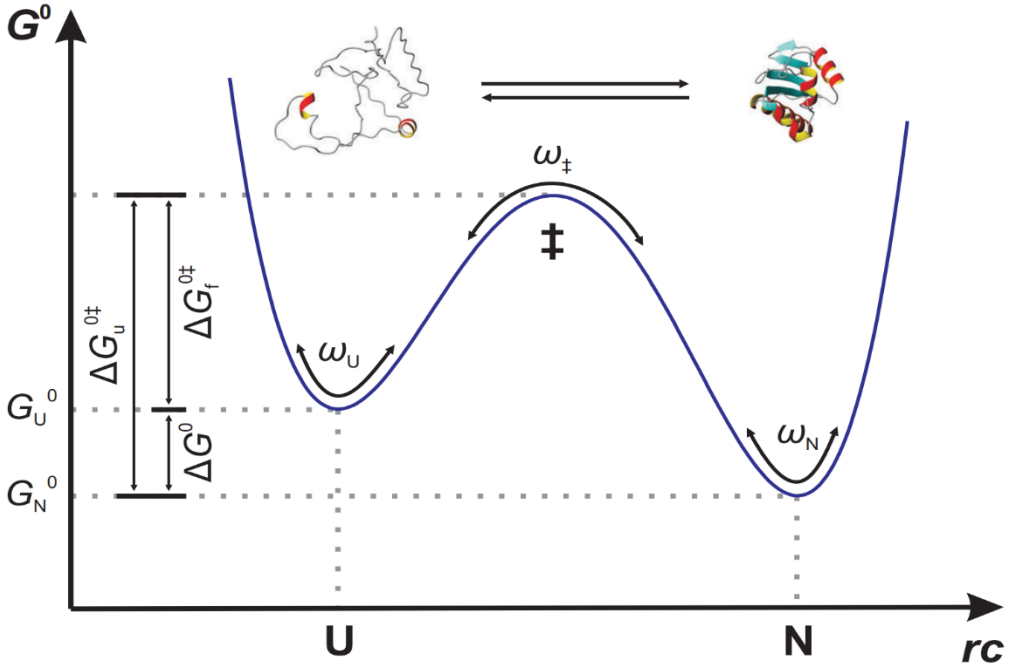
Free Energy Calculation I

Week 3 : What to Compute?



2024 Winter Molecular Simulation Seminar

February 23th, 2024

Presenter: Seonghyeon Kang

Advisor: Prof. Chang Yun Son

Week 3 : What to Compute?

1. Thermodynamics – Free Energy Calculation I

2. Thermodynamics – Free Energy Calculation II

3. Thermodynamics – Markov State Model

4. Kinetics – Rare Event Simulation

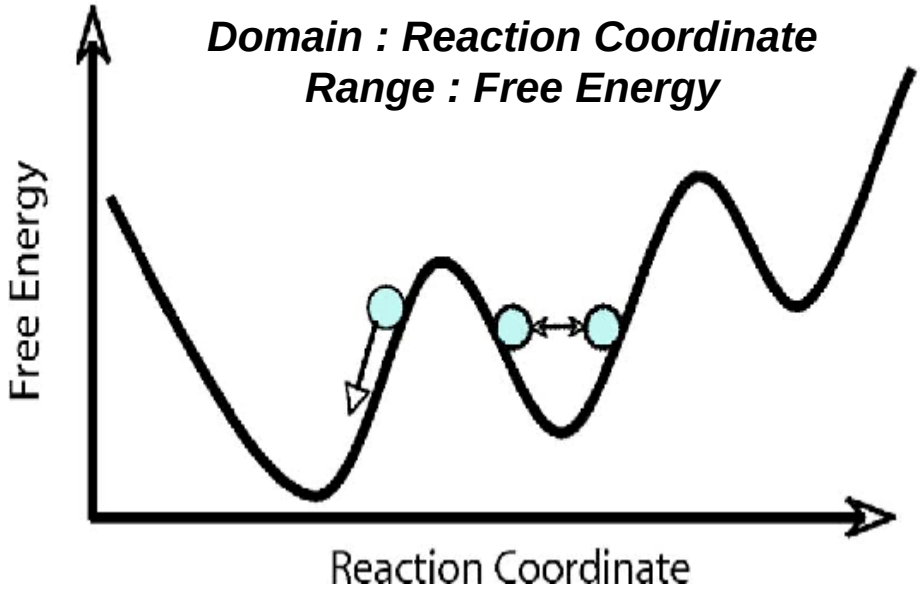
Contents

1. How to Define Proper Reaction Coordinate?

Reaction Coordinate

To make a good story for explaining chemical phenomena & mechanism,

Visualizing free energy diagram can be the powerful tool for it

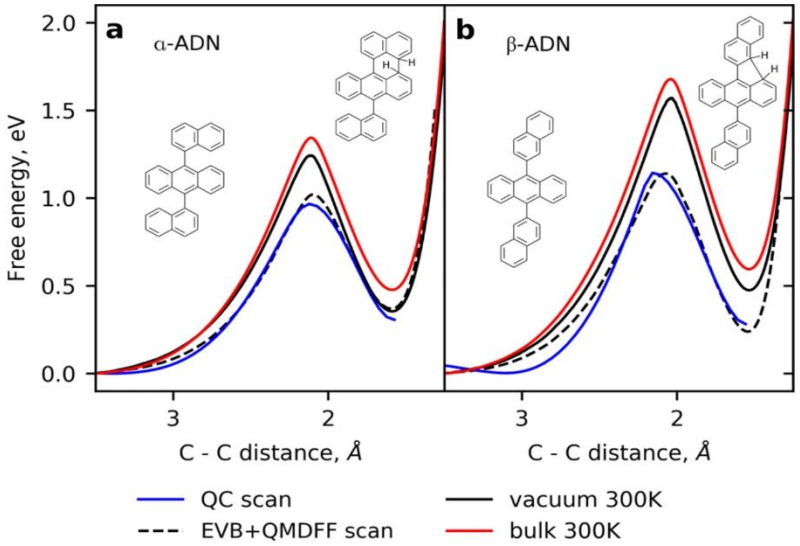


Yang, S., Onuchic, J. N., & Levine, H., Effective stochastic dynamics on a protein folding energy landscape. *J. Chem. Phys.*, **2006**, *125*(5).

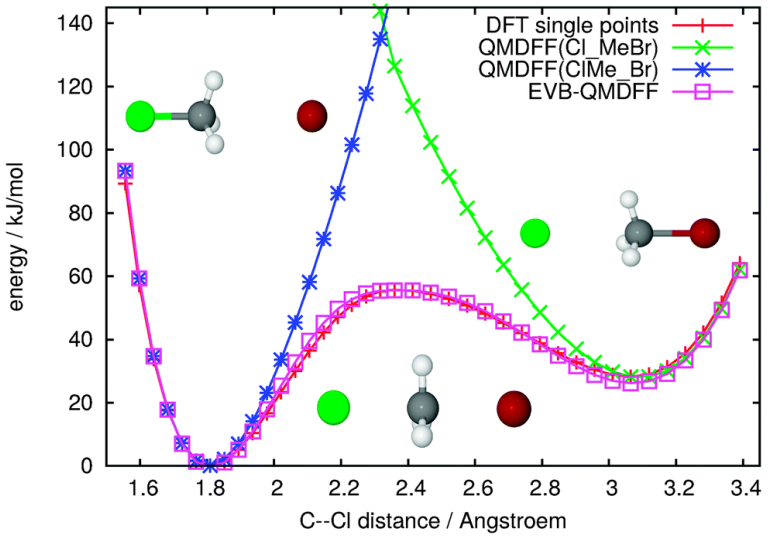
- ***But determining reaction coordinate with proper variables is quite tricky***
- ***Because various degree of freedom contribute in one process***
- ***We should define reaction coordinate in 1 or 2 dimensions***

1. How to Define Proper Reaction Coordinate?

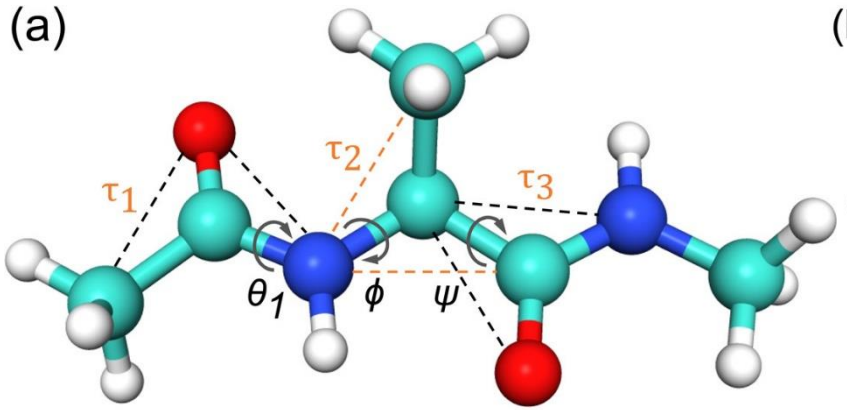
Examples of Reaction Coordinate

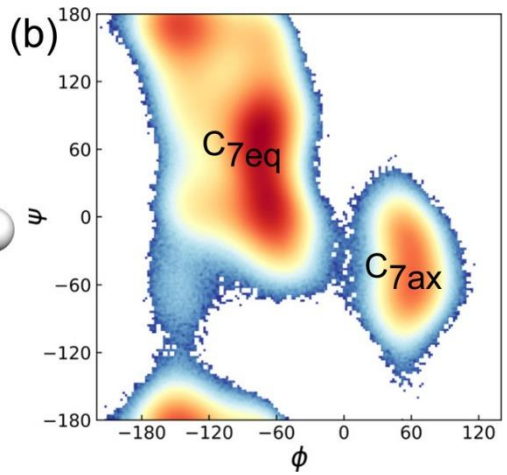


Odinokov, A., Yakubovich, A., Son, W. J., Jung, Y., & Choi, H., Exploiting the quantum mechanically derived force field for functional materials simulations. *npj Comput. Mater.*, **2021**, *7*(1), 155.



Hartke, B., & Grimme, S., Reactive force fields made simple. *Phys. Chem. Chem. Phys.*, **2015**,*17*(26), 16715-16718.





Wu, S., Li, H., & Ma, A., A rigorous method for identifying a one-dimensional reaction coordinate in complex molecules. *J. Chem. Theory. Comput.*, **2022**,*18*(5), 2836-2844.

1. How to Define Proper Reaction Coordinate?

Criteria of Good Reaction Coordinate

1. Depend only on the instantaneous configuration of the system

2. Must change monotonically as the system advances from reactants to products

3. Must be able to provide the fundamental details of the mechanism under study including all the relevant geometrical changes

4. Projection of the underlying multidimensional free energy surface (FES) onto the RC should include all the relevant features of the free energy landscape

Zinovjev, K., & Tuñón, I., Reaction coordinates and transition states in enzymatic catalysis. *Wiley Interdiscip. Rev. Comput. Mol. Sci.,* **2018**, *8(1), e1329.*

1. How to Define Proper Reaction Coordinate?

If the reaction coordinate poorly chosen…

1. An Incomplete Reaction Coordinate

- ***Not all the changes needed to produce the reaction are being captured.***
- ***Too low apparent free energy barriers***

2. If the coordinate is not able to drive the system efficiently

- ***When used in combination with enhanced sampling methods***
- ***Hysteresis and discontinuity problems can appear***
- ***Leading to artificially high barriers***

Zinovjev, K., & Tuñón, I., Reaction coordinates and transition states in enzymatic catalysis*. Wiley Interdiscip. Rev. Comput. Mol. Sci.*, **2018**, *8(1), e1329.*

1. How to Define Proper Reaction Coordinate?

If the reaction coordinate poorly chosen…

1. An Incomplete Reaction Coordinate

- ***Not all the changes needed to produce the reaction are being captured.***
- ***Too low apparent free energy barriers***

2. If the coordinate is not able to drive the system efficiently

- ***When used in combination with enhanced sampling methods***
- ***Hysteresis and discontinuity problems can appear***
- ***Leading to artificially high barriers***

Zinovjev, K., & Tuñón, I., Reaction coordinates and transition states in enzymatic catalysis*. Wiley Interdiscip. Rev. Comput. Mol. Sci.*, **2018**, *8(1), e1329.*

1. How to Define Proper Reaction Coordinate?

Two Kinds of Approach for Identifying Reaction Coordinate

1. Physics-Based Approach

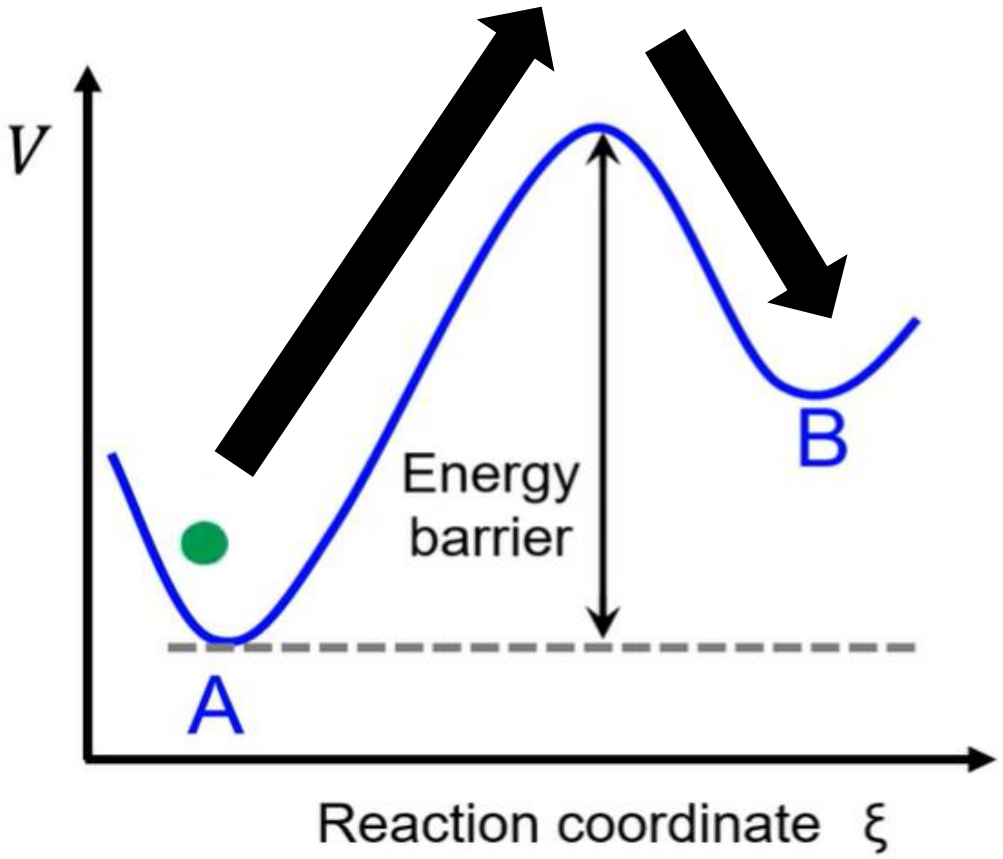
- ***Finding RC with quantities from physical intuition***
- ***Physically clear, but sometimes it is not the best option***

2. Machine-Learning Based Approach

- ***GNN, Likelihood maximization method, Kernel PCA, …***
- ***Finding reaction coordinate in perspective of informatics***
- ***Hard to find the physical meaning of the reaction coordinate***

2. Sampling over the Free Energy Barrier I : Umbrella Sampling

How to Generate Reaction Coordinate with High energy Barrier?



To generate the free energy diagram, we have to sample whole region which we want to probe

However, how is it possible when high energy barrier exists?

Image Source : https://www.youtube.com/watch?app=desktop&v=1tMd68ankiA

2. Sampling over the Free Energy Barrier I : Umbrella Sampling

How to Sample Trajectory with High energy Barrier

1. Biased Sampling

- Involves selecting samples in a manner that (un)intentionally favors certain characteristics or subsets of the population over others.

- ex) Umbrella Sampling, Metadynamics, …

- Treated in this talk(Free Energy Calculation I)

2. Unbiased Sampling

- Involves selecting samples from a population in such a way that each member of the population has an equal chance of being selected.

- ex) Accelerated Molecular Dynamics, Alchemy, …

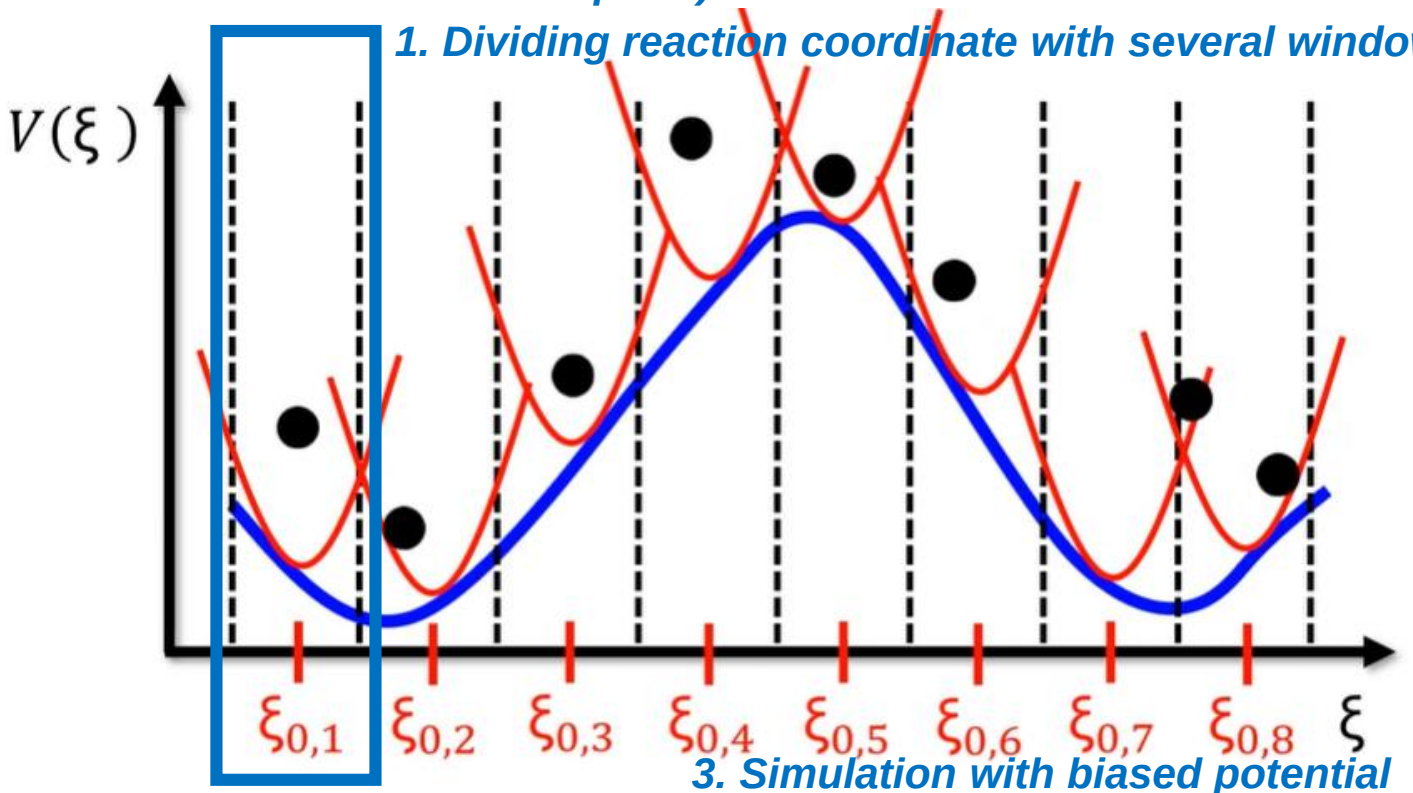
- Treated in next talk(Free Energy Calculation II)

2. Sampling over the Free Energy Barrier I : Umbrella Sampling

Umbrella Sampling

Assumption) We know the Reaction Coordinate

1. Dividing reaction coordinate with several windows



2. Applying Biased Potential

$$V^b = V^u + \omega(\xi) \quad (1)$$

$$V^b = V^u + k_i\left(\xi_i - \xi_{0,i}\right)^2 \quad (2)$$

3. Simulation with biased potential

$$q(t+\Delta t) = q(t) + \Delta t v(t) - \frac{\Delta t^2}{2m}\frac{dV^b(t)}{dq} \quad (3)$$

$$v(t+\Delta t) = v(t) - \frac{\Delta t}{2m}\left(\frac{dV^b(t+\Delta t)}{dq} + \frac{dV^b(t)}{dq}\right) \quad (4)$$

Image Source : https://www.youtube.com/watch?app=desktop&v=1tMd68ankiA

2. Sampling over the Free Energy Barrier I : Umbrella Sampling

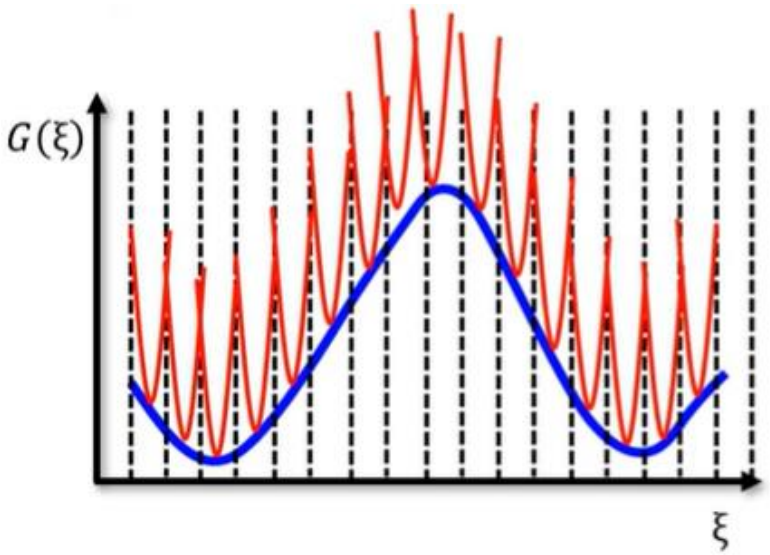
Umbrella Sampling – Spring Constant is Important!

$$V^b = V^u + k_i\left(\xi_i - \xi_{0,i}\right)^2$$

1) If k_i is too small

- ***It cannot help the system to overcome the energy barrier***
- ***Biased Potential ≈ Unbiased Potential***

2) If k_i is too large



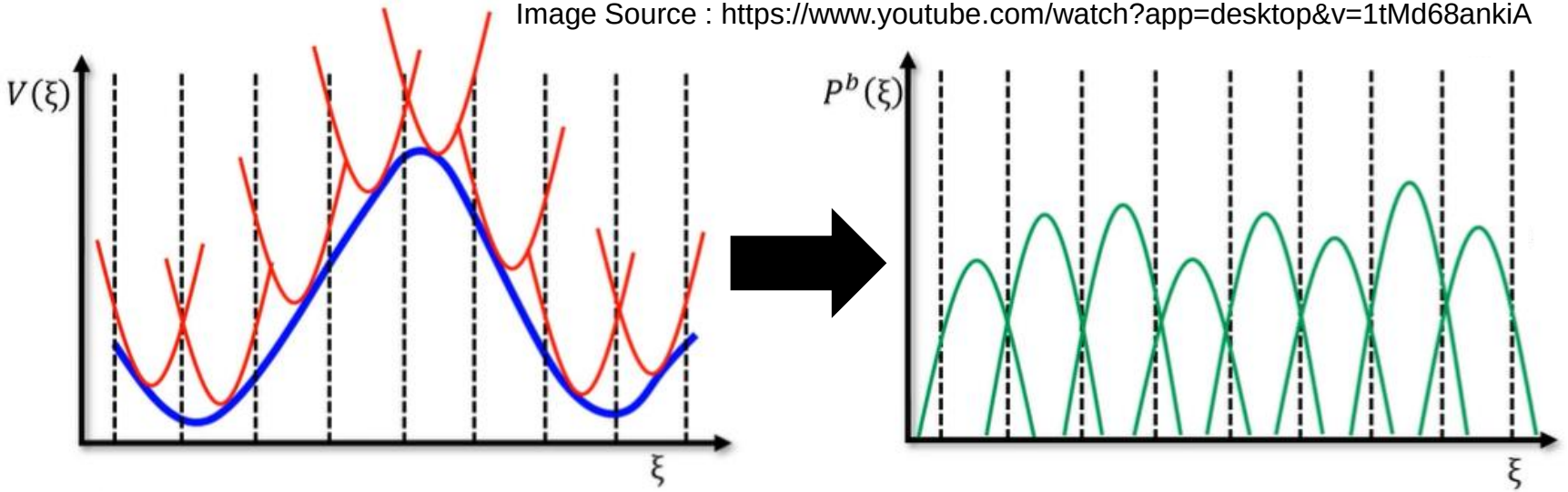
- ***Narrow distribution formed***
- ***Number of window increases***
- ***Computationally expensive***

Image Source : https://www.youtube.com/watch?app=desktop&v=1tMd68ankiA

2. Sampling over the Free Energy Barrier I : Umbrella Sampling

Umbrella Sampling

Image Source : https://www.youtube.com/watch?app=desktop&v=1tMd68ankiA



Biased Potential **Probability Distribution**

Probability Distribution to Free Energy

$$G^b(\xi) = -k_B T log P^b(\xi)$$

$$G^u(\xi) = -k_B T log P^u(\xi)$$

If converting $P^b(\xi)$ to $P^u(\xi)$ is possible, then we can get the free energy diagram with unbiased situation

Kästner, J., Umbrella sampling. *Wiley Interdiscip. Rev. Comput. Mol. Sci,* **2011**, *1(6), 932-942.*

2. Sampling over the Free Energy Barrier I : Umbrella Sampling

Umbrella Sampling

$$\begin{cases} P^u(\xi) = \dfrac{\int exp[-\beta E(r)]\delta[\xi'(r)-\xi]d^N r}{\int exp[-\beta E(r)]d^N r} \\ P^b(\xi) = \dfrac{\int exp\{-\beta[E(r)+\omega(\xi'(r))]\}\delta[\xi'(r)-\xi]d^N r}{\int exp\{-\beta[E(r)+\omega(\xi'(r))]\}d^N r} \end{cases}$$

$$\Rightarrow \quad P^u(\xi) = P^b(\xi) exp[\beta\omega(\xi)] < exp[-\beta\omega(\xi)] >$$

$$G^u(\xi) = -k_B T log\left(P^u(\xi)\right) = -k_B T log\left(P^b(\xi)\right) - \omega(\xi) - k_B T log < e^{-\beta\omega(\xi)} >$$

For Every Window,

$$G_i^u(\xi) = -k_B T log\left(P_i^b(\xi)\right) - \omega_i(\xi) + F_i$$

If we can derive the values of F_i, then we can get full, unbiased free energy diagram with respect to the reaction coordinate

Kästner, J., Umbrella sampling. *Wiley Interdiscip. Rev. Comput. Mol. Sci,* **2011**, *1(6), 932-942.*

2. Sampling over the Free Energy Barrier I : Umbrella Sampling

Umbrella Sampling

Weighted Histogram Analysis Method(WHAM)

$$P^u(\xi) = \sum_i^{N_w} p_i(\xi) P_i^u(\xi)$$

Global Distribution
= Weighted average of each window's distribution

$p_i(\xi)$ **is computed from by minimizing the statistical error of** $P_i^u(\xi)$

i,e, $\frac{\partial \sigma^2(P^u(\xi))}{\partial p_i(\xi)} = 0, \quad \sum_i^{N_w} p_i(\xi) = 1$

From these two conditions, we can get

$$p_i(\xi) = \frac{a_i(\xi)}{\sum_j^{N_w} a_j}, \quad a_i(\xi) = N_i e^{-\beta\omega_i(\xi) - \beta F_i}$$

N_i **: Number of the steps sampled for window i**

Kumar, S., Rosenberg, J. M., Bouzida, D., Swendsen, R. H., & Kollman, P. A., The weighted histogram analysis method for free-energy calculations on biomolecules. I. The method. *J. Comput. Chem.*, **1992**, *13*(8), 1011-1021.

Kästner, J., Umbrella sampling. *Wiley Interdiscip. Rev. Comput. Mol. Sci,* **2011***, 1(6), 932-942.*

2. Sampling over the Free Energy Barrier I : Umbrella Sampling

Umbrella Sampling

Weighted Histogram Analysis Method(WHAM)

$$P^u(\xi) = \sum_i^{N_w} p_i(\xi) P_i^u(\xi)$$

$$p_i(\xi) = \frac{a_i(\xi)}{\sum_j^{N_w} a_j},$$

$$a_i(\xi) = N_i e^{-\beta\omega_i(\xi) - \beta F_i}$$

$$e^{-\beta F_i} = \int P^u(\xi) e^{\beta\omega(\xi)} d\xi$$

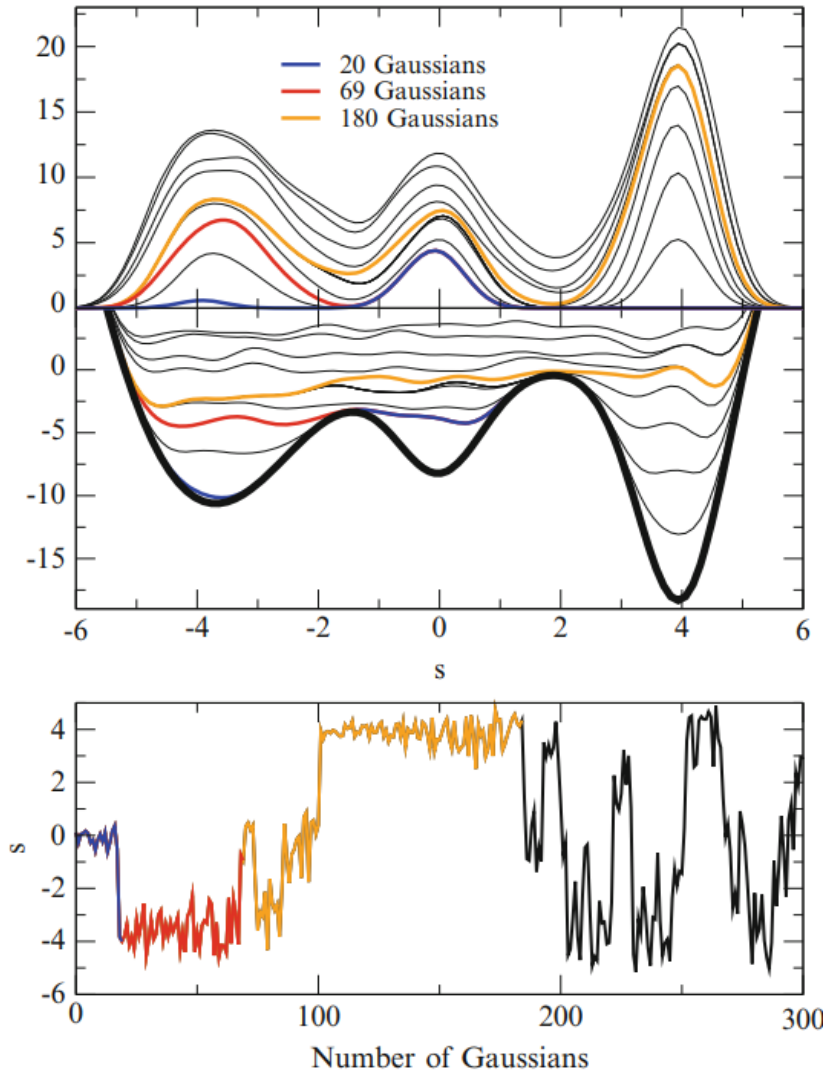
***By solving these equations until converge*, We can get the F_i *for all windows*.**

Then, we can make the free energy diagram for whole region which we want to probe

Kästner, J., Umbrella sampling. *Wiley Interdiscip. Rev. Comput. Mol. Sci*, **2011**, *1(6), 932-942.*

3. Sampling over the Free Energy Barrier II : Metadynamics

Metadynamics – Basic Principle



Adding external gaussian potential functions with certain period

$$V_G(S(x),t) = w \sum_{t'=\tau_G,2\tau_G,..} e^{-\frac{\left(S(x)-s(t')\right)^2}{2\delta s^2}}$$

1. The Gaussian height w
2. The Gaussian width δs.
3. The frequency τ_G at which the Gaussians are added.

Bussi, G., Laio, A., & Tiwary, P., Metadynamics: A unified framework for accelerating rare events and sampling thermodynamics and kinetics. *Handbook of materials modeling: Methods: theory and modeling*, 565-595., **2020**

3. Sampling over the Free Energy Barrier II : Metadynamics

Metadynamics – Basic Principles

Purpose & Benefits of using Metadynamics

- ***It accelerates the sampling of rare events by pushing the system away from local free energy minima.***

- ***It allows exploring new reaction pathways as the system tends to escape the minima passing through the lowest free-energy saddle point.***

- ***No a priori knowledge of the landscape is required***

- ***After a transient, the bias potential V_G provides an unbiased estimate of the underlying free energy***

$$V_G(s,t) \approx -F(s) + C(t)$$

($C(t)$ is only dependent on time, not the reaction coordinate)

Barducci, A., Bonomi, M., & Parrinello, M., Metadynamics. *Wiley Interdiscip. Rev. Comput. Mol. Sci.*, **2011**, *1*(5), 826-843.

3. Sampling over the Free Energy Barrier II : Metadynamics

Metadynamics – Drawbacks

1. In a single run, V_G does not converge modulo a constant to the free energy, but oscillates around it

- The bias potential overfills the underlying FES and pushes the system toward high energy regions of the CVs space

- It is not trivial to decide when to stop a simulation.

(it should be stopped when the motion of the RCs becomes diffusive in the region of interest)

- Solution : Well-Tempered Metadynamics

2. Identifying a set of RCs appropriate for describing complex processes is far from trivial

- Solution : Choosing proper reaction coordinate

Barducci, A., Bonomi, M., & Parrinello, M., Metadynamics. *Wiley Interdiscip. Rev. Comput. Mol. Sci.*, **2011**, *1*(5), 826-843.

3. Sampling over the Free Energy Barrier II : Metadynamics

Well-Temperated Metadynamics

$$V_G(S(x),t) = w \sum_{t'=\tau_G,2\tau_G,..} exp\left(-\frac{V_G(S(x),t)}{k_B \Delta T}\right) exp\left(-\frac{\left(S(x)-s(t')\right)^2}{2\delta s^2}\right)$$

Standard Metadynamics : $$V_G(S(x),t) = w \sum_{t'=\tau_G,2\tau_G,..} exp\left(-\frac{\left(S(x)-s(t')\right)^2}{2\delta s^2}\right)$$

Bussi, G., Laio, A., & Tiwary, P., Metadynamics: A unified framework for accelerating rare events and sampling thermodynamics and kinetics. *Handbook of materials modeling: Methods: theory and modeling*, 565-595., **2020**

Key Features

- ***Bias deposition rate decreases as 1/t, the dynamics of all the micro-scopic variables becomes progressively, closer to thermodynamic equilibrium as the simulation proceeds***

- ***Bias potential does not fully compensate the FES, but it converges to***

$$V_G(s,t) \approx -\frac{\Delta T}{T+\Delta T} F(s) + C(t)$$

Barducci, A., Bonomi, M., & Parrinello, M., Metadynamics. *Wiley Interdiscip. Rev. Comput. Mol. Sci.*, **2011**, *1*(5), 826-843.

3. Sampling over the Free Energy Barrier II : Metadynamics

Well-Temperated Metadynamics

$$V_G(S(x),t) = w \sum_{t'=\tau_G,2\tau_G,..} exp\left(-\frac{V_G(S(x),t)}{k_B \Delta T}\right) exp\left(-\frac{\left(S(x)-s(t')\right)^2}{2\delta s^2}\right)$$

Bussi, G., Laio, A., & Tiwary, P., Metadynamics: A unified framework for accelerating rare events and sampling thermodynamics and kinetics. *Handbook of materials modeling: Methods: theory and modeling*, 565-595., **2020**

Parameter ΔT

- ***Input parameter with the dimension of a temperature***
- ***Intermediate characteristic between MD and standard metadynamics***

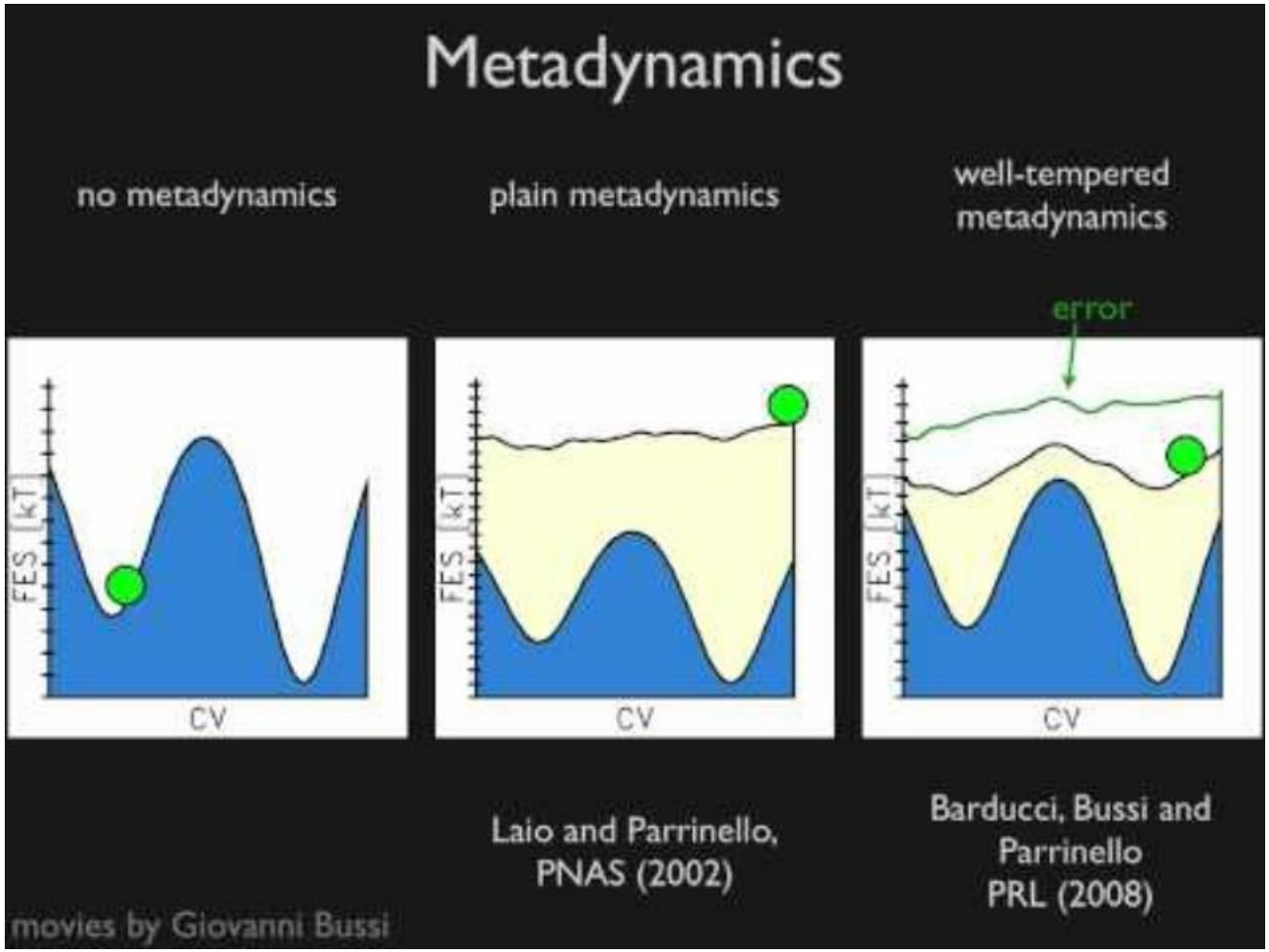
$$If\ \Delta T \approx 0 \rightarrow exp\left(-\frac{V_G(S(x),t)}{k_B \Delta T}\right) \approx 0 \rightarrow Classical\ Molecular\ Dynamics$$

$$If\ \Delta T \approx \infty \rightarrow exp\left(-\frac{V_G(S(x),t)}{k_B \Delta T}\right) \approx 1 \rightarrow Standard\ Metadynamics$$

Barducci, A., Bonomi, M., & Parrinello, M., Metadynamics. *Wiley Interdiscip. Rev. Comput. Mol. Sci.*, **2011**, *1*(5), 826-843.

3. Sampling over the Free Energy Barrier II : Metadynamics

Well-Temperated Metadynamics



Video Source : https://www.youtube.com/watch?v=IzEBpQ0c8TA

Takeaways

■ ***Identifying proper reaction coordinate** is important to visualize reasonable free energy surface(FES)*

■ ***Umbrella sampling** derives the free energy surfaces by using the **simulation with biased harmonic potentials** in several windows*

■ ***Metadynamics** adding **Gaussian biased potential periodically** to overcome the energy barrier*

Thank you for your kind attention / Q&A