

Statistical mechanics canonical ensemble

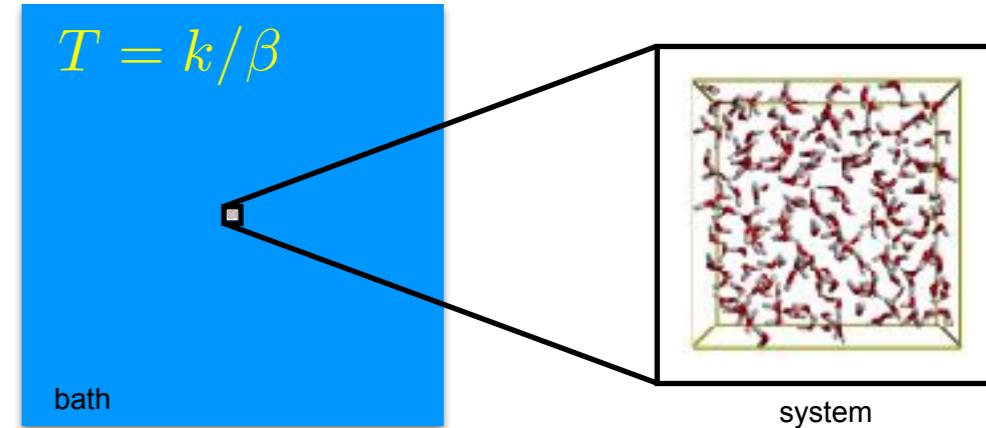
system in thermal equilibrium with bath
free energy of system

microscopic entropy

$$S = \frac{1}{T} \langle E \rangle + k \ln Z$$

microscopic free energy

$$-kT \ln Z = \langle E \rangle - TS$$



$$G = -kT \ln Z$$

macroscopic free energy

$$G = U - TS$$

from micro to macro: generate partition function

Monte Carlo

molecular dynamics simulations

Note that I use G instead of A....

Molecular dynamics simulations

classical nuclei

$$\mathbf{F}_i = -\nabla_{\mathbf{R}_i} V(\mathbf{R}_1, \mathbf{R}_2, \dots, \mathbf{R}_N) = m_i \frac{d^2}{dt^2} \mathbf{R}_i$$



Isaac Newton

Molecular dynamics simulations

classical nuclei

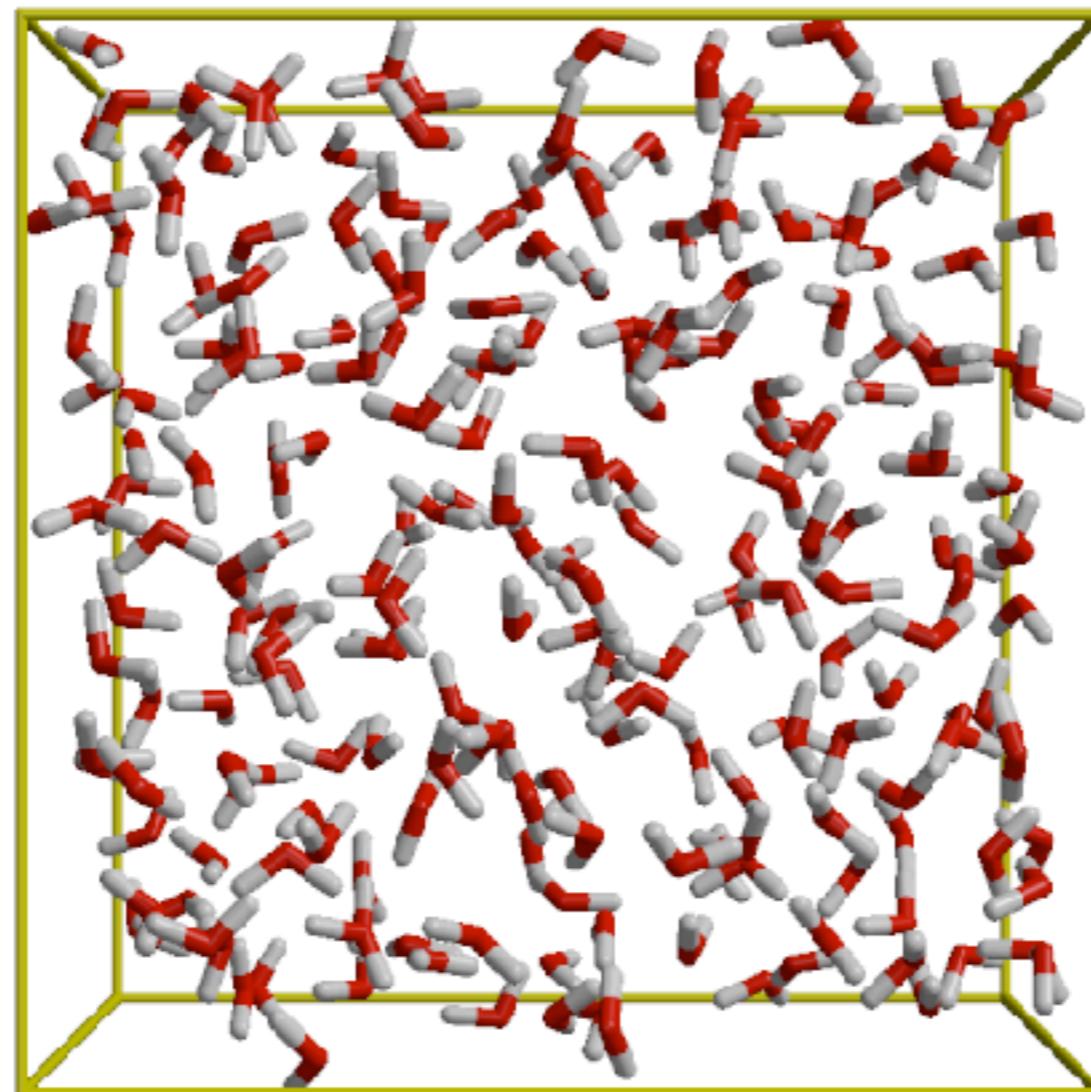
$$\mathbf{F}_i = -\nabla_{\mathbf{R}_i} V(\mathbf{R}_1, \mathbf{R}_2, \dots, \mathbf{R}_N) = m_i \frac{d^2}{dt^2} \mathbf{R}_i$$



Isaac Newton

trajectory

jiggling & wiggling



Molecular dynamics simulations

classical nuclei

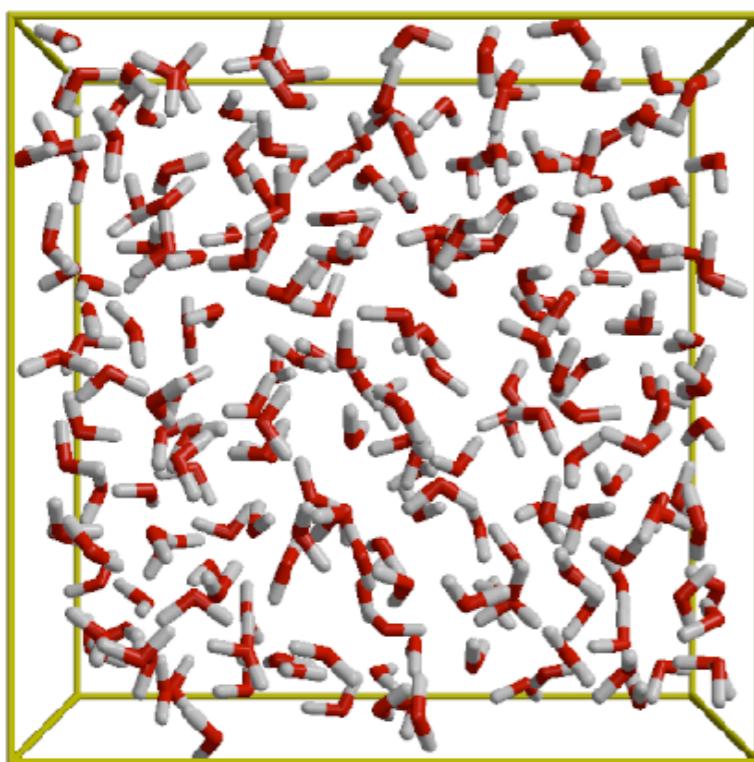
$$\mathbf{F}_i = -\nabla_{\mathbf{R}_i} V(\mathbf{R}_1, \mathbf{R}_2, \dots, \mathbf{R}_N) = m_i \frac{d^2}{dt^2} \mathbf{R}_i$$



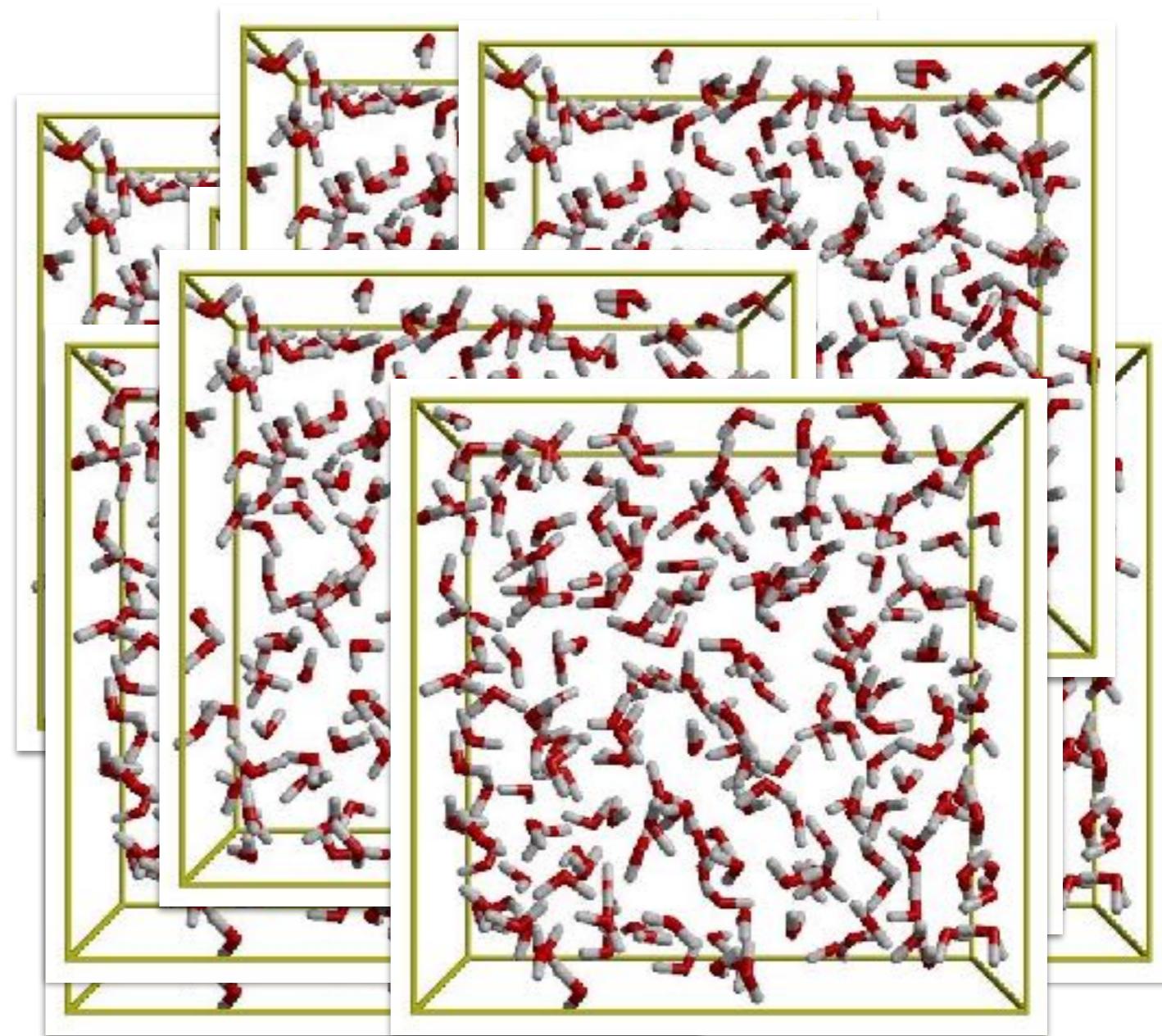
Isaac Newton

trajectory

ergodicity



=



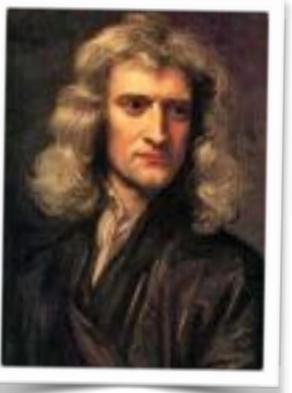
trajectory

ensemble

Molecular dynamics simulations

classical nuclei

$$\mathbf{F}_i = -\nabla_{\mathbf{R}_i} V(\mathbf{R}_1, \mathbf{R}_2, \dots, \mathbf{R}_N) = m_i \frac{d^2}{dt^2} \mathbf{R}_i$$

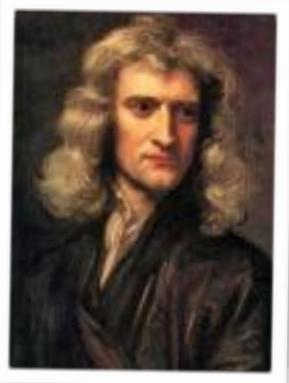


Isaac Newton

potential energy functions

Molecular dynamics simulations

classical nuclei



Isaac Newton

$$\mathbf{F}_i = -\nabla_{\mathbf{R}_i} V(\mathbf{R}_1, \mathbf{R}_2, \dots, \mathbf{R}_N) = m_i \frac{d^2}{dt^2} \mathbf{R}_i$$

potential energy functions

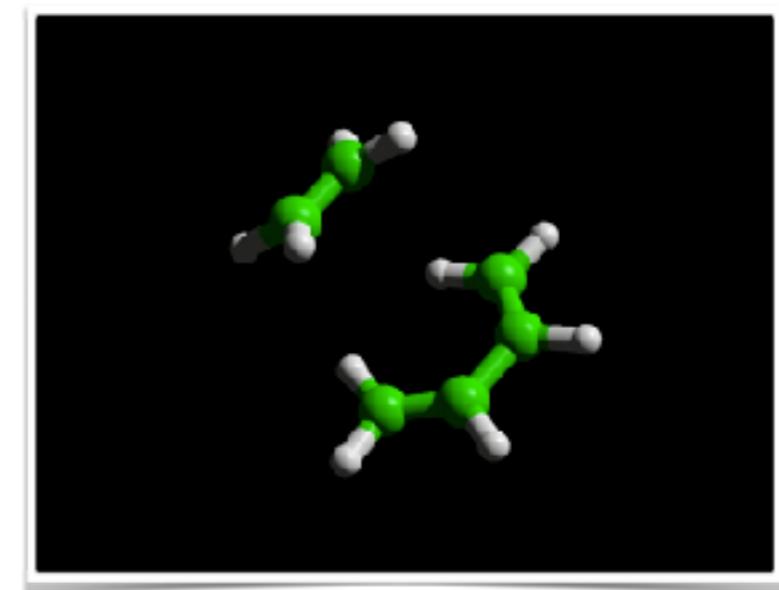
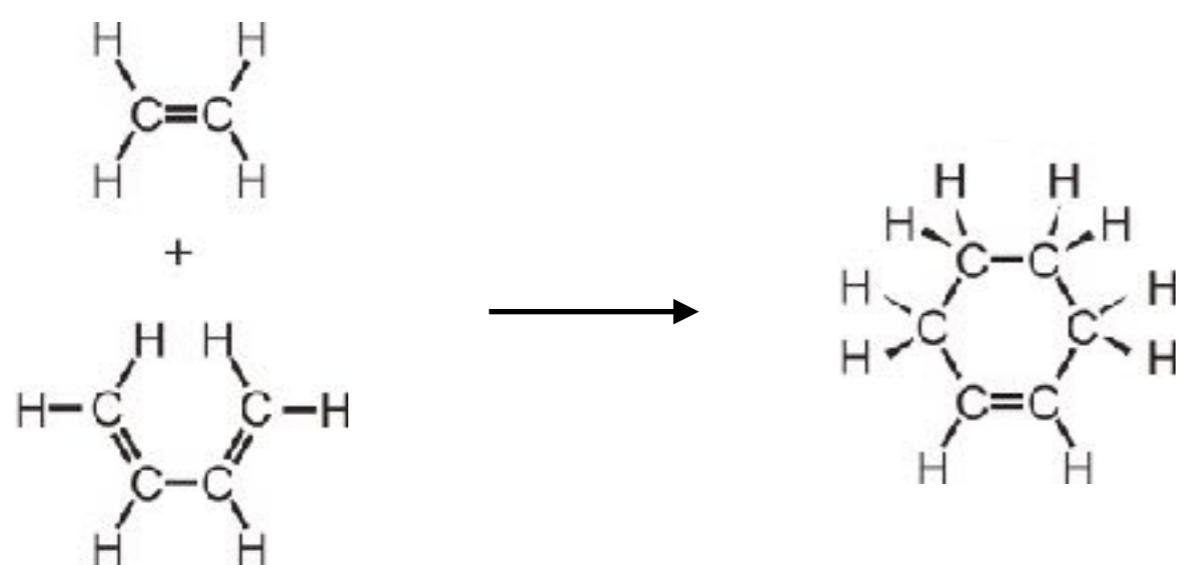


quantum chemistry (QM)

$$V^{\text{QM}}(\mathbf{R}_1, \mathbf{R}_2, \dots, \mathbf{R}_N) =$$

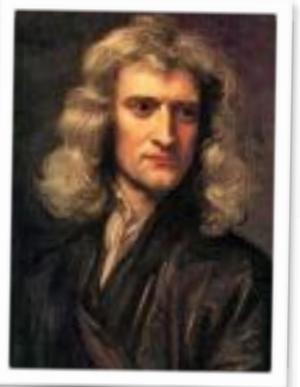
Erwin Schrödinger

$$\langle \Psi_e(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_n) | \hat{H}_e(\mathbf{R}_1, \mathbf{R}_2, \dots, \mathbf{R}_N) | \Psi_e(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_n) \rangle$$



Molecular dynamics simulations

classical nuclei



Isaac Newton

$$\mathbf{F}_i = -\nabla_{\mathbf{R}_i} V(\mathbf{R}_1, \mathbf{R}_2, \dots, \mathbf{R}_N) = m_i \frac{d^2}{dt^2} \mathbf{R}_i$$

potential energy functions

quantum chemistry (QM)



Erwin Schrödinger

$$V^{\text{QM}}(\mathbf{R}_1, \mathbf{R}_2, \dots, \mathbf{R}_N) =$$

$$\langle \Psi_e(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_n) | \hat{H}_e(\mathbf{R}_1, \mathbf{R}_2, \dots, \mathbf{R}_N) | \Psi_e(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_n) \rangle$$

molecular mechanics (MM)

empirical functions with parameters

$$V^{\text{MM}}(R_1, R_2, \dots, R_N) = \sum_k v_k(R_i, R_j, R_k, R_l; \{p_k\})$$

Molecular Mechanics Force Field

throw away the electrons!!!

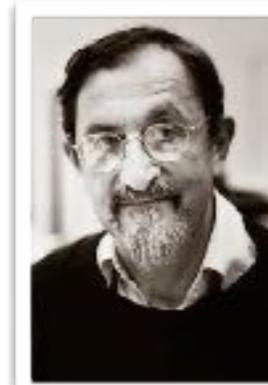
low-dimensional functions

empirical parameters

experimental data & *ab initio* calculations

$$V^{\text{MM}}(\mathbf{R}_1, \mathbf{R}_2, \dots, \mathbf{R}_N) = \sum_k v_k(\mathbf{R}_A, \mathbf{R}_B, \mathbf{R}_C, \mathbf{R}_D; \mathbf{p}_k)$$

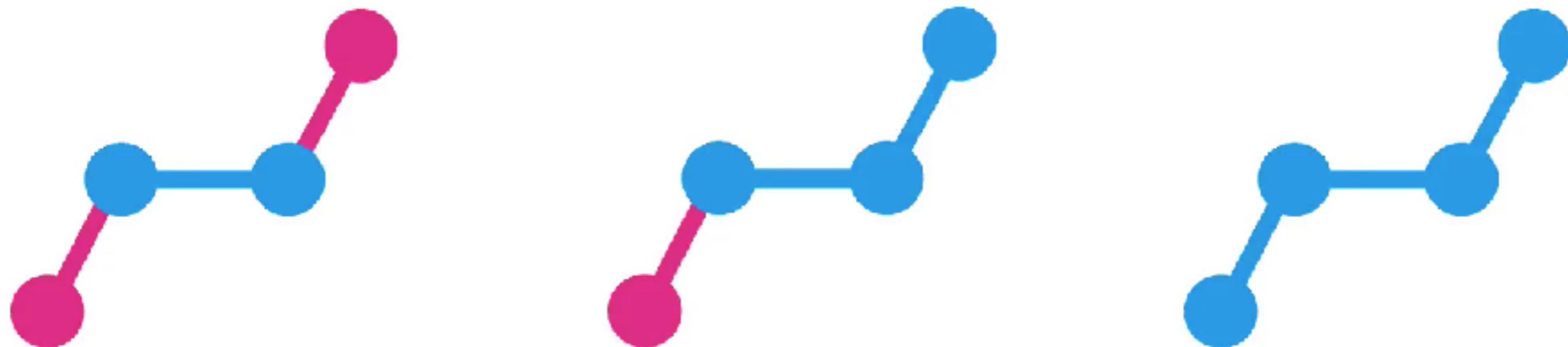
bonded interactions



Martin
Karplus

Michael
Levitt

Arieh
Warshel



$$v_b(\mathbf{R}_A, \mathbf{R}_B) = \frac{1}{2} k_b (R_{AB} - r_0)^2 \quad v_a(\mathbf{R}_A, \mathbf{R}_B, \mathbf{R}_C) = \frac{1}{2} k_a (\theta_{ABC} - \theta_0)^2 \quad v_d(\mathbf{R}_A, \mathbf{R}_B, \mathbf{R}_C, \mathbf{R}_D) = \sum_n^5 k_n (\cos(\varphi_{ABCD}))^n$$

Molecular Mechanics Force Field

throw away the electrons!!!

low-dimensional functions

empirical parameters

experimental data & *ab initio* calculations

$$V^{\text{MM}}(\mathbf{R}_1, \mathbf{R}_2, \dots, \mathbf{R}_N) = \sum_k v_k(\mathbf{R}_A, \mathbf{R}_B, \mathbf{R}_C, \mathbf{R}_D; \mathbf{p}_k)$$

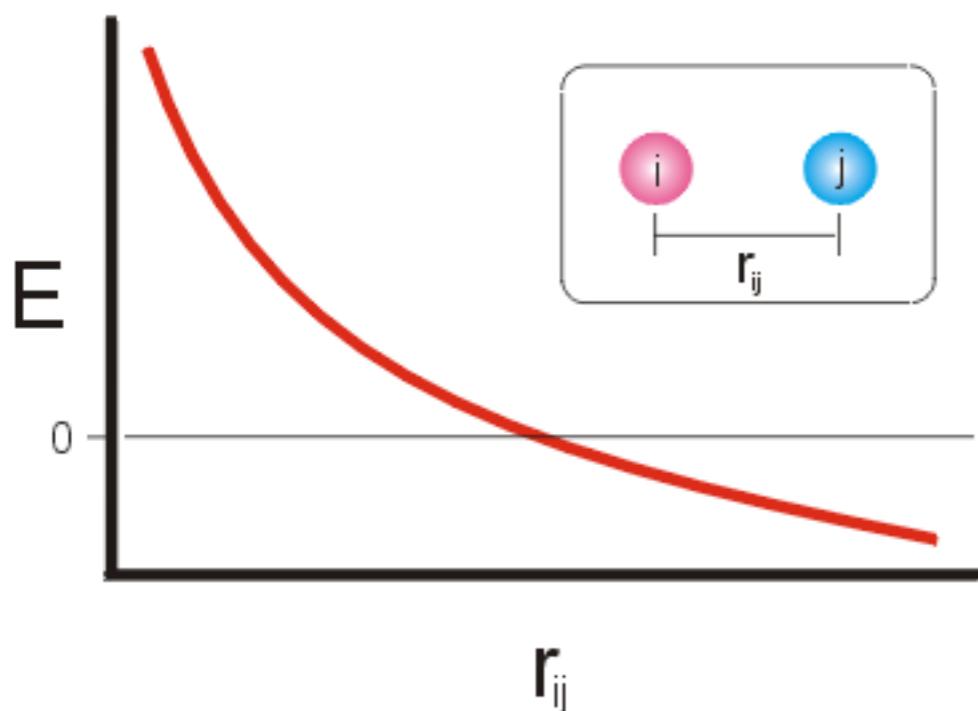
bonded interactions



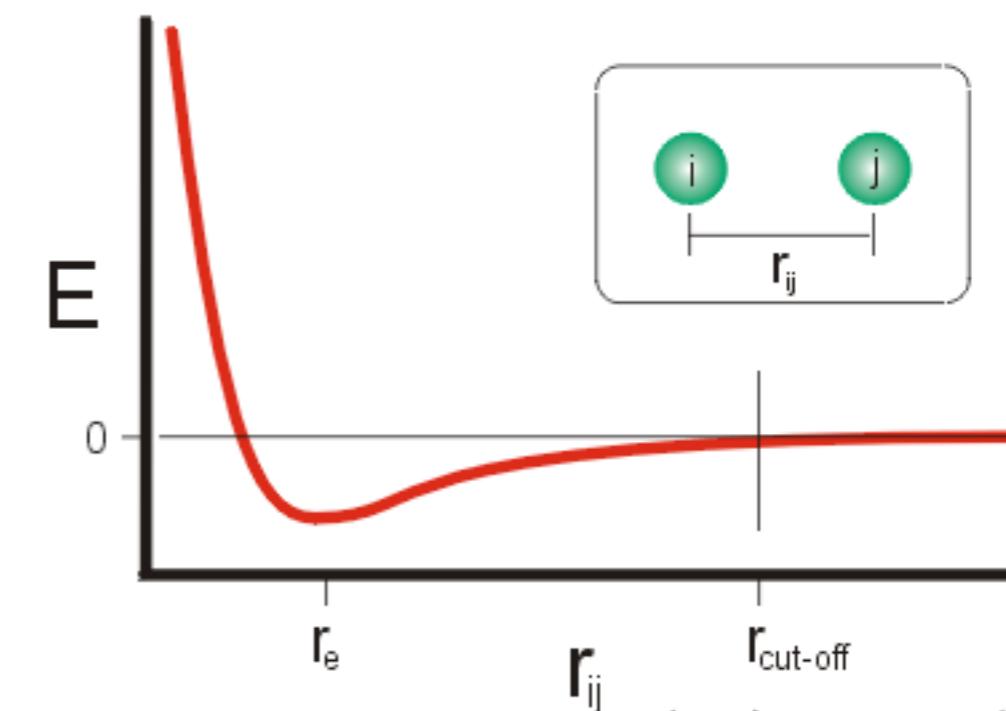
Martin
Karplus

Michael
Levitt

Arieh
Warshel



$$v_{\text{Coul}}(\mathbf{R}_A, \mathbf{R}_B) = \frac{e^2 q_A q_B}{4\pi\epsilon_0 R_{AB}}$$

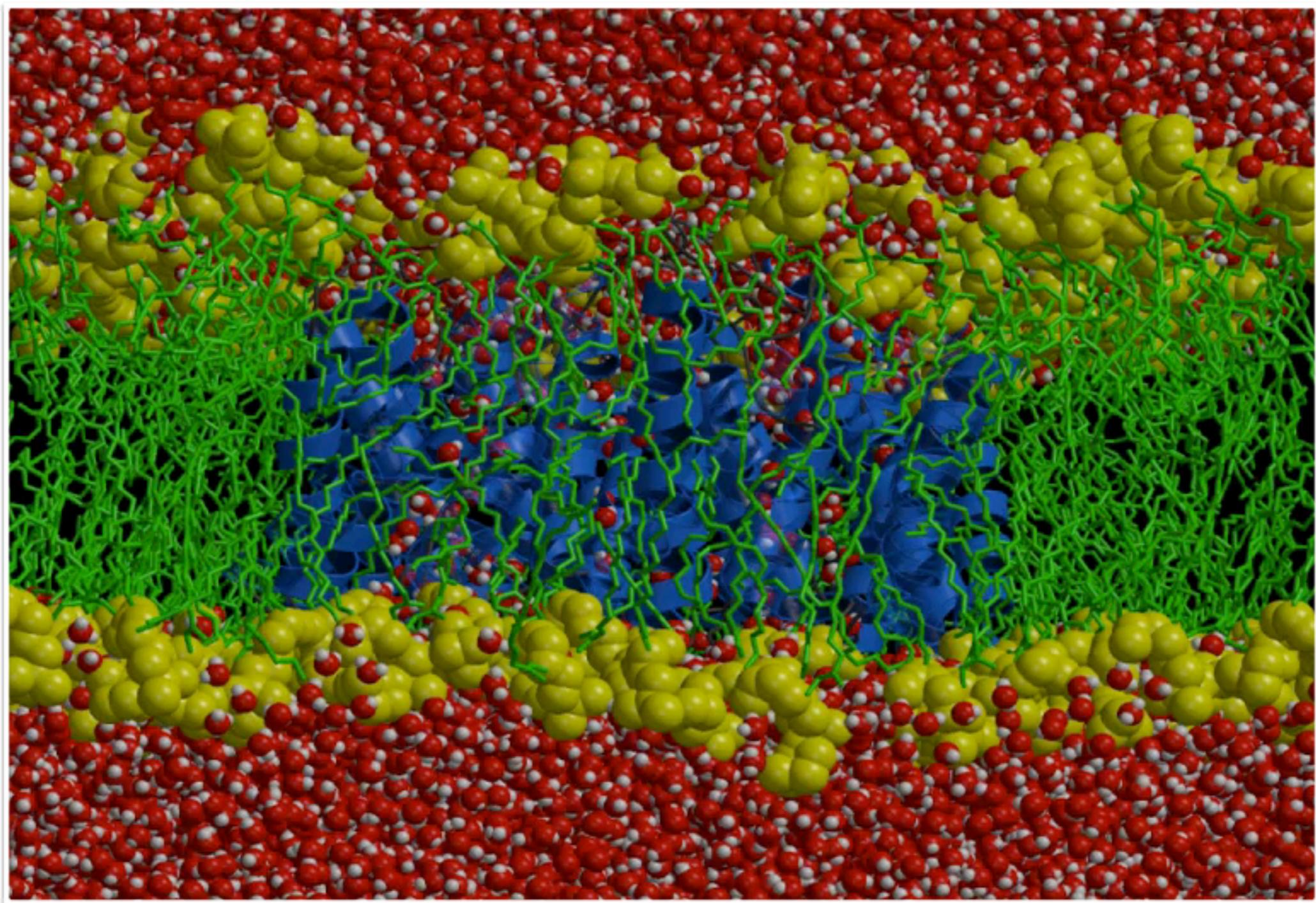


$$v_{\text{LJ}}(\mathbf{R}_A, \mathbf{R}_B) = \frac{C_{AB}^{(12)}}{R_{AB}^{12}} - \frac{C_{AB}^{(6)}}{R_{AB}^6}$$

Chemistry without test-tubes

observe jiggling and wiggling

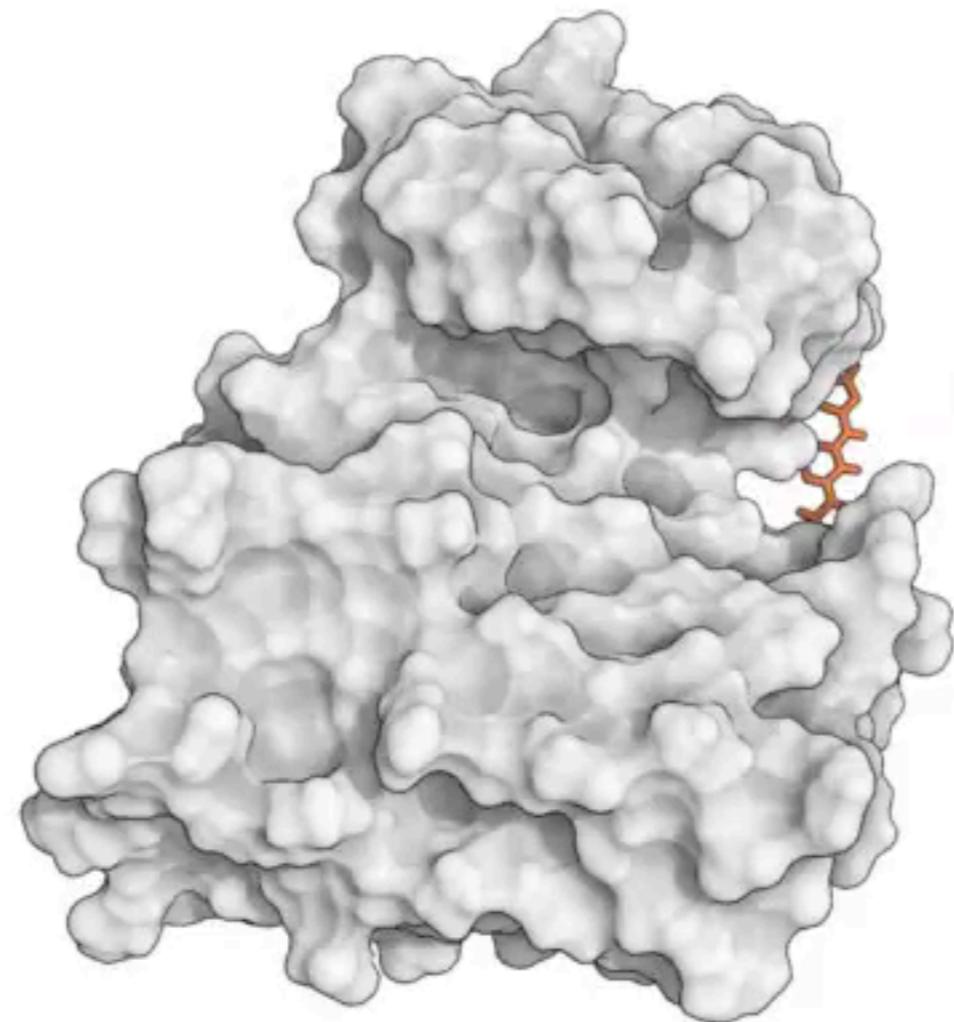
MD simulation of aqua-porin water channel



Chemistry without test-tubes

how drugs bind their targets

all-atom simulations of 2.3 μ s (in total 35 μ s)



Chemistry without test-tubes

starting structures

x-ray crystallography (www.rscb.org)

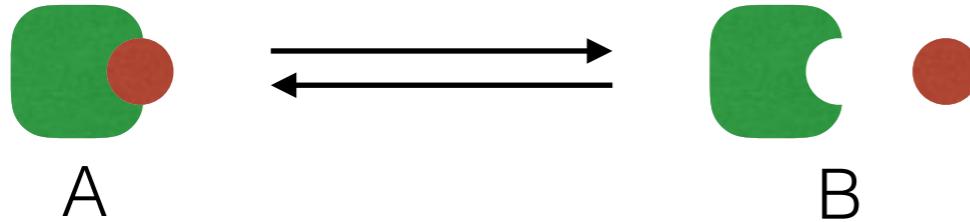
The screenshot shows the RCSB PDB homepage with a search bar at the top containing "gfp". Below the search bar, there are tabs for "386 Structures", "25 Unreleased Structures", "136 Citations", "136 Ligands", and "12 News & PDB-101 Articles". The main content area displays search parameters for "Text Search for: gfp" and "Refinements" for "ORGANISM" and "UNIPROT MOLECULE NAME". A search result for entry "4KF5" is shown, featuring a 3D ribbon model of a protein structure. The entry details include: "4KF5", "Crystal Structure of Split GFP complexed with engineered sfCherry with an insertion of GFP fragment", authors "Nguyen, H.B., Hung, L.W., Yeates, T.O., Terwilliger, T.C., Waldo, G.S.", publication "(2013) Acta Crystallogr D Biol Crystalogr 69: 2513-2523", release date "12/18/2013", method "X-ray Diffraction", resolution "2.6 Å", residue count "908", macromolecules "fluorescent protein GFP1-8 (protein)", "fluorescent protein sfCherry+GFP10 ... (protein)", unique ligands "CH6, CRO", and a search term match score of "330.10".

ready to go...

Chemistry without test-tubes

calculating free energies in MD simulations

e.g. binding



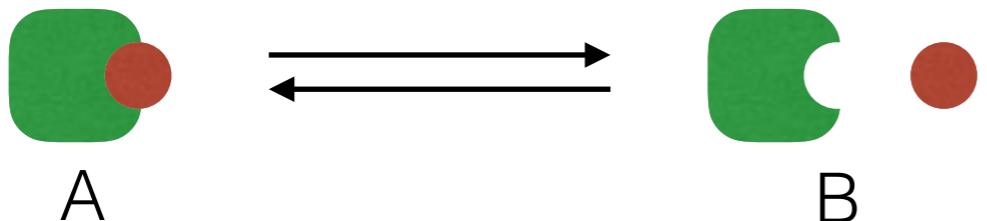
$$G_A = -kT \ln Z_A \quad G_B = -kT \ln Z_B$$

$$\Delta G = G_B - G_A = -kT \ln Z_B/Z_A$$

Chemistry without test-tubes

calculating free energies in MD simulations

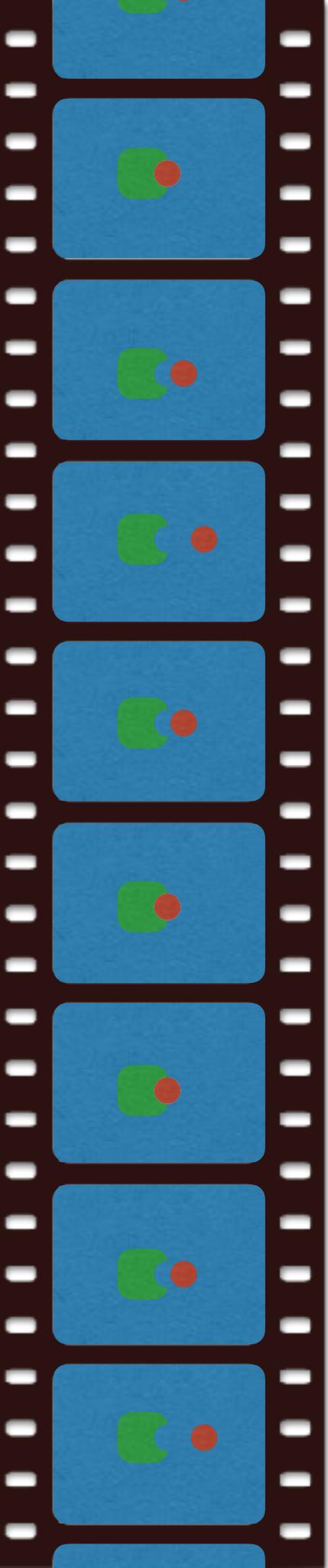
e.g. binding



$$G_A = -kT \ln Z_A$$

$$G_B = -kT \ln Z_B$$

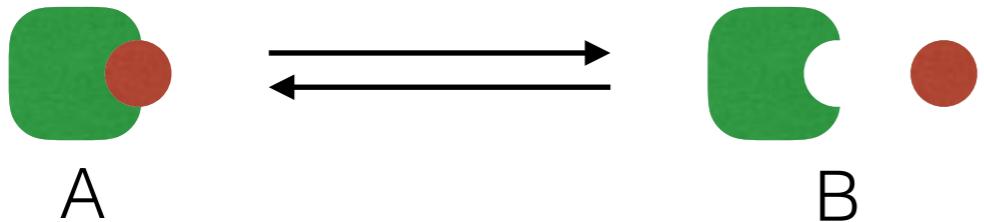
$$\Delta G = G_B - G_A = -kT \ln Z_B/Z_A$$



Chemistry without test-tubes

calculating free energies in MD simulations

e.g. binding



$$G_A = -kT \ln Z_A$$

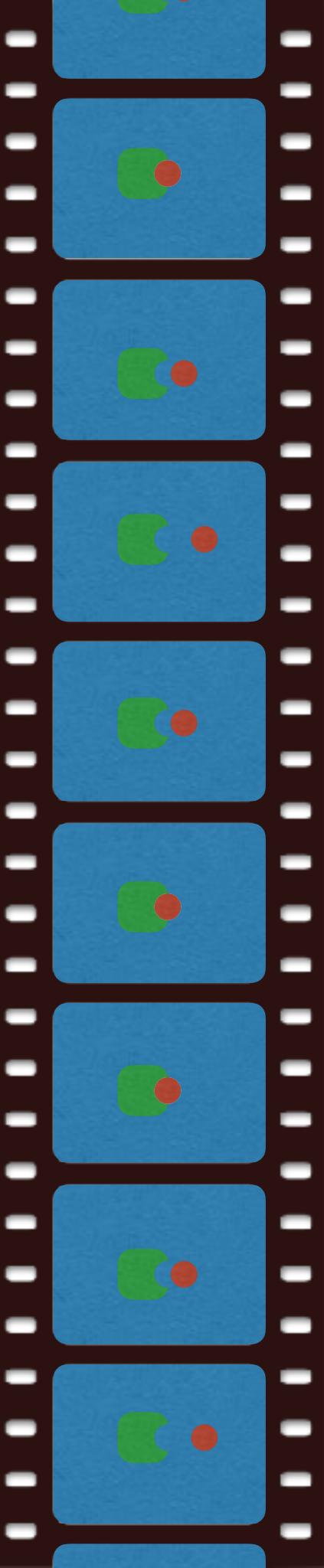
$$G_B = -kT \ln Z_B$$

$$\Delta G = G_B - G_A = -kT \ln Z_B/Z_A$$

most naive approach: counting

ergodicity: time average same as ensemble average

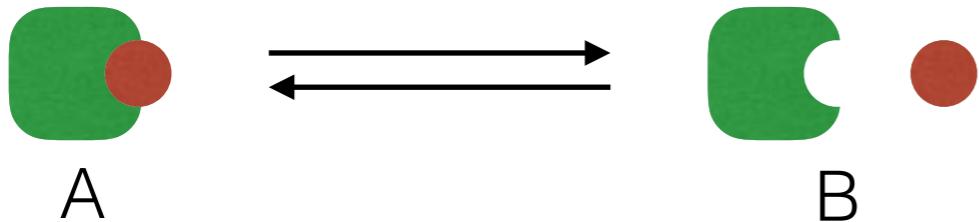
sort frames of a trajectory and compute probabilities



Chemistry without test-tubes

calculating free energies in MD simulations

e.g. binding



$$G_A = -kT \ln Z_A$$

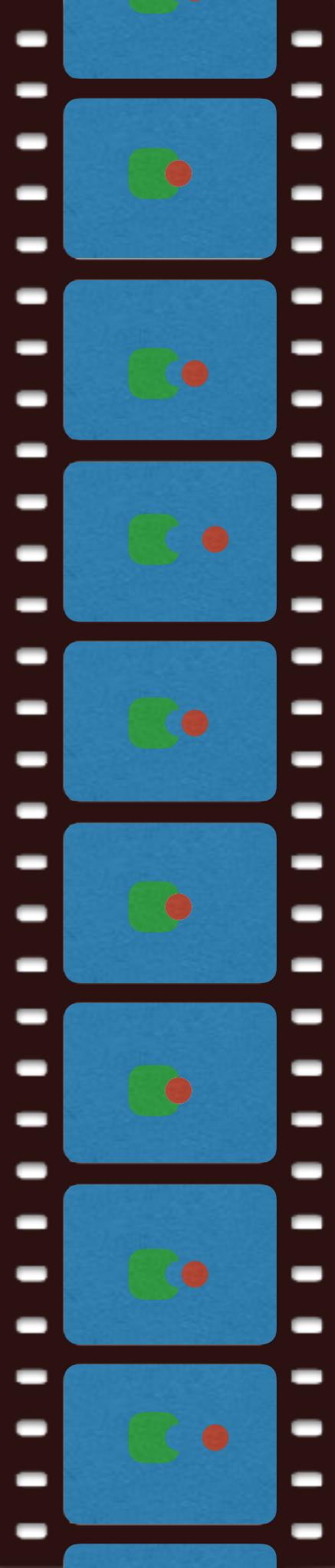
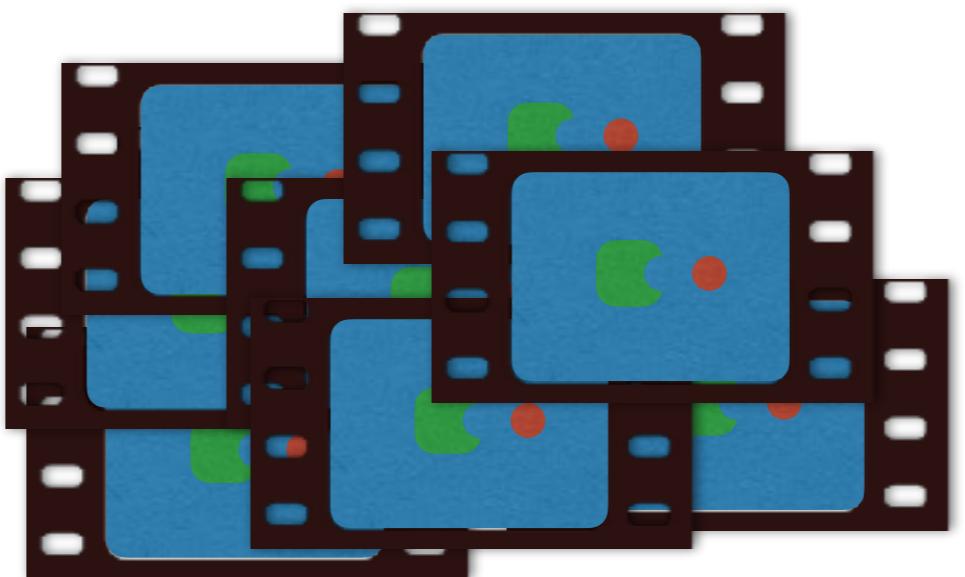
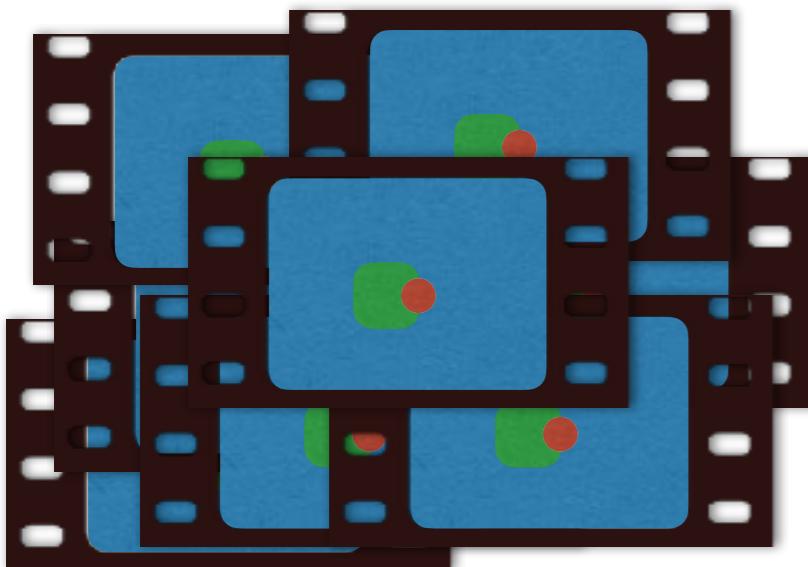
$$G_B = -kT \ln Z_B$$

$$\Delta G = G_B - G_A = -kT \ln Z_B/Z_A$$

most naive approach: counting

ergodicity: time average same as ensemble average

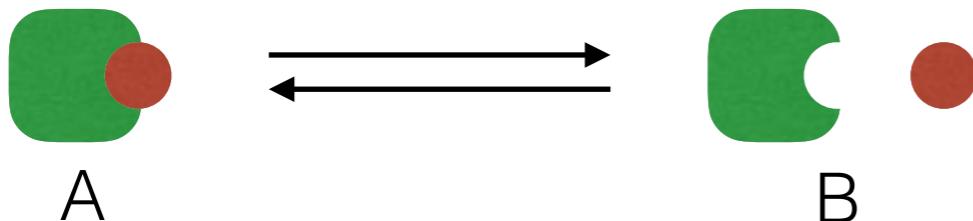
sort frames of a trajectory and compute probabilities



Chemistry without test-tubes

calculating free energies in MD simulations

e.g. binding



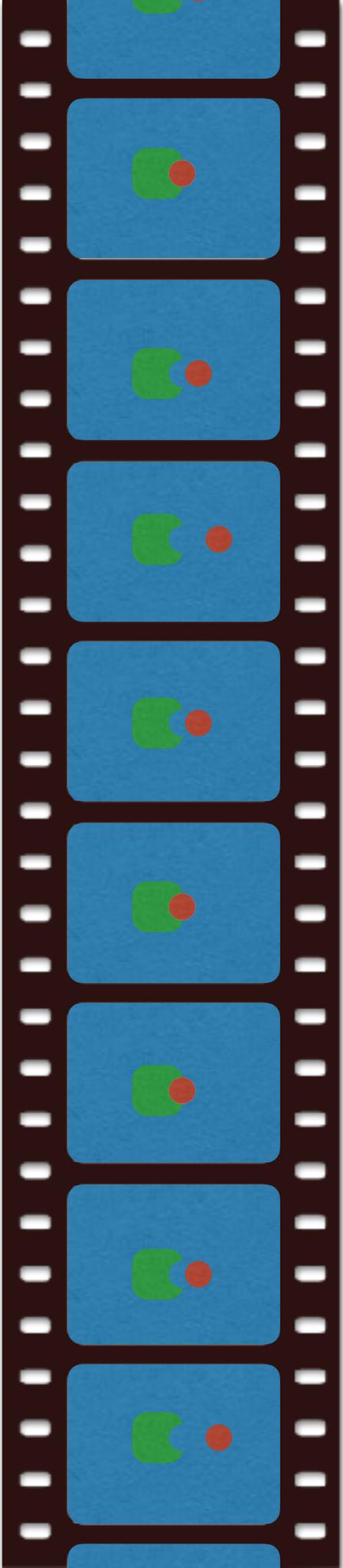
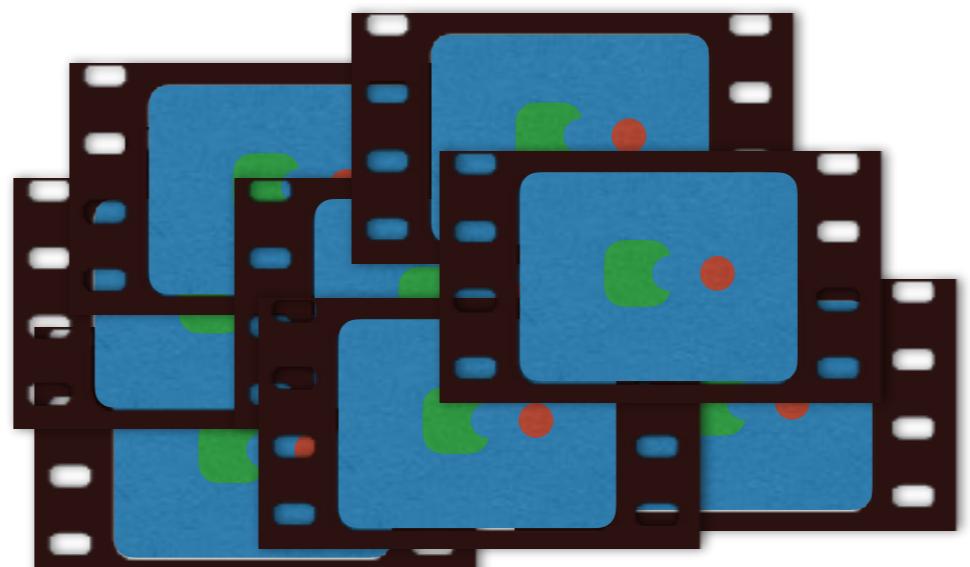
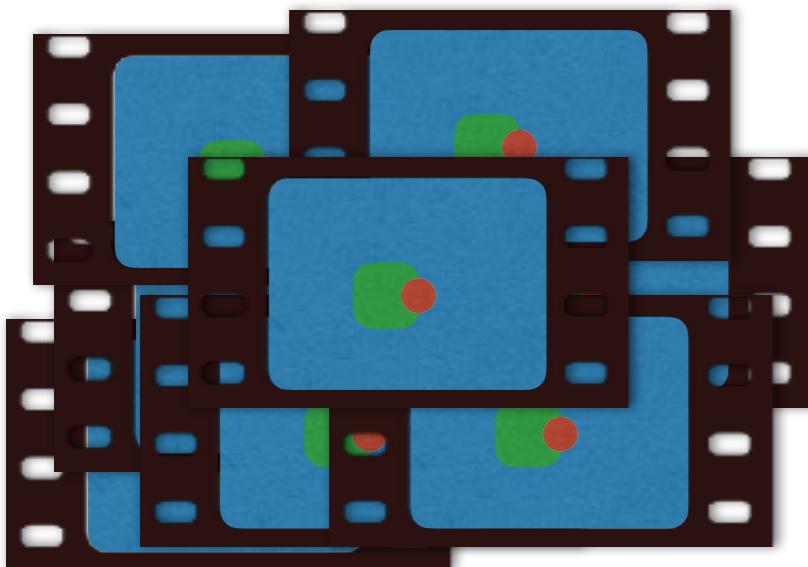
$$G_A = -kT \ln Z_A \quad G_B = -kT \ln Z_B$$

$$\Delta G = G_B - G_A = -kT \ln Z_B/Z_A$$

most naive approach: counting

ergodicity: time average same as ensemble average

sort frames of a trajectory and compute probabilities

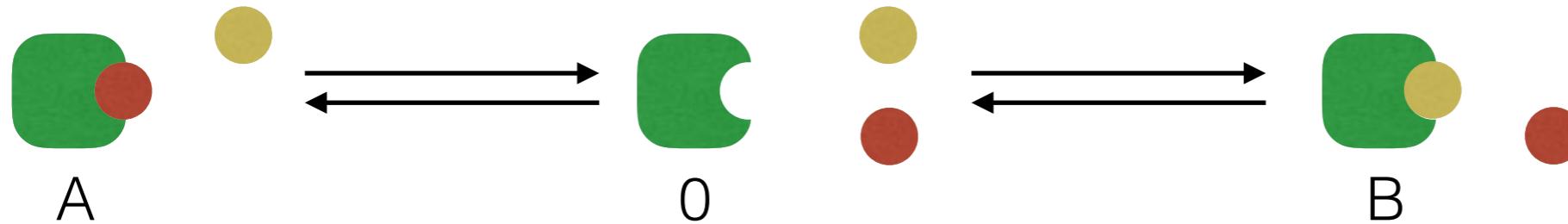


$$\frac{Z_B}{Z_A} = \frac{\sum_{i \in B} e^{-\beta E_i}}{\sum_{i \in A} e^{-\beta E_i}} = \frac{\sum_{i \in B} p_i}{\sum_{i \in A} p_i} = \frac{N_B^{\text{snapshot}}}{N_A^{\text{snapshot}}}$$

Chemistry without test-tubes

calculating free energy differences in MD simulations

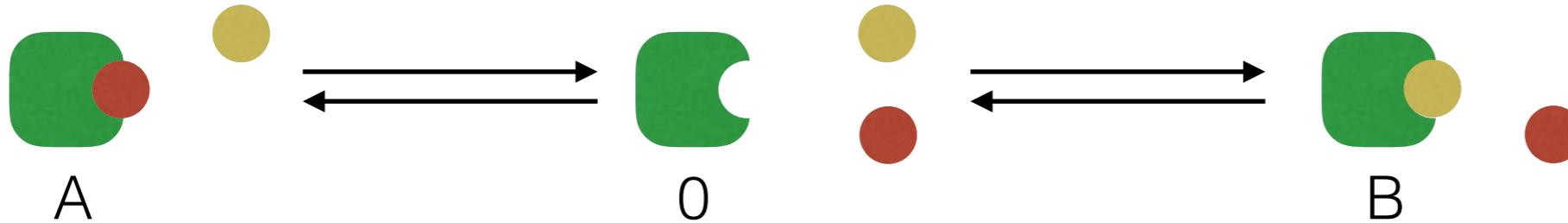
e.g. competitive binding



Chemistry without test-tubes

calculating free energy differences in MD simulations

e.g. competitive binding



$$G_A = -kT \ln Z_A$$

$$G_0 = -kT \ln Z_0$$

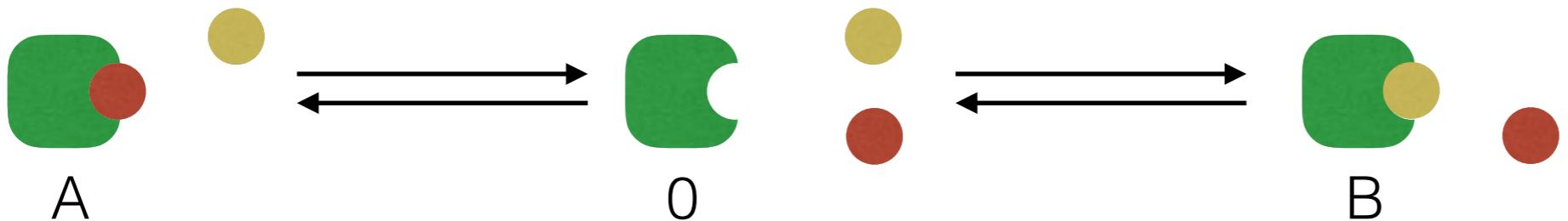
$$G_B = -kT \ln Z_B$$

$$\Delta\Delta G_{AB} = (G_B - G_0) - (G_A - G_0) = -kT \ln \left[\frac{Z_B}{Z_0} \frac{Z_0}{Z_A} \right] = -kT \ln Z_B / Z_A$$

Chemistry without test-tubes

calculating free energy differences in MD simulations

e.g. competitive binding



$$G_A = -kT \ln Z_A$$

$$G_0 = -kT \ln Z_0$$

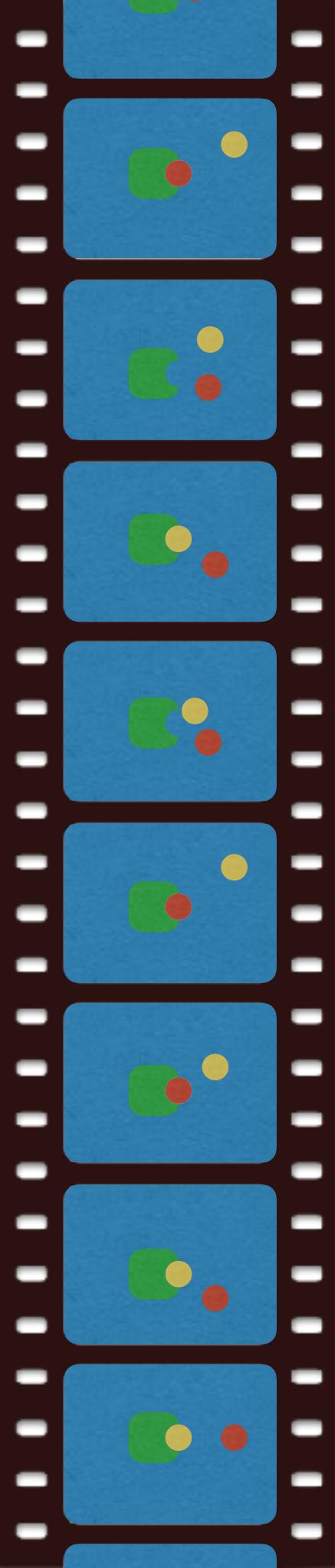
$$G_B = -kT \ln Z_B$$

$$\Delta\Delta G_{AB} = (G_B - G_0) - (G_A - G_0) = -kT \ln \left[\frac{Z_B}{Z_0} \frac{Z_0}{Z_A} \right] = -kT \ln Z_B / Z_A$$

most naive approach: counting

ergodicity: time average same as ensemble average

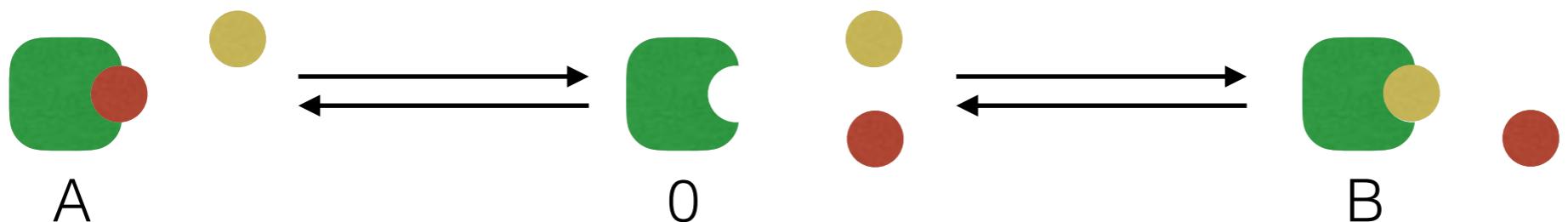
sort frames of a trajectory and compute probabilities



Chemistry without test-tubes

calculating free energy differences in MD simulations

e.g. competitive binding



$$G_A = -kT \ln Z_A$$

$$G_0 = -kT \ln Z_0$$

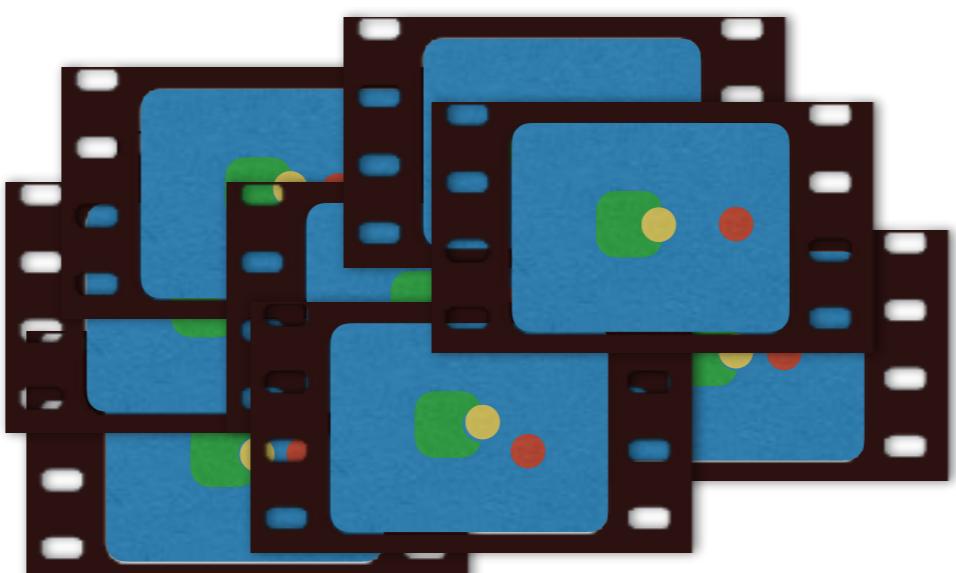
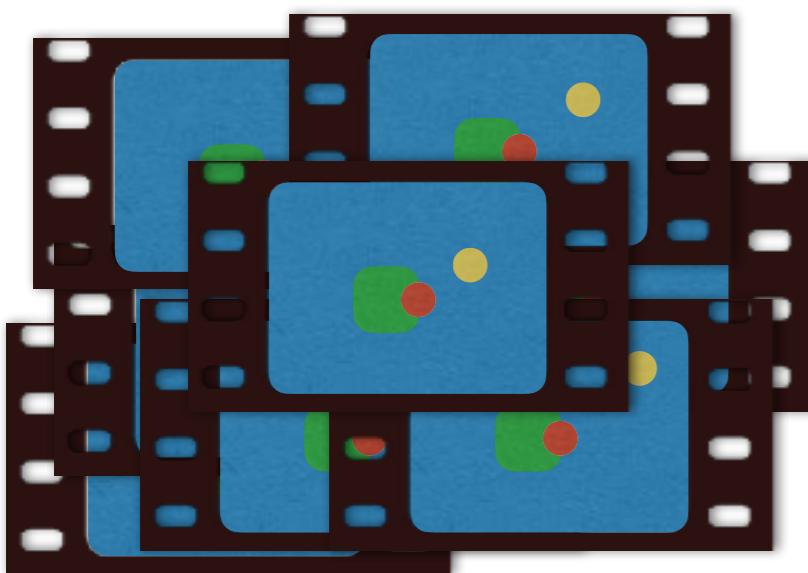
$$G_B = -kT \ln Z_B$$

$$\Delta\Delta G_{AB} = (G_B - G_0) - (G_A - G_0) = -kT \ln \left[\frac{Z_B}{Z_0} \frac{Z_0}{Z_A} \right] = -kT \ln Z_B / Z_A$$

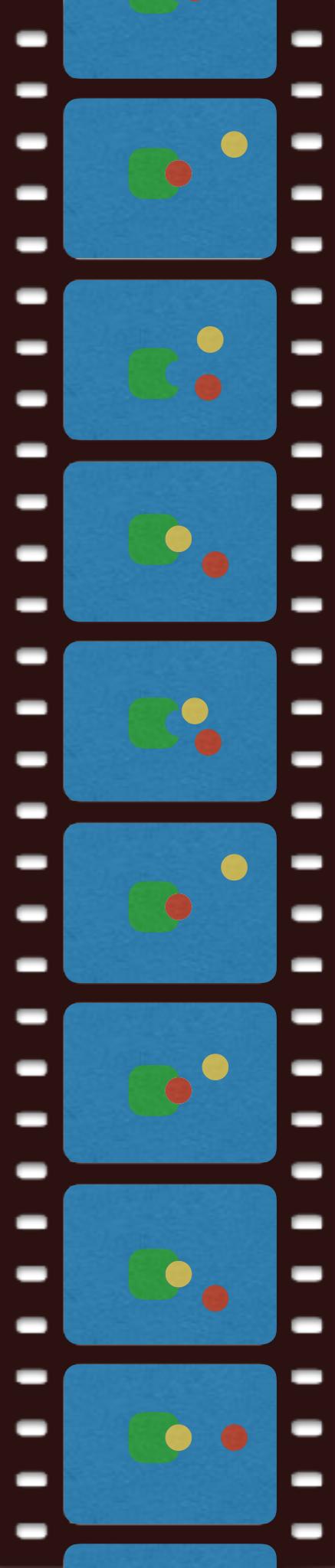
most naive approach: counting

ergodicity: time average same as ensemble average

sort frames of a trajectory and compute probabilities



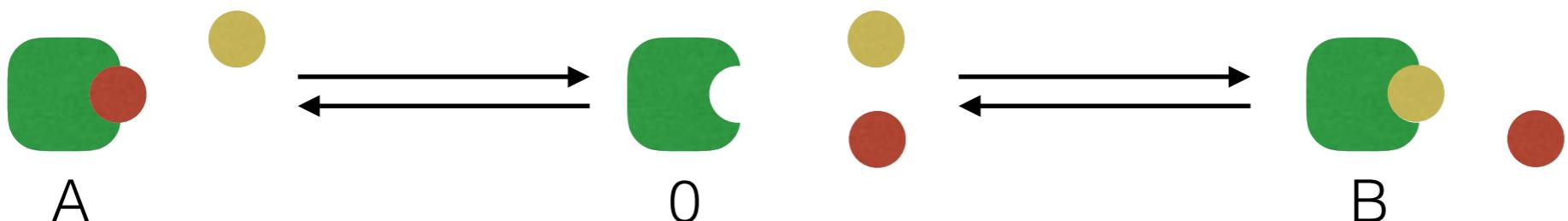
$$\frac{Z_B}{Z_A} = \frac{\sum_{i \in B} e^{-\beta E_i}}{\sum_{i \in A} e^{-\beta E_i}} = \frac{\sum_{i \in B} p_i}{\sum_{i \in A} p_i} = \frac{N_B^{\text{snapshot}}}{N_A^{\text{snapshot}}}$$



Chemistry without test-tubes

calculating free energy differences

e.g. competitive binding



$$G_A = -kT \ln Z_A$$

$$G_0 = -kT \ln Z_0$$

$$G_B = -kT \ln Z_B$$

$$\Delta\Delta G_{AB} = (G_B - G_0) - (G_A - G_0) = -kT \ln \left[\frac{Z_B}{Z_0} \frac{Z_0}{Z_A} \right] = -kT \ln Z_B / Z_A$$

most naive approach: counting

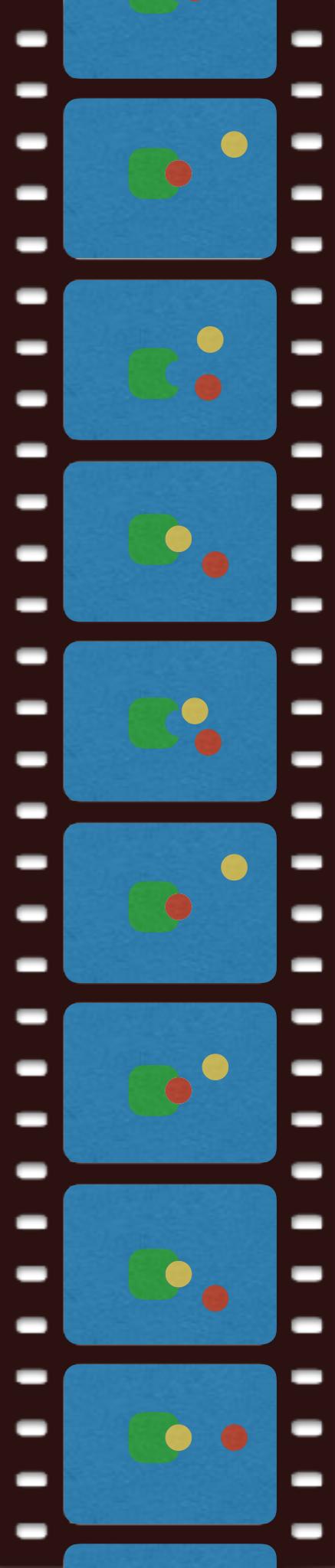
ergodicity: time average same as ensemble average

sort frames of a trajectory and compute probabilities

converged MD ensemble?

sufficiently long trajectory

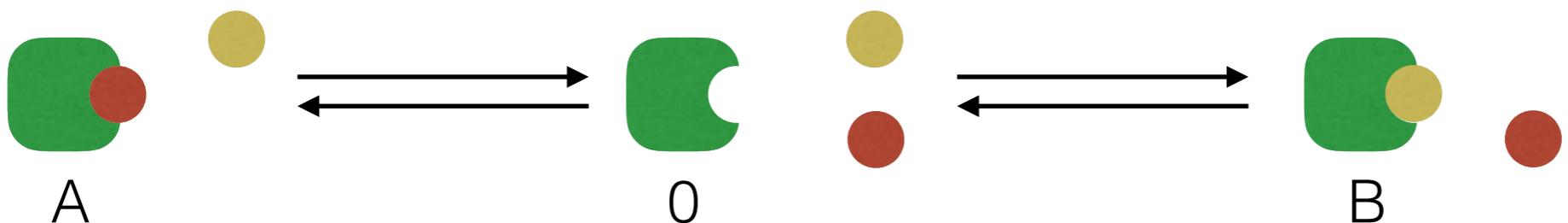
very difficult in practice



Chemistry without test-tubes

calculating free energy differences

e.g. competitive binding



$$G_A = -kT \ln Z_A$$

$$G_0 = -kT \ln Z_0$$

$$G_B = -kT \ln Z_B$$

$$\Delta\Delta G_{AB} = (G_B - G_0) - (G_A - G_0) = -kT \ln \left[\frac{Z_B}{Z_0} \frac{Z_0}{Z_A} \right] = -kT \ln Z_B / Z_A$$

most naive approach: counting

ergodicity: time average same as ensemble average

sort frames of a trajectory and compute probabilities

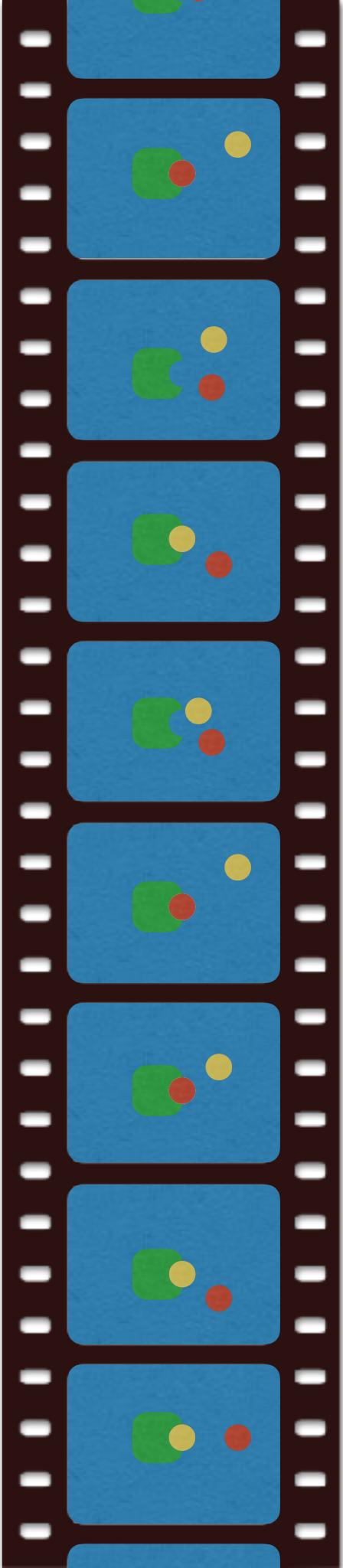
converged MD ensemble?

sufficiently long trajectory

very difficult in practice

alternative approaches

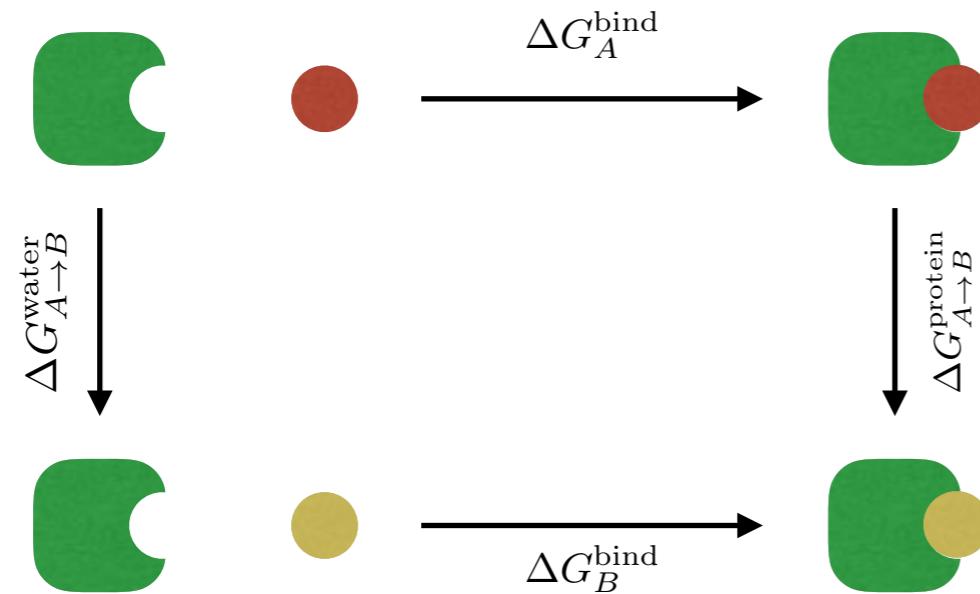
thermodynamic integration



Thermodynamic integration

calculating free energy differences

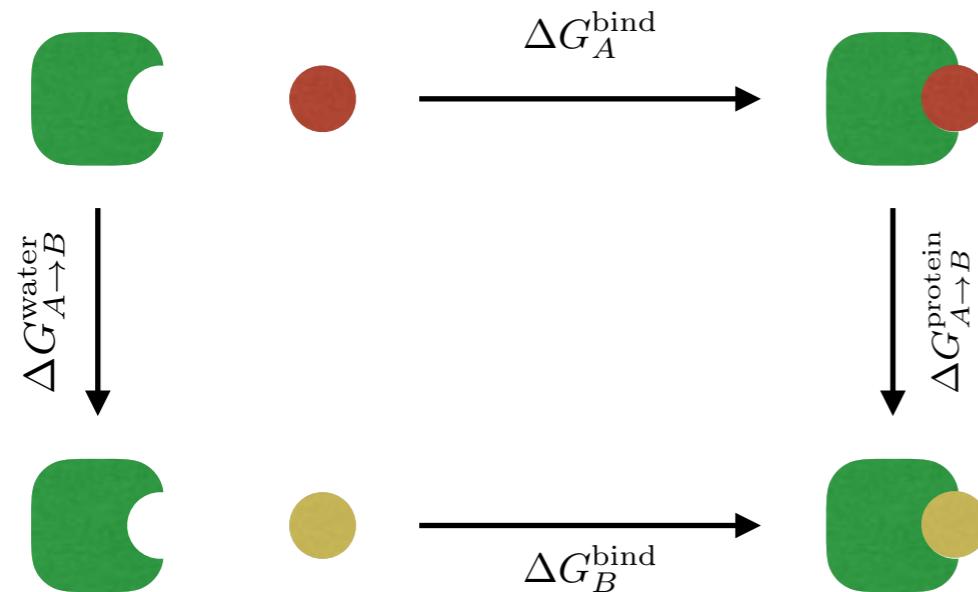
thermodynamic cycle



Thermodynamic integration

calculating free energy differences

thermodynamic cycle



$$\begin{aligned}\Delta\Delta G_{AB}^{\text{bind}} &= \Delta G_B^{\text{bind}} - \Delta G_A^{\text{bind}} \\ &= \Delta G_{A \rightarrow B}^{\text{protein}} - \Delta G_{A \rightarrow B}^{\text{water}}\end{aligned}$$

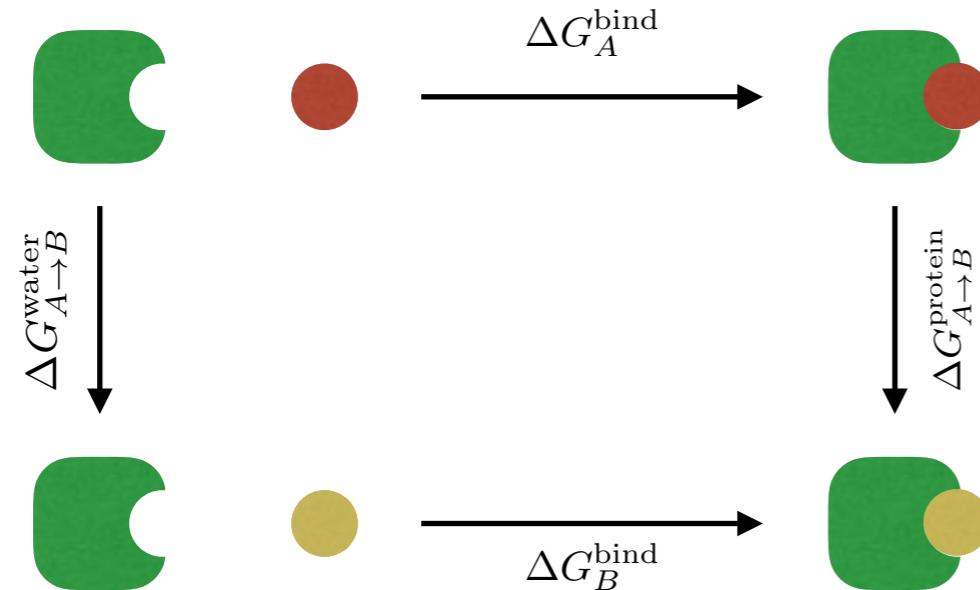
free energy is a state function: independent of path

$$\Delta G_A^{\text{bind}} + \Delta G_{A \rightarrow B}^{\text{protein}} = \Delta G_{A \rightarrow B}^{\text{water}} + \Delta G_B^{\text{bind}}$$

Thermodynamic integration

calculating free energy differences

thermodynamic cycle



$$\begin{aligned}\Delta\Delta G_{AB}^{\text{bind}} &= \Delta G_B^{\text{bind}} - \Delta G_A^{\text{bind}} \\ &= \Delta G_{A \rightarrow B}^{\text{protein}} - \Delta G_{A \rightarrow B}^{\text{water}}\end{aligned}$$

free energy is a state function: independent of path

$$\Delta G_A^{\text{bind}} + \Delta G_{A \rightarrow B}^{\text{protein}} = \Delta G_{A \rightarrow B}^{\text{water}} + \Delta G_B^{\text{bind}}$$

difficult to evaluate with MD simulations

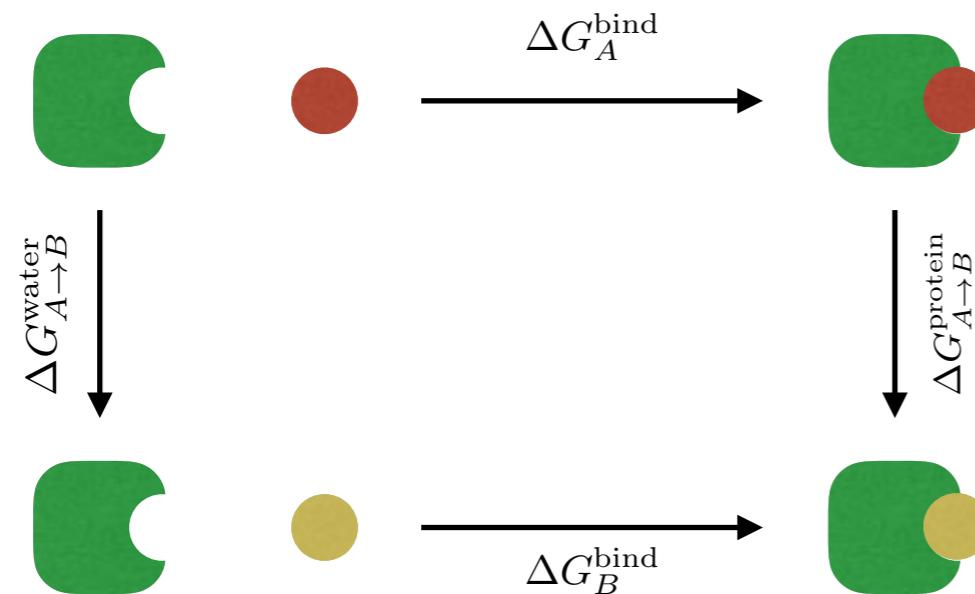
$$\Delta G_A^{\text{bind}}$$

$$\Delta G_B^{\text{bind}}$$

Thermodynamic integration

calculating free energy differences

thermodynamic cycle



$$\begin{aligned}\Delta\Delta G_{AB}^{\text{bind}} &= \Delta G_B^{\text{bind}} - \Delta G_A^{\text{bind}} \\ &= \Delta G_{A \rightarrow B}^{\text{protein}} - \Delta G_{A \rightarrow B}^{\text{water}}\end{aligned}$$

free energy is a state function: independent of path

$$\Delta G_A^{\text{bind}} + \Delta G_{A \rightarrow B}^{\text{protein}} = \Delta G_{A \rightarrow B}^{\text{water}} + \Delta G_B^{\text{bind}}$$

difficult to evaluate with MD simulations

$$\Delta G_A^{\text{bind}}$$

$$\Delta G_B^{\text{bind}}$$

'easy' (sometimes) to evaluate with MD simulations

$$\Delta G_{A \rightarrow B}^{\text{water}}$$

$$\Delta G_{A \rightarrow B}^{\text{protein}}$$

Thermodynamic integration

why ‘easy’?

interpolate the energy function

in silico everything is possible



Thermodynamic integration

why ‘easy’?

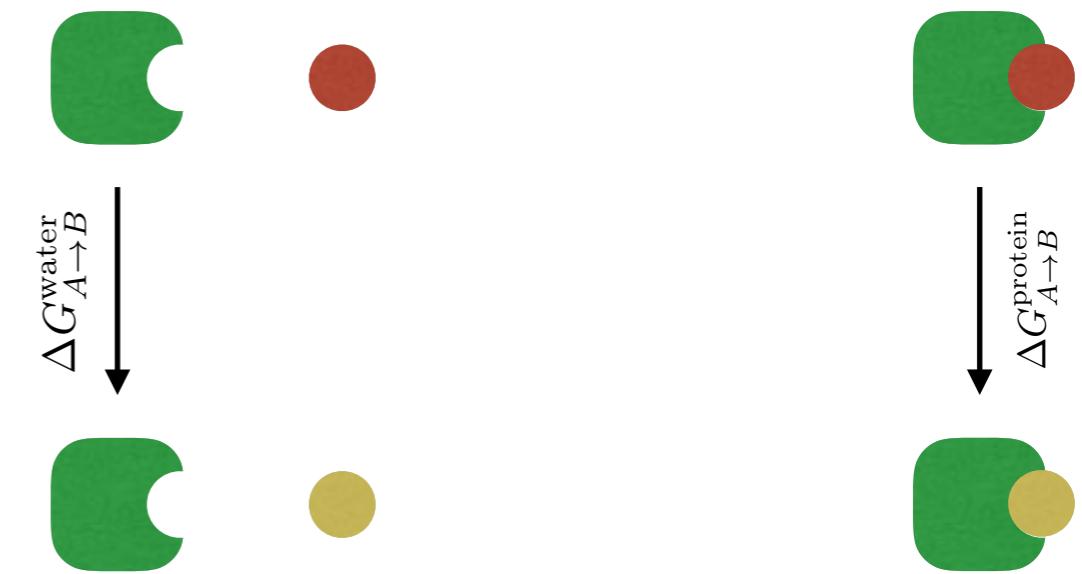
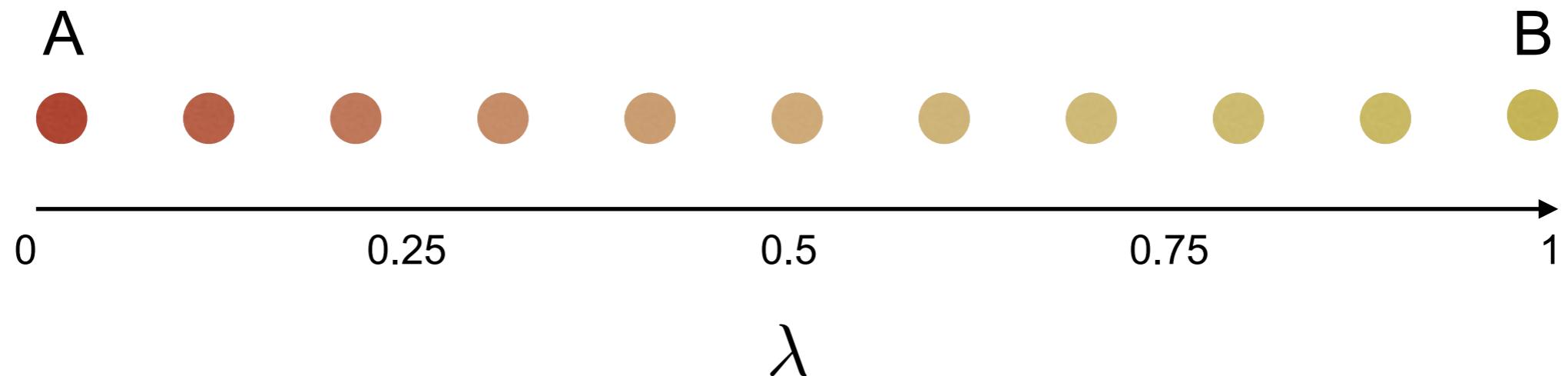
interpolate the energy function

in silico everything is possible

$$E^A(\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{r}_1, \mathbf{r}_2)$$

$$E^B(\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{r}_1, \mathbf{r}_2)$$

$$E(\lambda, \mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{r}_1, \mathbf{r}_2) = (1 - \lambda)E^A(\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{r}_1, \mathbf{r}_2) + \lambda E^B(\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{r}_1, \mathbf{r}_2)$$



Thermodynamic integration

why ‘easy’?

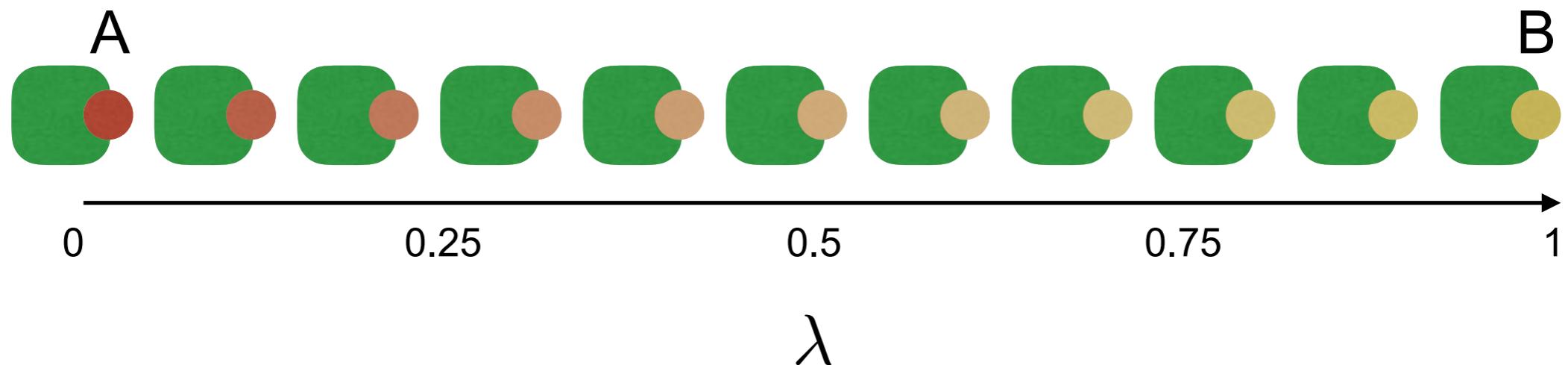
interpolate the energy function

in silico everything is possible

$$E^A(\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{r}_1, \mathbf{r}_2)$$

$$E^B(\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{r}_1, \mathbf{r}_2)$$

$$E(\lambda, \mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{r}_1, \mathbf{r}_2) = (1 - \lambda)E^A(\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{r}_1, \mathbf{r}_2) + \lambda E^B(\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{r}_1, \mathbf{r}_2)$$



Thermodynamic integration

why ‘easy’?

interpolate the energy function

in silico everything is possible

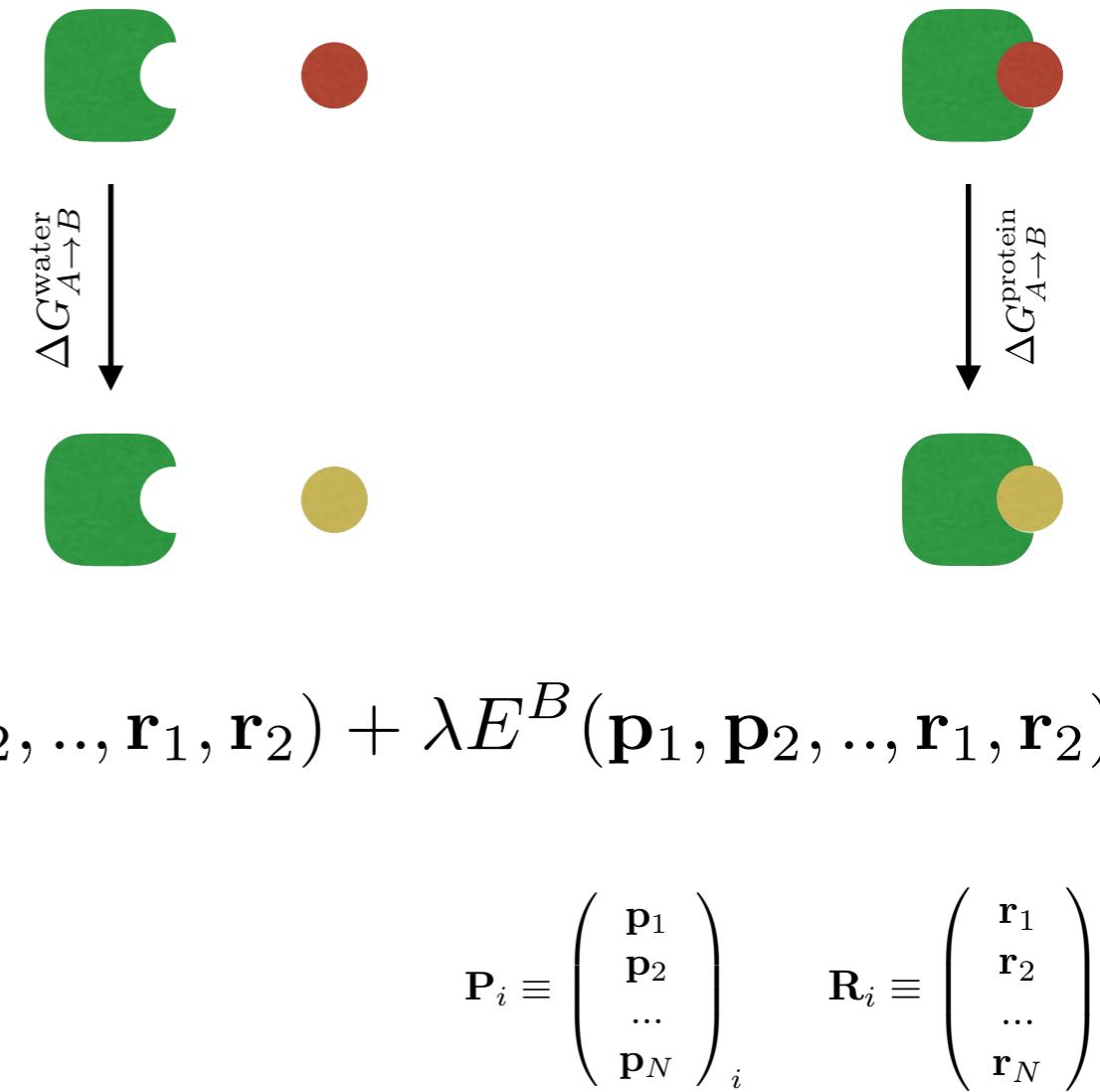
$$E^A(\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{r}_1, \mathbf{r}_2)$$

$$E^B(\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{r}_1, \mathbf{r}_2)$$

$$E(\lambda, \mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{r}_1, \mathbf{r}_2) = (1 - \lambda)E^A(\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{r}_1, \mathbf{r}_2) + \lambda E^B(\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{r}_1, \mathbf{r}_2)$$

partition function

$$\begin{aligned} Z(\lambda) &= \sum_i e^{-\beta E(\lambda, \mathbf{P}_i, \mathbf{R}_i)} \\ &= \sum_i e^{-\beta[(1-\lambda)E^A(\mathbf{P}_i, \mathbf{R}_i) + \lambda E^B(\mathbf{P}_i, \mathbf{R}_i)]} \end{aligned}$$



Thermodynamic integration

why ‘easy’?

interpolate the energy function

in silico everything is possible

$$E^A(\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{r}_1, \mathbf{r}_2)$$

$$E^B(\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{r}_1, \mathbf{r}_2)$$

$$E(\lambda, \mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{r}_1, \mathbf{r}_2) = (1 - \lambda)E^A(\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{r}_1, \mathbf{r}_2) + \lambda E^B(\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{r}_1, \mathbf{r}_2)$$

partition function

$$\begin{aligned} Z(\lambda) &= \sum_i e^{-\beta E(\lambda, \mathbf{P}_i, \mathbf{R}_i)} \\ &= \sum_i e^{-\beta[(1-\lambda)E^A(\mathbf{P}_i, \mathbf{R}_i) + \lambda E^B(\mathbf{P}_i, \mathbf{R}_i)]} \end{aligned}$$

free energy

$$G(\lambda) = -kT \ln Z(\lambda)$$

how does this help?



$$\mathbf{P}_i \equiv \begin{pmatrix} \mathbf{p}_1 \\ \mathbf{p}_2 \\ \dots \\ \mathbf{p}_N \end{pmatrix}_i \quad \mathbf{R}_i \equiv \begin{pmatrix} \mathbf{r}_1 \\ \mathbf{r}_2 \\ \dots \\ \mathbf{r}_N \end{pmatrix}_i$$

Thermodynamic integration

why ‘easy’?

interpolate the energy function

partition function

replace sum by integration (classical)

$$\begin{aligned} Z(\lambda) &= \sum_i e^{-\beta E(\lambda, \mathbf{P}_i, \mathbf{R}_i)} \\ &= \int \int e^{-\beta E(\lambda, \mathbf{P}, \mathbf{R})} d\mathbf{P} d\mathbf{R} \end{aligned}$$



$$\mathbf{P} \equiv \begin{pmatrix} \mathbf{p}_1 \\ \mathbf{p}_2 \\ \dots \\ \mathbf{p}_N \end{pmatrix} \quad \mathbf{R} \equiv \begin{pmatrix} \mathbf{r}_1 \\ \mathbf{r}_2 \\ \dots \\ \mathbf{r}_N \end{pmatrix}$$

Thermodynamic integration

why ‘easy’?

interpolate the energy function

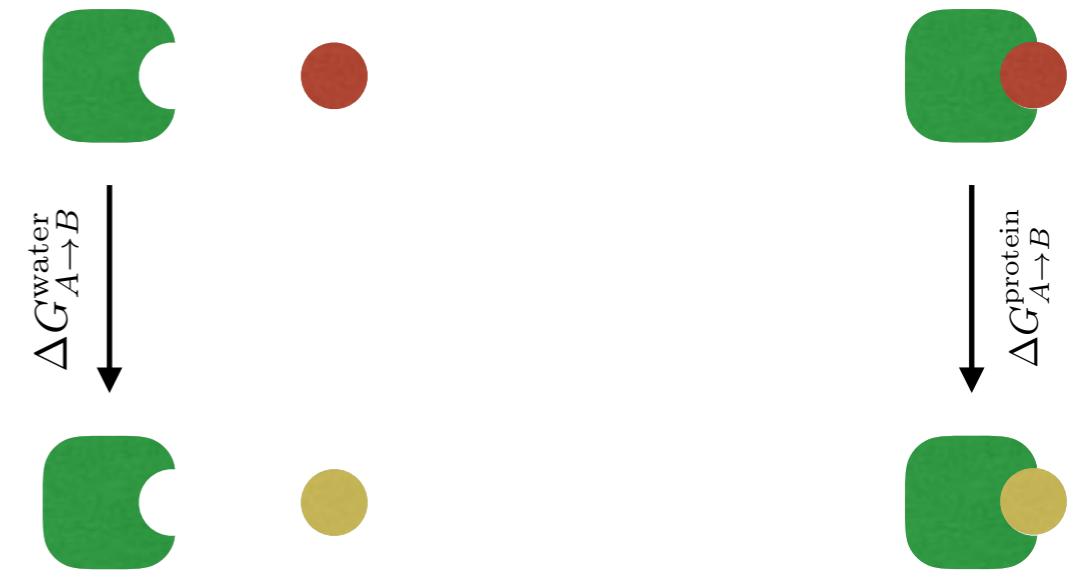
partition function

replace sum by integration (classical)

$$\begin{aligned} Z(\lambda) &= \sum_i e^{-\beta E(\lambda, \mathbf{P}_i, \mathbf{R}_i)} \\ &= \int \int e^{-\beta E(\lambda, \mathbf{P}, \mathbf{R})} d\mathbf{P} d\mathbf{R} \end{aligned}$$

free energy

$$G(\lambda) = -kT \ln Z(\lambda)$$



$$\mathbf{P} \equiv \begin{pmatrix} \mathbf{p}_1 \\ \mathbf{p}_2 \\ \dots \\ \mathbf{p}_N \end{pmatrix} \quad \mathbf{R} \equiv \begin{pmatrix} \mathbf{r}_1 \\ \mathbf{r}_2 \\ \dots \\ \mathbf{r}_N \end{pmatrix}$$

Thermodynamic integration

why ‘easy’?

interpolate the energy function

partition function

replace sum by integration (classical)

$$\begin{aligned} Z(\lambda) &= \sum_i e^{-\beta E(\lambda, \mathbf{P}_i, \mathbf{R}_i)} \\ &= \int \int e^{-\beta E(\lambda, \mathbf{P}, \mathbf{R})} d\mathbf{P} d\mathbf{R} \end{aligned}$$

free energy

$$G(\lambda) = -kT \ln Z(\lambda)$$

derivative of free energy

$$\begin{aligned} \frac{\partial G}{\partial \lambda} &= -\frac{kT}{Z(\lambda)} \frac{\partial Z}{\partial \lambda} \\ &= \int \int \frac{\partial E(\lambda)}{\partial \lambda} \frac{e^{-\beta E(\lambda, \mathbf{P}, \mathbf{R})}}{Z} d\mathbf{P} d\mathbf{R} \\ &= \left\langle \frac{\partial E(\lambda)}{\partial \lambda} \right\rangle_\lambda \end{aligned}$$



$$\mathbf{P} \equiv \begin{pmatrix} \mathbf{p}_1 \\ \mathbf{p}_2 \\ \dots \\ \mathbf{p}_N \end{pmatrix} \quad \mathbf{R} \equiv \begin{pmatrix} \mathbf{r}_1 \\ \mathbf{r}_2 \\ \dots \\ \mathbf{r}_N \end{pmatrix}$$

Thermodynamic integration

why ‘easy’?

interpolate the energy function

derivative of free energy

$$\begin{aligned}\frac{\partial G}{\partial \lambda} &= -\frac{kT}{Z(\lambda)} \frac{\partial Z}{\partial \lambda} \\ &= \int \int \frac{\partial E(\lambda)}{\partial \lambda} \frac{e^{-\beta E(\lambda, \mathbf{P}, \mathbf{R})}}{Z} d\mathbf{P} d\mathbf{R} \\ &= \left\langle \frac{\partial E(\lambda)}{\partial \lambda} \right\rangle_{\lambda}\end{aligned}$$



Thermodynamic integration

why ‘easy’?

interpolate the energy function

derivative of free energy

$$\begin{aligned}\frac{\partial G}{\partial \lambda} &= -\frac{kT}{Z(\lambda)} \frac{\partial Z}{\partial \lambda} \\ &= \int \int \frac{\partial E(\lambda)}{\partial \lambda} \frac{e^{-\beta E(\lambda, \mathbf{P}, \mathbf{R})}}{Z} d\mathbf{P} d\mathbf{R} \\ &= \left\langle \frac{\partial E(\lambda)}{\partial \lambda} \right\rangle_{\lambda}\end{aligned}$$

free energy

$$\Delta G_{A \rightarrow B} = \int_0^1 \left\langle \frac{\partial E}{\partial \lambda} \right\rangle_{\lambda} d\lambda$$



Thermodynamic integration

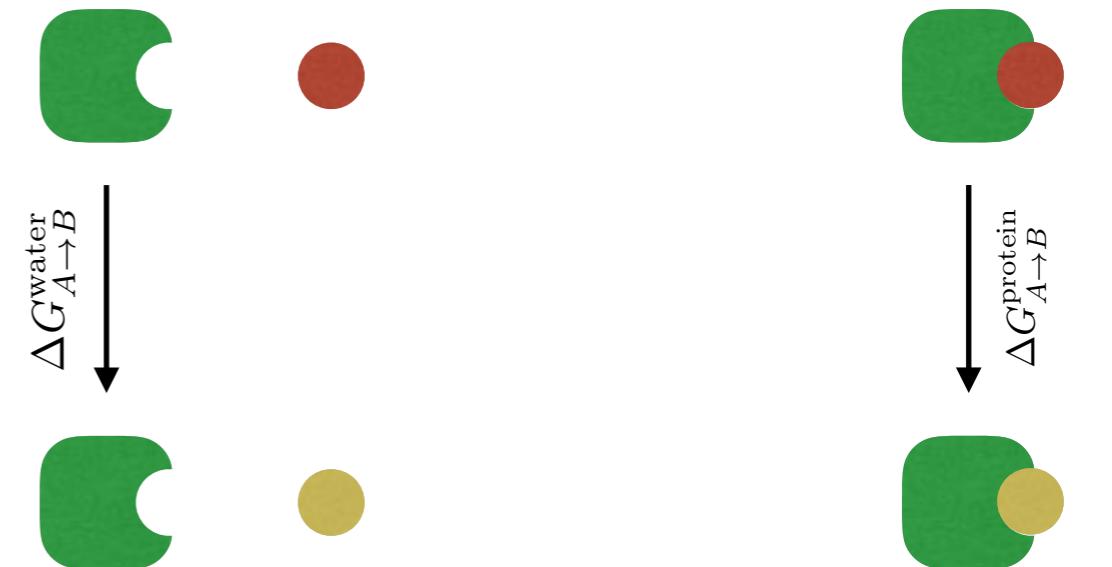
why ‘easy’?

interpolate the energy function

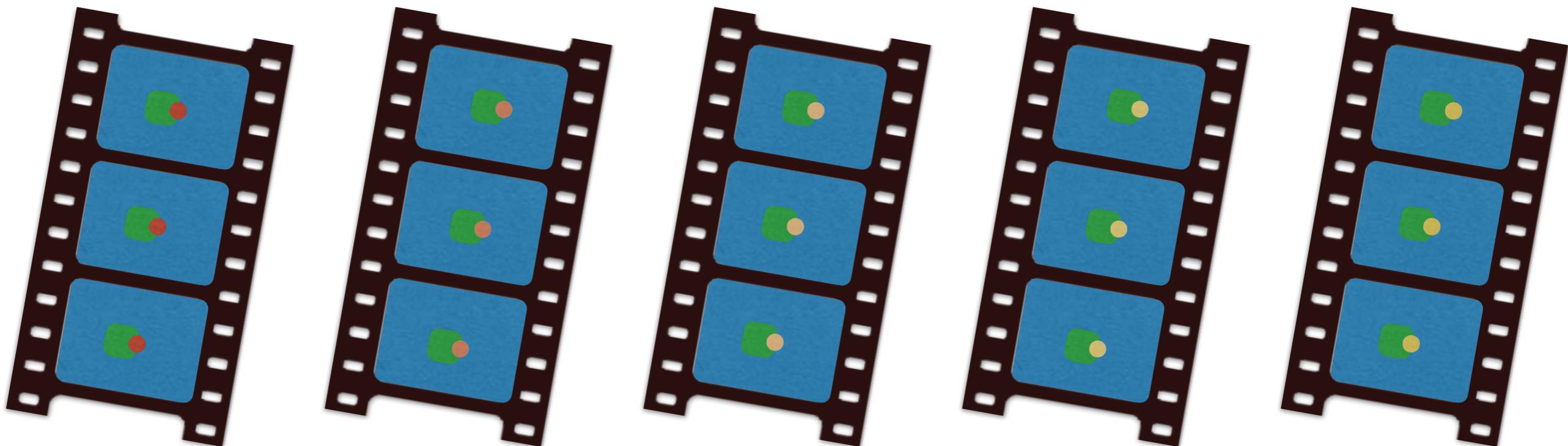
free energy

$$\Delta G_{A \rightarrow B} = \int_0^1 \left\langle \frac{\partial E}{\partial \lambda} \right\rangle_\lambda d\lambda$$

multiple MD trajectories



$$\Delta G_{A \rightarrow B}^{\text{protein}}$$



$$\left\langle \frac{\partial E}{\partial \lambda} \right\rangle_{\lambda=0.0}$$

$$\left\langle \frac{\partial E}{\partial \lambda} \right\rangle_{\lambda=0.25}$$

$$\left\langle \frac{\partial E}{\partial \lambda} \right\rangle_{\lambda=0.50}$$

$$\left\langle \frac{\partial E}{\partial \lambda} \right\rangle_{\lambda=0.75}$$

$$\left\langle \frac{\partial E}{\partial \lambda} \right\rangle_{\lambda=1.0}$$

Thermodynamic integration

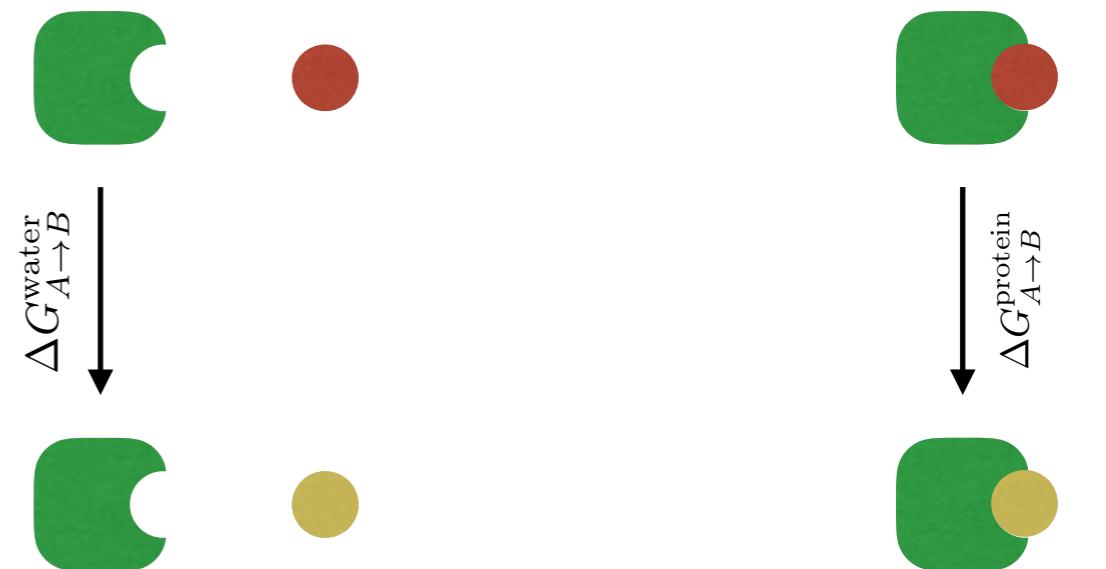
why ‘easy’?

interpolate the energy function

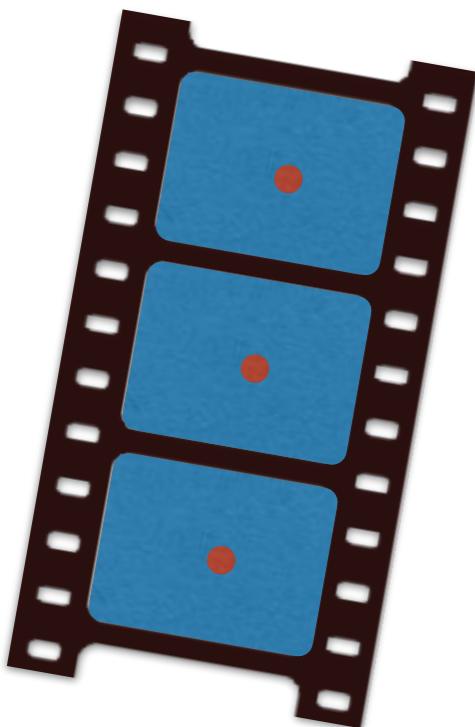
free energy

$$\Delta G_{A \rightarrow B} = \int_0^1 \left\langle \frac{\partial E}{\partial \lambda} \right\rangle_\lambda d\lambda$$

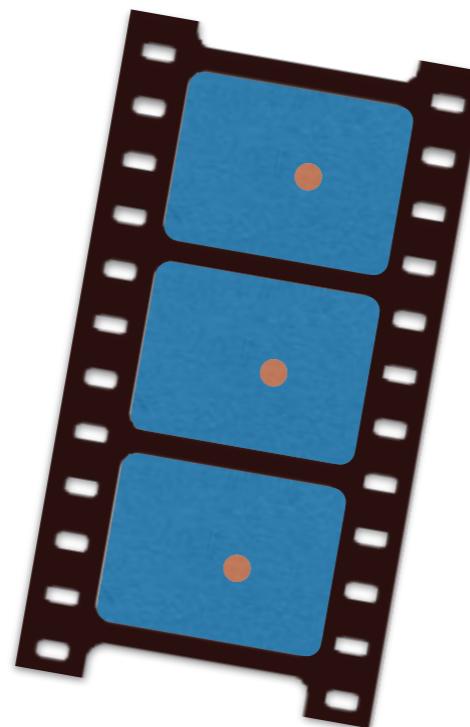
multiple MD trajectories



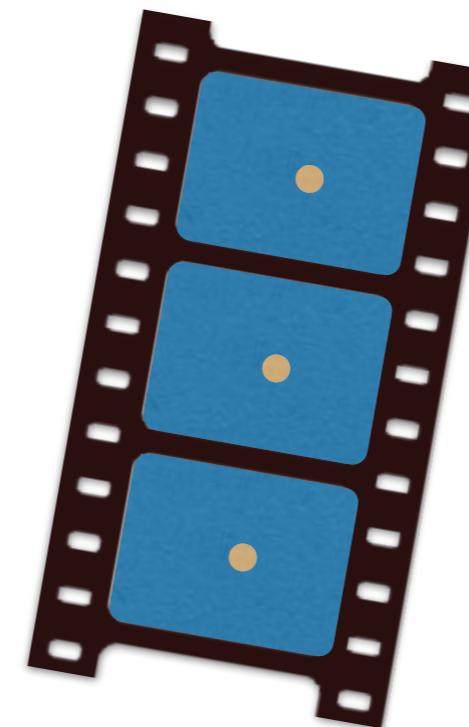
$$\Delta G_{A \rightarrow B}^{\text{water}}$$



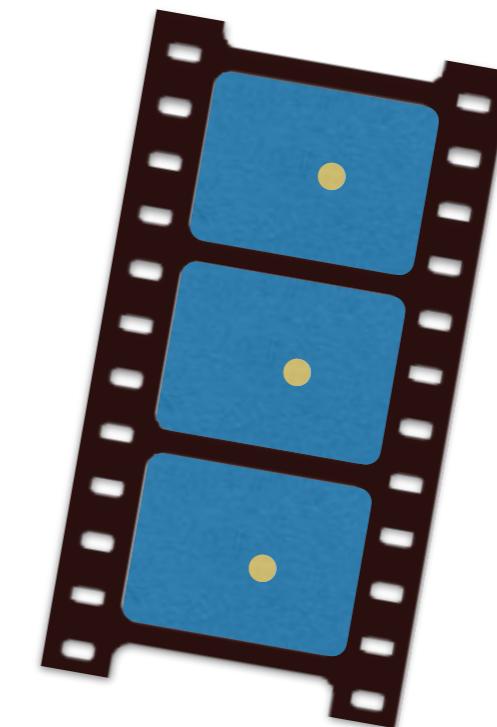
$$\left\langle \frac{\partial E}{\partial \lambda} \right\rangle_{\lambda=0.0}$$



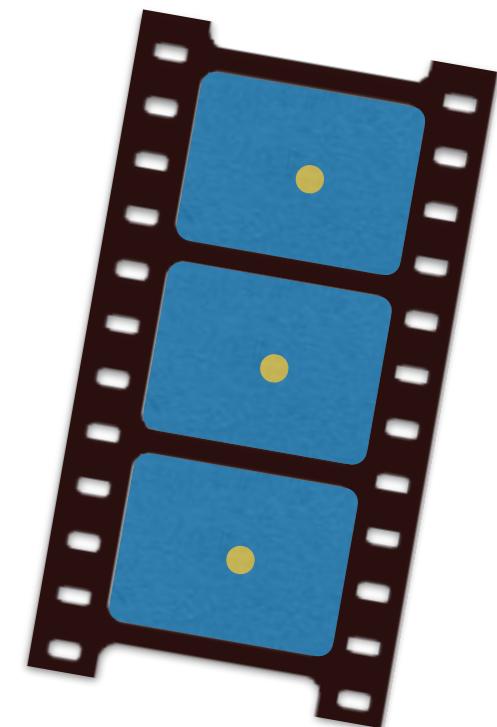
$$\left\langle \frac{\partial E}{\partial \lambda} \right\rangle_{\lambda=0.25}$$



$$\left\langle \frac{\partial E}{\partial \lambda} \right\rangle_{\lambda=0.50}$$



$$\left\langle \frac{\partial E}{\partial \lambda} \right\rangle_{\lambda=0.75}$$



$$\left\langle \frac{\partial E}{\partial \lambda} \right\rangle_{\lambda=1.0}$$

Thermodynamic integration

in practice

energy function (Hamiltonian)

$$E = H = E_{\text{kin}} + E_{\text{pot}}$$



Thermodynamic integration

in practice

energy function (Hamiltonian)

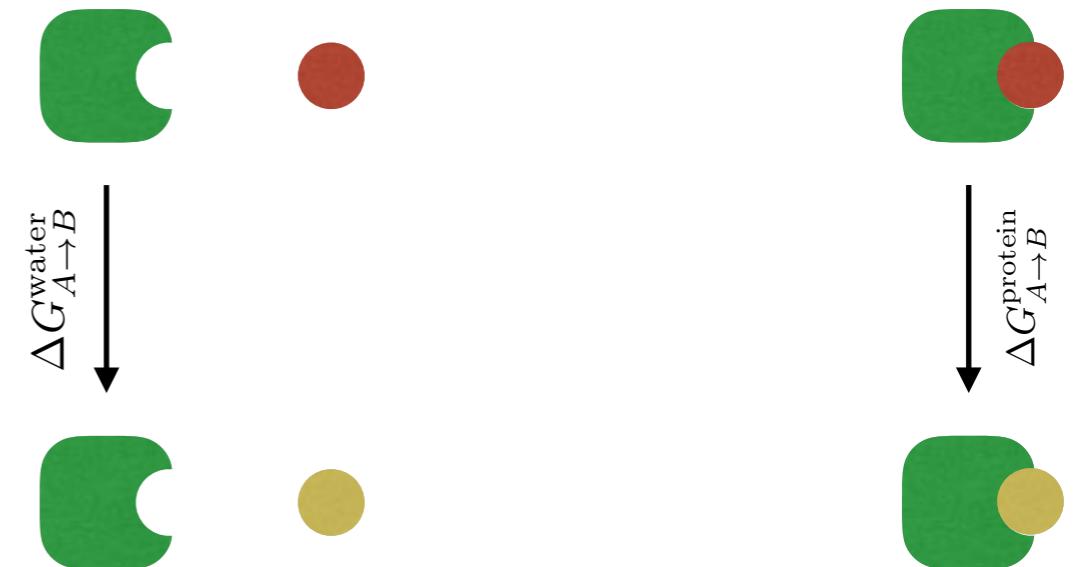
$$E = H = E_{\text{kin}} + E_{\text{pot}}$$

leave kinetic energy untouched

$$E_{\text{kin}} = \frac{1}{2} \sum_i^N \frac{p_i^2}{m_i}$$

interpolate only potential energy

$$E_{\text{pot}} = V(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N, \lambda)$$



Thermodynamic integration

in practice

energy function (Hamiltonian)

$$E = H = E_{\text{kin}} + E_{\text{pot}}$$

leave kinetic energy untouched

$$E_{\text{kin}} = \frac{1}{2} \sum_i^N \frac{p_i^2}{m_i}$$

interpolate only potential energy

$$E_{\text{pot}} = V(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N, \lambda)$$

partition function

$$Z(\lambda) = \int e^{-\beta E_{\text{kin}}(\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_N)} d\mathbf{p}_1 d\mathbf{p}_2 \dots d\mathbf{p}_N \cdot \int e^{-\beta E_{\text{pot}}(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N, \lambda)} d\mathbf{r}_1 d\mathbf{r}_2 \dots d\mathbf{r}_N$$

always same!



Thermodynamic integration

in practice

energy function (Hamiltonian)

$$E = H = E_{\text{kin}} + E_{\text{pot}}$$

leave kinetic energy untouched

$$E_{\text{kin}} = \frac{1}{2} \sum_i^N \frac{p_i^2}{m_i}$$

interpolate only potential energy

$$E_{\text{pot}} = V(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N, \lambda)$$

partition function

$$Z(\lambda) = \int e^{-\beta E_{\text{kin}}(\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_N)} d\mathbf{p}_1 d\mathbf{p}_2 \dots d\mathbf{p}_N \cdot \int e^{-\beta E_{\text{pot}}(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N, \lambda)} d\mathbf{r}_1 d\mathbf{r}_2 \dots d\mathbf{r}_N$$

always same!

derivative of free energy

$$\frac{\partial G}{\partial \lambda} = \left\langle \frac{\partial V(\lambda)}{\partial \lambda} \right\rangle_\lambda$$

