

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

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

\$ NIH

Paris, June, 2012

# Plan

- Milestoning: Theory and algorithm
- Example
  - Helix unfolding under stress.
- Retention of intramolecular interactions in a thermodynamic cycle (Mauro L Mugnai)

## MLST references

Steven Kreuzer, Ron Elber and Tess J Moon, “Early Events in **Helix Unfolding Under External Forces**: A Milestoning Analysis”, J. Phys. Chem. B, accepted

Serdal Kirmizialtin, Virginia Nguyen, Kenneth A Johnson, and Ron Elber, “How Conformational Dynamics of **DNA Polymerase** Select Correct Substrates: Experiments and Simulations”, Structure, 20,618-627(2012)

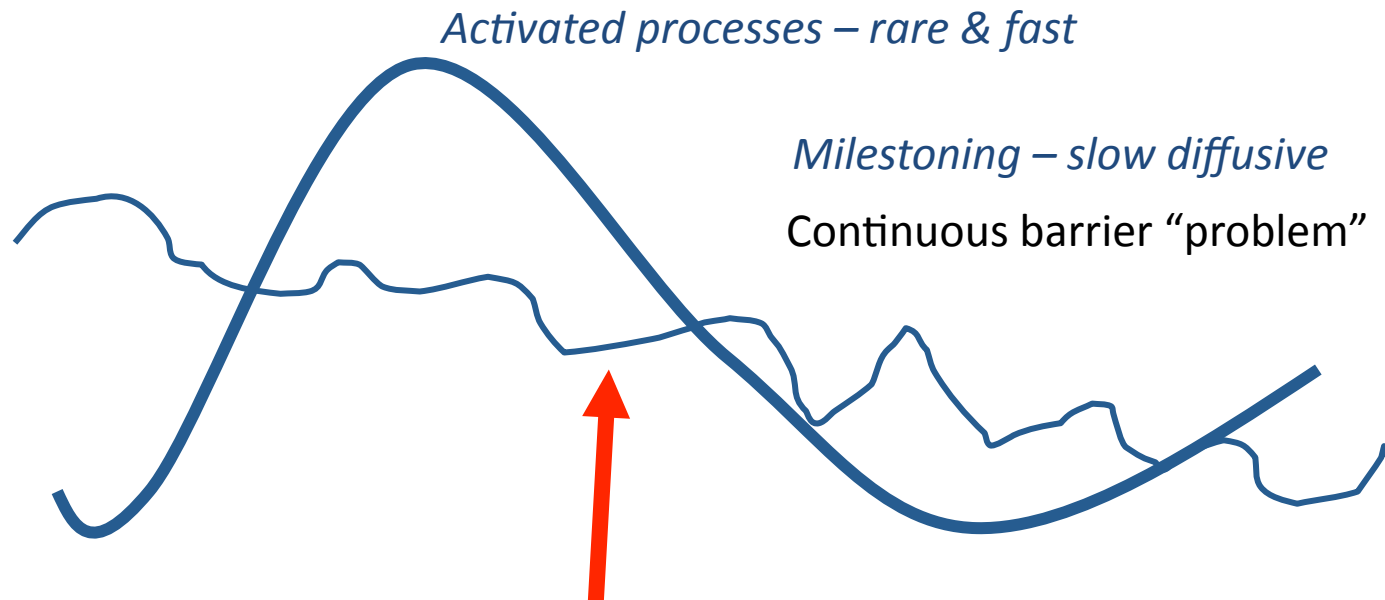
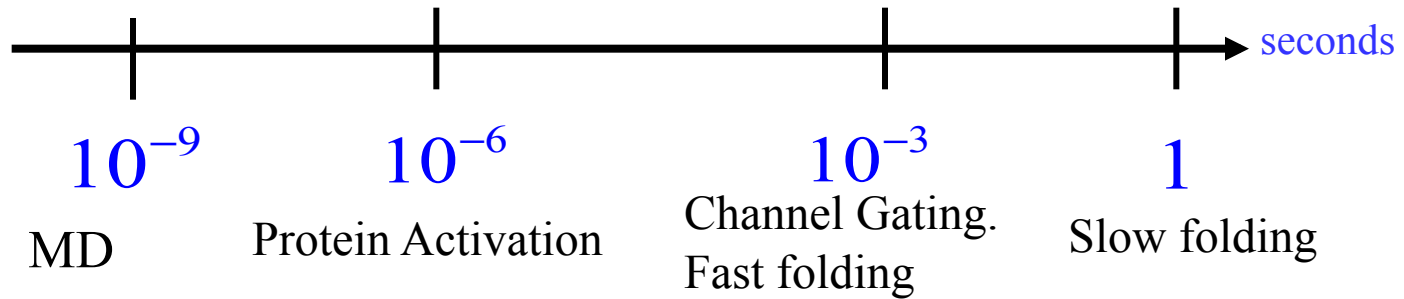
Gouri S. Jas, Wendy Hegefeld, Peter Májek, Krzysztof Kuczera, and Ron Elber, “Experiments and comprehensive simulations of the **formation of a helical turn**”, J. Phys. Chem. B, accepted.

Alfredo E. Cardenas, Gouri S. Jas, Krzysztof Kuczera, and Ron Elber, “**Unassisted transport of N-acety-L-tryptophanamide through DOPC membrane**: Experiment and simulation”, J. Phys. Chem. B, 116,2739-2750(2012).

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)

Ron Elber, "A milestoning study of the kinetics of an allosteric transition: Atomically detailed simulations of deoxy **Scapharca hemoglobin**", Biophysical J.,2007 92: L85-L87

# Long time processes in biophysics: activation or long range diffusion





## Efficiency (more Milestones the better!)

- 
- *Diffusive speedup:*

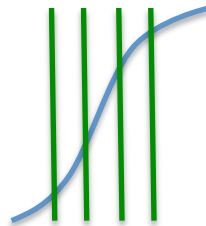
$$t \sim L^2 \rightarrow t \sim M(L/M)^2 = L^2/M$$

- *Parallelization speedup:*

$$t \sim L^2/M^2$$

- *Exponential bootstrapping at large barrier:*

$$\frac{1}{qq'} \rightarrow \frac{1}{q} + \frac{1}{q'}$$



- Results on myosin for the recovery stroke (240 milestones) predict submillisecond timescale using nanosecond simulations
- Membrane permeation (computed with ~30 Milestones) predicts time scale of hours using microsecond trajectories

# Overcoming time scale and sampling gaps with Milestoning

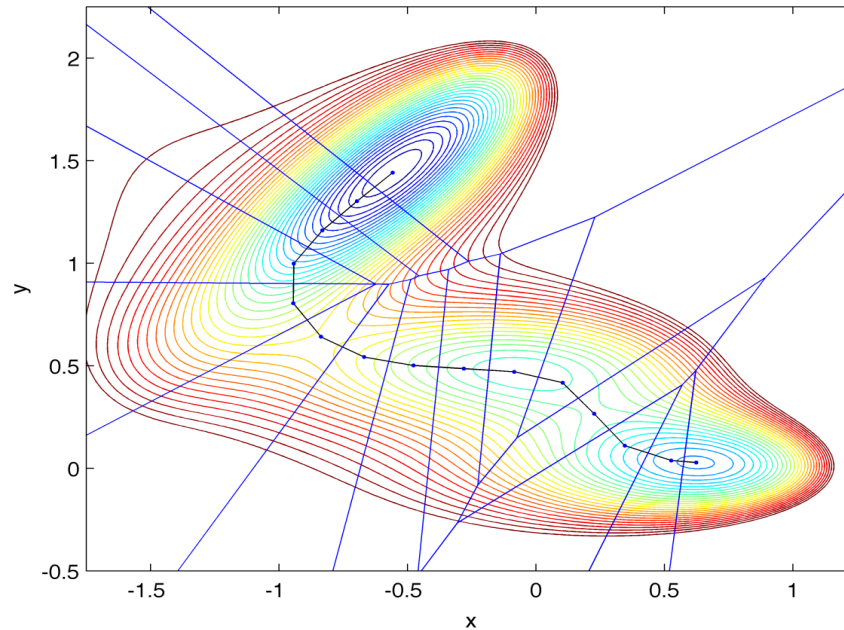
- Provide coarse sampling of the system and define cell representation
- Compute local fluxes between the cells
- Compute overall kinetics and thermodynamics (and reaction coordinates) at cell resolution
  - Serdal Kirmizialtin and Ron Elber, “Revisiting and Computing Reaction Coordinates with Directional Milestoning”, J. Phys. Chem. A, 115,6137-6148(2011)

# 1. Define cells/anchors/milestones in coarse space

(Note : Milestoning can handle a reasonable large number of coarse variables. Peptide, or RNA folding includes 10-20 coarse variables)



Rough description is possible with the help of a minimum (free) energy coordinate



$$S[X(l)] = \int_{X_R}^{X_P} |F| dl$$

Olender and R. Elber, J. Mol. Struct. 398-399, 63-72 (1997).

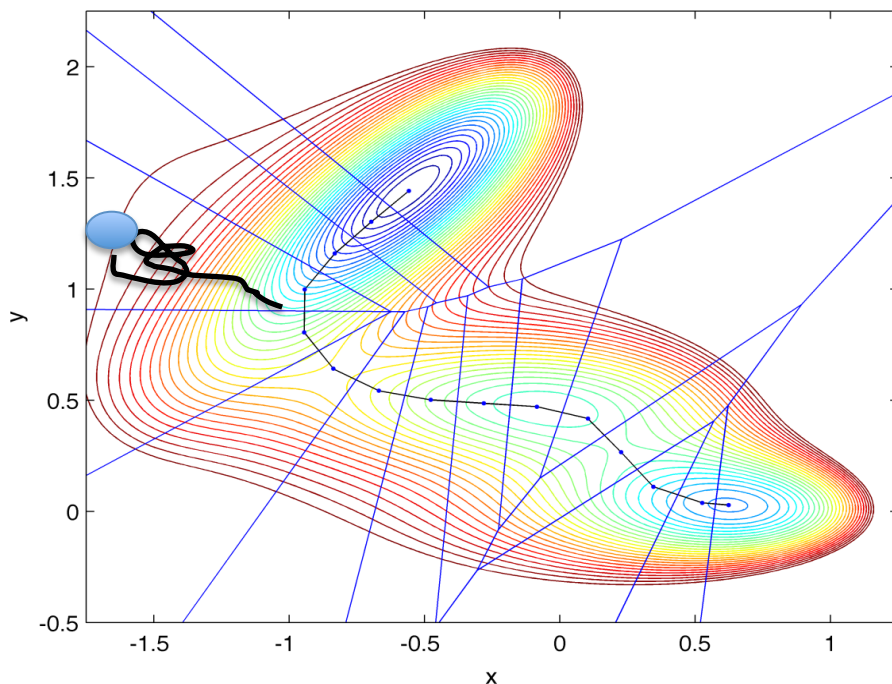
Alternatives: Pratt,

LUP / String - Elber, Vanden Eijnden

NEB - Johansson

MaxFlux – Berkowitz, Straub, Skeel

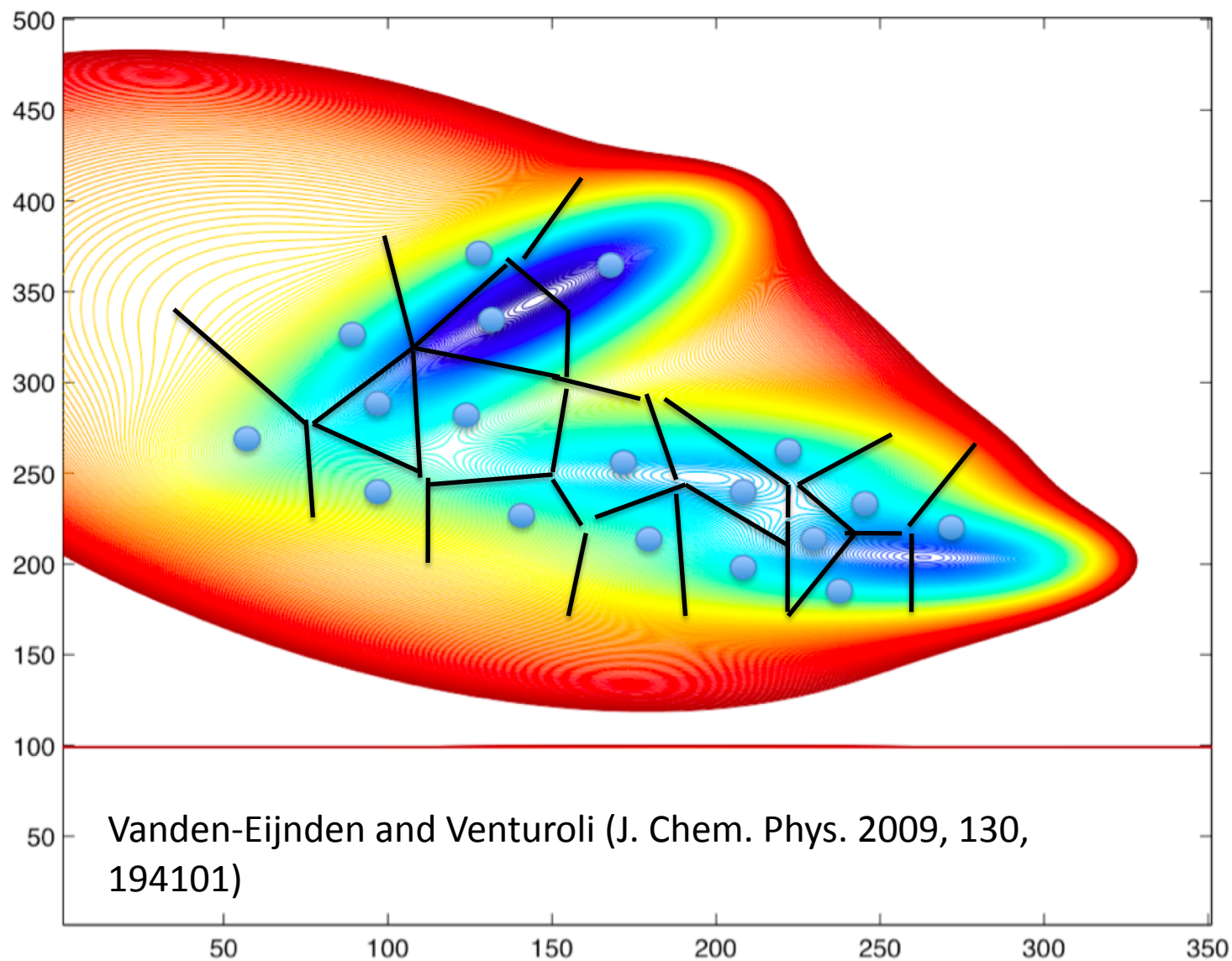
Compute trajectories between interfaces (milestones). If one of the trajectories wanders far from other anchors, add a new one



Similar to searches in the Markov State Model

Chodera, J. D.; Singhal, N.; Pande, V. S.; Dill, K. A.; Swope, W. C. *J. Chem. Phys.* **2007**, *126*, 155101-155117.

Provide “rough” description of energy landscape using Voronoi-like cells and anchors in coarse space



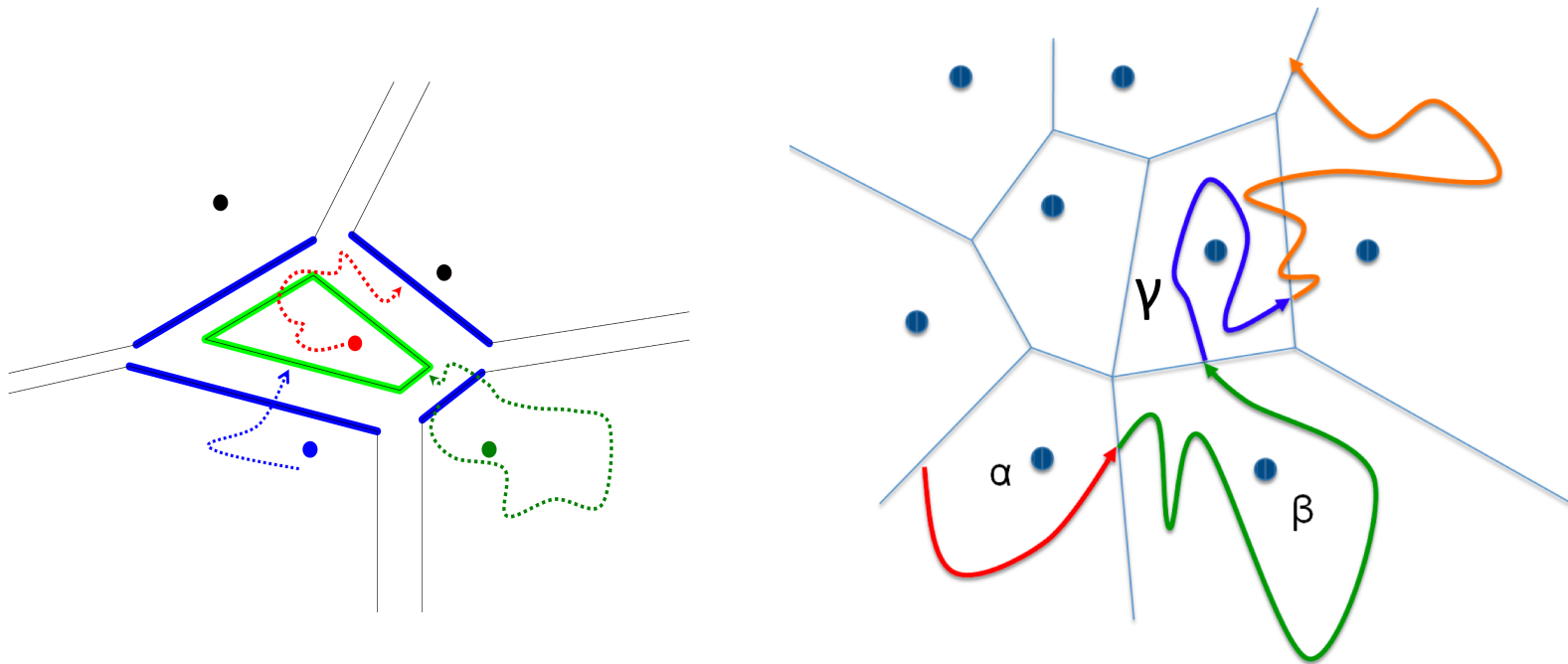
# Identify cells and coarse variables

- Besides reaction coordinates
  - Replica exchange
  - High temperature MD
  - Chemical intuition

2. Compute local transition times between milestones (interfaces). We assume loss of correlation between milestones.

Compute (short) trajectories between interfaces and estimate flux  $K_{\alpha\beta,\beta\gamma}(t)$  between interfaces  $\alpha\beta$  and  $\beta\gamma$

$K_{\alpha\beta,\beta\gamma}(t)$  is the probability density for a trajectory conditioned to start at  $\alpha\beta$  and to end at  $\beta\gamma$  exactly at time  $t$  (green trajectory). We initiate the trajectory by sampling at the interface  $\alpha\beta$  from first hitting point distribution.



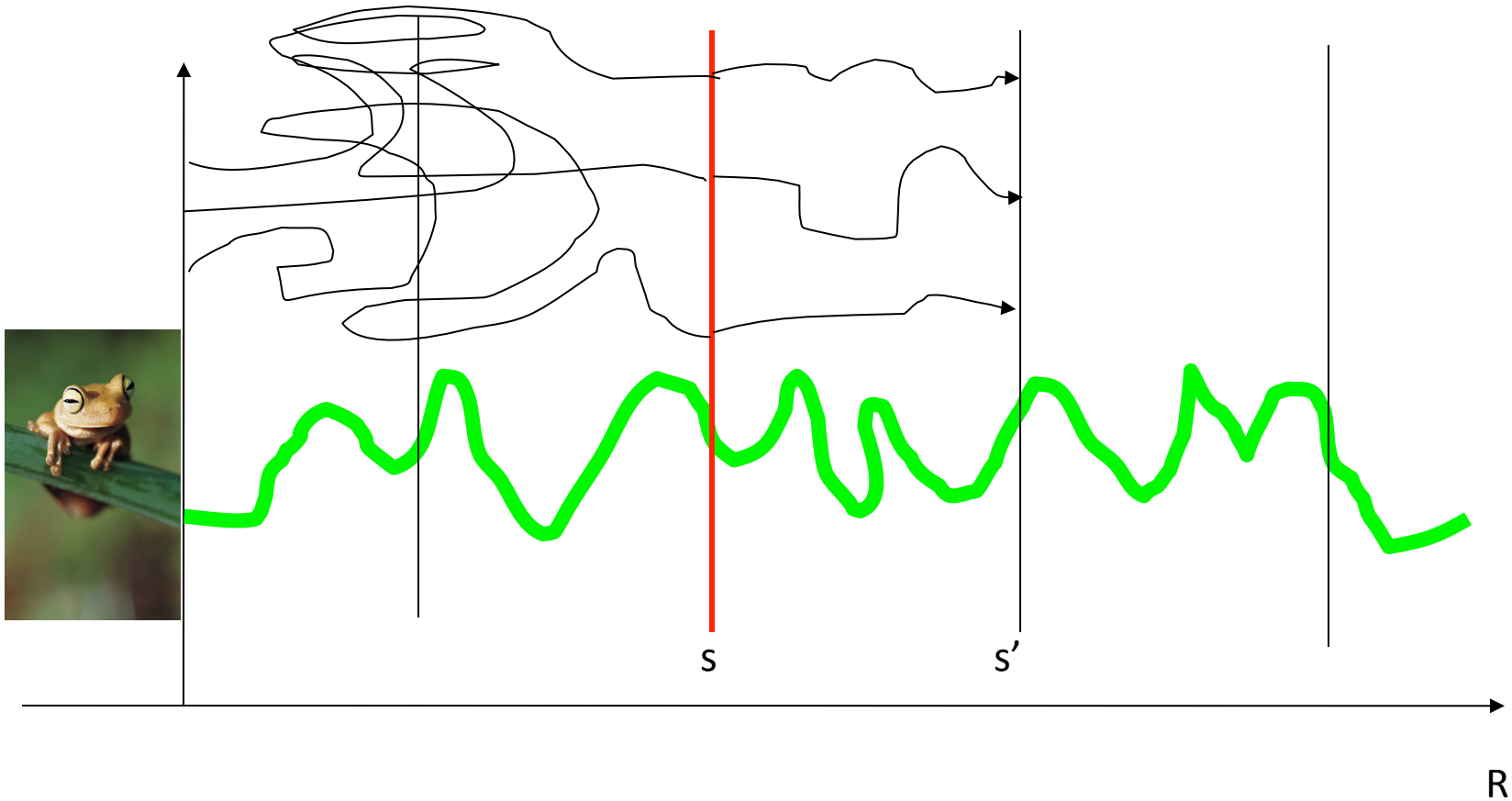
System close to equilibrium, loss of correlation between interfaces

# How to compute K?

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

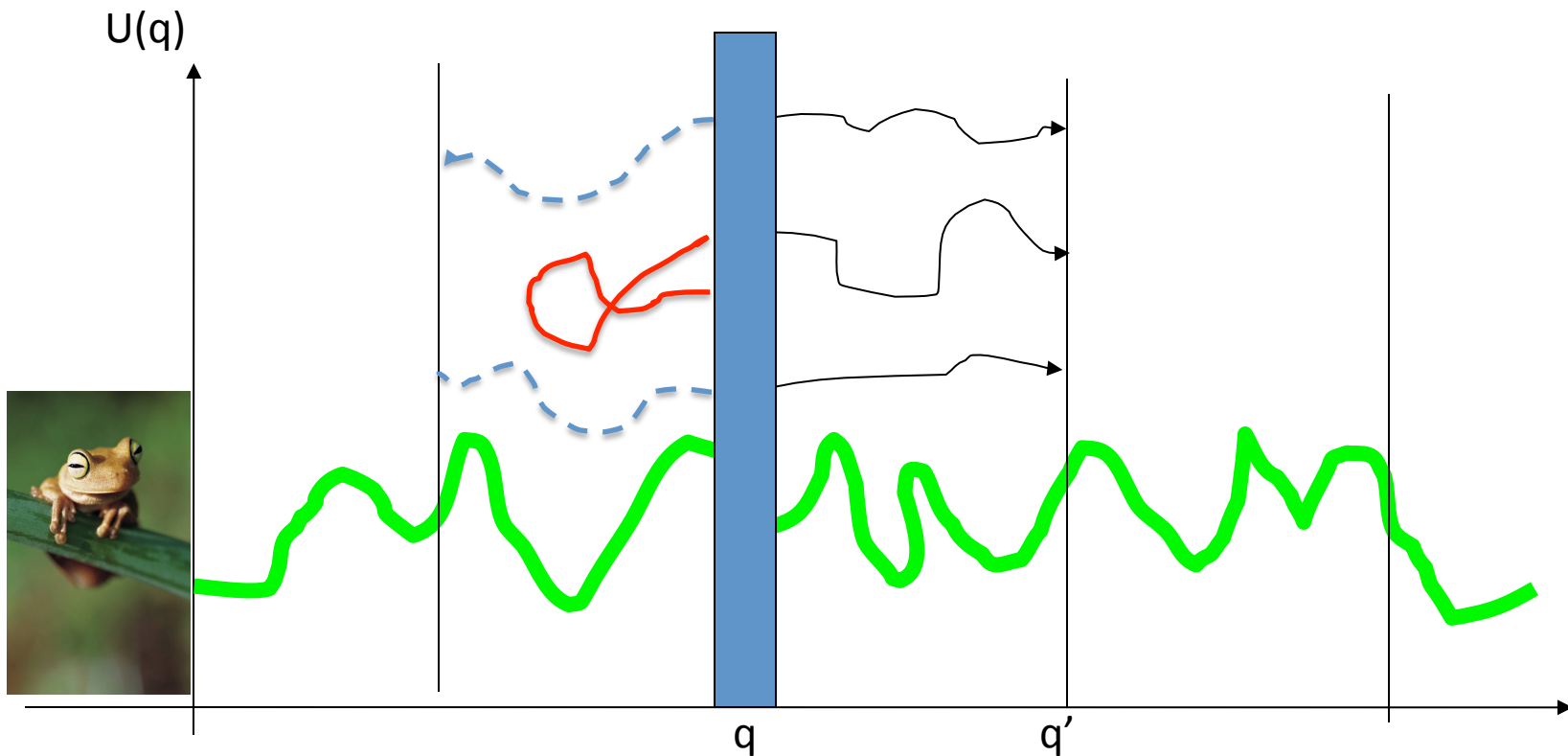
How to obtain the “appropriate” initial distribution?

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

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



Directional Milestoning: With Majek, JCTC 2010; with Kirmizialtin JPC 2011  
Tests similar to PPTIS D. Moroni, P. G. Bolhuis, and T. S. van Erp JCP, 2004



3. Compute kinetics and thermodynamics using the local transition kernel

With  $K_{s,s'}(\tau)$  determined, the Milestoning Equation is:

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

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

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

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

$s, s'$  -- 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 (with Vanden Eijnden, JCP 2008)

# Equivalent to Generalized Master Equation (GLE)

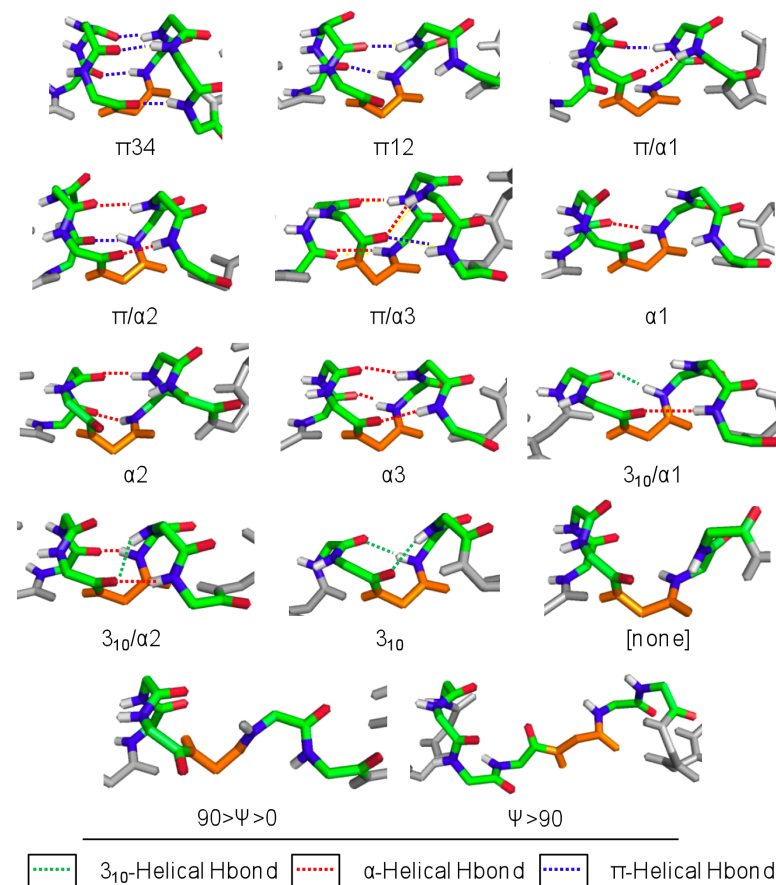
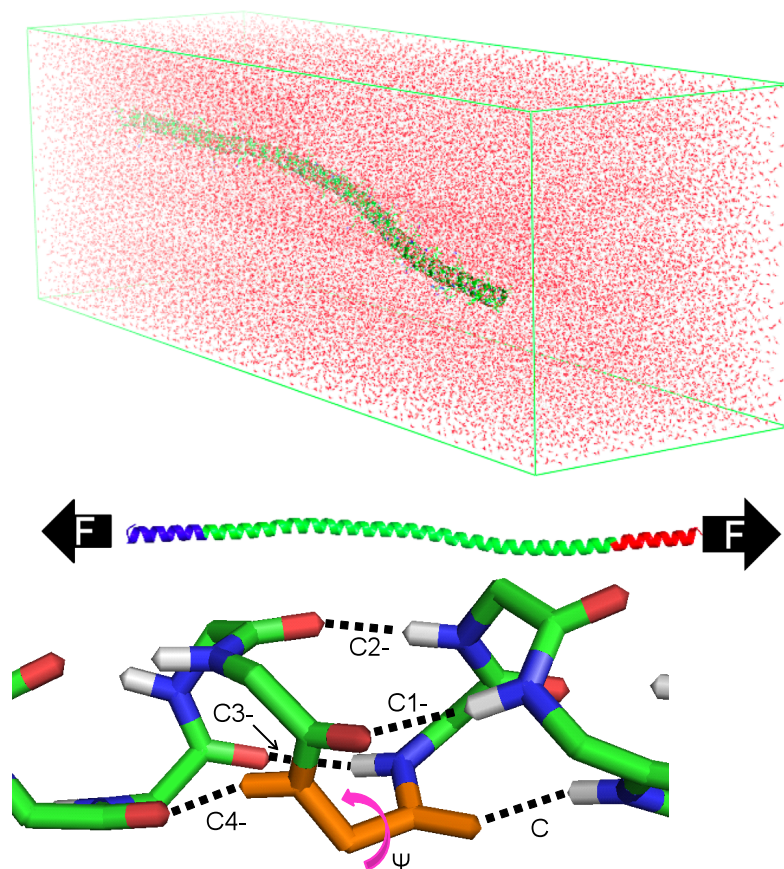
- **GLE has time dependent rate coefficients**

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

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

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

# Early Events in Helix Unfolding



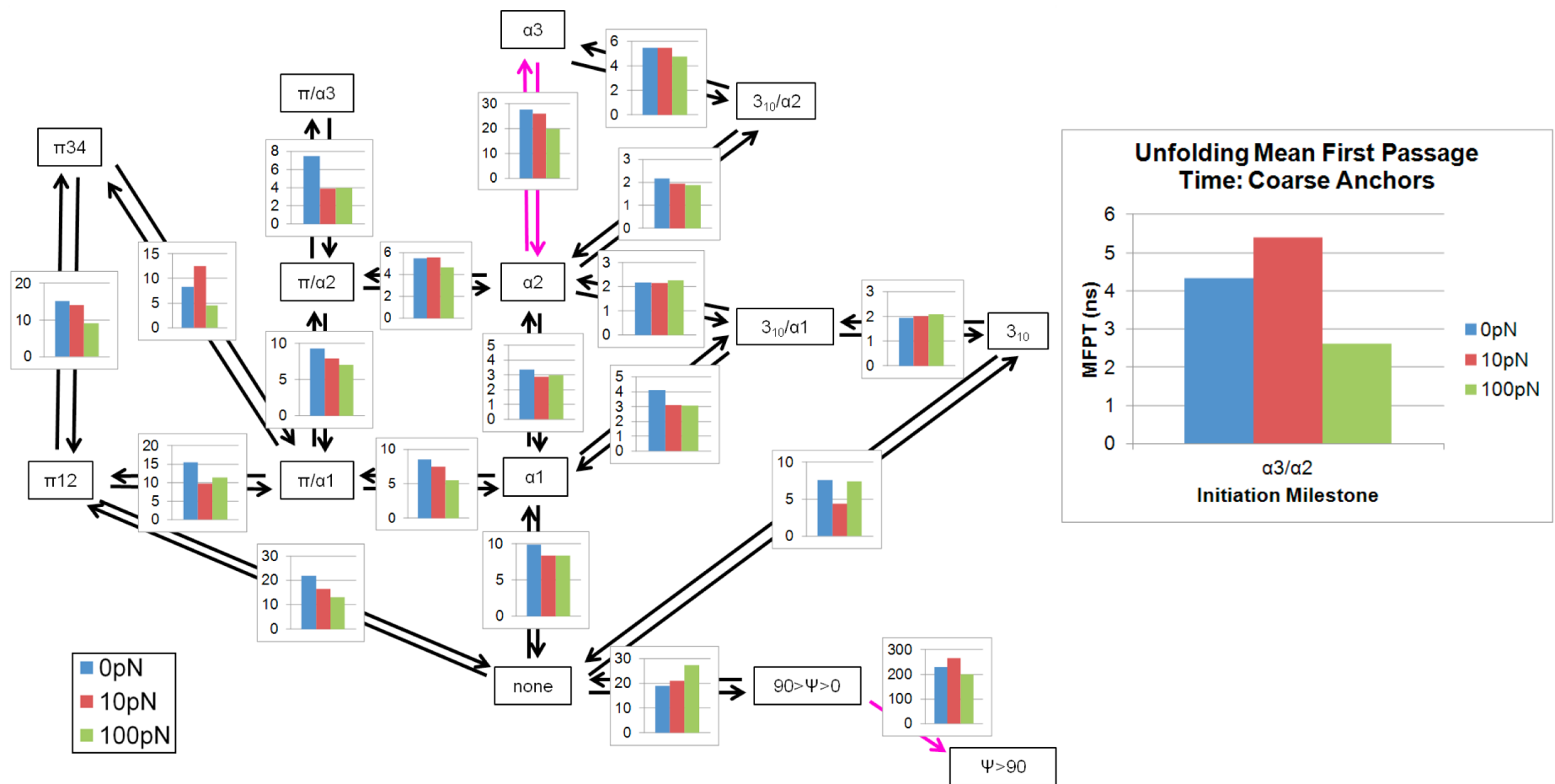
Catch bond: Increase the applied force,  
The rate **decreases**.

HB  $d < 3.5 \text{ \AA}$   $\theta > 150^\circ$

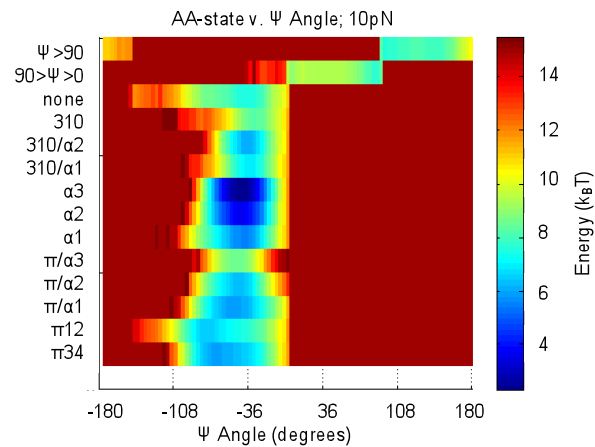
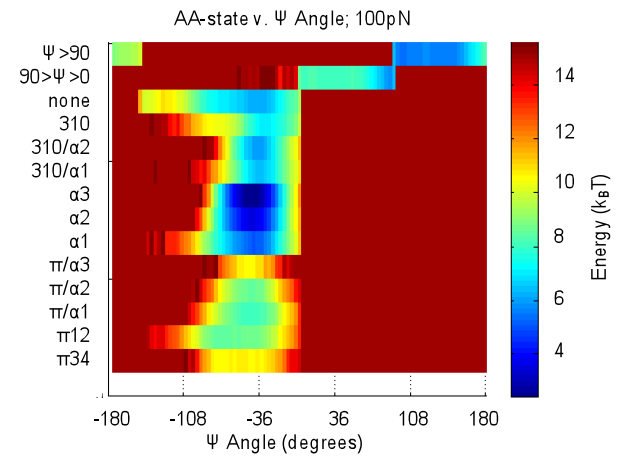
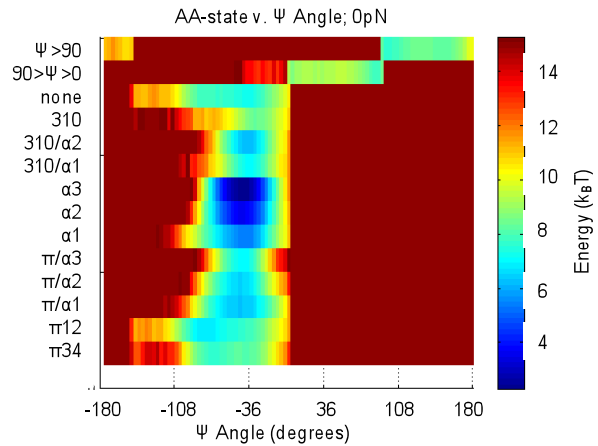
Six coarse variables: torsion + 5 hydrogen bonds

with Tess Moon and Steve Kreuzer, JPC B in press

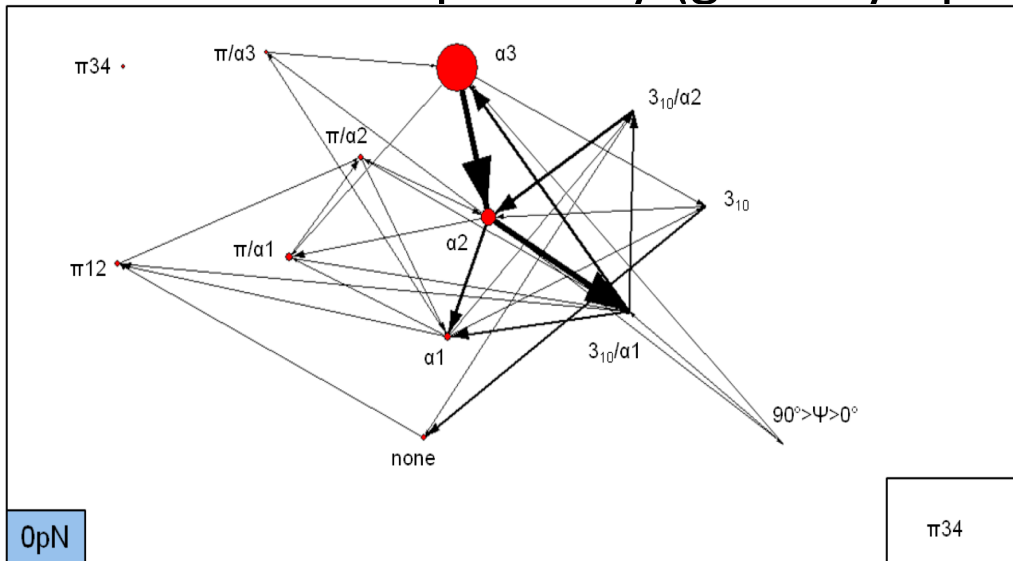
# Early events in helix unfolding (network)



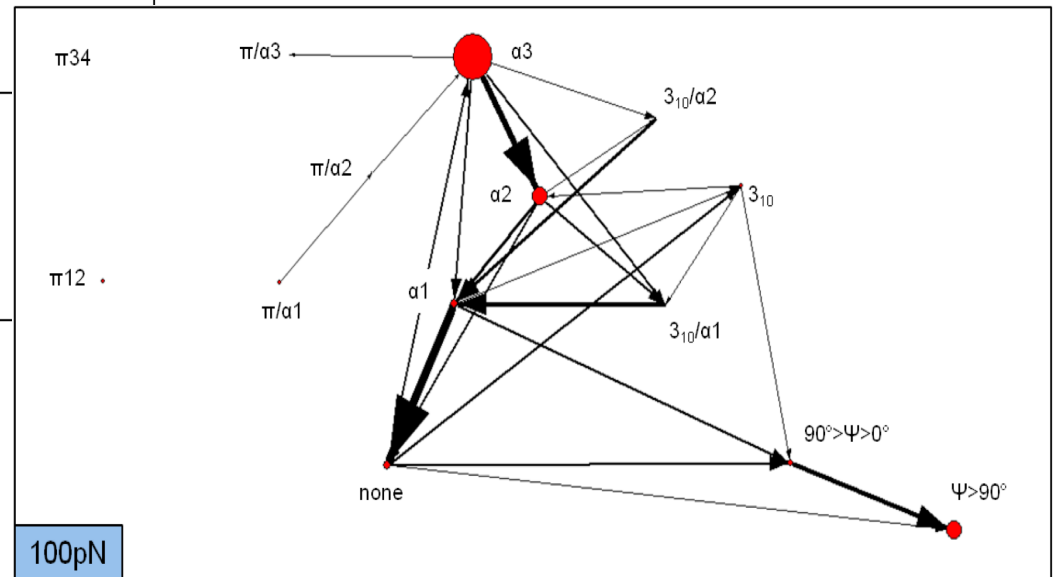
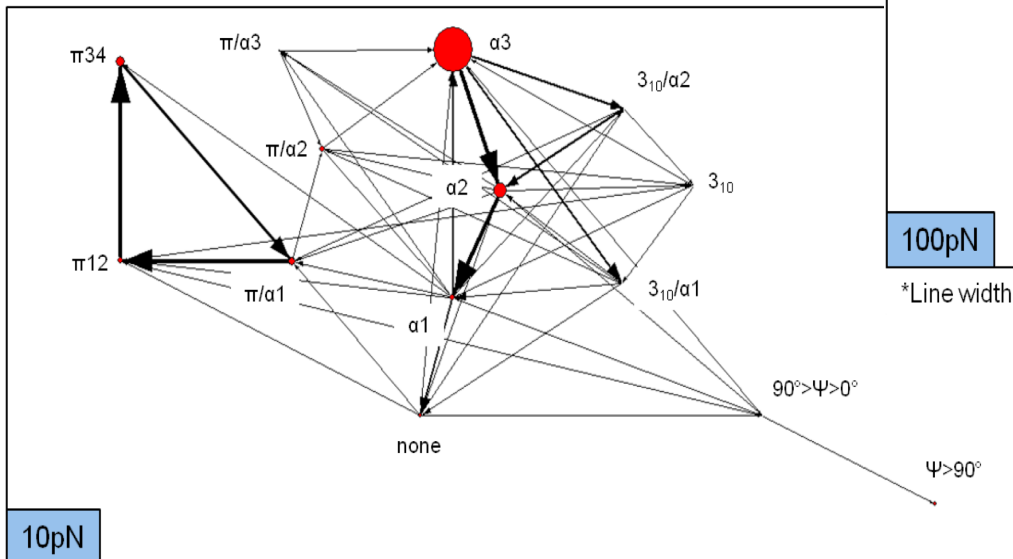
# Steady State Distributions



# Max flux pathway (globally optimized reaction coordinates)

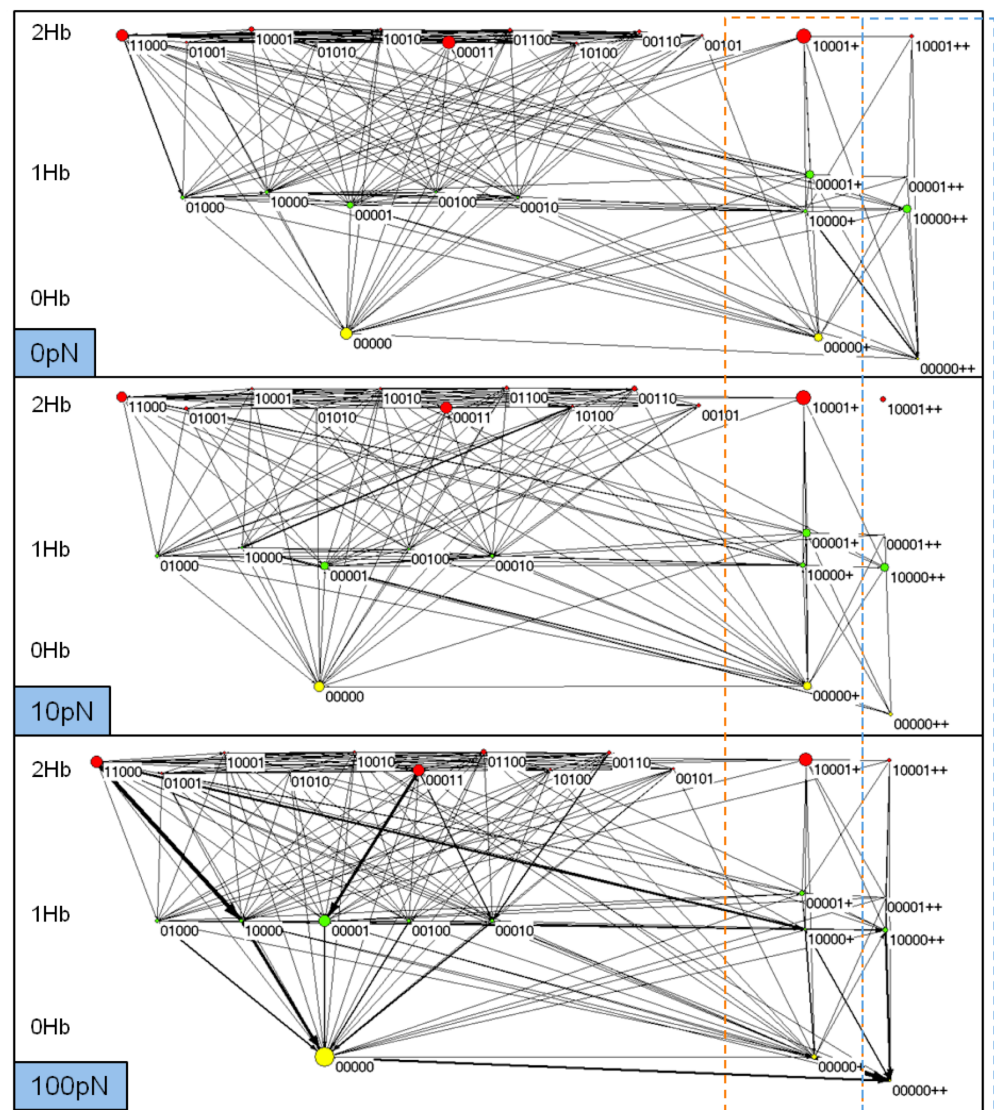


$\pi$  helix off the pathway  
intermediate at 10pN load:  
Slower unfolding



\*Line width normalized within each load.

# More detailed states



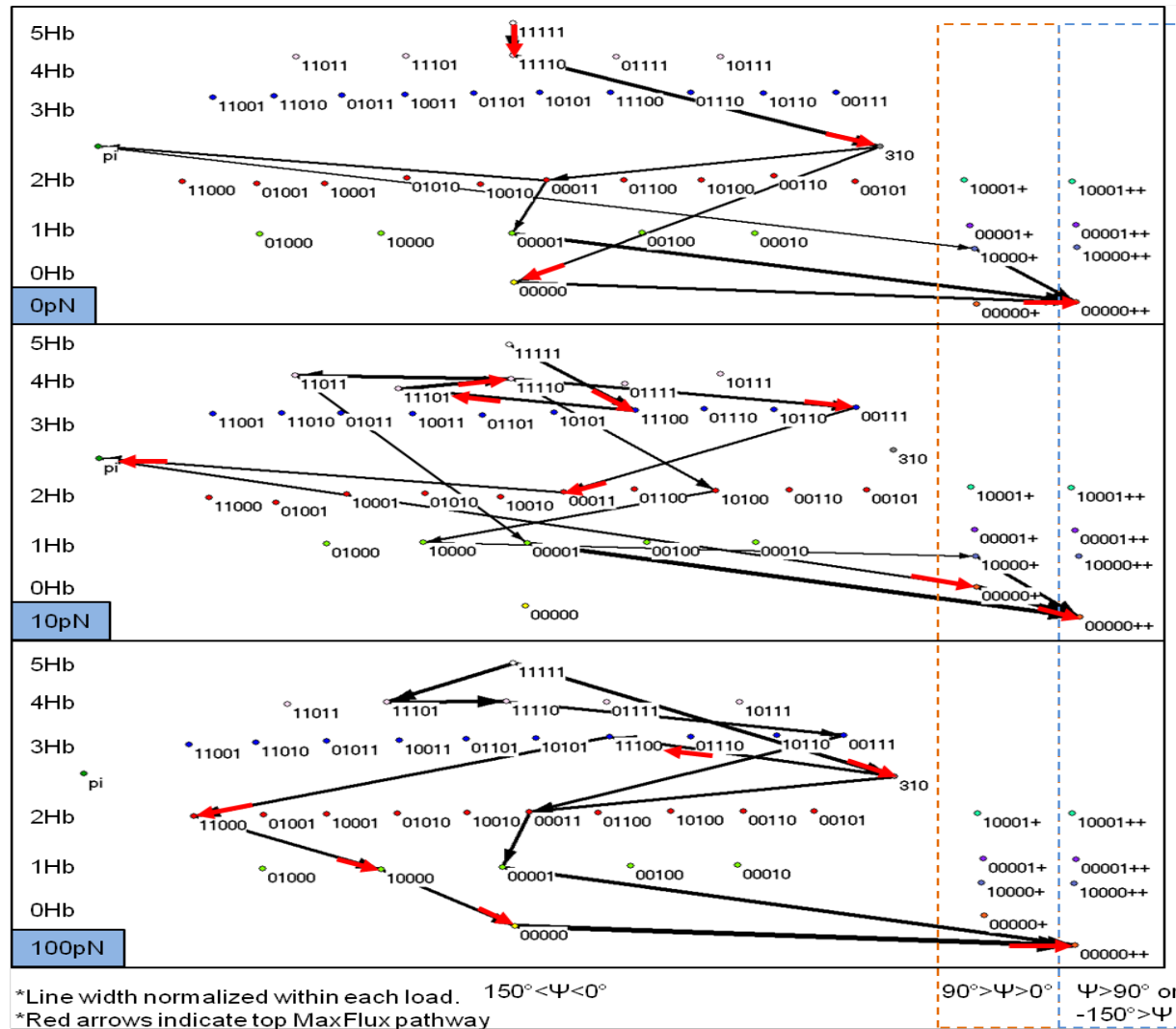
$150^\circ < \Psi < 0^\circ$        $90^\circ > \Psi > 0^\circ$        $\Psi > 90^\circ$  or  $-150^\circ > \Psi$

\*Line width normalized within each load.

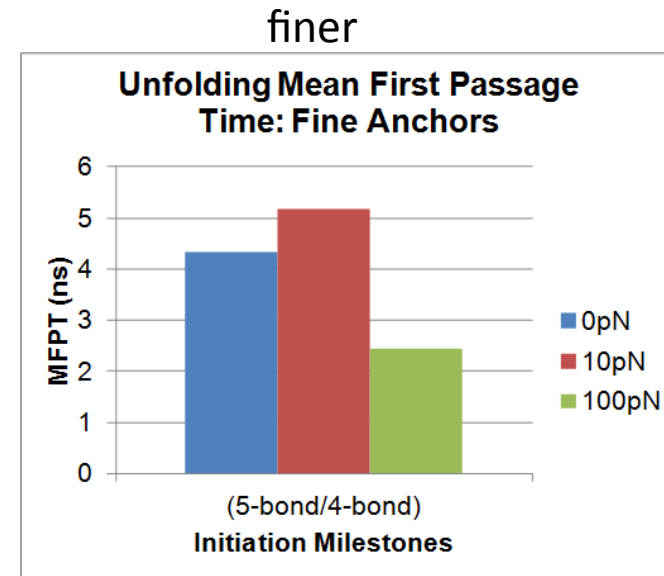
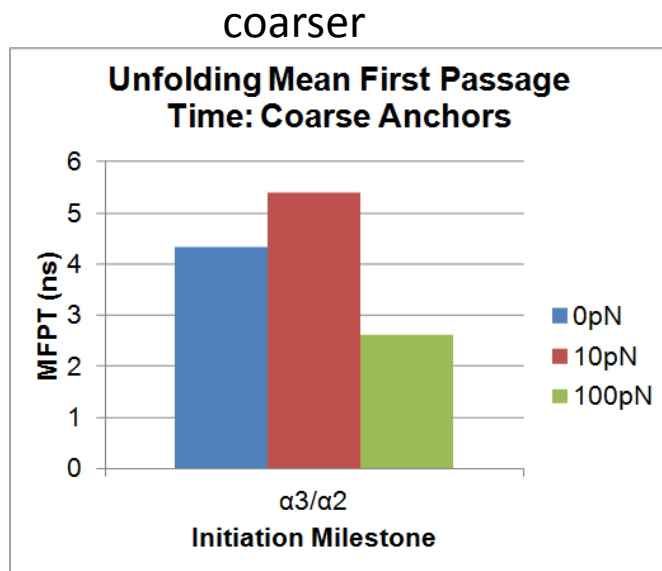
Fine Anchor Designation	Quantity of $\alpha$ -Helical Hbonds	Hydrogen Bond Status (1=Intact; 0=Broken)					$\Psi$ Angle Values
		C4-	C3-	C2-	C-	C	
(# Indicates Hydrogen Bond Pattern)							
32	5	1	1	1	1	1	all
31	4	1	0	1	1	1	all
30	4	0	1	1	1	1	all
29	4	1	1	1	1	0	all
28	4	1	1	1	0	1	all
27	4	1	1	0	1	1	all
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°
$\pi$	NOTE: this fine anchor incorporates any residues with a $\pi$ -helical hydrogen bond (and therefore not in a pure $\alpha$ -helical state)						
310	NOTE: this fine anchor incorporates any residues with a 310-helical hydrogen bond (and therefore not in a pure $\alpha$ -helical state)						



# Pathways with more refined anchors for alpha helix



# Unfolding times at different coarsening



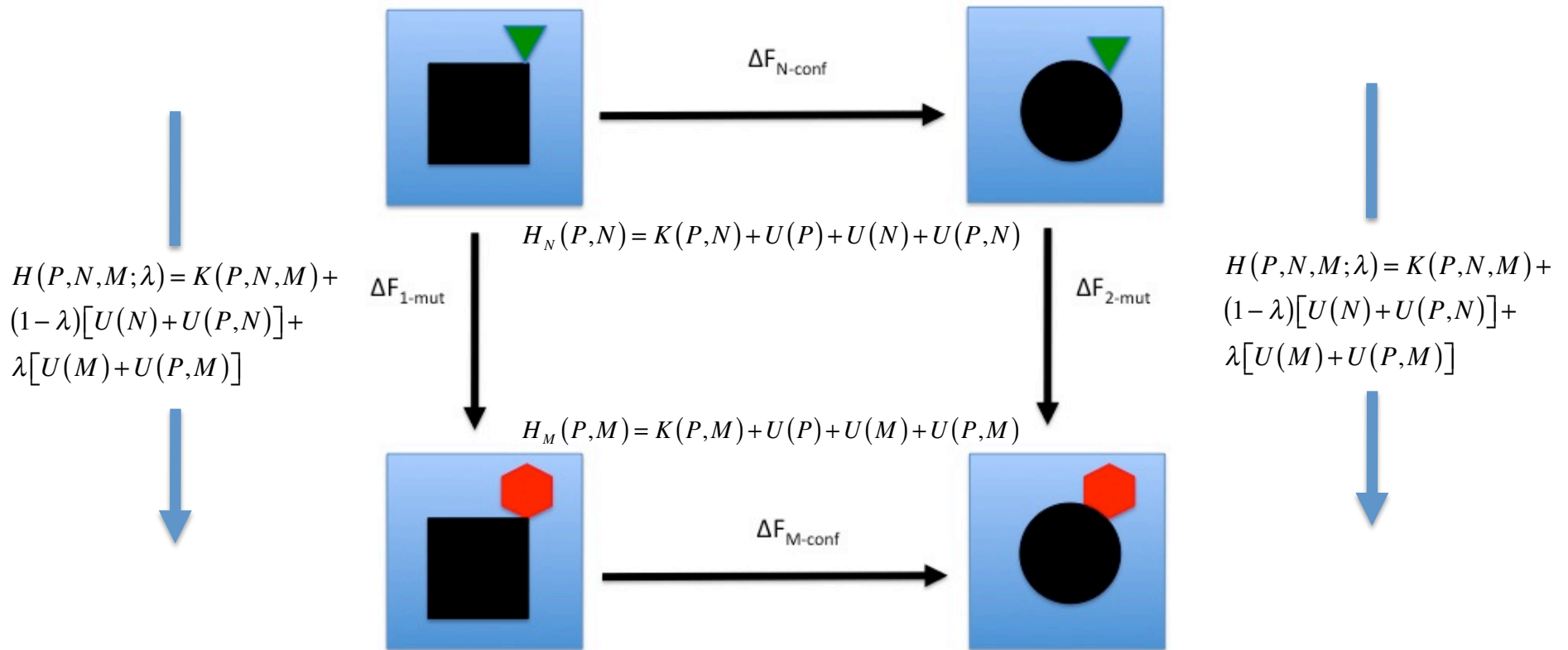
Load Level (pN)	Mean First Passage Time (ns)		Percent Difference
	Coarse Anchors (n=1)	Fine Anchors (n=5)	
0	4.3	4.3	0%
10	5.4	5.2	4%
100	2.6	2.5	7%

# Summary

- Milestoning exploits trajectory fragments between milestones that partition the reaction space. These fragments make it possible to study quantitatively and efficiently the kinetics and thermodynamics of complex systems.
- Code of MOIL:
  - <http://clsb.ices.utexas.edu/prebuilt/>

# Thermodynamic Cycles Retaining the Self Interactions

P = solvent and protein  
 N = native residue  
 M = mutant residue



$$\Delta\Delta F_{\text{exp}} = \Delta F_{M\text{-conf}} - \Delta F_{N\text{-conf}} = \Delta F_{1\text{-mut}} - \Delta F_{2\text{-mut}} = \Delta\Delta F_{\text{mut}}$$

The role of self interactions has been debated in the past...  
 Can we retain ALL the self interactions U(M) U(N) [1-2]?

[1]Boresch, Karplus, *J. Phys. Chem. A* **1999**, 103, 103

[2]Shobana et al, *J. Phys. Chem. B* **2000**, 104, 5179

# Retaining the self-interactions U (N) and U(M)

$$\Delta F_{1-mut} = C_{MN} - \beta^{-1} \ln \frac{\int dX_{PM} \Theta_{P1} \exp\{-\beta[U(P)+U(M)+U(P,M)]\}}{\int dX_{PN} \Theta_{P1} \exp\{-\beta[U(P)+U(N)+U(P,N)]\}} = C_{MN} - \beta^{-1} \ln \frac{Z_{P1M}}{Z_{P1N}}$$

An alchemical Hamiltonian that retains the self-interactions is

$$H(P,N,M;\lambda) = K(P,N,M) + U(P) + (1-\lambda)U(P,N) + U(N) + \lambda U(P,M) + U(M)$$

$$\Delta f_{1-mut} = -\beta^{-1} \ln \frac{\int dX_N \exp[-\beta U(N)]}{\int dX_M \exp[-\beta U(M)]} - \beta^{-1} \ln \frac{\int dX_{PM} \Theta_{P1} \exp\{-\beta[U(P)+U(M)+U(P,M)]\}}{\int dX_{PN} \Theta_{P1} \exp\{-\beta[U(P)+U(N)+U(P,N)]\}} = -\beta^{-1} \ln \frac{Z_N}{Z_M} - \beta^{-1} \ln \frac{Z_{P1M}}{Z_{P1N}}$$

$$\Delta F_{1-mut} = \Delta f_{1-mut} + \Delta F_{MN}$$

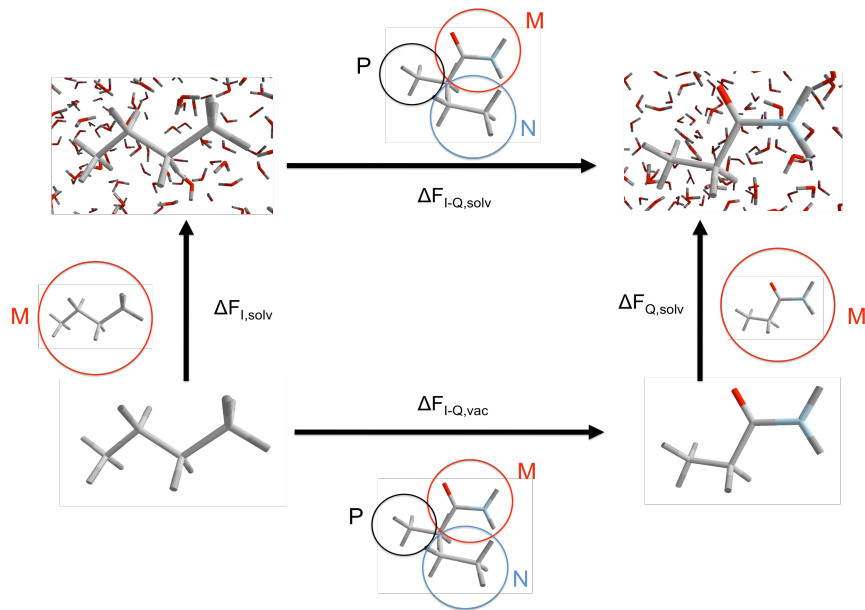
$\Delta F_{MN}$  DOES NOT DEPEND ON "P"

Therefore:

$$\Delta\Delta F_{exp} = \Delta\Delta F_{mut} = \Delta\Delta f_{mut}$$

**WE CAN RETAIN ALL THE SELF-INTERACTIONS, BONDED AND NON-BONDED, AND STILL OBTAIN THE CORRECT RELATIVE FREE ENERGY DIFFERENCE**

# Numerical Illustration: Relative Solvation Free Energy of ILE and GLN side chain analogs[3]



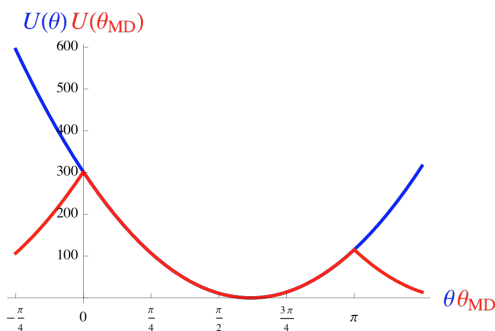
## GENERAL SIMULATION DETAILS

- NVT (T rescaling or Langevin)
- OPLS-AAL and TIP3P (with corrections [3])
- Box Side 34.45Å
- 1355 water molecules
- $\lambda^4$  and soft interactions for VdW [4]
- Separate decoupling Ele and VdW [5]
- Bond "P" - "M" / "N" never scaled [6]
- Total simulation time  $\sim 500$ ns

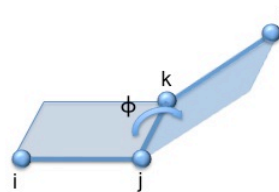
## SPECIAL SIMULATION DETAILS

ANGLES... Use UB

$$U(\theta_{MD}) = k [\arccos(\cos(\theta)) - \theta_0]^2$$



Torsions... remove after angles



Ewald ... multiple calls

1	$\lambda U(P)$
2	$(1-\lambda)[U(P)+U(P,N)+U(N)]$
3	$U(M)$
4	$\lambda U(N)$
TOT	$U(P)+U(N)+U(M)+(1-\lambda)U(P,N)$

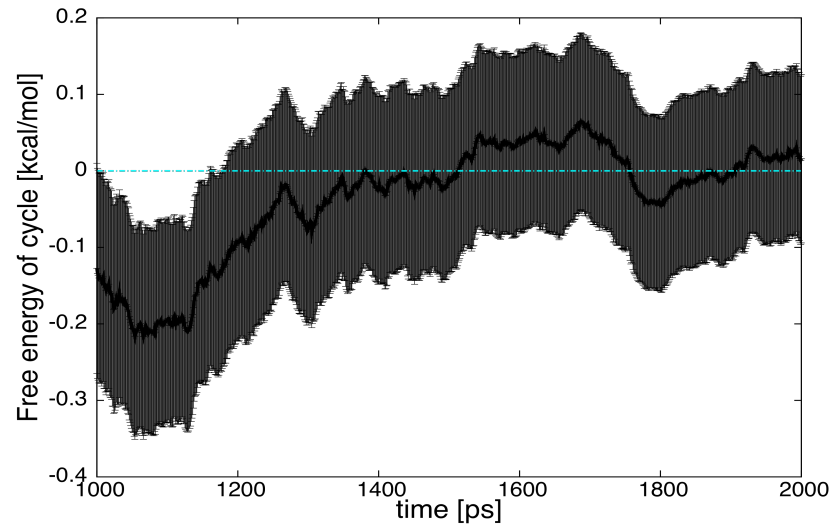
[3] Shirts et al, *J. Chem. Phys.* **2003**, 119, 5740

[4] Beutler et al, *Chem. Phys. Lett.* **1994**, 222, 529

[5] Pohorille et al *J. Phys. Chem. B* **2010**, 110, 10235

[6] Boresch et al, *J. Phys. Chem. B* **2003**, 107, 9535

# Results



	<b>OURS</b>	<b>SIM [3]</b>	<b>Exp[7]</b>
$\Delta F_{I-Q,solv}$	$-14.07 \pm 0.07$		
$\Delta F_{I-Q,vac}$	$-2.85 \pm 0.07$		
$\Delta F_{I,solv}$	$2.89 \pm 0.05$	$2.73 \pm 0.03$	2.15
$\Delta F_{Q,solv}$	$-8.35 \pm 0.05$	$-8.40 \pm 0.04$	-9.38

$$\oint dF = (0.02 \pm 0.12) \text{ kcal / mol}$$

[3] Shirts et al, *J. Chem. Phys.* **2003**, 119, 5740

[7] Wolfenden et al, *Biochemistry* **1981**, 20, 849