Weighted-Ensemble Brownian Dynamics Simulations for Protein Association Reactions

Gary A. Huber and Sangtae Kim
Biophysical Journal, Volume 70, January 1996, pp. 97-110

Presented by Amanda McCoy
Mentor: Dr. Daniel M. Zuckerman
Reviewed by Ryan Cheng

The Problem of Analyzing Protein Structure

- Proteins are dynamic molecules and rarely remain in one conformation, but move between different ones in order to function.
- Documentation of all the possible conformational structures a protein might exist in is not yet available.
  - We cannot determine the structures of rare protein conformations because it is difficult to isolate proteins in their high energy states for a period long enough to acquire an accurate structure.
- Since the structures of these high energy conformations cannot be determined experimentally, a dynamics simulation is next best option for resolving these mysteries.
- A variety of dynamics simulation methods have been proposed to capture proteins in their high energy states, yet none have been able to give a satisfactory solution.

The Problem of Protein Dynamics Simulations

- Current algorithms (standard Brownian) for dynamical modeling fail in their inability to simulate a protein for a period long enough to observe a large-scale conformational change.
- The massive number of parameters and iterations required to accurately model a reasonably-sized protein at atomic resolution are beyond a supercomputer's processing abilities, and scientists are often "waiting a very long time for an unlikely event to occur" (Huber and Kim).

One-Dimensional Example of Brownian Motion

time-dependent position of the i^{th} particle is defined as:

$$x_i(t)$$

which satisfies the Langevin Equation:

$$m_i \frac{d^2 x_i}{dt^2} = F_{\text{external}} - \beta \frac{dx_i}{dt} + \eta_i(t)$$

where:

$$m_i = \text{mass of } i^{\text{th}} \text{ particle}$$

$$F_{\text{external}} = F_{\text{bonds}} + F_{\text{angles}} + F_{\text{torsions}} + F_{\text{intermolecular}}$$

$$\beta = \text{damping constant}$$

$$\eta_i(t) = \text{time-dependent thermal noise of } i^{\text{th}} \text{ particle}$$

More Detailed: External Forces

$$F_{\text{bonds}} = \sum_{\text{bonds}} \frac{k_i}{2}\left(l_i - l_{i,0}\right)^2$$

$$F_{\text{angles}} = \sum_{\text{angles}} \frac{k_i}{2}\left(\theta_i - \theta_{i,0}\right)^2$$

$$F_{\text{torsions}} = \sum_{\text{torsions}} \frac{V_n}{2}\left(1 + \cos\left(n\omega - \gamma\right)\right)$$

$$F_{\text{intermolecular}} = \sum_{i=1}^{N} \sum_{j=i+1}^{N} \left(4\varepsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right] + \frac{q_i q_j}{4\pi\varepsilon_0 r_{ij}^2} \right)$$

Example Quick and Dirty Solution to Langevin Equation

set:

$$F_{\text{external}} = 0$$

multiply both sides of equation by x_i to get:

$$x_i\left(m_i \frac{d^2 x_i}{dt^2}\right) = x_i\left(\eta_i(t) - \beta \frac{dx_i}{dt}\right)$$

by definition:

$$\frac{1}{2} m_i \left(\frac{dx_i}{dt}\right)^2 = \frac{1}{2} kT$$

$$\frac{dx_i}{dt} = \dot{x}_i$$

$$\left\langle x_i \eta_i(t) \right\rangle = 0$$

Example Quick and Dirty Solution to Langevin Equation

Langevin equation simplifies to:

$$m_i \frac{d}{dt}\left(\dot{x}_i x_i\right) = kT - \beta \dot{x}_i x_i$$

after integration, we find the root-mean squared distance the i[th] particle travels:

$$\left\langle x_i^2 \right\rangle = \frac{2kT}{\beta}\left[t - \gamma_i^{-1}\left(1 - e^{-\gamma_i t}\right)\right] \quad \text{where} \quad \gamma_i = \frac{\beta}{m_i}$$

which simplifies to:

$$\left\langle x_i^2 \right\rangle = \frac{2kT}{\beta} t \quad \text{for} \quad t >> \gamma_i^{-1}$$

* note : γ_i is different from the γ on slide 5

Standard Brownian Dynamics Simulation

- Define External Force Field
- Integrate Langevin Equation to calculate the movement of the i^{th} particle for one time-step (iteration)
- Complete integration for all particles in the system
- Repeat iterations to produce dynamics simulation using the length of one time-step as the time-scale:
- This method can be expanded to multiple dimensions using vector notation

$$F_{\text{external}} = F_{\text{bonds}} + F_{\text{angles}} + F_{\text{torsions}} + F_{\text{intramolecular}}$$

$$m_i \frac{d^2 x_i}{dt^2} = F_{\text{external}} - \beta \frac{dx_i}{dt} + \eta_i(t)$$

$$\Delta t = \text{length of one time-step}$$

$$m_i \frac{d^2 \vec{r}_i}{dt^2} = \vec{F}_{\text{external}} - \beta \frac{d\vec{r}_i}{dt} + \vec{\eta}_i(t)$$

Ermak and McCammon Variant of Standard Brownian Algorithm

- Two-Dimensional Configuration Space
- One Free Energy Barrier
- Defined Reactant and Product Surfaces
- Base on the 1978 Ermak and McCammon Algorithm:

$$\vec{\mathbf{r}}_{i+1} = \vec{\mathbf{r}}_i - \frac{1}{kT} \cdot \mathbf{D} \frac{\partial V}{\partial \vec{\mathbf{r}}} \Delta t + \sqrt{2}\sqrt{\mathbf{D}} \cdot \Delta \mathbf{W} + \frac{\partial}{\partial \vec{\mathbf{r}}} \cdot \mathbf{D} \Delta t$$

$\vec{\mathbf{r}}$ = position in configuration space

$\mathbf{D}$ = general diffusion matrix

$\Delta \mathbf{W}$ = vector of independent random Gaussian numbers

V = potential of $\overline{F}_{\text{external}}$

Δt = time-step

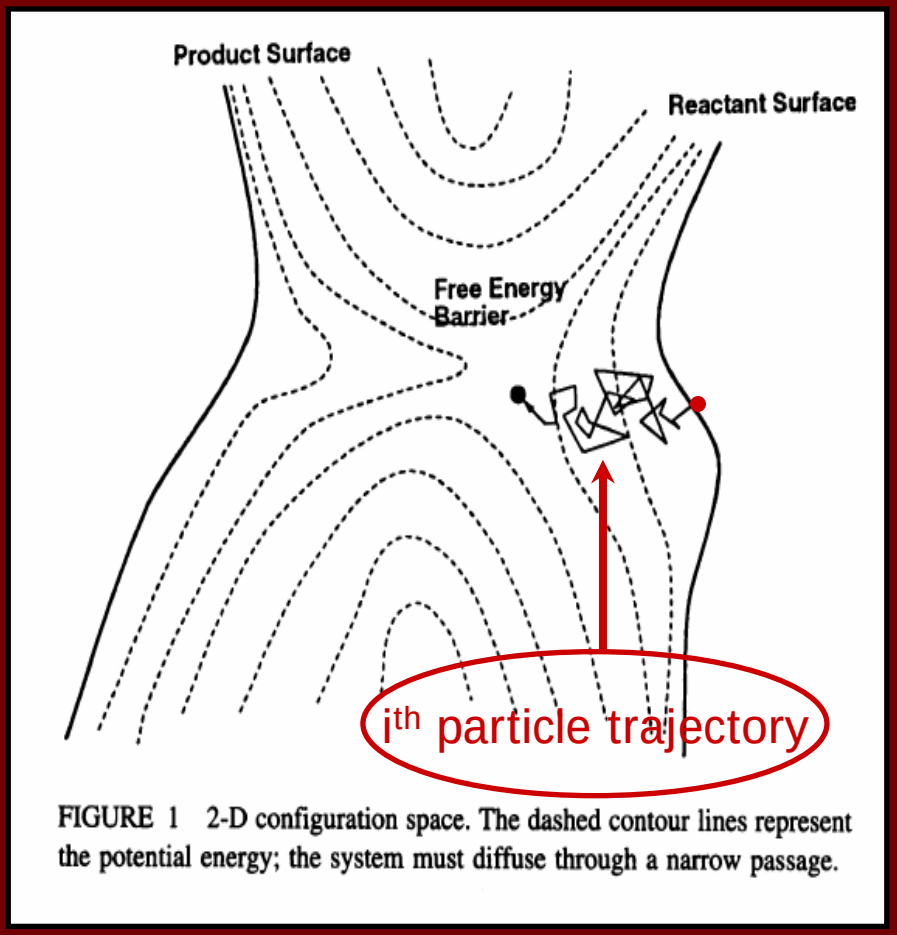


FIGURE 1 2-D configuration space. The dashed contour lines represent the potential energy; the system must diffuse through a narrow passage.

* note : V is different from V_n on slide 5

Calculation of Interest

- Mean Passage Time (MPT)
  - Expected time for a particle to surmount the energy barrier and reach the product surface.

$$t_{\mathrm{MPT}} = \frac{t_1 + \ldots t_i + \ldots t_n}{n}$$

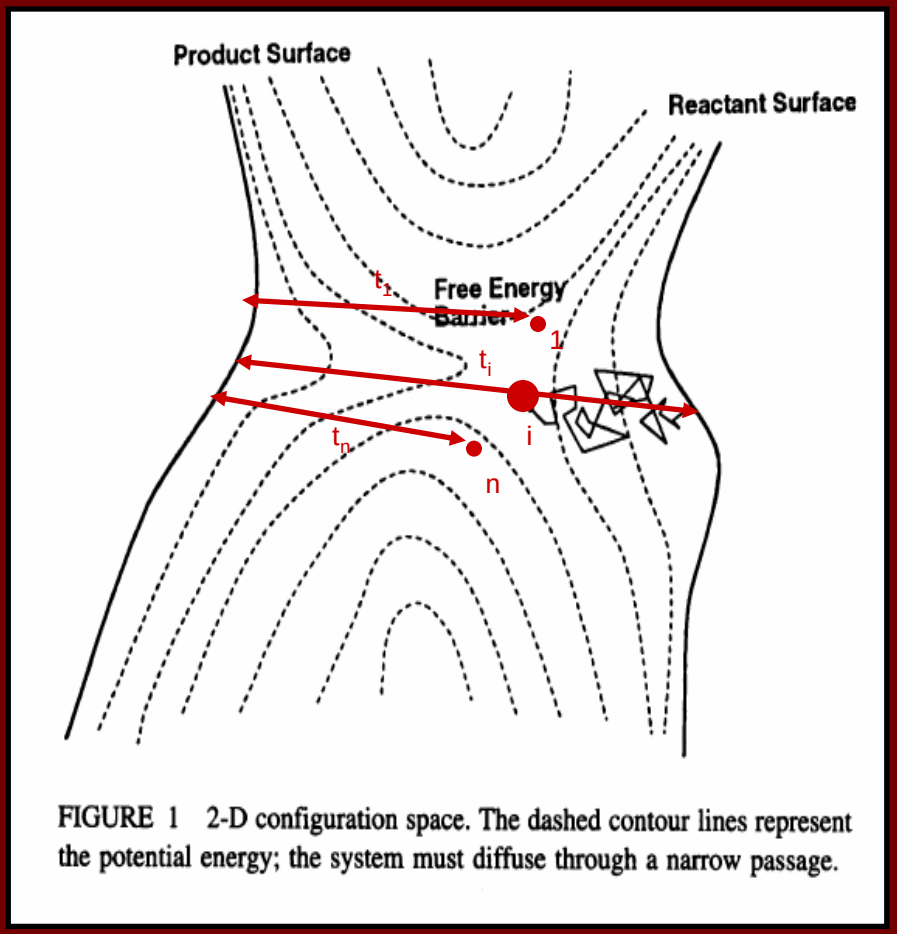


FIGURE 1 2-D configuration space. The dashed contour lines represent the potential energy; the system must diffuse through a narrow passage.

Disadvantages of a Standard Brownian Method

- Numerous degrees of freedom
- Difficult to calculate necessary number of iterations for all particles in the system to accurately model a system overcoming a energy barrier and transitioning from one equilibrium energy state to another
- Not cost effective to use a vast amount of computing time for very little simulation time
- Not all particles will reach the product surface within a reasonable amount of time
  - MPT near impossible to calculate

The Solution: Weighted-Ensemble Brownian Dynamics (WEB) Simulation

- Configuration space is divided into a number of bins along reaction coordinate.
  - Each bin is assigned a probability based on a Boltzmann distribution:

$$P(\vec{\mathbf{r}}) = e^{\frac{-V(\vec{\mathbf{r}})}{kT}} \quad \text{and} \quad \sum_{i=1}^{16} P_i = 1$$

- Each bin is initially filled with a number of weighted particles.
- Simulation is run by stepping particles forward on the Ermak and McCammon algorithm

$$\vec{\mathbf{r}}_{i+1} = \vec{\mathbf{r}}_i - \frac{1}{kT} \cdot \mathbf{D} \frac{\partial V}{\partial \vec{\mathbf{r}}} \Delta t + \sqrt{2}\sqrt{\mathbf{D}} \cdot \Delta \mathbf{W} + \frac{\partial}{\partial \vec{\mathbf{r}}} \cdot \mathbf{D} \Delta t$$

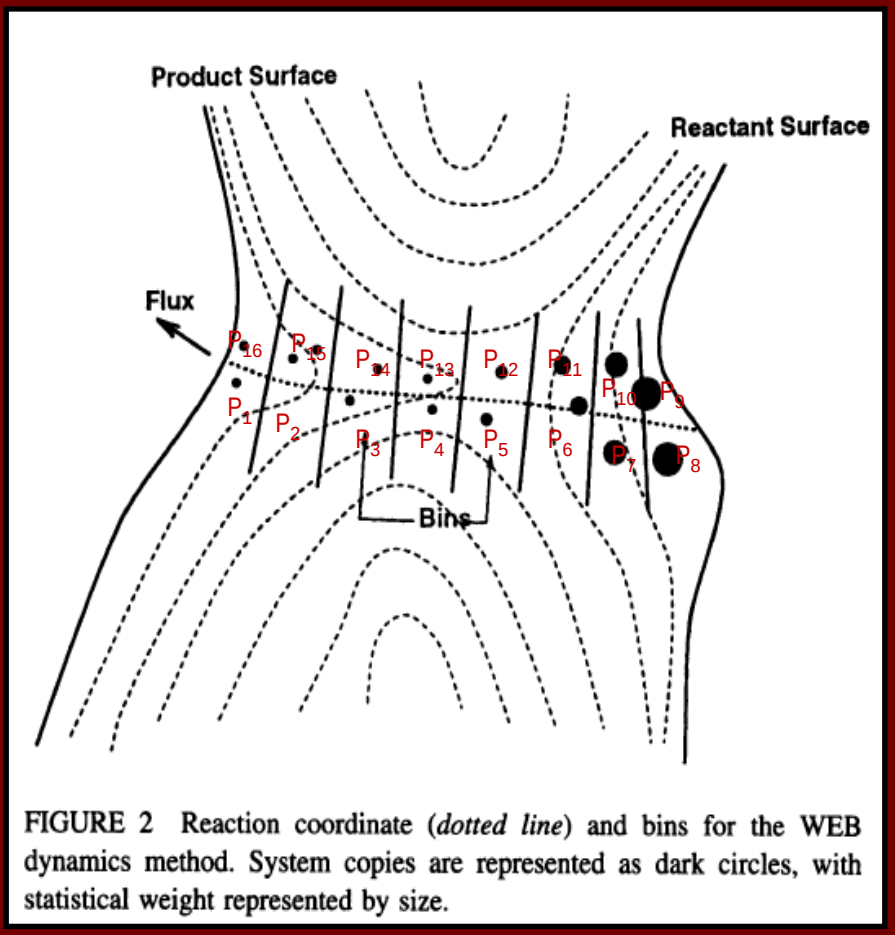


FIGURE 2 Reaction coordinate (*dotted line*) and bins for the WEB dynamics method. System copies are represented as dark circles, with statistical weight represented by size.

Some Considerations

- As each particle surmounts the energy barrier and reaches the product surface, it is automatically reintroduced back into the system on the reactant surface
- At each iteration, the probability space is reevaluated
  - Reemerging particles on the reaction surface statistical weights are modified (split, combined, or destroyed) to conserve probability.

for n particles per the i^{th} bin: $\frac{P_i}{n} = \text{ideal particle weight}$

Splitting

$$\text{if } w_{\text{particle}}^{\text{old}} > \frac{2P_i}{n}$$

Then the old particle is split into m new particles such that:

$$\overline{w}_{\text{particles}}^{\text{new}} = \frac{w_{\text{particle}}^{\text{old}}}{m} \quad \text{and} \quad \frac{P_i}{n} < w_{\text{particle}}^{\text{new}} < \frac{2P_i}{n}$$

Combining

$$\text{if } w_{\text{particle}}^{\text{old}} < \frac{P_i}{2n}$$

Then the old particle is combined with other small particles so that:

$$\frac{P_i}{n} < w_{conglomerate} < \frac{2P_i}{n}$$

The conglomerate particle will begin at the position of one of the substituent particles: dependent on the statistical weights of the individual particles in their original positions

Mini-WEB Simulation

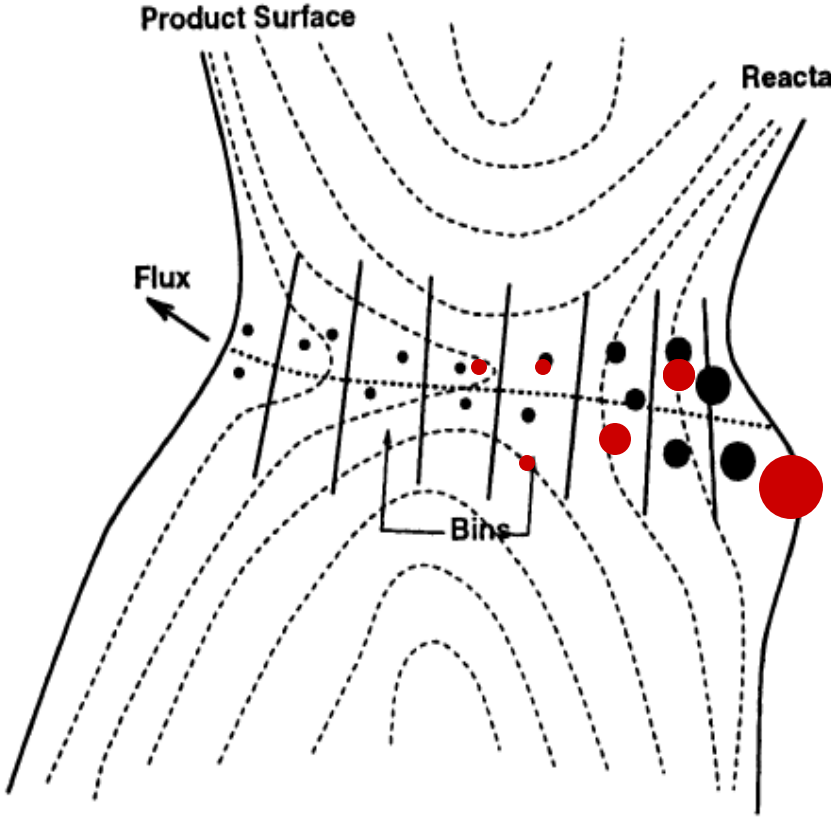


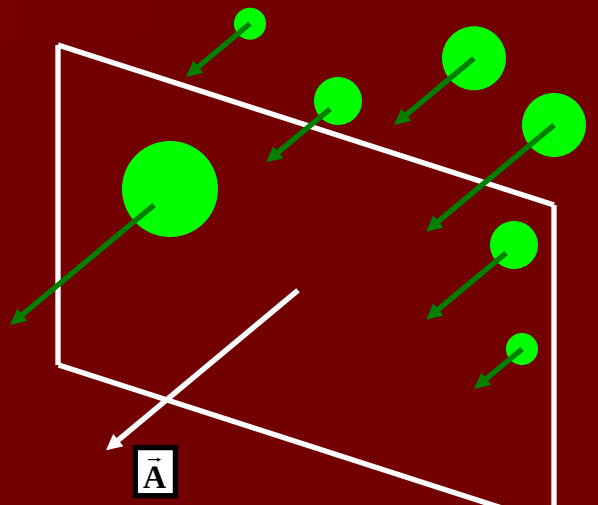
FIGURE 2 Reaction coordinate (*dotted line*) and bins for the WEB dynamics method. System copies are represented as dark circles, with statistical weight represented by size.

Calculation of Interest (WEB)

- Calculation of interest is the steady-state flux of the particles across the product surface.

$$\Phi_{\text{steady-state}} \propto \frac{1}{t_{\text{MPT}}}$$

- To avoid the problem of not enough evenly-weighted particles arriving at the product surface to determine a flux:
  - They endowed particles with variable statistical weights
  - Particles with smaller statistical weights are able to surmount the reaction energy barrier more frequently
- Steady-state flux can be measured on a more realistic time-scale.



$$\Phi_{\text{steady-state}} = \sum_{i=1}^{N} P(i)\left(\vec{\mathbf{A}} \cdot \vec{\mathbf{v}}_i\right)$$

Which is equivalent to:

$$\frac{\text{weight of particles reemerging at reactant surface}}{\text{time}}$$

$$t_{\text{MPT}} = \frac{1}{\Phi_{\text{steady-state}}}$$

Conclusion

- From the WEB dynamics steady-state flux we can calculate the MPT at a fraction (10^{-4} to 10^{-3}) of the computing time used for a standard Brownian dynamics simulation
  - Huber and Kim had results are consistent with Northrup and Erickson's
- The WEB method can be expanded far beyond 2-D particle systems
  - Paper also discusses using WEB for protein-protein association simulations with similar results as the 2-D simulation
- Refinement of the WEB dynamics simulation method will provide for in-depth exploration the dynamics of complicated proteins and systems
  - Ex: Myosin, muscle contraction protein, that adopts many different conformations as it traverses its reaction pathway

References

Hofmann, W.A. and P. de Lanerolle. 2006. "Nuclear Actin: To Polymerize or Not to Polymerize." *Journal of Cellular Biology*. Vol 172 (4): 541-552.

Huber, Gary A. and Sangtae Kim. 1996. "Weighted-Ensemble Brownian Dynamics Simulations for Protein Association Reactions." *Biophysical Journal*. Vol 70: 97-110.

Leach, Andrew R. *Molecular Modelling Principles and Applications*. 2nd ed. New York: Prentice Hall, 2001.

Northrup, S. H. and H. P. Erickson. 1992. "Kinetics of a Protein-Protein Association Explained by Brownian Dynamics Computer Simulation." *Proc. Nat. Acad. Sci. USA*. Vol 89: 3338-3342.

Northrup, S.H. and J. A. McCammon. 1980. Efficient Trajectory Simulation Methods for Diffusional Barrier Crossing Processes. *Journal of Chemical Physics*. Vol 72: 4569-4578.

Reif, F. *Fundamentals of Statistical and Thermal Physics*. San Francisco: McGraw-Hill. 560-573.

Weiss, J.N., Z. Qu, A. Garfinkel. 2003. "Understanding Biological Complexity: Lessons from the Past." *Journal of the Federation of American Societies for Experimental Biology*. Vol 17 (1): 1-6.