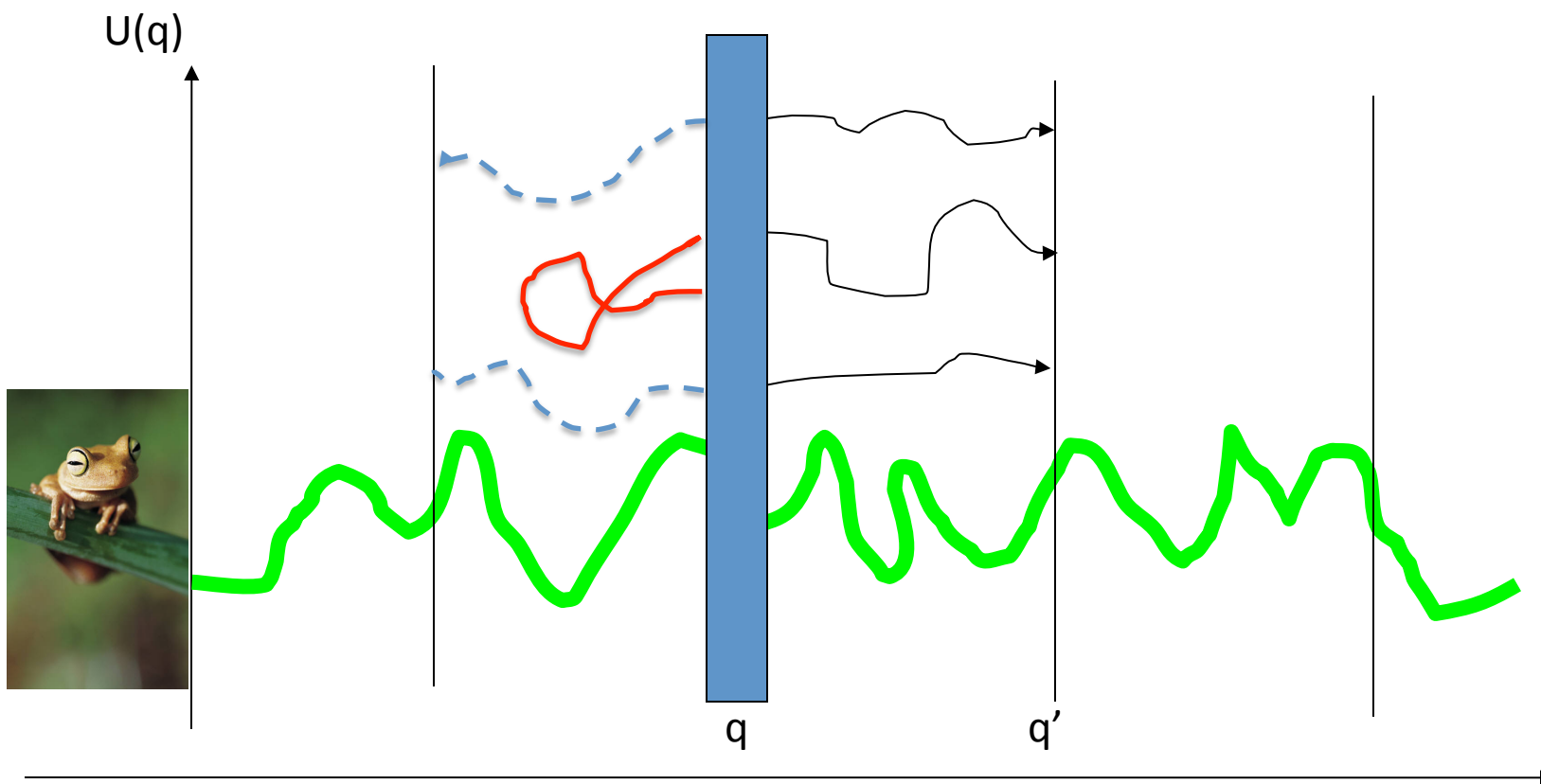
Run exact trajectories to s (TIS, Bolhuis; FFS, Allen, WE, Kim)



Initiate traj. at the Milestone from Equilibrium Check first time hit by running backward

Directional Milestoning: With Majek, JCTC 2010; with Kirmizialtin JPC 2011

$$\sim c(x)\exp(-\beta U(x,q))$$



First hitting point trajectories inside cells with cell as small as $\sim 0.1\text{\AA}$: Especially short $\sim \text{ps}$

Milestoning assumption (more milestones not always better):

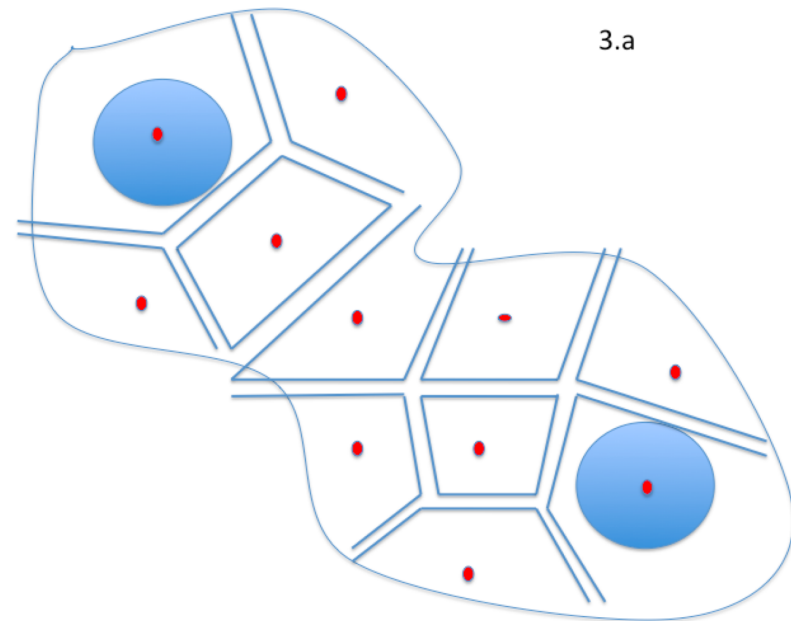
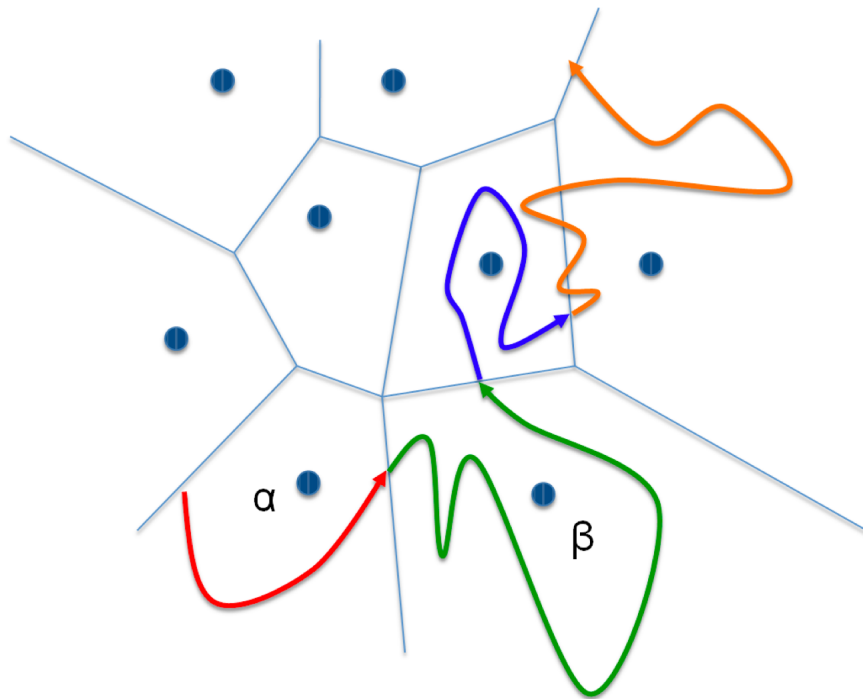
1. Loss of flux correlation between Milestones (interfaces)

$$K_{s,s'}(X_s, X_{s'}, t) \approx K_{s,s'}(X_{s'}, t)$$

2. Good approximation if the Milestones are sufficiently separated to allow loss of correlation between sequential “hits” (West et al., JCP 2007; Majek & Elber JCTC 2010).
3. Exact MFPT if the Milestones are iso-committors (with Vanden Eijnden, 2008).
4. Velocity de-correlation useful measure
5. Taking in and out interfaces for convergence check
6. Comparing sampling at interfaces and terminating distributions
7. Extension to next-next-... interface (Hawk and Makarov JCP 2011)

Directional Milestoning works in high dimension using transitions between interfaces of cells (Majek & Elber, JCTC 2010).

The idea of using Voronoi cell for Milestoning was of Vanden Eijnden and Venturoli, JCP, 2009



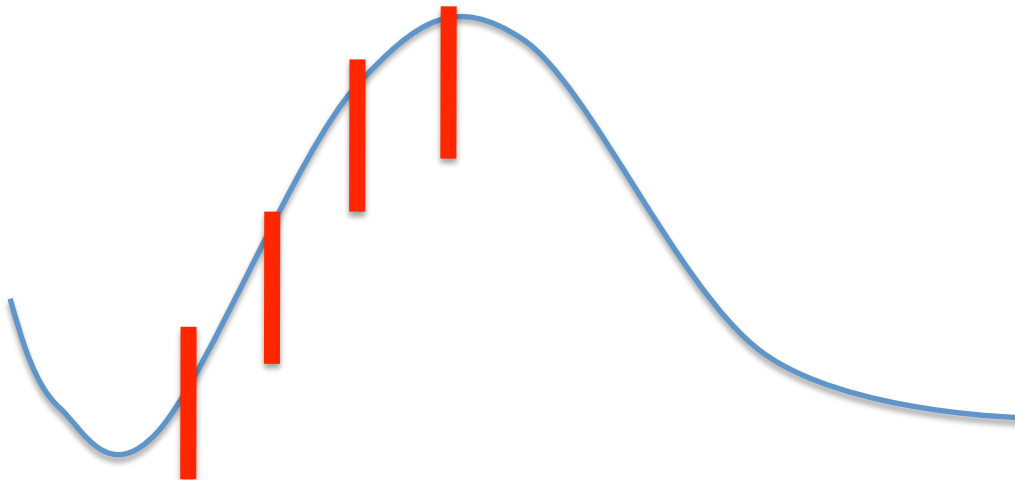
Efficiency (more Milestones the better!)

- *Diffusive speedup:* $t \sim L^2 \rightarrow t \sim M(L/M)^2 = L^2 / M$



- *Parallelization speedup:* $t \sim L^2 / M^2$

- *Exponential bootstrapping at large barrier:* $\frac{1}{qq'} \rightarrow \frac{1}{q} + \frac{1}{q'}$



Speed-up in practice

Microsecond allosteric transition rate predicted for *Scapharca* (in accord with experiment) based on an ensemble of picosecond trajs totaling 10 ns

Results on myosin for the recovery stroke predict *submillisecond* timescale (similar to experiment) using nanosecond simulations

HIV reverse transcriptase *millisecond* nanosecond simulations

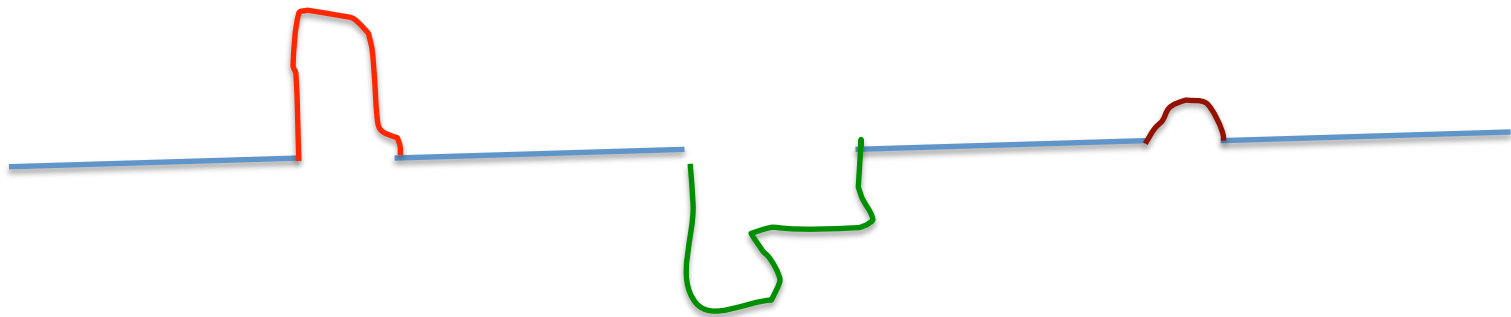
Membrane permeation: *hours* – nanosecond simulations

W-AAA-H (WH5):,

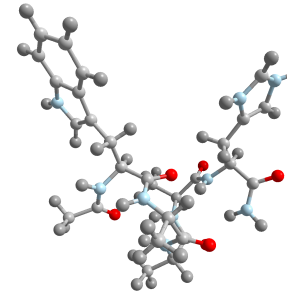
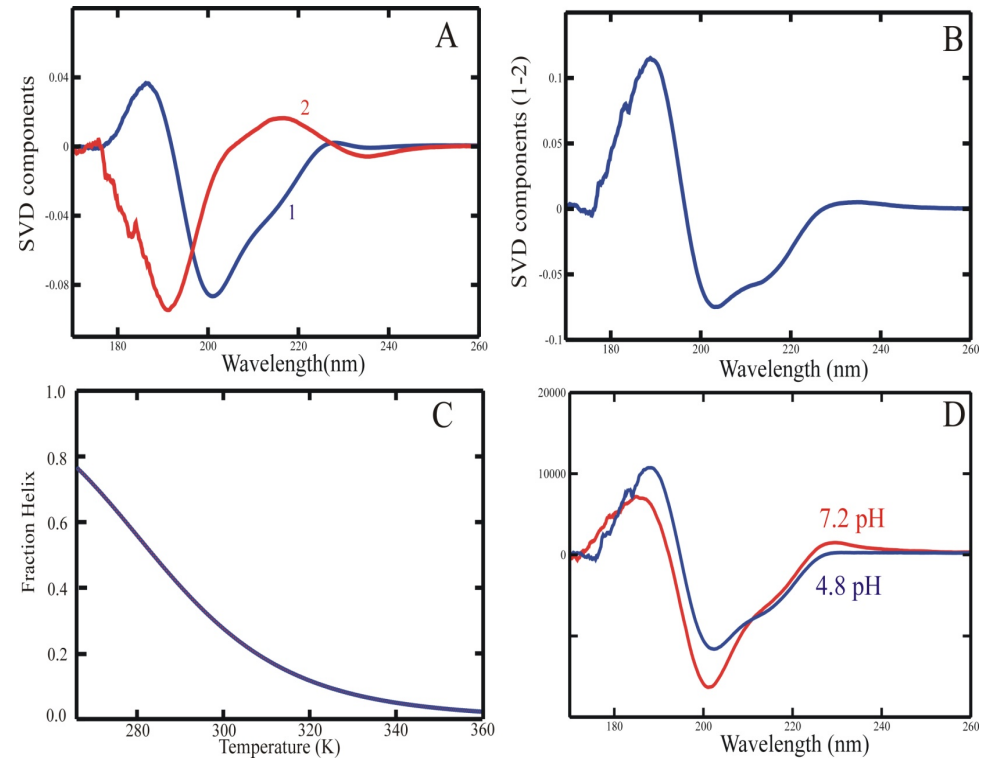
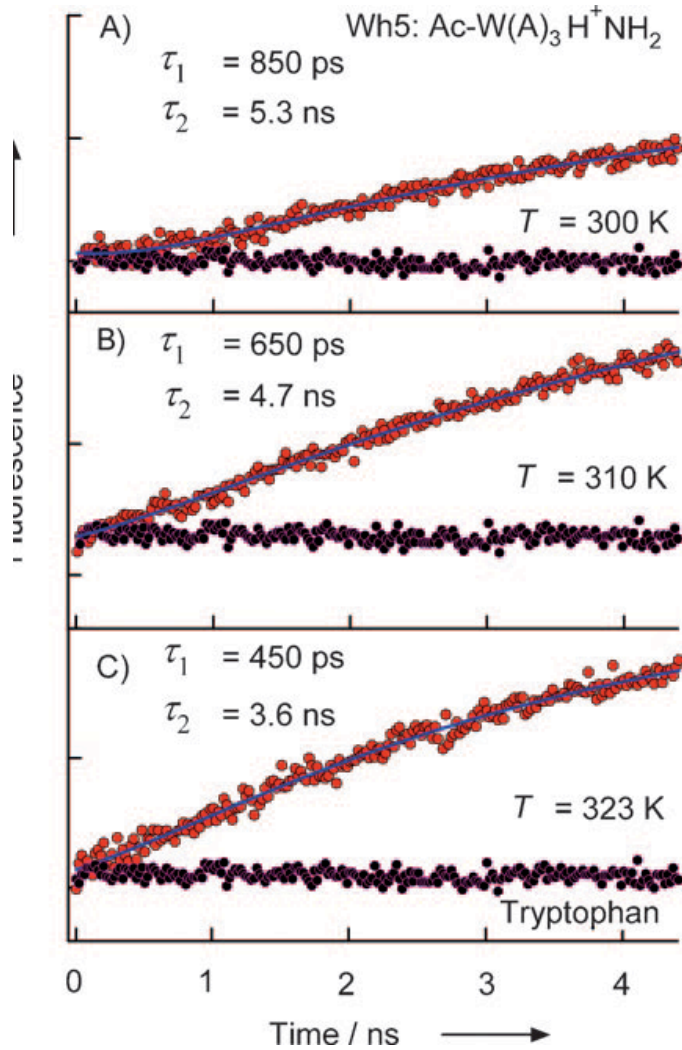
Peter Majek, Gouri Jas, Krzysztof Kuczera, Ron Elber, submitted

A nucleus: A significant inclination to structure in a **short** peptide segment(s), can speed up folding rate (Local Go model, Zwanzig, Brungelson & Wolynes, Thirumalai)

What is the shortest peptide chain that still has significant tendency to structure?



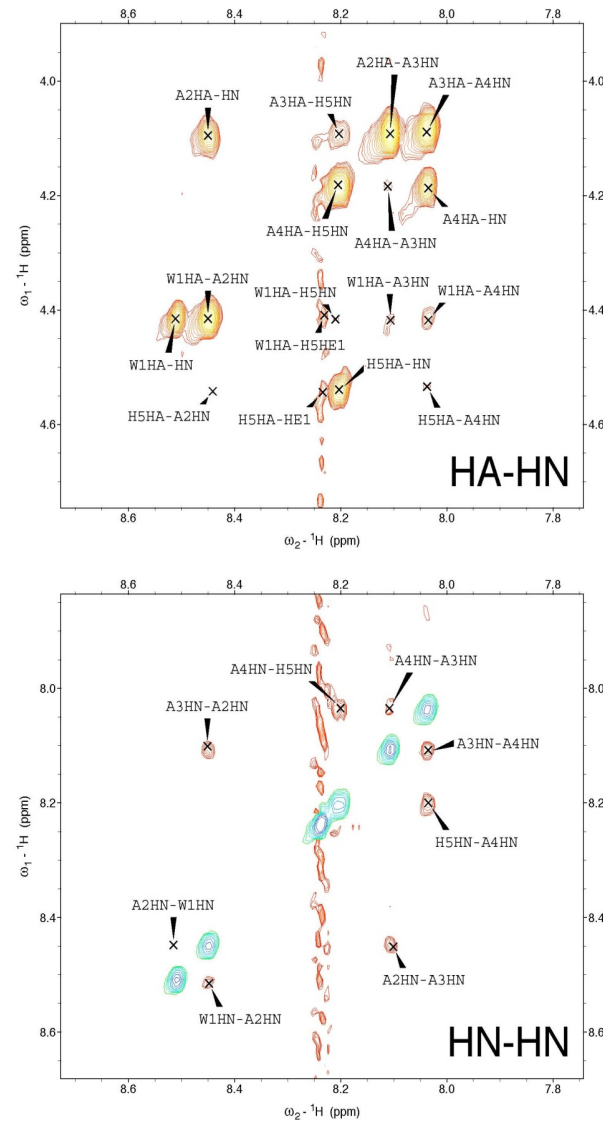
W-AAA-H (WH5): multiple evidence for clear structure in a short peptide, Gouri Jas



Measurements by Ad Bax

2D ^1H - ^1H ROESY and 2D ^1H - ^{13}C HSQC measurements were employed to measure the ^1H and ^{13}C chemical shifts and obtain ^1H - ^1H distance information. The $^3J_{\text{HNHa}}$ coupling constants were also determined. The ROESY and HSQC experiments were performed on a 5 mM sample of the peptide, WAAAH, at pH 4.2 and 5 °C (the $^3J_{\text{HNHa}}$ constants were also measured at 20 °C).

Regions of 2D ^1H - ^1H ROESY on WH-5 (pH 4.2, 5 °C) showing the HA-HN and HN-HN correlations



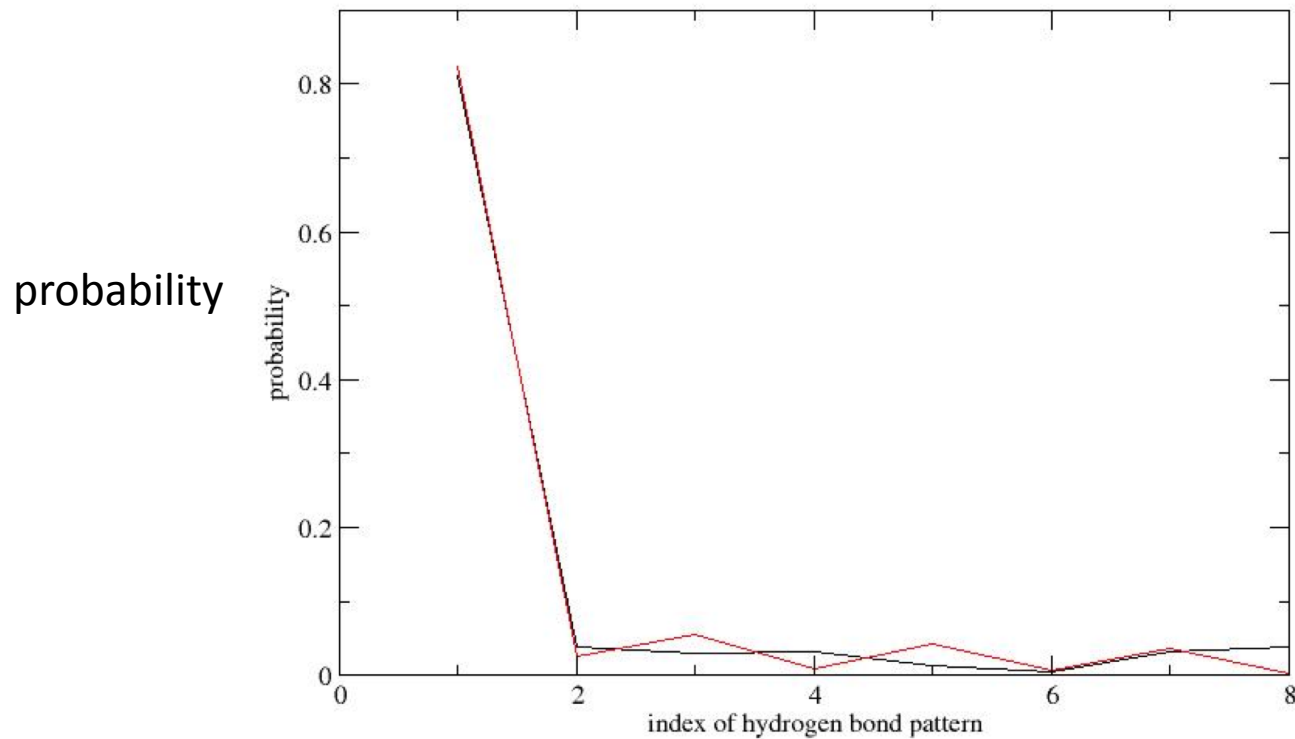
Milestoning calculations: Peter Majek (with Krzysztof Kuczera and Gouri Jas), JPC B, in press

- Peptide solvated in water box: 1 μ s traj. by Kuczera
- 10 coarse variables (all ϕ, ψ torsions)
- An interface ij is the set of points X with distance $d(i)$ and $d(j)$ from interfaces i and j plus a shift DELTA.

$$M_{i \rightarrow j} \equiv \left\{ X \mid d(X, X_i)^2 = d(X, X_j)^2 + \Delta_i^2 \text{ and } \forall k \ d(X, X_j) \leq d(X, X_k) \right\}$$

- 153 images from clustering 1 μ s traj. conf. reduced to 90
- 6186 reachable interfaces at 300K
- 50 traj at each interface
- ~310K traj of ~10ps each
- On a computer with 30K cores ~ 20min. On our system 2 months

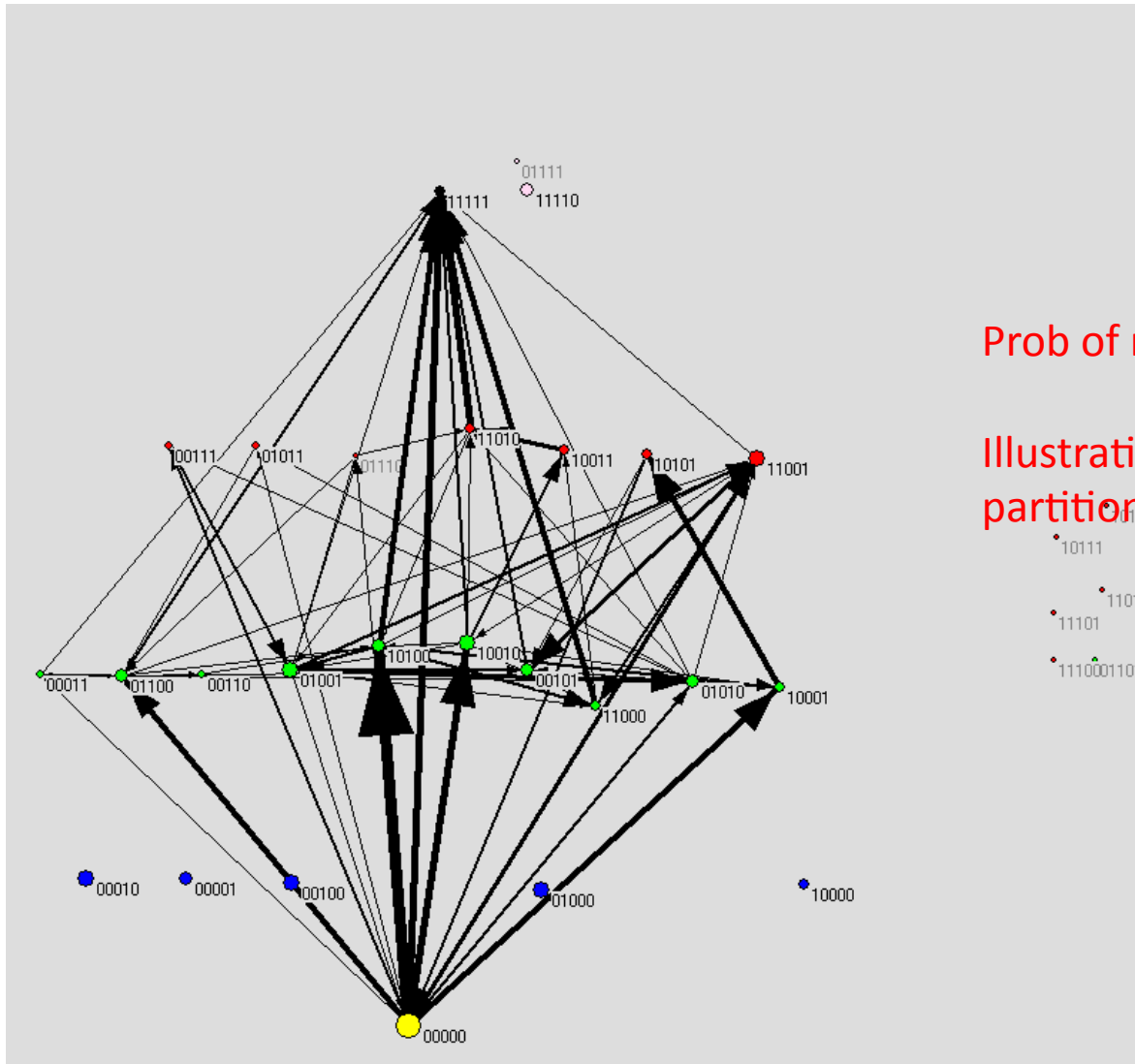
Comparison of MD and Directional Milestoning: Eq. probability



Hydrogen bond pattern (000, 001,...,111) enrichment in HB ~20

Folding network (from Milestoning)

Dominance of direct path and illustration of dead ends

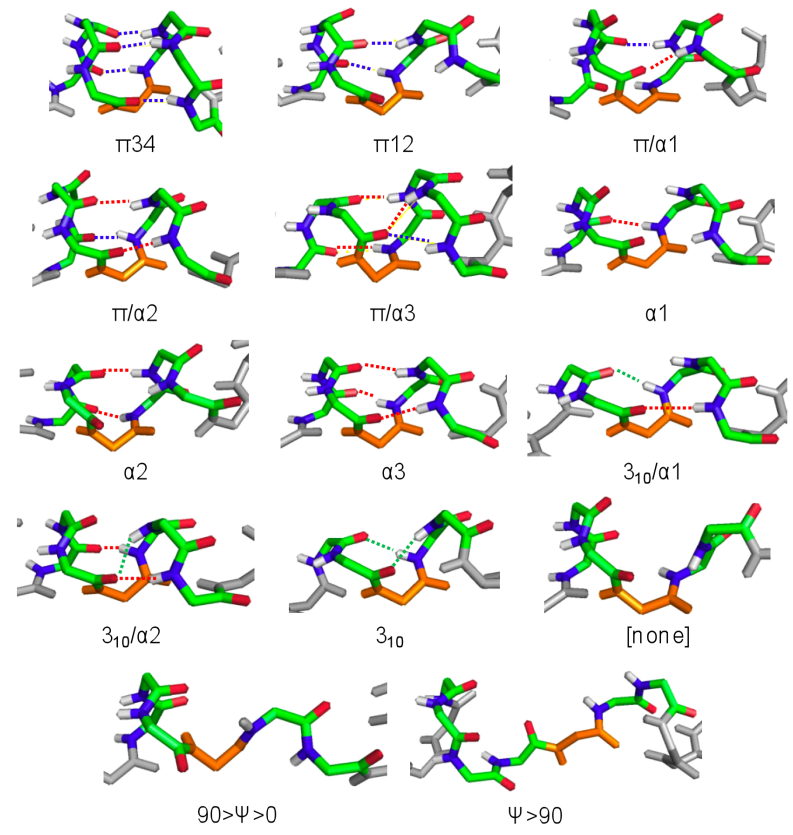
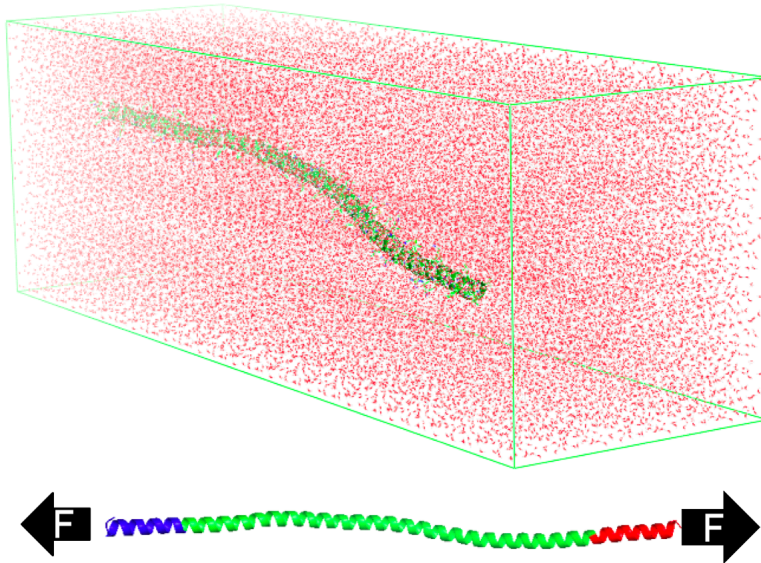


Prob of residue in helix ~20%

Illustration of kinetic partitioning (Thirumalai)



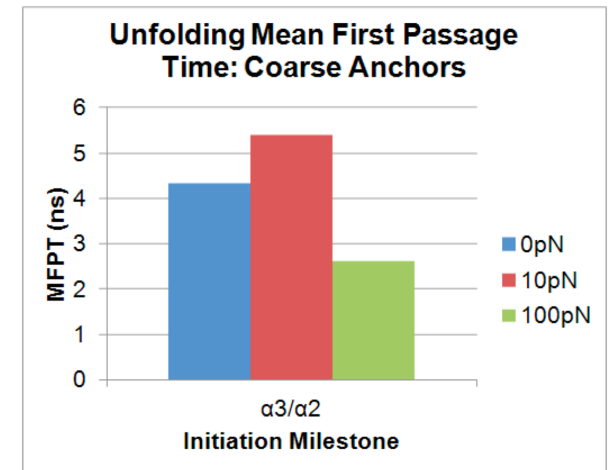
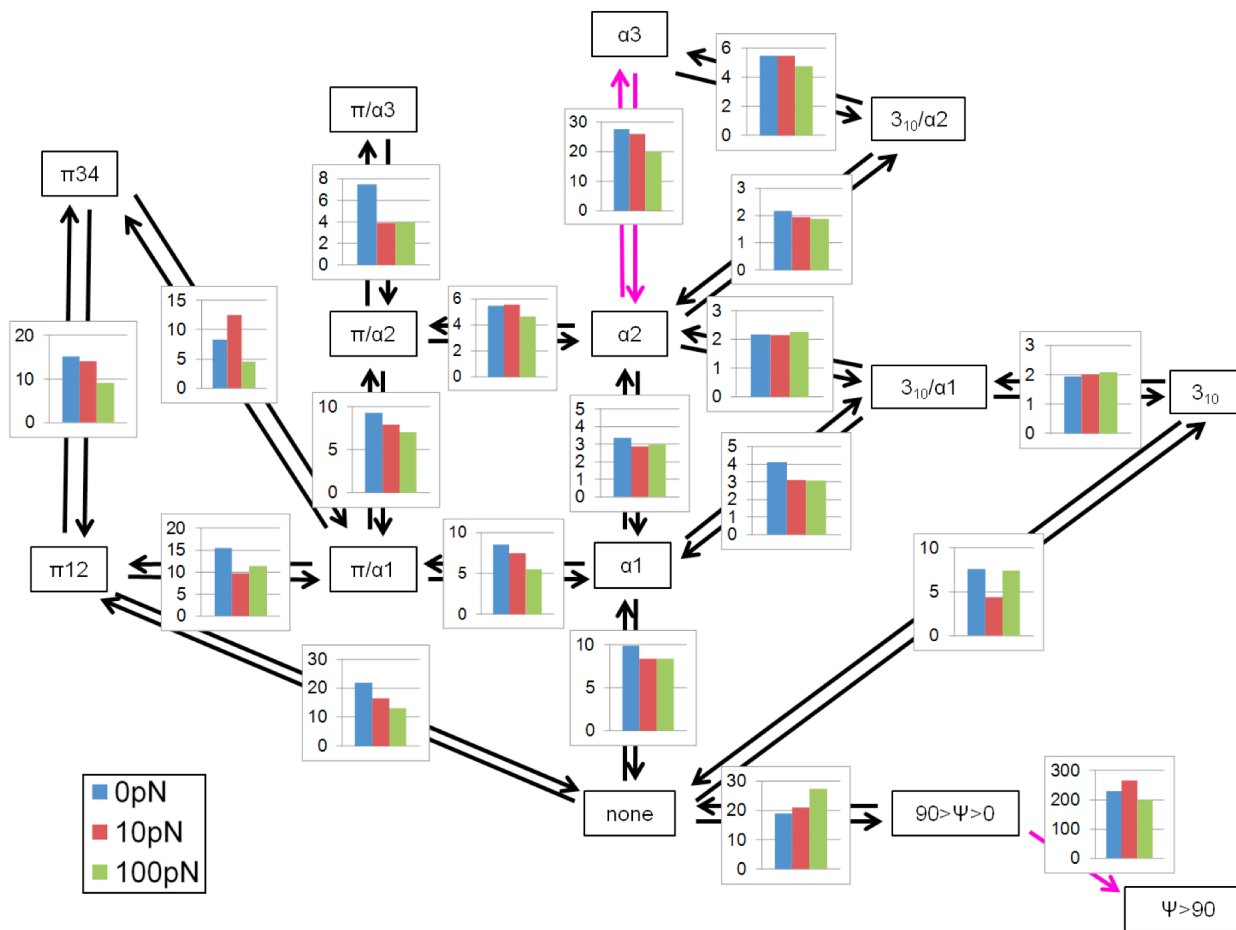
Early Events in Helix Unfolding



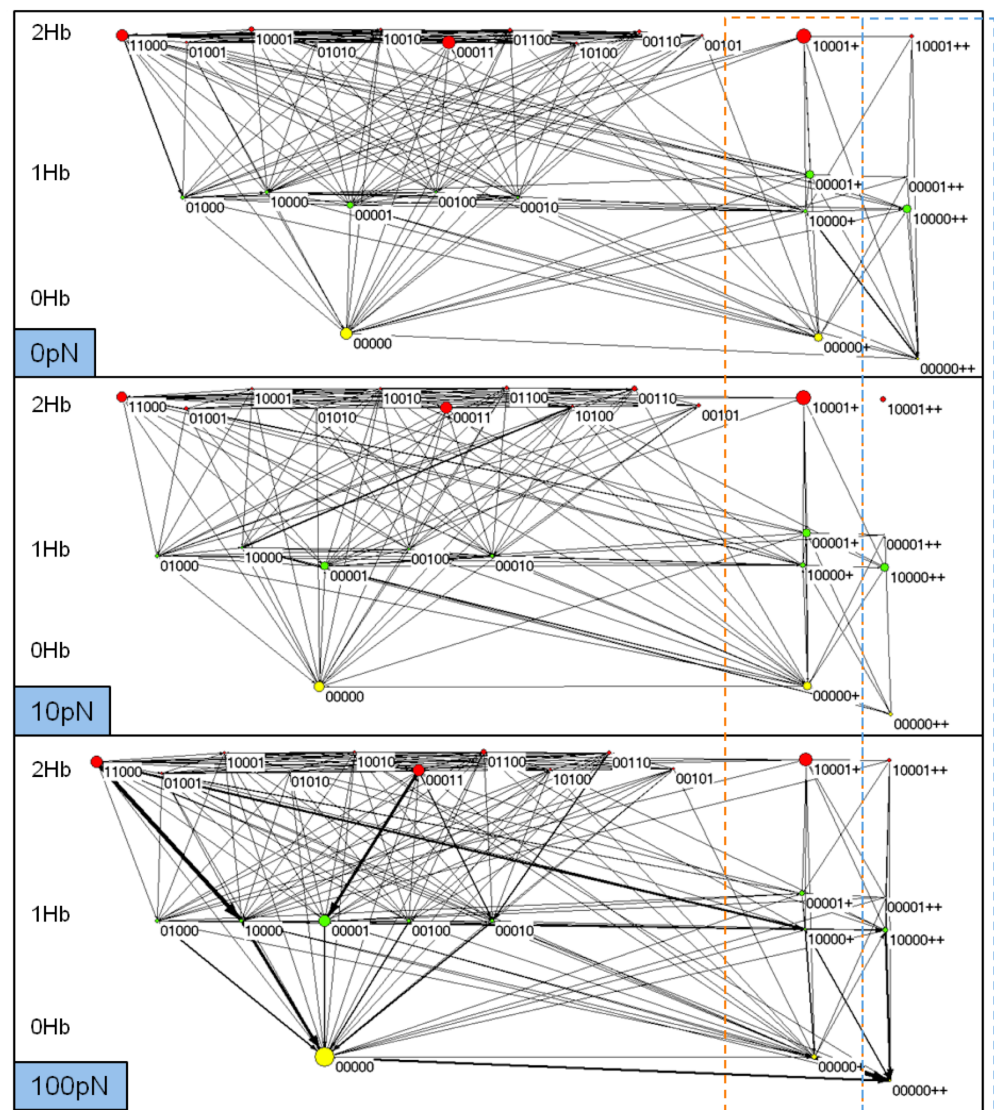
..... 3_{10} -Helical Hbond
 α -Helical Hbond
 π -Helical Hbond

with Tess Moon and Steve Kreuzer

Early events in helix unfolding (network)



More detailed states



*Line width normalized within each load.

Fine Anchor Designation	Quantity of α -Helical Hbonds	Hydrogen Bond Status (1=Intact; 0=Broken)					Ψ Angle Values
		C4-	C3-	C2-	C-	C	
(# Indicates Hydrogen Bond Pattern)							
32	5	1	1	1	1	1	all
31	4	1	0	1	1	1	all
30	4	0	1	1	1	1	all
29	4	1	1	1	1	0	all
28	4	1	1	1	0	1	all
27	4	1	1	0	1	1	all
26	3	0	0	1	1	1	all
25	3	1	0	1	1	0	all
24	3	0	1	1	1	0	all
23	3	1	1	1	0	0	all
22	3	1	0	1	0	1	all
21	3	0	1	1	0	1	all
20	3	1	0	0	1	1	all
19	3	0	1	0	1	1	all
18	3	1	1	0	1	0	all
17	3	1	1	0	0	1	all
16	2	0	0	1	0	1	all
15	2	0	0	1	1	0	all
14	2	1	0	1	0	0	all
13	2	0	1	1	0	0	all
12	2	0	0	0	1	1	all
11	2	1	0	0	1	0	all
10	2	0	1	0	1	0	all
9	2	1	0	0	0	1	-150°:0°
9+	2	1	0	0	0	1	0°:90°
9++	2	1	0	0	0	1	90°:180° or -180°:-150°
8	2	0	1	0	0	1	all
7	2	1	1	0	0	0	all
6	1	0	0	0	1	0	all
5	1	0	0	1	0	0	all
4	1	0	0	0	0	1	-150°:0°
4+	1	0	0	0	0	1	0°:90°
4++	1	0	0	0	0	1	90°:180° or -180°:-150°
3	1	1	0	0	0	0	-150°:0°
3+	1	1	0	0	0	0	0°:90°
3++	1	1	0	0	0	0	90°:180° or -180°:-150°
2	1	0	1	0	0	0	all
1	0	0	0	0	0	0	-150°:0°
1+	0	0	0	0	0	0	0°:90°
1++	0	0	0	0	0	0	90°:180° or -180°:-150°
π	NOTE: this fine anchor incorporates any residues with a π -helical hydrogen bond (and therefore not in a pure α -helical state)						
310	NOTE: this fine anchor incorporates any residues with a 310-helical hydrogen bond (and therefore not in a pure α -helical state)						

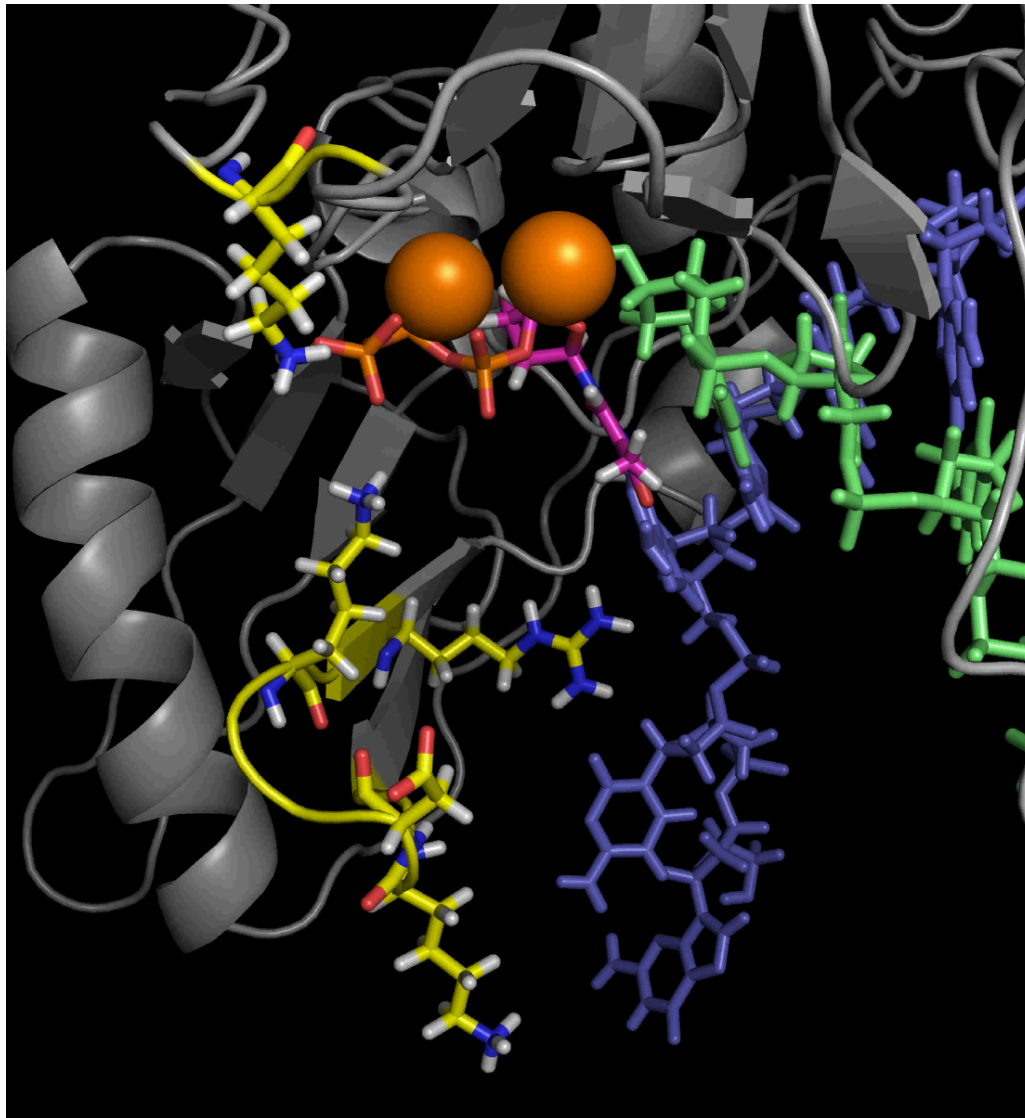
HIV reverse transcriptase
Avoiding errors in the genetic code
with KA Johnsson, V Nguyen, and S Kirmizialtin

- Enzyme generates DNA from RNA sequence
- Non equilibrium steady state system: Pictorially, nucleotide binds weakly in an open form, protein changes to close form, chemistry...
- How does the protein select the correct substrate??

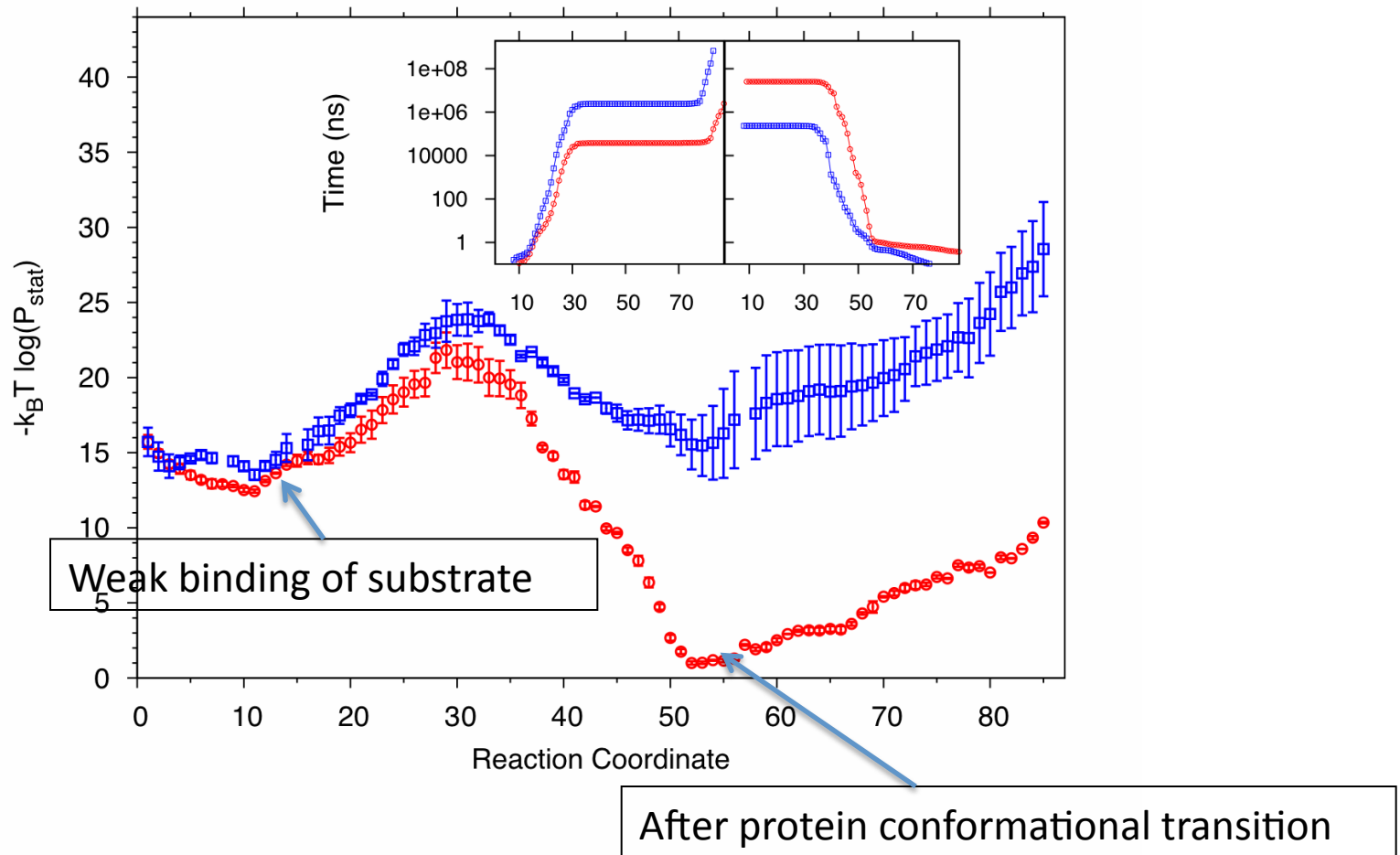
And a movie (a molecular process,
time scale – millisecond)



And another...

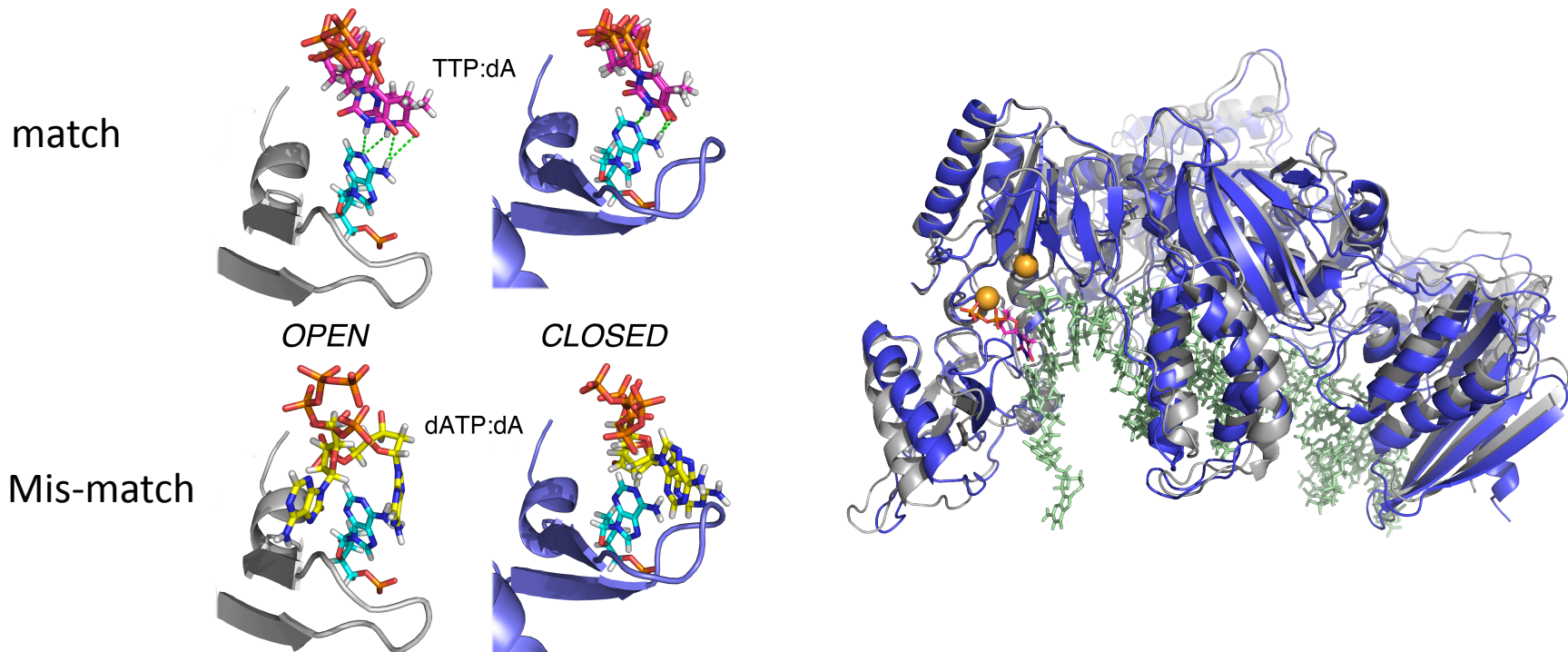
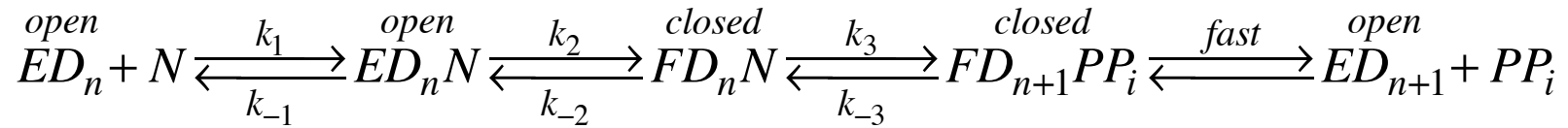


HIV reverse transcriptase synthesizes DNA: Selection by an induced fit, red correct substrate, blue incorrect. Calculation by Serdal Kirmizialtin



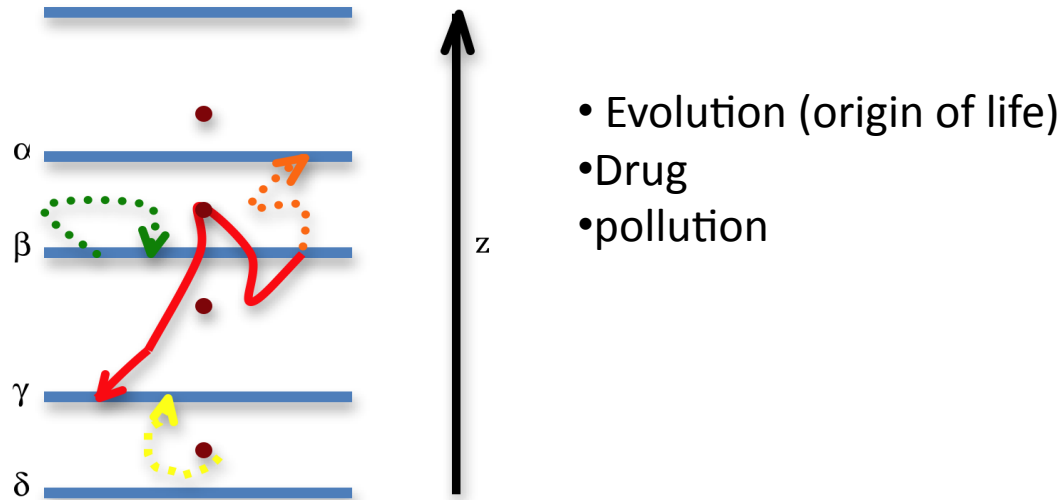
With KA Johnson

	TTP		dATP	
Source	k_2	k_{-2}	k_2	k_{-2}
Experiment	2500	4	>500	>1200
Theory	2500	40	200-400	~4000



Milestoning for membranes

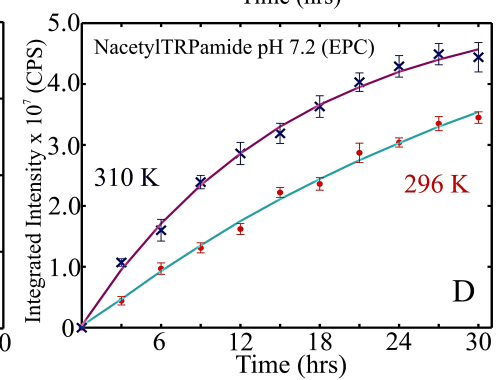
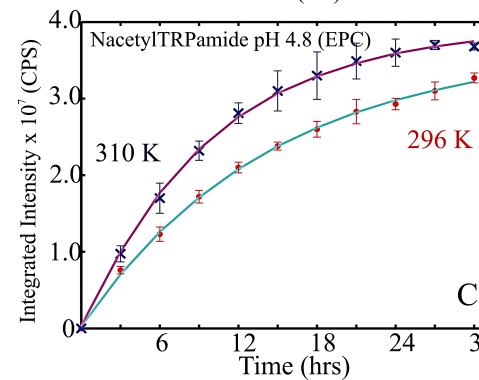
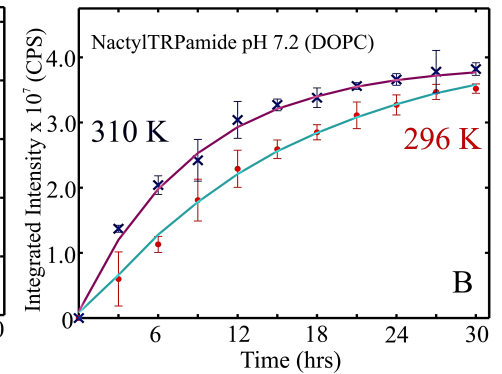
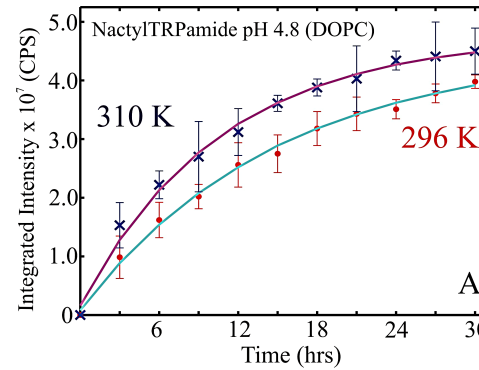
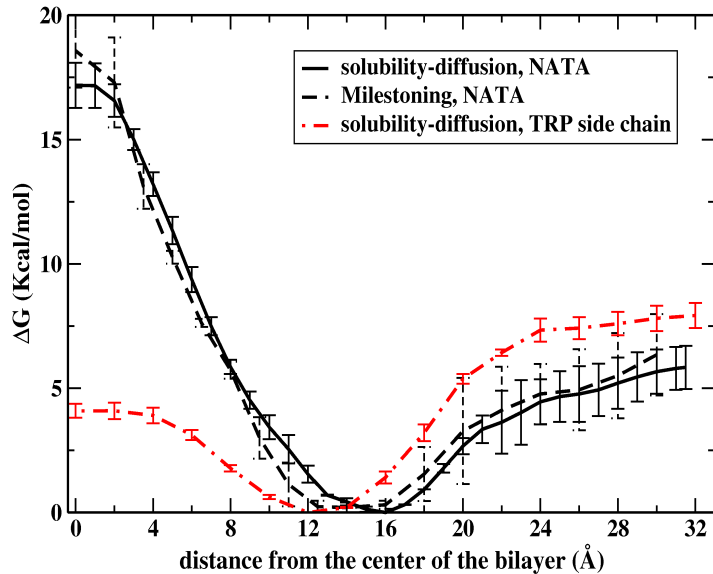
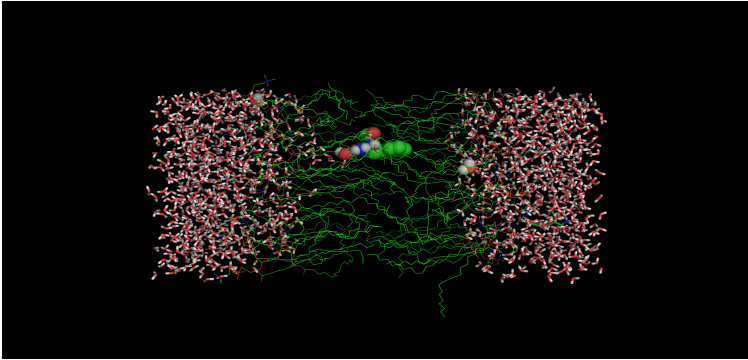
- Compute trajectory fragments between interfaces (Milestones) to compute the $K_{\alpha\beta}(t)$ Kernel



$$K_{\beta\gamma}(t) \cong n_{\beta\gamma}(t)/n_{\beta}$$

Membrane permeation

Dogma: Charged species no permeation, Neutral, yes.



Dogma incorrect

Summary

- Code of MOIL:
 - <http://clsb.ices.utexas.edu/prebuilt/>
- Milestoning is a method based on rigorous theory that builds on short trajectory fragments to obtain long time kinetic and thermodynamic properties of the system.
- Studies of biophysical systems along a reaction coordinate (or not)
 - Allosteric transition in Scapharca hemoglobin
 - Myosin recovery stroke
 - HIV reverse transcriptase
 - Membrane permeation
 - Helix folding/unfolding

Thanks!

Thanks to

- my collaborators: David Shalloway, Eric Vanden Eijnden, Giovanni Ciccotti, Kent Johnson, Tess Moon, Jas Gouri, Krzysztof Kuczera
- Students and postdocs: Tony West, Peter Majek, Serdal Kirmizialtin, Alfredo Cardenas, Steve Kreuzer

MLST references

Serdal Kirmizialtin, Virginia Nguyen, Kenneth A Johnson, and Ron Elber, "Molecular mechanisms by which induced-fit determines enzyme specificity", *Structure*, accepted
Gouri S. Jas, Wendy Hegefeld, Peter Májek, Krzysztof Kuczera, and Ron Elber, "Experiments and comprehensive simulations of the formation of a helical turn", *J. Phys. Chem. B*, accepted.

Alfredo E. Cardenas, Gouri S. Jas, Krzysztof Kuczera, and Ron Elber, "Unassisted transport of N-acetyl-L-tryptophanamide through DOPC membrane: Experiment and simulation", *J. Phys. Chem. B*, accepted.

Serdal Kirmizialtin and Ron Elber, "Revisiting and Computing Reaction Coordinates with Directional Milestoning", *J. Phys. Chem. A*, 115,6137-6148(2011)

Peter Májek and Ron Elber, "Milestoning without a reaction coordinate", *Journal of Chemical Theory and Computations* 6,1805–1817(2010)

Ron Elber and Anthony West, "Atomically Detailed Simulation of the Recovery Stroke in Myosin by Milestoning", *Proceeding of the National Academy of Sciences USA*, 107, 5001-5005 ,(2010)

Vanden-Eijnden Eric, Maddalena Venturoli, Giovanni Ciccotti, and Ron Elber, "On the assumptions underlying Milestoning", *J. Chem. Phys.* 129,174102(2008)

Anthony M.A. West, Ron Elber, and David Shalloway, "Extending molecular dynamics timescales with milestoning: Example of complex kinetics in a solvated peptide", *J. Chem. Phys.* 126,145104, (2007)

Anton K. Faradjian and Ron Elber, "Computing time scales from reaction coordinates by milestoning", *J. Chem. Phys.* 120:10880-10889(2004)