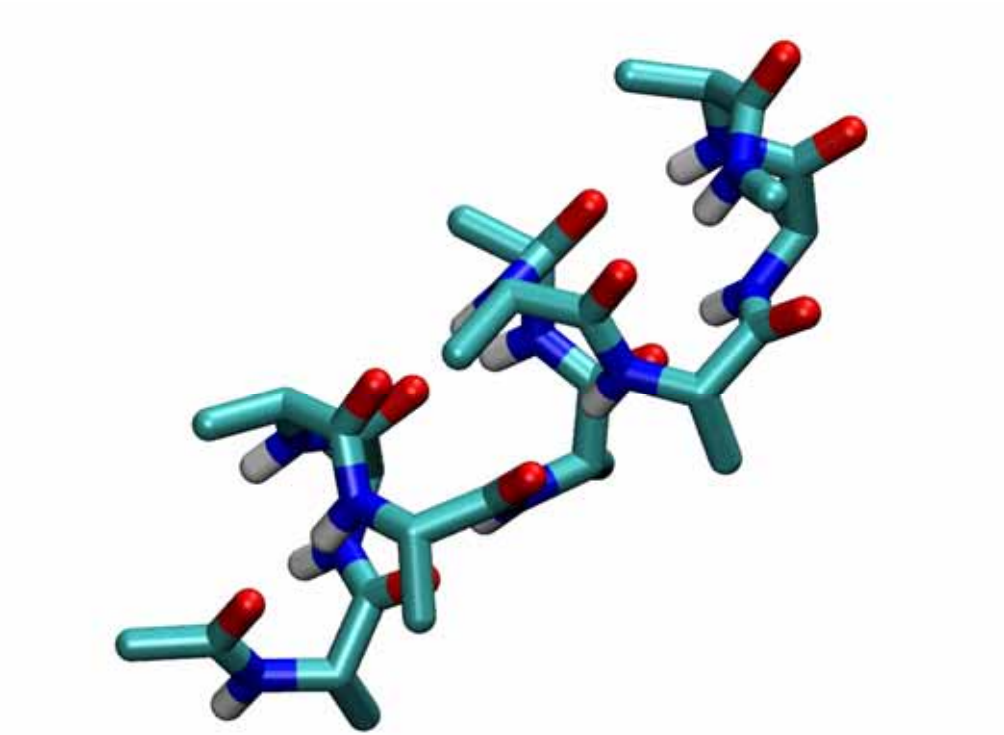
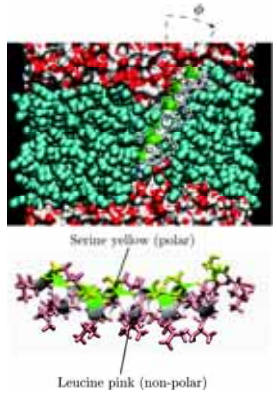
Adaptive Methods for Free Energy Computation and Coarse-graining Strategies

Eric Darve, Stanford University

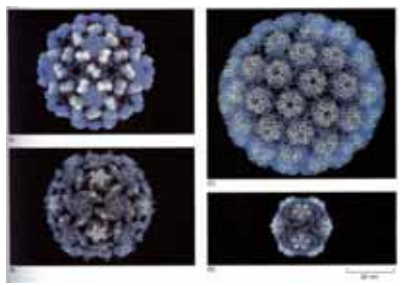


Outline

Free energy computation



Coarse graining strategies



Classical model of proteins

Proteins are modeled using empirical force fields which approximate the true quantum interaction field.

The Hamiltonian function has the form:

$$H = \frac{p^T \cdot p}{2} + U(q)$$

$U(q)$ is decomposed as a sum of potentials modeling bond stretch, bond angles, dihedral angles, Lennard-Jones and electrostatic forces.

Canonical ensemble

We will assume that we are dealing with an *NVT* system.

In that case, the probability density function, i.e., probability of being at (p,q), is given by:

$$\frac{e^{-H(p,q)}}{Z} = \frac{e^{-H(p,q)}}{\int e^{-H(p,q)}\, dpdq}$$

Other ensembles are often used: *NVE*, *NPT*, etc.

Free energy is related to reaction equilibrium constants

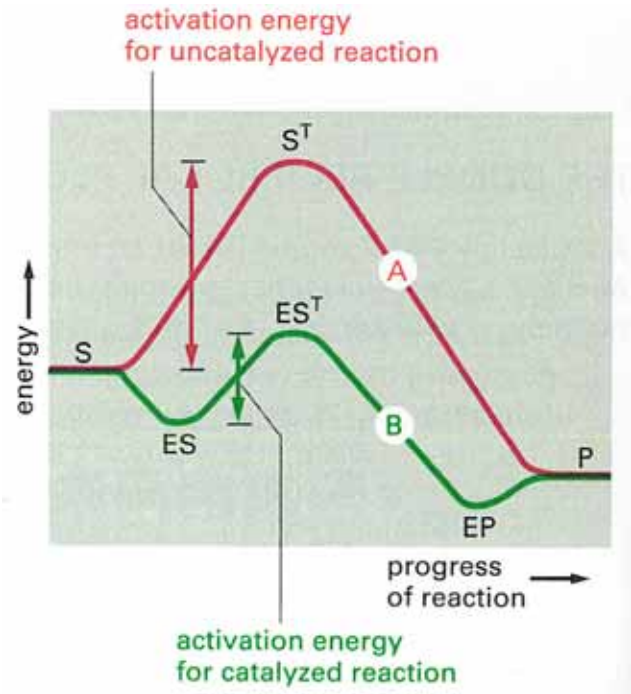
$$\frac{[\mathrm{AB}]}{[\mathrm{A}][\mathrm{B}]} = K = \text{equilibrium constant} = \mathrm{e}^{-\Delta F}$$

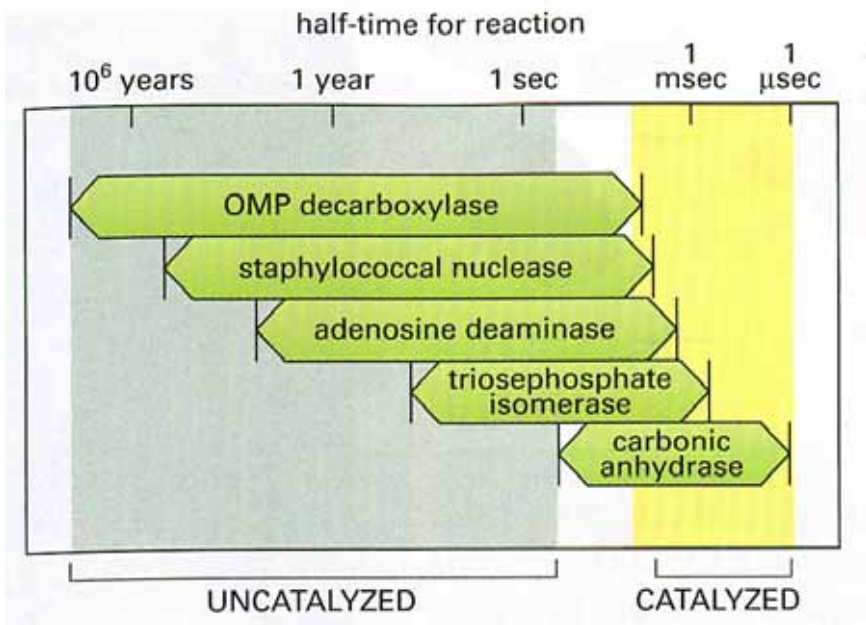
Example: 1000 molecules of A and B in a cell = concentration of 10^{-9} M.

Assume ΔF = – 14.2 kcal/mole.

At equilibrium:	**270**	**270**	**730**
	A	**B**	**AB**
Increase ΔF by 2.8 kcal/mole:	**915**	**915**	**85**
	A	**B**	**AB**

Principles of catalysis





Without catalysis some reactions essentially do not occur.

Definition of reaction coordinates and order parameters

Function of atom coordinates: $\xi(q)$

Characterizes the conformation of the system

Examples:

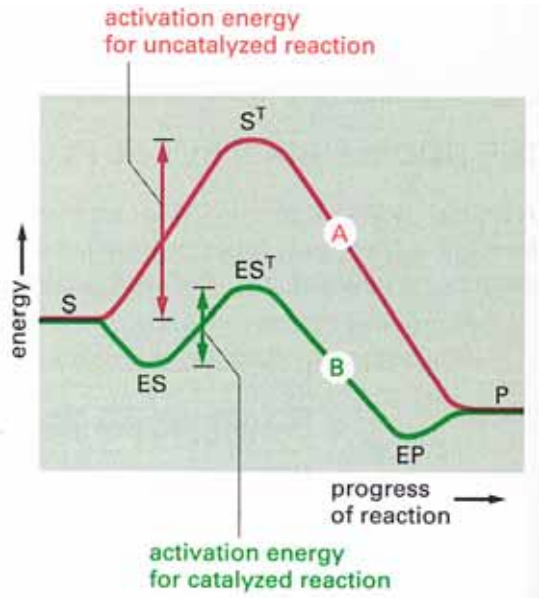
- **Torsion angle along the backbone**
- **End-to-end distance in a protein**
- **Distance between two molecules**
- **Relative orientation of two molecules**
- **Hydration number**

Mathematical definition of free energy

Definition:

$$e^{-A(\xi^*)} = \int \delta(\xi(q) - \xi^*) e^{-H}\, dpdq = Z\, P(\xi^*)$$

From the free energy, equilibrium constants can be obtained and reaction rates can be approximated (transition state theory).

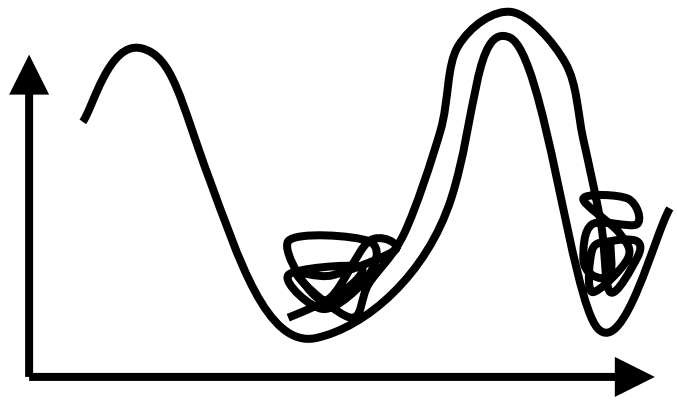


The histogram method

Run a long MD simulation, bin ξ and estimate $P(\xi)$:

$$A(\xi^*) = -\ln P(\xi^*) + \text{constant}$$

Difficulty: at the top of a barrier, sampling will be insufficient $\rightarrow$ large statistical error



References on thermodynamic integration

Method of constraints

E. Carter, G. Ciccotti, J. Hynes, R. Kapral,1989

W. K. den Otter, W. J. Briels, 1998

Adaptive Biasing Force

E. Darve, A. Pohorille, 2001

D. Rodriguez-Gomez, E. Darve, A. Pohorille, 2004

C. Chipot, J. Hénin, 2005

E. H. Lee, J. Hsin, O. Mayans, K. Schulten, 2007

C. Chipot, A. Pohorille, Free energy calculation: theory and applications in chemistry and biology, 2006

Thermodynamic integration

Instead of computing $A(\xi)$ directly, we compute $dA/d\xi$:

$$\frac{dA}{d\xi} = \left\langle \frac{\partial H}{\partial \xi} \middle| \xi \right\rangle = \left\langle \frac{\partial U}{\partial \xi} + \frac{\partial \ln |J|}{\partial \xi} \middle| \xi \right\rangle$$

Requires defining generalized coordinates.

Better:

$$\frac{dA}{d\xi} = -\left\langle \frac{d}{dt}\left(m_\xi \frac{d\xi}{dt} \right) \middle| \xi \right\rangle \stackrel{\text{def}}{=} -\left\langle F_\xi \middle| \xi \right\rangle$$

$$m_\xi^{-1} = |\nabla \xi|^2$$

Adaptive method to improve the efficiency of the sampling strategy

By itself, the previous formula does not improve sampling.

The adaptive biasing force scheme can be applied to improve sampling.

Key idea: add an external biasing force, opposing the mean force.

Define the following running average:

$$F^{\mathbf{ABF}}(t) = -\frac{\int_0^t F(\tau)\ \delta(\xi(t) - \xi(\tau))\, d\tau}{\int_0^t \delta(\xi(t) - \xi(\tau))\, d\tau}$$

$$F(t) = F_\xi(p(t), q(t)) - F^{\mathbf{ABF}}(t)$$

Applying the biasing force

The new equations of motion become:

$$\frac{dq}{dt} = p$$

$$\frac{dp}{dt} = -\nabla U + F^{\mathbf{ABF}}(t)\,\nabla\xi$$

Upon convergence:

$$F^{\mathbf{ABF}}(t) \;\rightarrow \frac{dA}{d\xi}(\xi(t))$$

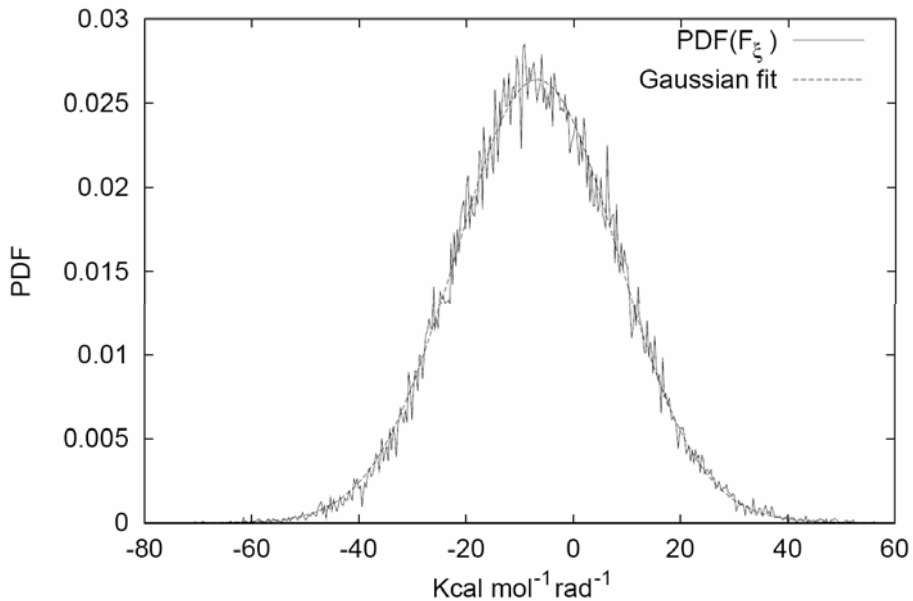
ξ has a diffusive motion in a flat free energy profile $\rightarrow$ fast decay of statistical error

Statistical error and distribution of force

The statistical error decays like $1/\sqrt{N}$:

$$\delta(A^{\mathbf{ABF}}) \approx \Delta\xi \frac{\delta(F_\xi)}{\sqrt{N}}(1+2\kappa)^{1/2}$$

$$\kappa : \text{correlation length of } F_\xi(t)$$



Algorithm

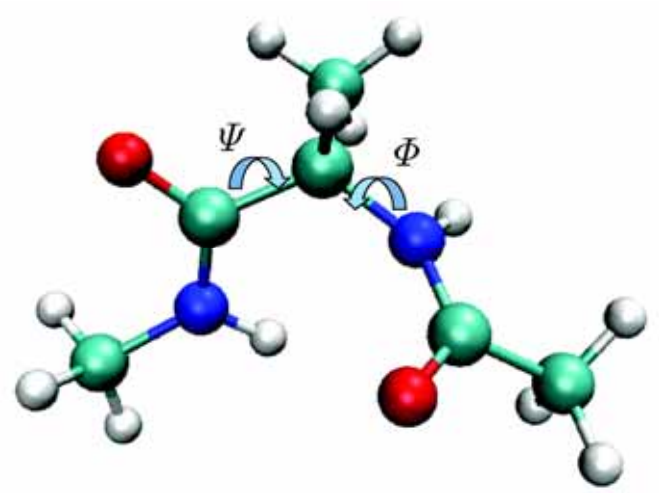
$\boldsymbol{q} \leftarrow \boldsymbol{q}_0$ // Position
$\boldsymbol{v} \leftarrow \boldsymbol{v}_{-1/2}$ // Velocity
Loop over time steps $i = 1, \cdots, N$
  $\boldsymbol{a} \leftarrow -\nabla U(\boldsymbol{q})$
  $\boldsymbol{a} \leftarrow \mathrm{ABF}(i-1, \Delta t, \boldsymbol{a}, \boldsymbol{q}, \boldsymbol{v})$
  $\boldsymbol{v} \leftarrow \boldsymbol{v} + \Delta t\, \boldsymbol{a}$ // Advance the velocity
  $\boldsymbol{q} \leftarrow \boldsymbol{q} + \Delta t\, \boldsymbol{v}$ // Advance the position
End loop

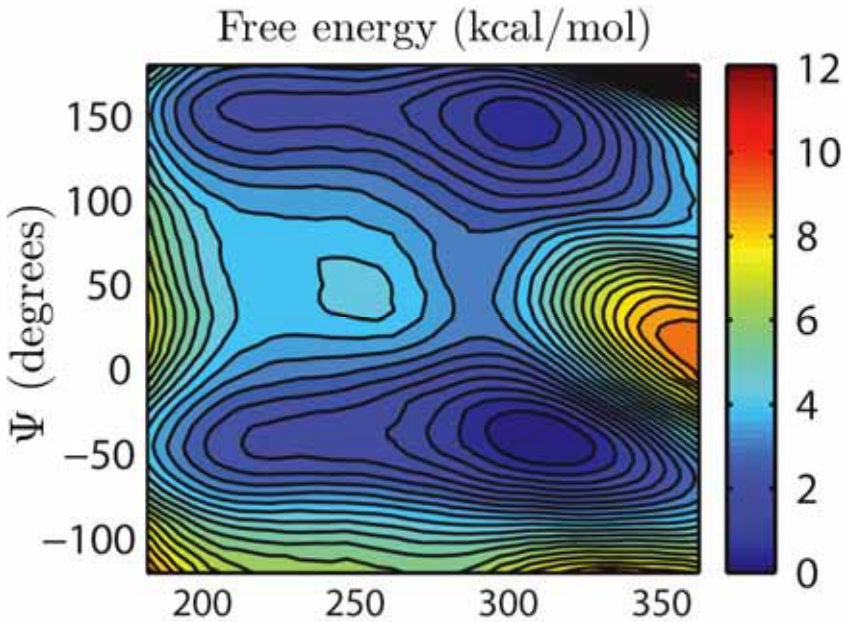
Subroutine ABF(i, Δt, $\boldsymbol{a}$, $\boldsymbol{q}$, $\boldsymbol{v}$).

Save n, F, F_0, $\boldsymbol{g}_0$, $\boldsymbol{a}_0$

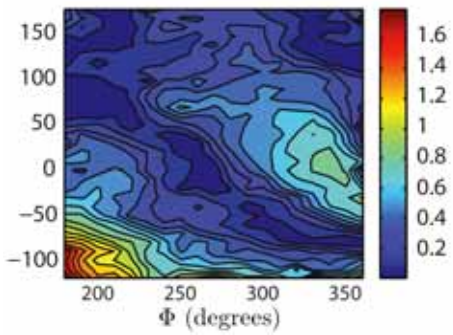
$k \leftarrow$ bin corresponding to $\xi(\boldsymbol{q})$
$F^{\mathsf{ABF}} \leftarrow F(k)/n(k)$ // Compute current estimate of $dA/d\xi$
$\boldsymbol{a} \leftarrow \boldsymbol{a} + F^{\mathsf{ABF}} \nabla\xi(\boldsymbol{q})$ // Apply biasing force
$\boldsymbol{g} \leftarrow m_\xi(\boldsymbol{q})\nabla\xi(\boldsymbol{q})$ // Save this product
if $i \geq 1$:
  // Estimate of force for step $i - 1/2$
  $F_0 \leftarrow (1/\Delta t)(\boldsymbol{g} - \boldsymbol{g}_0) \cdot \boldsymbol{v} + 1/4\,(\boldsymbol{g} + \boldsymbol{g}_0) \cdot (\boldsymbol{a} + \boldsymbol{a}_0) - 1/2\,(F_0 + F^{\mathsf{ABF}})$
  $k_0 \leftarrow$ bin at step $i - 1/2$
  Increment $n(k_0)$ by 1 // Increment counter
  Increment $F(k_0)$ by $-F_0$ // Add new sample to array $F(k)$
End if
$F_0 \leftarrow F^{\mathsf{ABF}}$ // Save F^{ABF} for next step
$\boldsymbol{g}_0 \leftarrow \boldsymbol{g}$ // Save $\boldsymbol{g}$ for next step
$\boldsymbol{a}_0 \leftarrow \boldsymbol{a}$ // Save $\boldsymbol{a}$ for next step

Example 1: alanine dipeptide

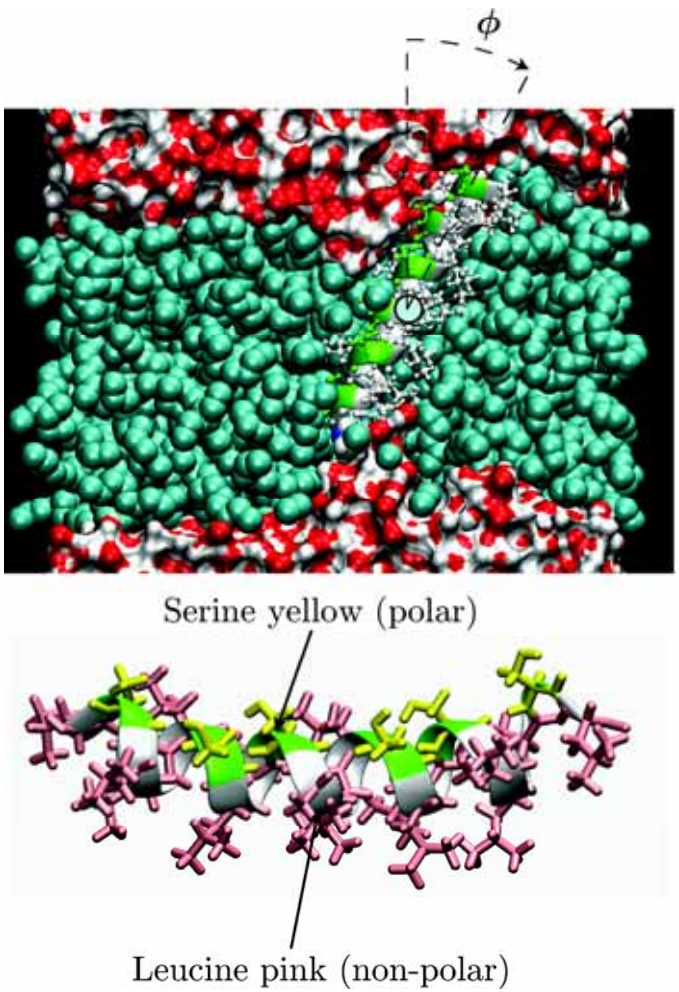


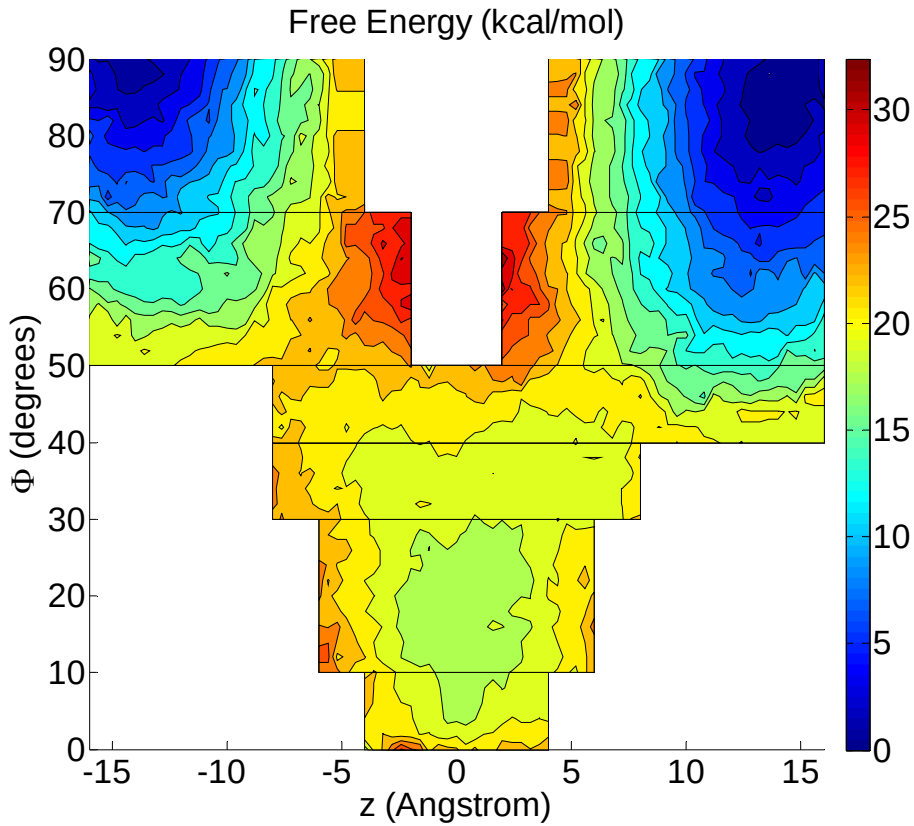


Error



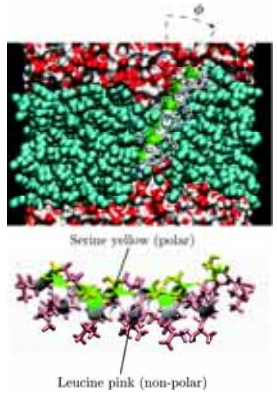
Example 2: LS peptide



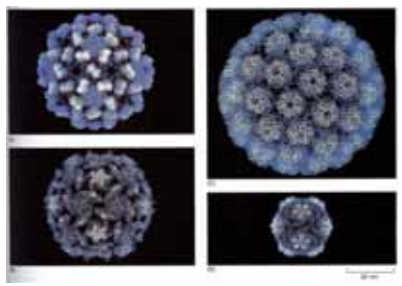


Outline

Free energy computation

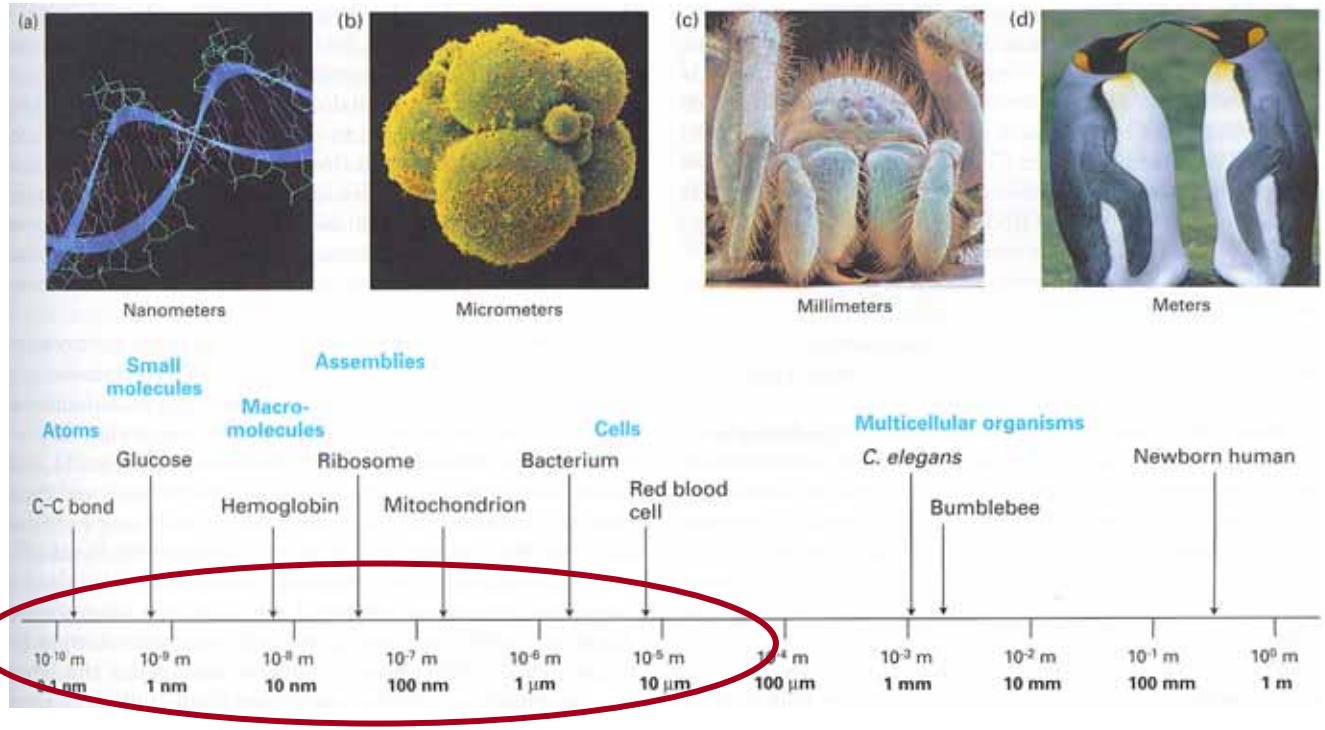


Coarse graining strategies



Biology is a multiscale problem

Atomistic computer models



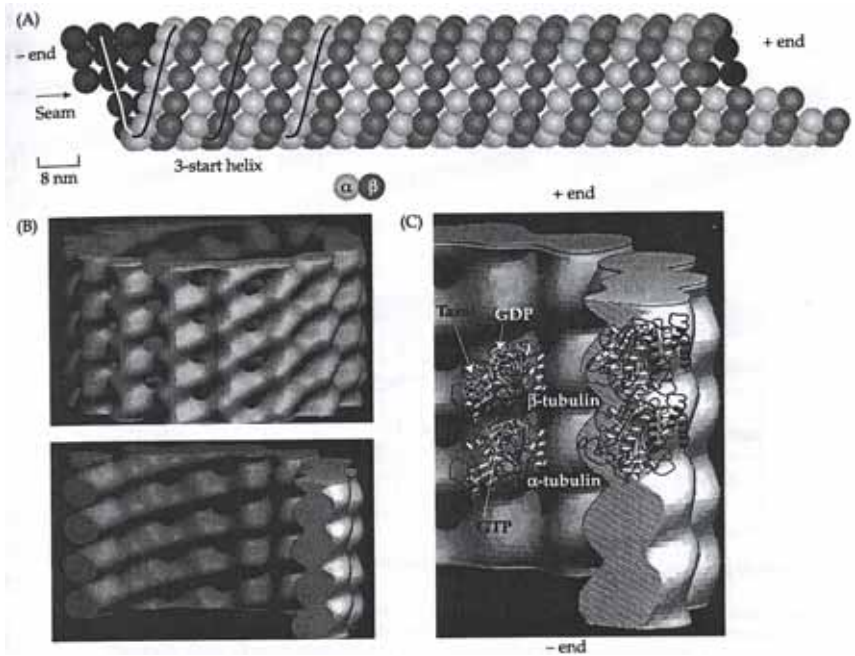
DNA double helix: 2 nm
Eight cells in an embryo: 200 μm
Wolf spider: 15 mm
Emperor penguin: 1 m

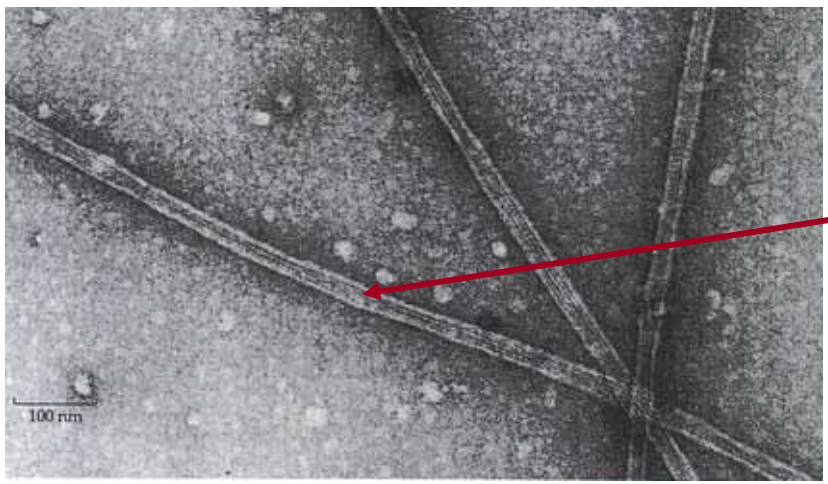
The coarse graining problem

Example: $\alpha-$ *and* $\beta-$ tubulin dimer.

See VMD of 1JFF.pdb

Dimers assemble to form a long microtubule.



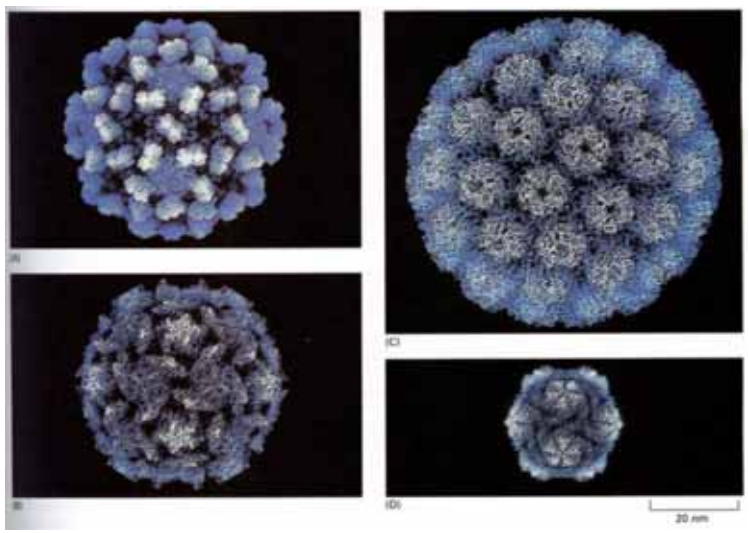


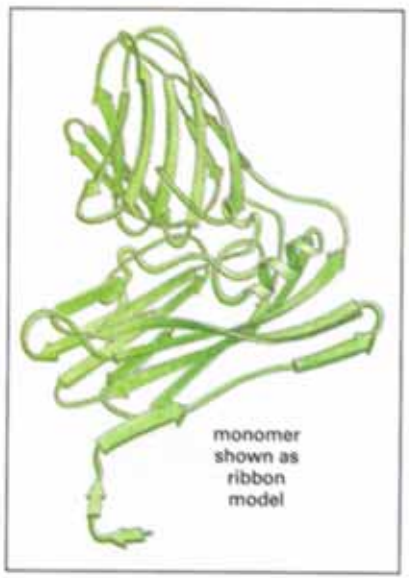
Individual protofilament

Electron micrographs of cytoskeletal filaments.

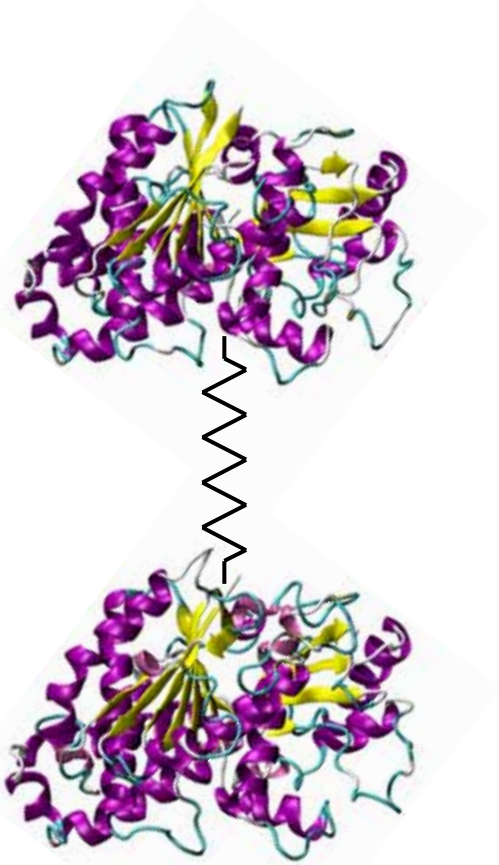
Many structures are formed by assembling many identical motifs

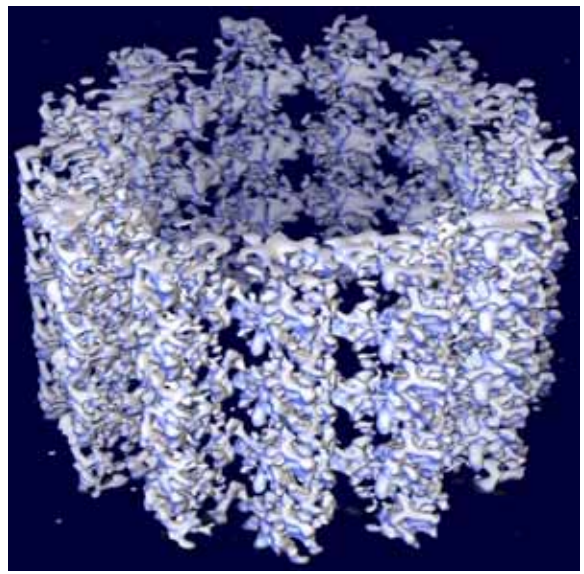
Example: virus capsid. Hollow sphere formed by hundreds of identical structures that enclose the viral nucleic acid.





Building a stochastic model of a large molecular assembly





K. Downing, cryo-EM of microtubule

References on Mori-Zwanzig projector

Mori-Zwanzig theory

H. Mori, 1965
J. M. Deutch, I. Oppenheim, 1971
R. Zwanzig, 1980
D. Evans, G. Morriss, 1990

Recent work and optimal prediction

A. J. Chorin, A. P. Kast, R. Kupferman, 1998
O. H. Hald, 1999
A. J. Chorin, O. H. Hald, R. Kupferman, 2000
R. L. C. Akkermans, W. J. Briels, 2000
E. Vanden-Eijnden, 2001
W. Just, H. Kantz, C. Rödenbeck, M. Helm, 2001
W. Just, K. Gelfert, N. Baba, A. Riegert, H. Kantz, 2003
P. Stinis, 2004
Review: D. Givon, R. Kupferman, A. Stuart, 2004
R. L. C. Akkermans, 2005

Mori-Zwanzig formalism

Problem statement: is it possible to calculate the evolution of a set of variables ("resolved" variables) "without" evolving all the variables?

Answer: yes, if we agree to some approximations. This is the coarse graining problem.

Mori-Zwanzig is a framework to derive stochastic equations for certain variables by averaging "out" the other variables in the system.

Based on the projector formalism: write all quantities in terms of the resolved variables Γ :

$$f(p,q) \;\rightarrow\; \mathcal{P}f = \langle f|\Gamma\rangle$$

Projector decomposition

Any evolution can be decomposed in the following way:

$\mathcal{L}$: Liouvillian (linear differential operator)

$$\mathcal{P}^{\perp} = \mathcal{I} - \mathcal{P}, \qquad \mathcal{L}^{\perp} = (\mathcal{I} - \mathcal{P})\mathcal{L}$$

$$e^{\mathcal{L}t} = e^{\mathcal{L}t}\mathcal{P} + \int_0^t e^{\mathcal{L}(t-s)}\,\mathcal{P}\,\mathcal{L}\,e^{\mathcal{L}^{\perp}s}\mathcal{P}^{\perp}\,ds + e^{\mathcal{L}^{\perp}t}\mathcal{P}^{\perp}$$

Drift term **“Friction” with memory** **Fluctuating term**

Decomposition of rate of change of resolved variables

Using the previous projector formalism, we can decompose any function $f(p,q)$ into 3 parts:

$$f(t) = f_{\mathrm{d}}(t) + f_{\mathrm{f}}(t) + f_{\mathrm{r}}(t)$$

$$f_{\mathrm{d}}(t) = \langle f | \, \mathbf{\Gamma}(t) \rangle$$

$$f_{\mathrm{f}}(t) = \int_0^t e^{\mathcal{L}(t-s)} \langle \dot{f}_{\mathrm{r}} | \, \mathbf{\Gamma} \rangle \, ds$$

$$f_{\mathrm{r}}(t) = e^{(\mathcal{I}-\mathcal{P})\mathcal{L}t} (f - \langle f | \, \mathbf{\Gamma} \rangle)$$

$$\text{Choose } f = \dot{\Gamma}$$

Construction of reduced model

Using the previous decomposition and integration by parts, we can prove in general that:

$$\dot{\Gamma}(t) =$$

$$\langle \dot{\Gamma} | \Gamma \rangle + \int_0^t e^{\mathcal{L}(t-s)} \left[\left(\nabla_\Gamma - \nabla_\Gamma A \right) \cdot \langle \dot{\Gamma}_f(0)\, \dot{\Gamma}_f(s)^T | \Gamma \rangle \right]^T ds$$

$$+ \dot{\Gamma}_f(t)$$

where $e^{-A(\Gamma^*)} = \int e^{-H} \delta(\Gamma(q) - \Gamma^*) dp dq$

Markovian approximation

For slow variables:

$$\dot{\Gamma}(t) = \langle \dot{\Gamma} | \Gamma \rangle + \left[\left(\nabla_\Gamma - \nabla_\Gamma A \right) \cdot D(\Gamma) \right]^T + \dot{\Gamma}_f(t)$$

$$D(\Gamma) = \int_0^\infty \left\langle \dot{\Gamma}_f(0)\, \dot{\Gamma}_f(s)^T \,|\, \Gamma \right\rangle ds$$

Replacing the fluctuation force by a random noise of the form

$$\dot{\Gamma}_f(t) \approx (D + D^T)^{1/2}\, \eta(t)$$

leads to the same Fokker-Planck equation.

Brownian dynamics

Consider as resolved variable a single order parameter ξ. The previous equation simplifies to Brownian dynamics:

$$\frac{d\xi}{dt} = \frac{dD}{d\xi} - D(\xi)\,\frac{dA}{d\xi} + \sqrt{2D}\,\eta(t)$$

$$D(\xi) = \int_0^\infty \left\langle \dot{\xi}(0)\,\dot{\xi}(s) \middle| \xi \right\rangle ds$$

A and D can be computed efficiently using ABF.

Langevin equation

To obtain a Langevin equation, we need to keep ξ and p_ξ as resolved variables.

In this case, the drift term takes a special form:

$$\langle \dot{\xi} |\, p_\xi, \xi \rangle = \frac{\partial A}{\partial p_\xi}, \qquad \langle \dot{p}_\xi |\, p_\xi, \xi \rangle = -\frac{\partial A}{\partial \xi},$$

$$A(p_\xi^*, \xi^*) = -\ln \int e^{-H}\, \delta(p_\xi(p,q) - p_\xi^*)\, \delta(\xi(q) - \xi^*)\, dpdq.$$

$$\dot{\Gamma}(t) = \langle \dot{\Gamma} | \Gamma \rangle + \left[\left(\nabla_\Gamma - \nabla_\Gamma A \right) \cdot D(\Gamma) \right]^T + \dot{\Gamma}_f(t)$$

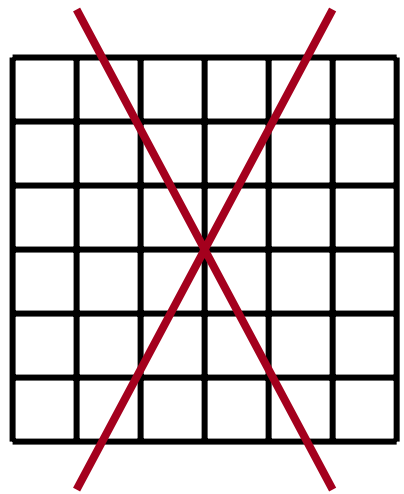
Factoring out the momentum

The free energy A needs to be computed as a function of ξ and p_ξ: $\langle \bullet \mid p_\xi, \xi \rangle$

This can be a difficulty. In practice, the free energy can be computed only as a function of one, two or at most 3 parameters.

1 variable $\xi_1 \rightarrow$ 2 parameters with momentum

2 variables ξ_1, $\xi_2 \rightarrow$ 4 parameters with momentum



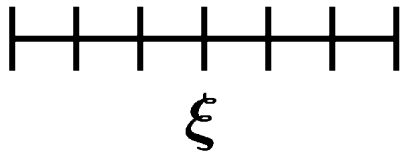
Question: $\langle \bullet \mid p_\xi, \xi \rangle \rightarrow \langle \bullet \mid \xi \rangle$ **?**

Example: expansion of the drift term

$$\langle \dot{p}_\xi | p_\xi, \xi \rangle = -\frac{\partial A(p_\xi, \xi)}{\partial \xi}$$

$$= -\frac{\partial A(0, \xi)}{\partial \xi} - \frac{p_\xi^2}{2} \frac{\partial \bar{\mathrm{m}}_\xi^{-1}}{\partial \xi}$$

$$+ \frac{\sum_{k \geq 2} \frac{(-1)^k}{k!} \left(\frac{p_\xi^2}{2} \right)^k \frac{\partial}{\partial \xi} \langle \eta^k | \, \xi, \dot{\xi} = 0 \rangle}{\sum_k \frac{(-1)^k}{k!} \left(\frac{p_\xi^2}{2} \right)^k \langle \eta^k | \, \xi, \dot{\xi} = 0 \rangle}$$



$$\bar{\mathrm{m}}_\xi^{-1} = \langle m_\xi^{-1} | \xi, \dot{\xi} = 0 \rangle, \quad \eta = m_\xi^{-1} - \bar{\mathrm{m}}_\xi^{-1}$$

Nearly constant mass

For a nearly constant mass m_ξ, the equations reduce to a well known form:

$$\langle \dot{\xi} |\, p_\xi, \xi \rangle = \frac{p_\xi}{\bar{\mathrm{m}}_\xi}$$

$$\langle \dot{p}_\xi |\, p_\xi, \xi \rangle = \lambda - \frac{\partial \bar{\mathrm{m}}_\xi^{-1}}{\partial \xi} \frac{p_\xi^2}{2}$$

$$H^{\mathrm{c}} = \frac{p^T \cdot p}{2} + U(q) + \lambda(p,q)\left(\xi(q) - \xi^*\right)$$

Friction coefficient

$$\dot{\Gamma}(t) = \langle \dot{\Gamma} | \Gamma \rangle + \left[\left(\nabla_\Gamma - \nabla_\Gamma A \right) \cdot D(\Gamma) \right]^T + \dot{\Gamma}_f(t)$$

$$D(\Gamma) = \int_0^\infty \left\langle \dot{\Gamma}_f(0)\, \dot{\Gamma}_f(s)^T \,|\, \Gamma \right\rangle ds$$

Under two assumptions:

1. ξ and p_ξ are slow variables
2. the mass m_ξ is nearly constant

the friction term takes a simple form.

Simplified form of Langevin equation

$$\dot{\xi} = \frac{\partial A}{\partial p_\xi}$$

$$\dot{p}_\xi = -\frac{\partial A}{\partial \xi} - \zeta(\xi)\,\frac{p_\xi}{\bar{\mathbf{m}}_\xi} + \sqrt{2\zeta}\,\eta(t)$$

$$\zeta(\xi) = \int_0^\infty \left\langle \lambda(t)\lambda(0) \middle| \xi, \dot{\xi} = 0 \right\rangle - \left\langle \lambda^2 \middle| \xi, \dot{\xi} = 0 \right\rangle dt$$

Computational cost: simulation using constraints can be used to calculate the mean of λ, the auto-correlation of $\lambda(t)$, and the moments of m_ξ.

Conclusion

ABF: efficient adaptive technique to compute free energy.

Mori-Zwanzig formalism allows deriving generalized Langevin equations to coarse grain a macro-molecular system.

Under certain assumptions, the equations take a familiar form and can be computed relatively easily.