Milestoneing: 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

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

- $10^{-9}$ seconds: MD
- $10^{-6}$ seconds: Protein Activation
- $10^{-3}$ seconds: Channel Gating, Fast folding
- 1 second: Slow folding

Activated processes – rare & fast

Milestoneing – slow diffusive

Continuous barrier “problem”
Consider exact trajectories. Can we compute a statistically equivalent picture using trajectory fragments?

Approximate trajectories and “rough sampling” (Leimkuhler)
Many short trajectories are more efficient (Spichty)
Efficiency (more Milestones the better!)

- **Diffusive speedup:** 
  \[ t \sim L^2 \rightarrow t \sim M \left( \frac{L}{M} \right)^2 = \frac{L^2}{M} \]

- **Parallelization speedup:** 
  \[ t \sim \frac{L^2}{M^2} \]

- **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

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

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


Alternatives: Pratt,
LUP / String - Elber, Vanden Eijnden
NEB - Johansson
MaxFlux – Berkowitz, Straub, Skeel
Compute trajectories between interfaces (milestones). If one of the trajectories wonders far from other anchors, add a new one.

Similar to searches in the Markov State Model
Provide “rough” description of energy landscape using Voronoi-like cells and anchors in coarse space

Vanden-Eijnden and Venturoli (J. Chem. Phys. 2009, 130, 194101)
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:

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

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

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

\[
Q_{stat} \left( I - \int_0^\infty K(\tau) d\tau \right) = 0 \quad P_{s,stat} = Q_{s,stat} t_s
\]

\( s,s' \) -- 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**

\[
\tilde{R}_{s,s'} = \frac{u\tilde{K}_{s,s'}(u)}{1 - \sum_{s'} \tilde{K}_{s',s}(u)}
\]
Early Events in Helix Unfolding

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

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

![AA-state v. $\Psi$ Angle: 0pN](image1)

![AA-state v. $\Psi$ Angle: 100pN](image2)

![AA-state v. $\Psi$ Angle: 10pN](image3)
Max flux pathway (globally optimized reaction coordinates)

π helix off the pathway intermediate at 10pN load: Slower unfolding

*Line width normalized within each load.*
More detailed states

*Line width normalized within each load.

<table>
<thead>
<tr>
<th>Fine Anchor Designation</th>
<th>Quantity of α-Helical Hbonds</th>
<th>Hydrogen Bond Status (1=Intact; 0=Broken)</th>
<th>Ψ Angle Values</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(# Indicates Hydrogen Bond Pattern)</td>
<td>C4-</td>
<td>C3-</td>
</tr>
<tr>
<td>32</td>
<td>5</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>31</td>
<td>4</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>30</td>
<td>4</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>29</td>
<td>4</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>28</td>
<td>4</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>27</td>
<td>4</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>26</td>
<td>3</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>25</td>
<td>3</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>24</td>
<td>3</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>23</td>
<td>3</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>22</td>
<td>3</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>21</td>
<td>3</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>20</td>
<td>3</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>19</td>
<td>3</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>18</td>
<td>3</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>17</td>
<td>3</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>16</td>
<td>2</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>15</td>
<td>2</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>14</td>
<td>2</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>13</td>
<td>2</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>12</td>
<td>2</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>11</td>
<td>2</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>10</td>
<td>2</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>9</td>
<td>2</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>9+</td>
<td>2</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>9++</td>
<td>2</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>8</td>
<td>2</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>7</td>
<td>2</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>6</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>4+</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>4++</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>3+</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>3++</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1+</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1++</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

| TT                       | NOTE: this fine anchor incorporates any residues with a τ-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) |
Pathways with more refined anchors for alpha helix
Unfolding times at different coarsening

### Coarser

**Unfolding Mean First Passage Time: Coarse Anchors**

<table>
<thead>
<tr>
<th>Load Level (pN)</th>
<th>Mean First Passage Time (ns)</th>
<th>Percent Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>4.3</td>
<td>0%</td>
</tr>
<tr>
<td>10</td>
<td>5.4</td>
<td>4%</td>
</tr>
<tr>
<td>100</td>
<td>2.6</td>
<td>7%</td>
</tr>
</tbody>
</table>

### Finer

**Unfolding Mean First Passage Time: Fine Anchors**

**Initiation Milestone**

<table>
<thead>
<tr>
<th>Load Level (pN)</th>
<th>Mean First Passage Time (ns)</th>
<th>Percent Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>4.3</td>
<td>0%</td>
</tr>
<tr>
<td>10</td>
<td>5.2</td>
<td>4%</td>
</tr>
<tr>
<td>100</td>
<td>2.5</td>
<td>7%</td>
</tr>
</tbody>
</table>
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

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

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

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

\[
H_N(P,N) = K(P,N) + U(P) + U(N) + U(P,N)
\]

\[
H_M(P,M) = K(P,M) + U(P) + U(M) + U(P,M)
\]

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

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

An alchemical Hamiltonian that retains the self-interactions is

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

$$\Delta F_{1-mut} = C_{MN} - \beta^{-1} \ln \int dX_{PM} \Theta_{P_1} \exp \{-\beta [U(P) + U(M) + U(P,M)]\}$$

$$= C_{MN} - \beta^{-1} \ln \frac{Z_{PM}}{Z_{P,N}}$$

$$\Delta f_{1-mut} = -\beta^{-1} \ln \int dX_N \exp [-\beta U(N)] - \beta^{-1} \ln \int dX_M \exp [-\beta U(M)]$$

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

Therefore:

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

$$\Delta F_{MN}$$

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 $\sim$ 500ns

SPECIAL SIMULATION DETAILS

ANGLES... Use UB

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

Torsions... remove after angles

Ewald ... multiple calls

$\lambda U(P)$

$(1-\lambda)[U(P)+U(P,N)+U(N)]$

$U(M)$

$\lambda U(N)$

$U(P)+U(N)+U(M)+(1-\lambda)U(P,N)$

Results

\[ \int dF = (0.02 \pm 0.12) \text{ kcal/mol} \]

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>( \Delta F_{l-Q,\text{solv}} )</td>
<td>-14.07±0.07</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \Delta F_{l-Q,\text{vac}} )</td>
<td>-2.85±0.07</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \Delta F_{l,\text{solv}} )</td>
<td>2.89±0.05</td>
<td>2.73±0.03</td>
<td>2.15</td>
</tr>
<tr>
<td>( \Delta F_{Q,\text{solv}} )</td>
<td>-8.35±0.05</td>
<td>-8.40±0.04</td>
<td>-9.38</td>
</tr>
</tbody>
</table>