Temperature Accelerated Molecular Dynamics simulations of Biological Molecules

Luca Maragliano
Italian Institute of Technology
luca.maragliano@iit.it

Eric Vanden-Eijnden

Giovanni Ciccotti
Grazia Cottone
Cameron F. Abrams
Harish Vashisth
Therese Malliavin
Edithe Selwa

Outline

Illustration of the method

Applications:

Modeling the CaM-free conformation of adenylate cyclase

CO diffusion in myoglobin

Activation loop conformational transition in insuline receptor kinase

The Free Energy (potential of mean force)

Consider a system $\boldsymbol{x}$ subject to a potential $V(\boldsymbol{x})$

$$\rho(\boldsymbol{x}) \propto e^{-V(\boldsymbol{x})/k_B T}$$

Introduce N functions of the system's coordinates (collective variables)

$$\boldsymbol{\theta}(\boldsymbol{x}) = (\theta_1(\boldsymbol{x}), \ldots, \theta_N(\boldsymbol{x}))$$

The PMF $A_T(\boldsymbol{z})$ associated to the $\boldsymbol{\theta}(\boldsymbol{x})$ variables is defined via their probability density function

$$A_T(\boldsymbol{z}) = -k_B T \ln Z^{-1} \int e^{-V(\boldsymbol{x})/k_B T} \prod \delta(\theta(\boldsymbol{x}) - \boldsymbol{z}) d\boldsymbol{x}$$

Temperature Accelerated Molecular Dynamics

Suppose we could simulate:

$$\bar{\gamma}\dot{z}_i = -\frac{\partial A_T(\boldsymbol{z})}{\partial z_i} + \sqrt{2k_B\bar{T}\bar{\gamma}}\boldsymbol{\eta}(t)$$

The trajectory would sample

$$\rho(\boldsymbol{z}) \propto e^{-A_T(\boldsymbol{z})/k_B\bar{T}}$$

Then to cross over free energy barriers we could take

$$k_B\bar{T} \gtrsim \Delta A_T(\boldsymbol{z})$$

Temperature Accelerated Molecular Dynamics

Suppose we could simulate:

$$\bar{\gamma}\dot{z}_i = -\frac{\partial A_T(\boldsymbol{z})}{\partial z_i} + \sqrt{2k_B\bar{T}\bar{\gamma}}\boldsymbol{\eta}(t)$$

The negative gradient of the PMF (mean force) can be computed locally via an expectation on x, conditional on $\theta(x) = z$.

Multi-scale approach:

- given a point z, compute the mean force locally from MD of x variables constrained (or restrained) at $\theta(x) = z$

- evolve z

$\longrightarrow$ re-initialization problems for the x

Temperature Accelerated Molecular Dynamics

Suppose we could simulate:

$$\bar{\gamma}\dot{z}_i = -\frac{\partial A_T(\boldsymbol{z})}{\partial z_i} + \sqrt{2k_B\bar{T}\bar{\gamma}}\boldsymbol{\eta}(t)$$

Seamless scheme (E, Ren, Vanden-Eijnden J. Comput. Physics (2009)):

evolve z and x concurrently, with two different time-scales, exchanging data at every step.

Temperature Accelerated Molecular Dynamics

L. M. & E. Vanden-Eijnden *Chem. Phys. Lett.*, 426, 168 (2006)

Consider the system of equations

$$\begin{cases} \gamma \dot{x}_i = -\dfrac{\partial V(\boldsymbol{x})}{\partial x_i} - \kappa \displaystyle\sum_{j=1}^{N} (\theta_j(\boldsymbol{x}) - z_j) \frac{\partial \theta_j(\boldsymbol{x})}{\partial x_i} + \sqrt{2\beta^{-1}\gamma}\, \eta_i^x(t), \\ \bar{\gamma} \dot{z}_j = \kappa(\theta_j(\boldsymbol{x}) - z_j) + \sqrt{2\bar{\beta}^{-1}\bar{\gamma}}\, \eta_j^z(t) \end{cases}$$

Where $\boldsymbol{\eta}^x$ and $\boldsymbol{\eta}^z$ are independent white noises, γ and $\bar{\gamma}$ are different frictions, and β^{-1}, $\bar{\beta}^{-1}$ are different temperatures.

They describe the evolution of the system $(\boldsymbol{x}, \boldsymbol{z}) \in \mathbb{R}^n \times \mathbb{R}^N$ under the potential

$$U_\kappa(\boldsymbol{x}, \boldsymbol{z}) = V(\boldsymbol{x}) + \frac{1}{2}\kappa \sum_{j=1}^{N} (z_j - \theta_j(\boldsymbol{x}))^2.$$

Temperature Accelerated Molecular Dynamics

L. M. & E. Vanden-Eijnden *Chem. Phys. Lett.*, 426, 168 (2006)

Consider the system of equations

$$\begin{cases} \gamma \dot{x}_i = -\dfrac{\partial V(\boldsymbol{x})}{\partial x_i} - \kappa \displaystyle\sum_{j=1}^{N} (\theta_j(\boldsymbol{x}) - z_j) \dfrac{\partial \theta_j(\boldsymbol{x})}{\partial x_i} + \sqrt{2\beta^{-1}\gamma}\, \eta_i^x(t), \\ \bar{\gamma} \dot{z}_j = \kappa(\theta_j(\boldsymbol{x}) - z_j) + \sqrt{2\bar{\beta}^{-1}\bar{\gamma}}\, \eta_j^z(t) \end{cases}$$

When $\gamma/\bar{\gamma} \longrightarrow 0$, the dynamics of $\boldsymbol{z}(t)$ is approximately

$$\bar{\gamma} \dot{z}_j = -\frac{\partial A_{\kappa,T}(\boldsymbol{z})}{\partial z_j} + \sqrt{2\bar{\beta}^{-1}\bar{\gamma}} \eta_j^z(t)$$

Where

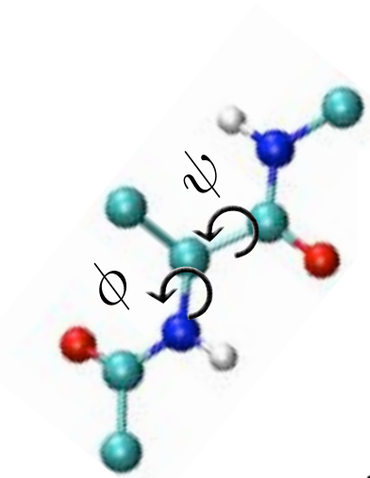
$$A_{\kappa,T}(\boldsymbol{z}) = -\beta^{-1} \ln Z^{-1} \int \exp\left(-\beta V(x) - \frac{1}{2}\beta\kappa \sum_{j=1}^{N} (z_j - \theta(\boldsymbol{x}))^2 \right) dx$$

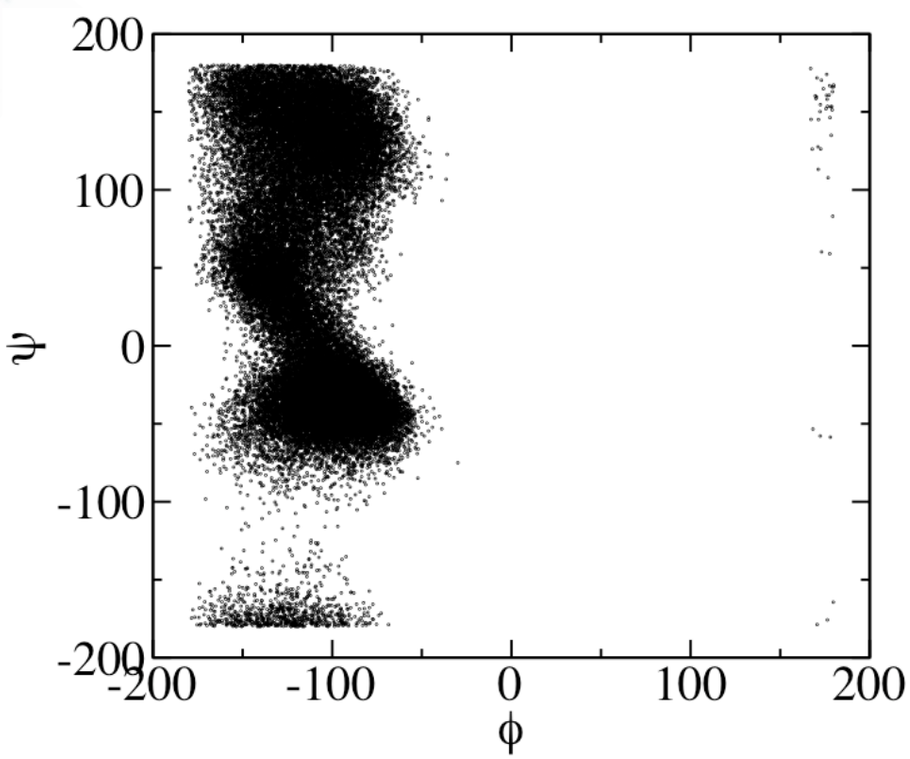
$$\longrightarrow A_T(\boldsymbol{z}) \quad \text{as} \quad \kappa \longrightarrow \infty$$

The limiting equation is a result of standard averaging theorems (G.C. Papanicolaou, Rocky Mt. J. Math. 6, 653 (1976); G.C. Papanicolaou, in: R.C. Di Prima (Ed.), Lect. Appl. Math., 16, American Mathematical Society, 1977; E. Vanden-Eijnden, Comm. Math. Sci. 1, 2003)

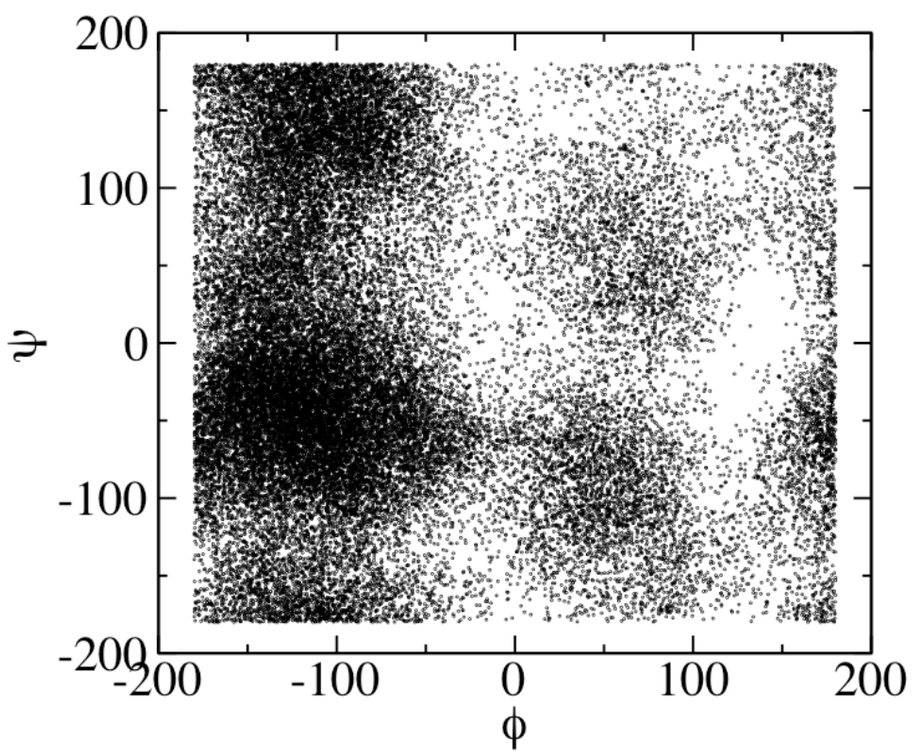
Temperature Accelerated Molecular Dynamics

Simulate the coupled system with $\gamma/\bar{\gamma}$ small and κ large to explore the PMF landscape $A_T(\boldsymbol{z})$. Use $\bar{\beta}^{-1} > \beta^{-1}$ to cross energy barriers prohibitive for uncoupled $\boldsymbol{\theta}(\boldsymbol{x}(t))$





40ns of standard MD simulation



40ps of TAMD with $\bar{T}/T \approx 17$

Temperature Accelerated Molecular Dynamics

Advantages of TAMD:

Untargeted exploration

Can be used with many collective variables

TAMD borrows ideas from other enhanced sampling methods:

Metadynamics (Laio & Parrinello, PNAS 2002 ...)

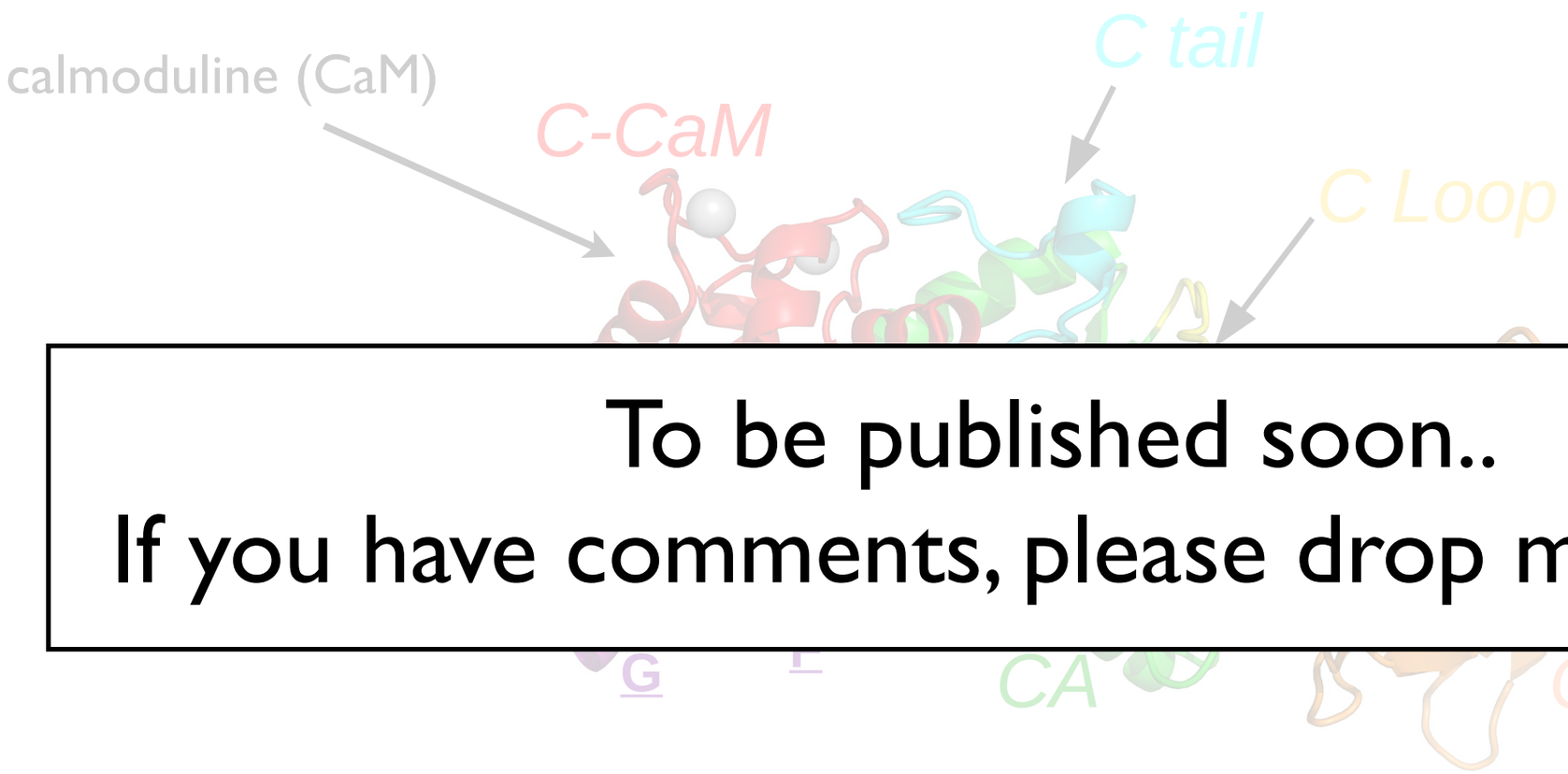
AFED (Tuckerman and coworkers: Rosso et al. JCP 2002 ...)

Methods thought to be used to reconstruct $A_T(\boldsymbol{z})$ by direct sampling.

Rather: use TAMD for a quick exploration of the PMF surface, and then reconstruct it at a second stage with different methods.

TAMD simulation of *Bordetella pertussis* CyaA conformational transition

with E. Selwa, G. Ciccotti and T. Malliavin



To be published soon..
If you have comments, please drop me an email

CyaA is a key virulence factor of B.p., the bacterium causing whooping cough. Its toxic activity is regulated by calmoduline binding (in red).

In order to understand the molecular basis of activation, it is important to inspect the conformation of CyaA before interaction with CaM, but a crystal structure of the isolated CyaA is still missing.

Reconstruction of PMF surfaces and reaction pathways

Combine TAMD with other methods

- Single Sweep for PMF surface reconstruction

- String method for finding reaction pathways

Single-sweep method for PMF calculations

L. M. & E. Vanden-Eijnden *J. Chem. Phys.*, 128, 184110 (2008)

Three separate, independent stages:

1) Use TAMD to rapidly explore the unknown PMF landscape

2) Compute the gradient of the PMF (a.k.a mean force) at points selected from the TAMD trajectory

3) Use an interpolation/variational method to reconstruct globally the PMF from the mean force data (no more MD)

Advantages: does not rely on histograms; the computational effort is concentrated on the mean force calculations: simulations independent from each other that are *distributed* on clusters.

Stage 2: Computing the mean force

Extract points z_k from TAMD trajectory and simulate

$$\gamma \dot{x}_i = -\frac{\partial V(\boldsymbol{x})}{\partial x_i} - \bar{\kappa} \sum_{j=1}^{N} (\theta_j(\boldsymbol{x}) - z_{k,j}) \frac{\partial \theta_j(\boldsymbol{x})}{\partial x_i} + \sqrt{2\beta^{-1}\gamma}\, \eta_i^x(t)$$

Now with z_k fixed!

Estimate the mean force from

$$\boxed{\boldsymbol{f}_k = \frac{1}{T} \int_0^T \bar{\kappa} \left(\boldsymbol{\theta}(\boldsymbol{x}(t)) - \boldsymbol{z}_k \right) dt} \approx -\nabla_z A(\boldsymbol{z})$$

$\bar{\kappa}$ can be chosen in such a way that one gets accurate estimate but without making the system too stiff

Interpolation using Radial Basis Functions (Single-Sweep)

L. M. & E. Vanden-Eijnden *J. Chem. Phys.*, 128, 184110 (2008)

We introduce a RBF representation of the PMF

$$\tilde{A}(\boldsymbol{z}) = \sum_{k=1}^{K} a_k \varphi_\sigma(|\boldsymbol{z} - \boldsymbol{z}_k|)$$

and determine the unknown parameters by optimizing an objective function defined as the difference between the calculated gradients and those from the representation

$$E(a,\sigma) = \sum_{k=1}^{K} \left| \nabla_z \tilde{A}(\boldsymbol{z}_k) + \boldsymbol{f}_k \right|^2 = \sum_{k=1}^{K} \left| \sum_{k'=1}^{K} a_{k'} \nabla_z \varphi_\sigma(|\boldsymbol{z}_k - \boldsymbol{z}_{k'}|) + \boldsymbol{f}_k \right|^2$$

Different RBF can be used. A typical choice are gaussian functions $\varphi_\sigma(u) = \exp\left(-\frac{u^2}{2\sigma^2}\right)$

(For the case of different σ see M. Monteferrante, S. Bonella, S. Meloni, G. Ciccotti *Mol. Sim.*, 35, 1116 (2009))

The centers do not have to lie on a regular grid: the method can be used in more than 2 dimensions. In our case, centers are extracted from the TAMD trajectory.