

CHEM 436 / 630

Molecular modelling of proteins

Winter 2018 Term

Instructor:

Guillaume Lamoureux
Concordia University, Montréal, Canada

Molecular dynamics

“Putting the thermal fluctuations
back into a protein structure”

Basic techniques :

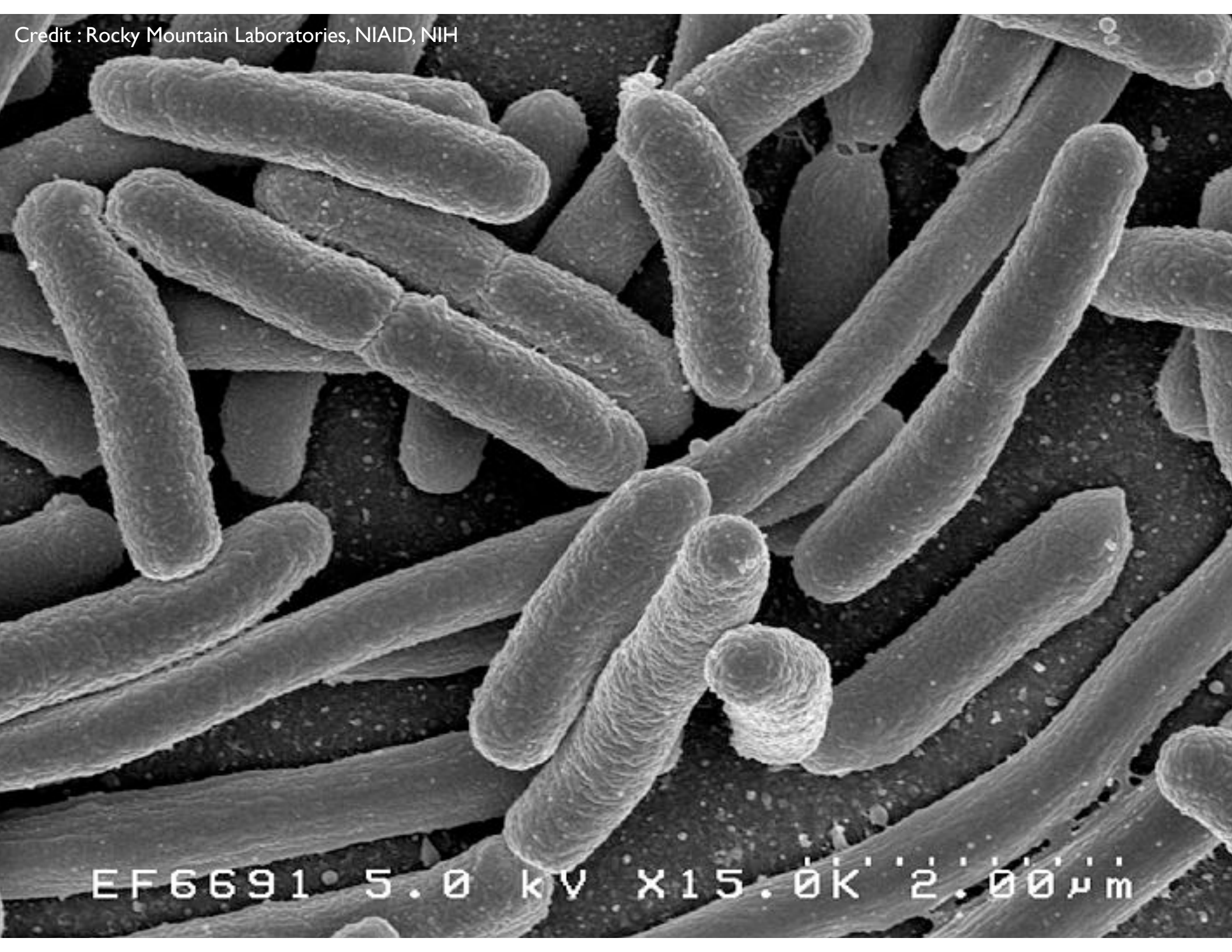
Simulation of a system
undergoing thermal fluctuations

- Conformational sampling / searching
- Exploration of dynamics
- Debye–Waller factors (x-ray)
- Diffusion coefficients
- IR spectra
- NMR observables
- Raman spectra

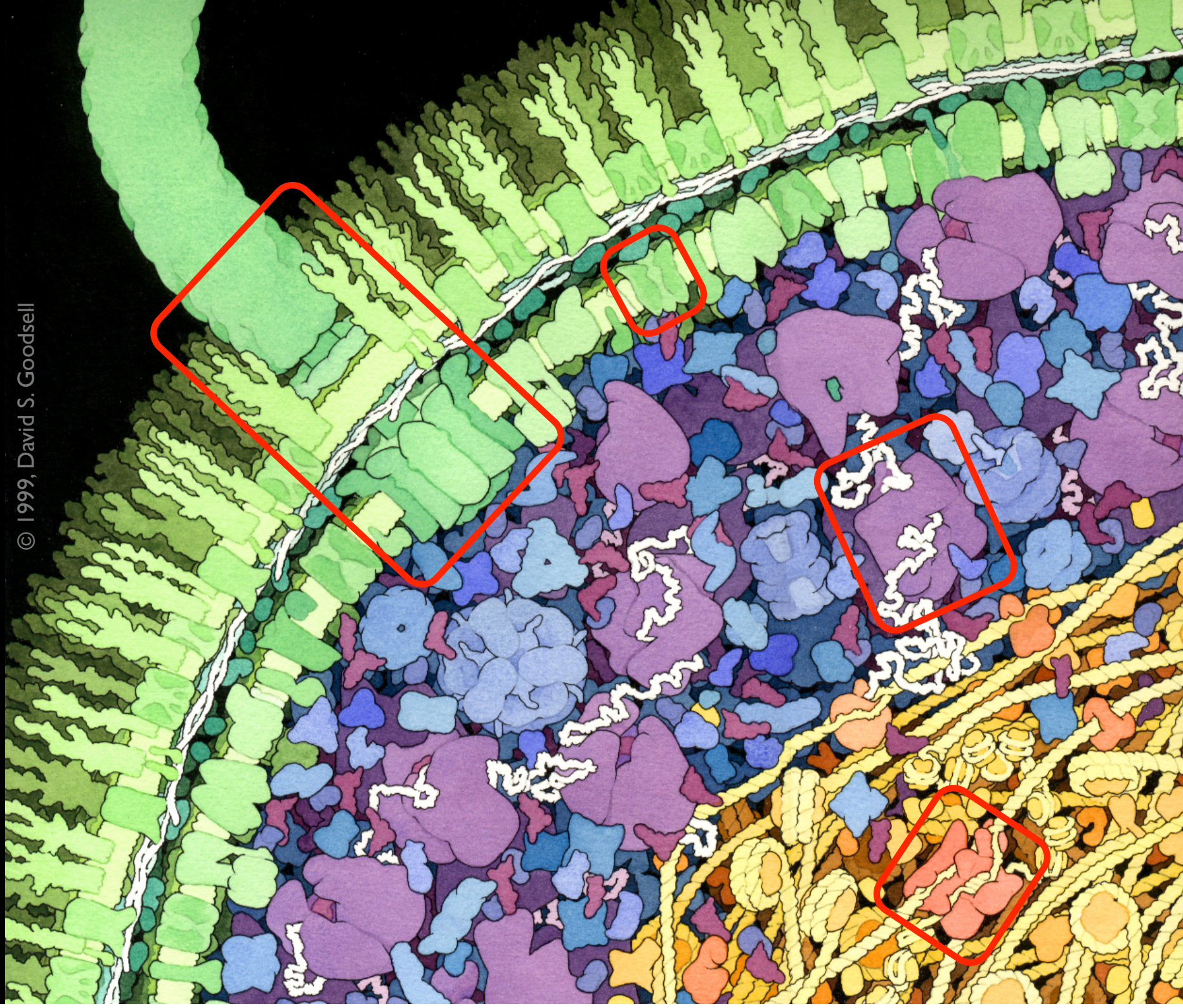
Advanced techniques :

Free energy calculations

- Reaction profiles (conformational change, substrate permeation, etc.)
- Binding free energies of ligands
- Partition constants (“log P ”)
- Thermodynamic effects of a mutation
- pK_a calculations
- Rate constants / kinetics



EF6691 5.0 kV X15.0K 2.00µm



Structure Summary

3D View

Annotations

Sequence

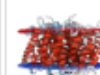
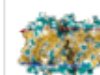
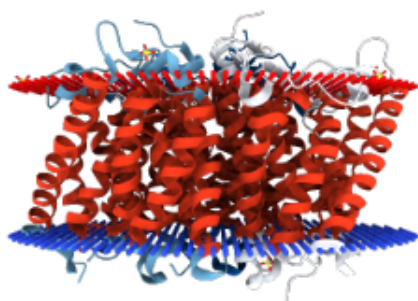
Sequence Similarity

Structure Similarity

Experiment

Transmembrane View

transmembrane regions OPM



3D View: [Structure](#) | [Electron Density](#) |
[Ligand Interaction](#)

Standalone Viewers

[Protein Workshop](#) | [Ligand Explorer](#)

Macromolecule Content

- Total Structure Weight: 43970.08 ⓘ
- Atom Count: 2679 ⓘ
- Residue Count: 418 ⓘ
- Unique protein chains: 1

1XQE

The mechanism of ammonia transport based on the crystal structure of AmtB of *E. coli*.

DOI: [10.2210/pdb1XQE/pdb](https://doi.org/10.2210/pdb1XQE/pdb)

Classification: [TRANSPORT PROTEIN](#)

Organism(s): [Escherichia coli \(strain K12\)](#)

Expression System: [Escherichia coli](#)

Deposited: 2004-10-12 Released: 2004-10-26

Deposition Author(s): [Zheng, L.](#), [Kostrewa, D.](#), [Berneche, S.](#), [Winkler, F.K.](#), [Li, X.-D.](#)

Experimental Data Snapshot

Method: X-RAY DIFFRACTION

Resolution: 2.1 Å

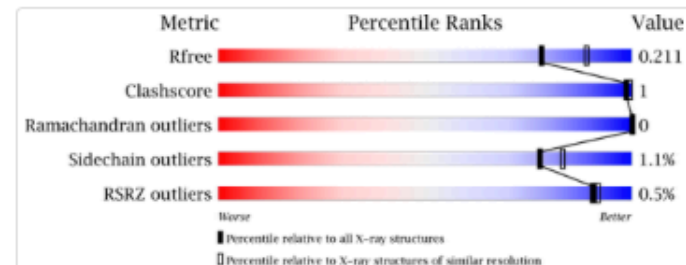
R-Value Free: 0.186

R-Value Work: 0.159

wwPDB Validation

3D Report

Full Report



This is version 1.3 of the entry. See complete [history](#).

Literature

[Download Primary Citation](#)

The mechanism of ammonia transport based on the crystal structure of AmtB of *Escherichia coli*

[Zheng, L.](#), [Kostrewa, D.](#), [Berneche, S.](#), [Winkler, F.K.](#), [Li, X.-D.](#)

(2004) Proc.Natl.Acad.Sci.USA **101**: 17090-17095

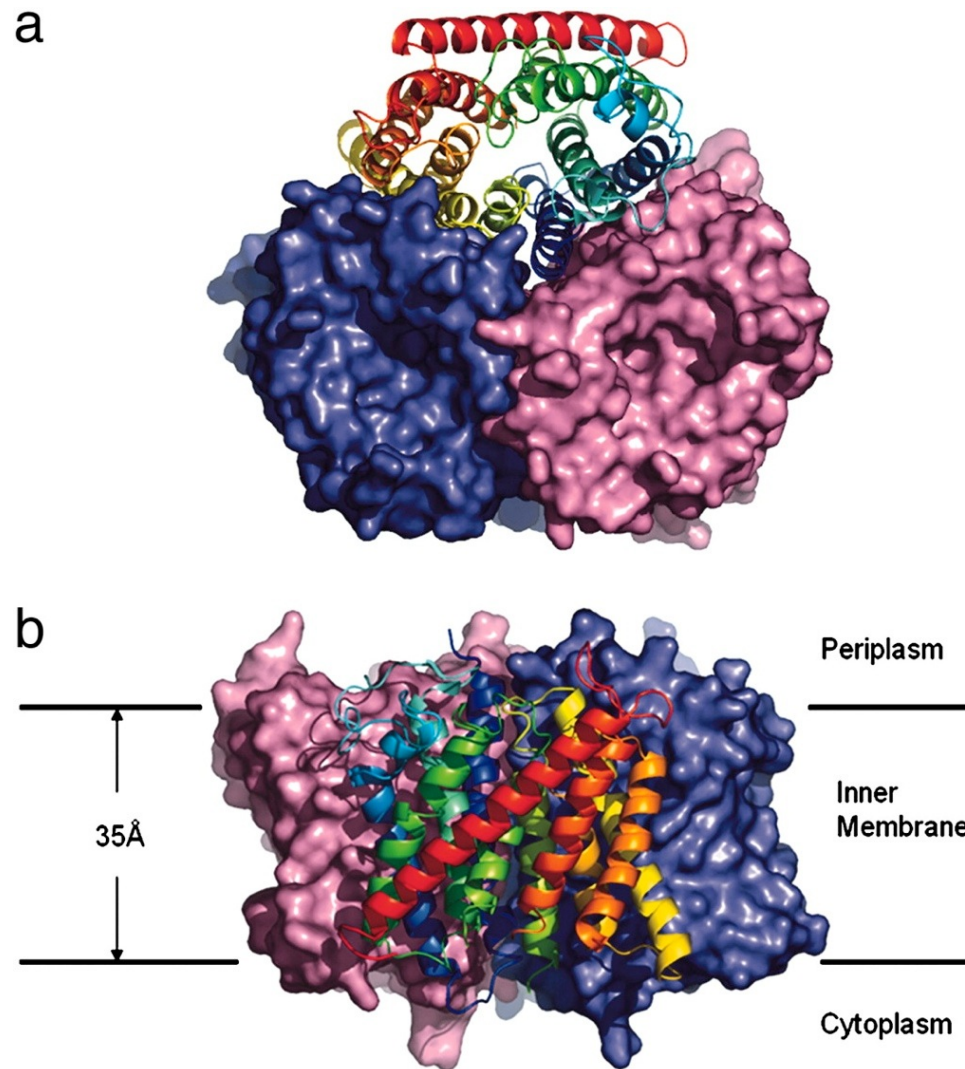
PubMed: [15563598](#)

[Search on PubMed](#)

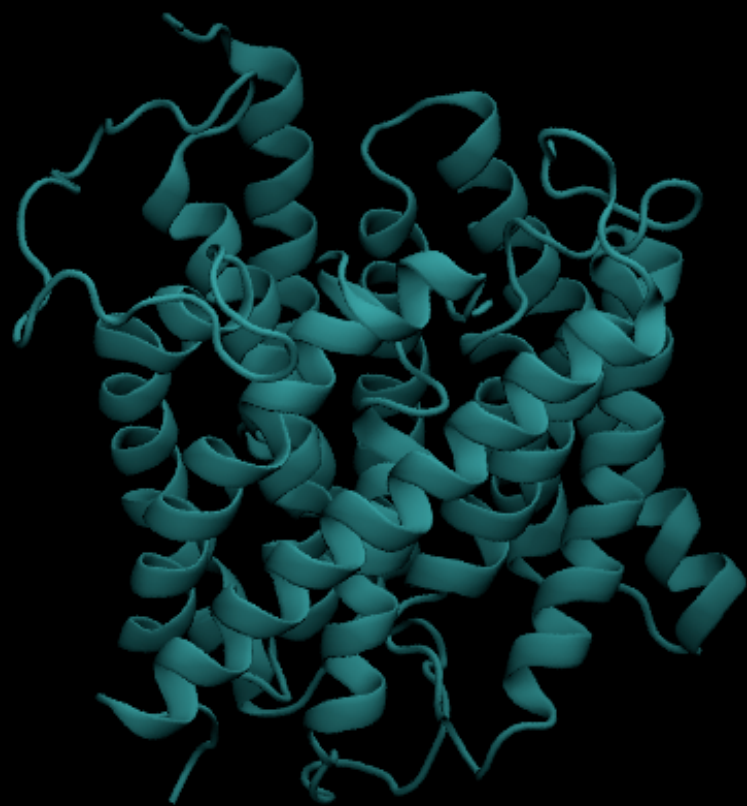
[Search on PubMed Central](#)

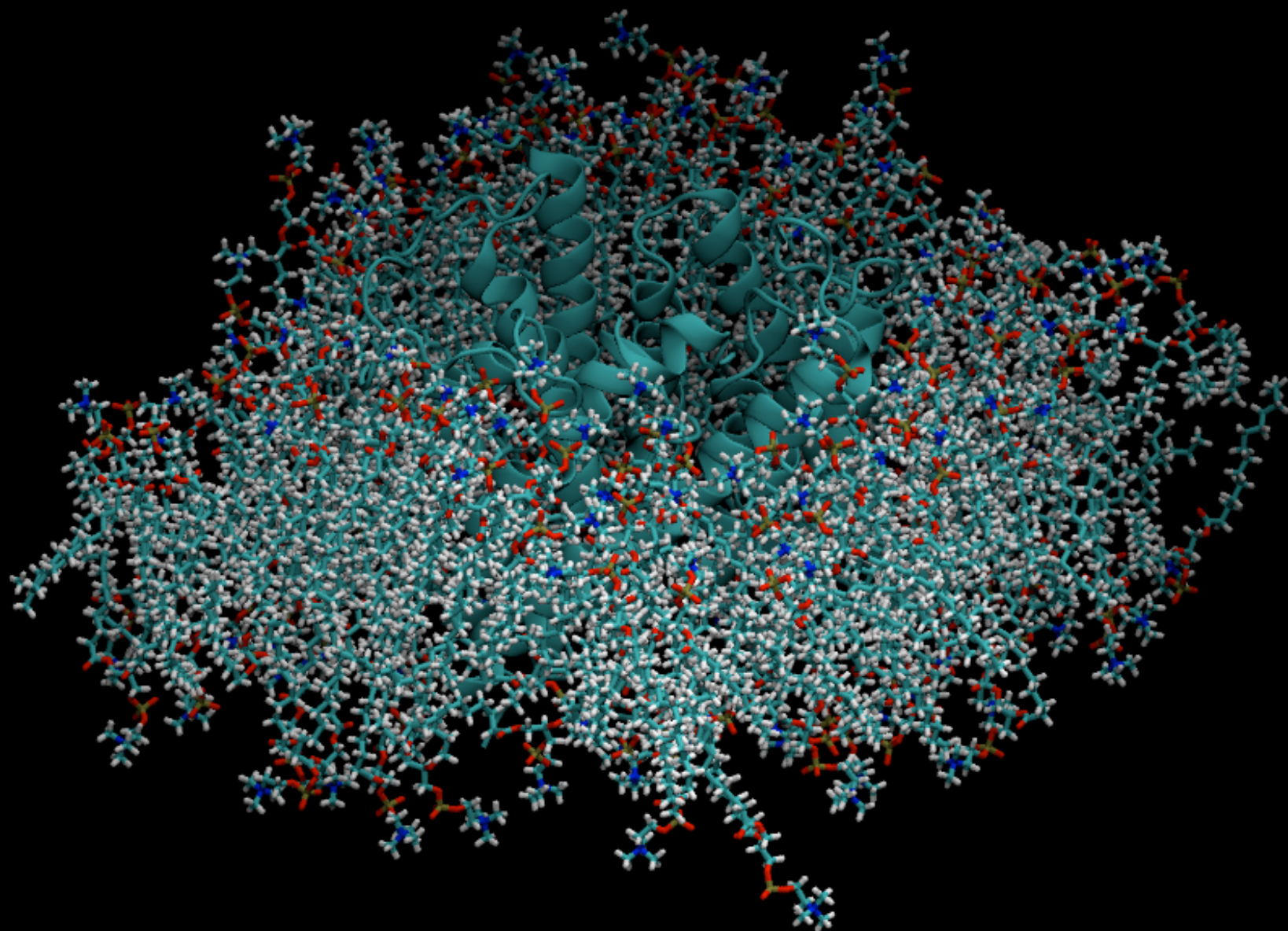
DOI: [10.1073/pnas.0406475101](https://doi.org/10.1073/pnas.0406475101)

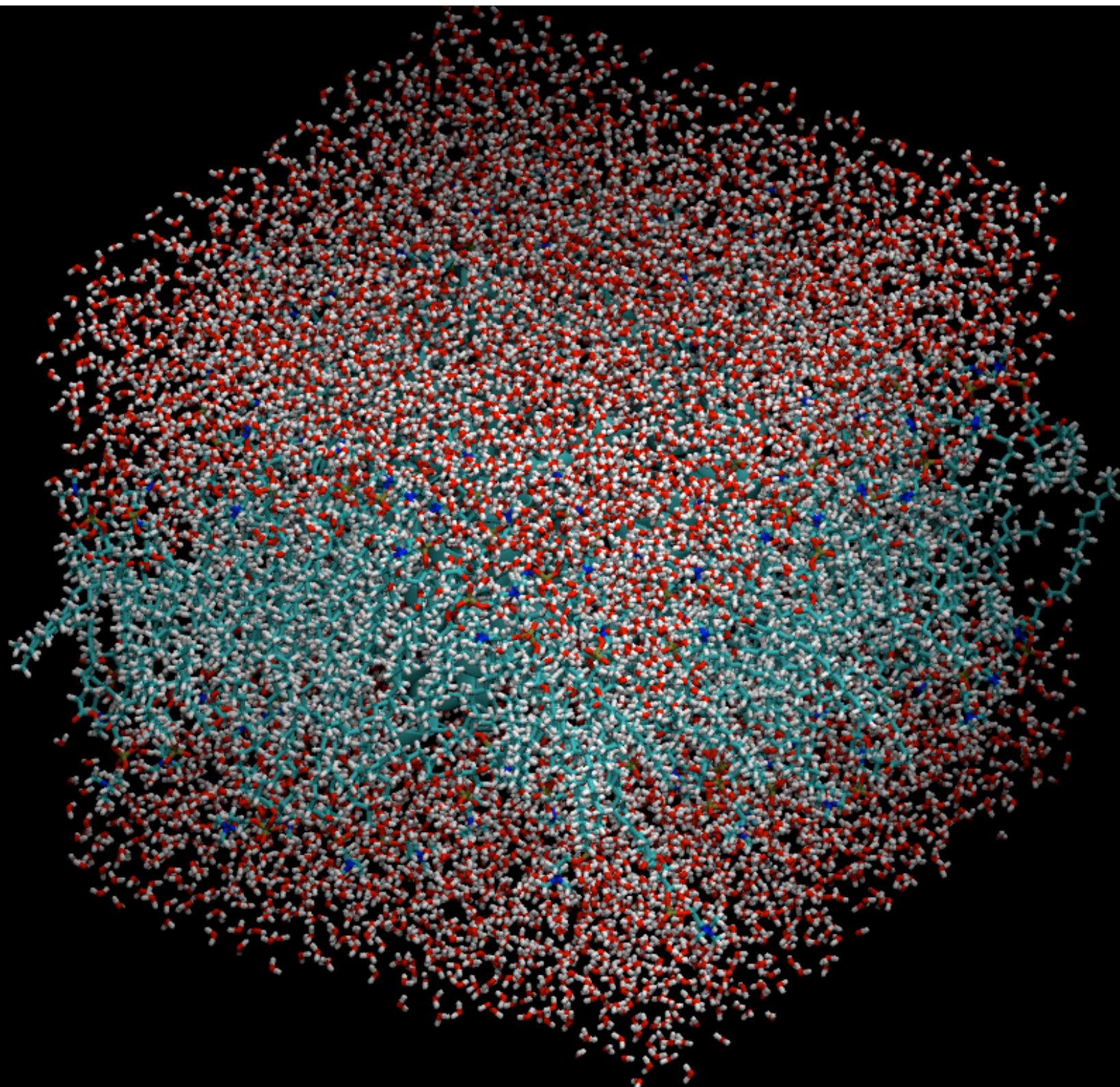
AmtB structure

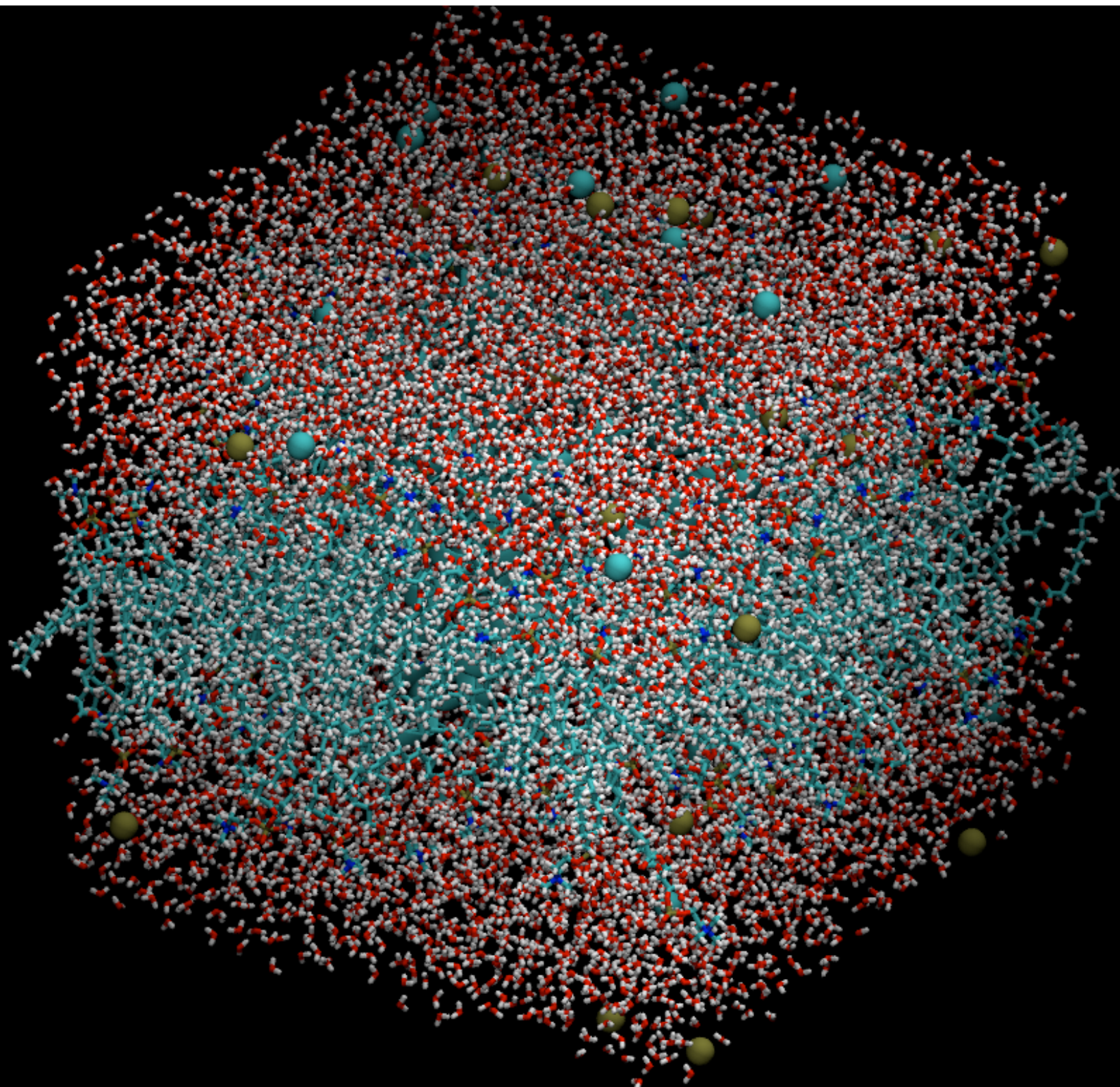


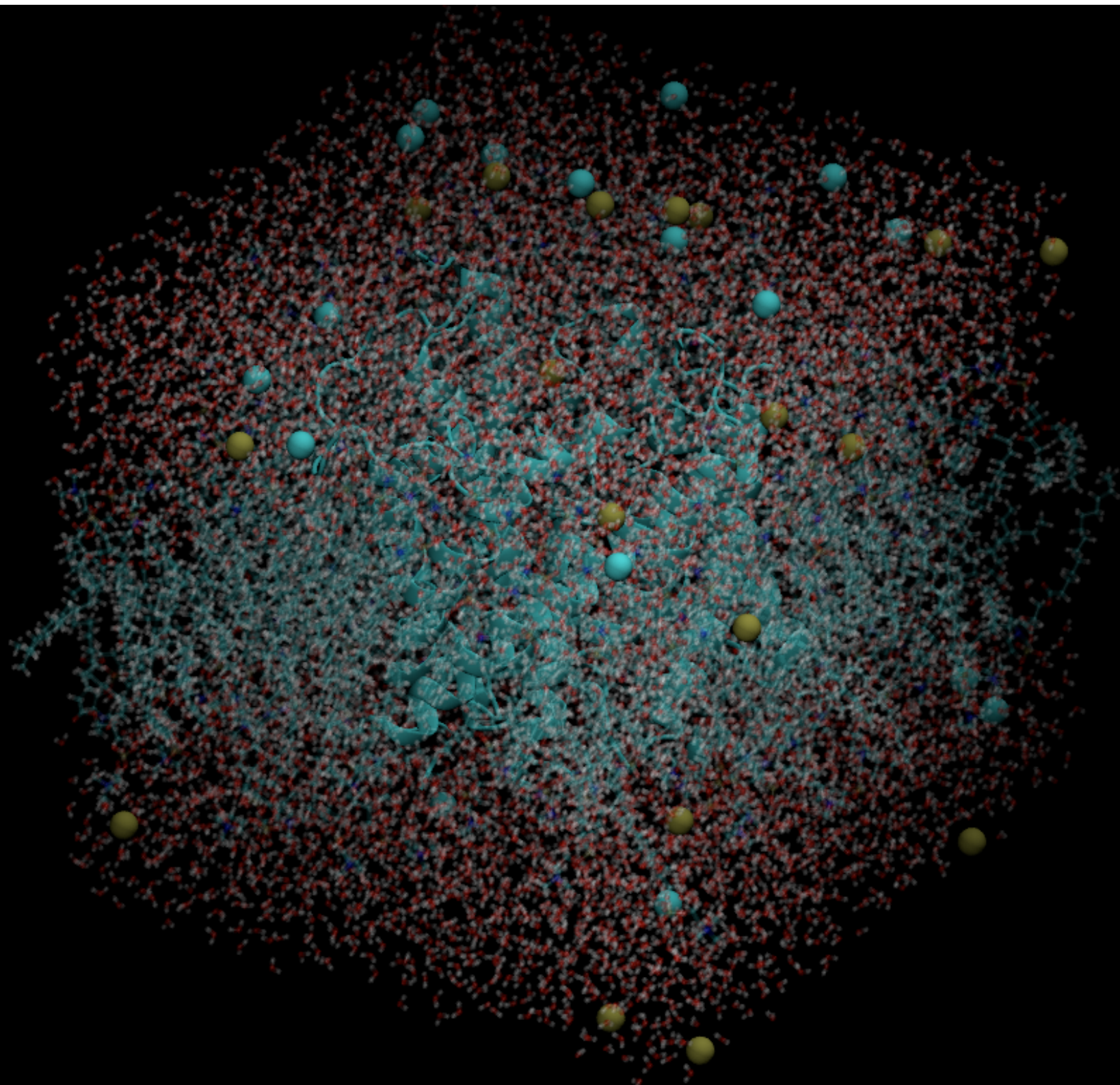
AmtB crystal structure
(Zheng et al., *PNAS* 2004, **101**, 17090)

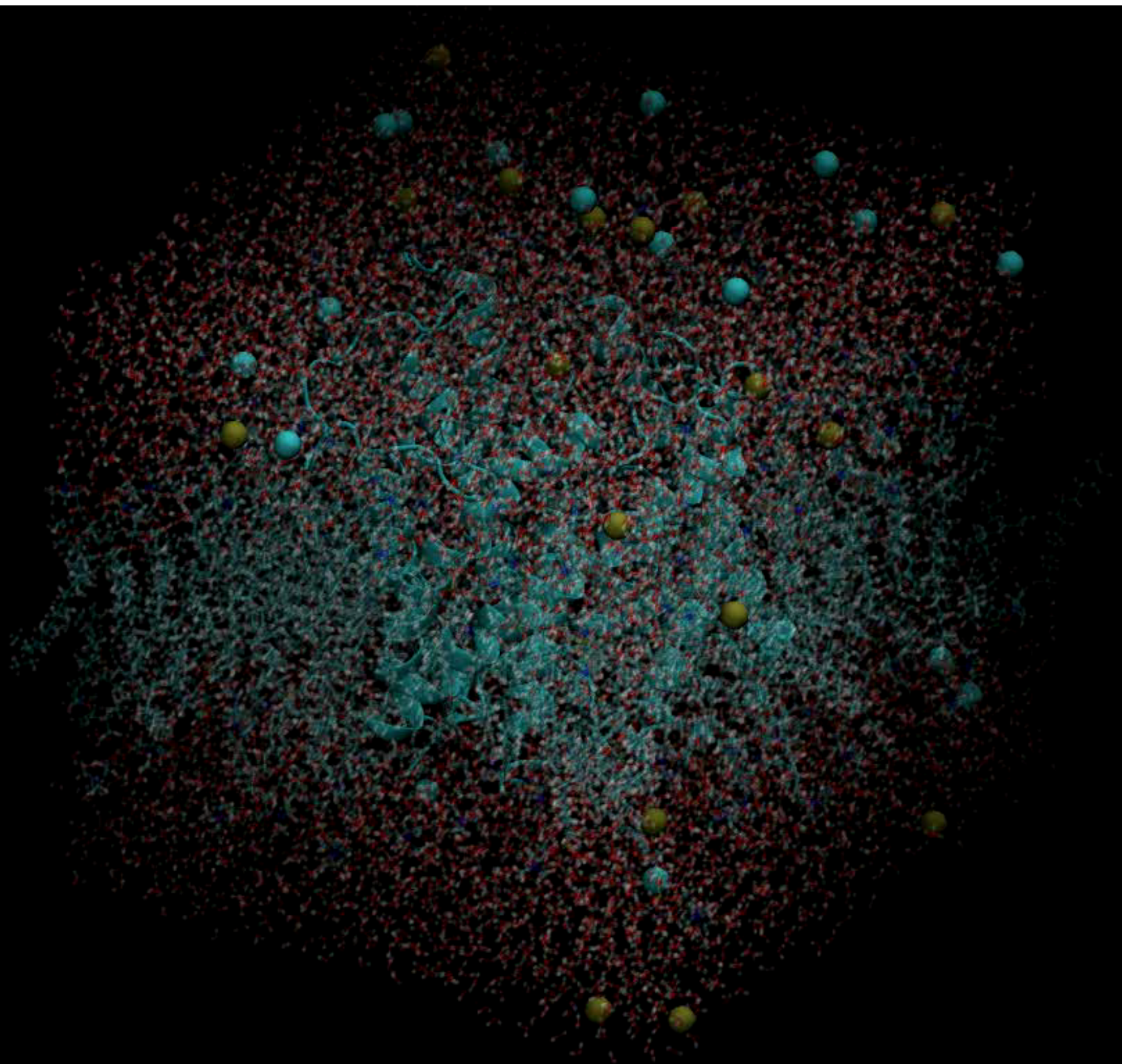






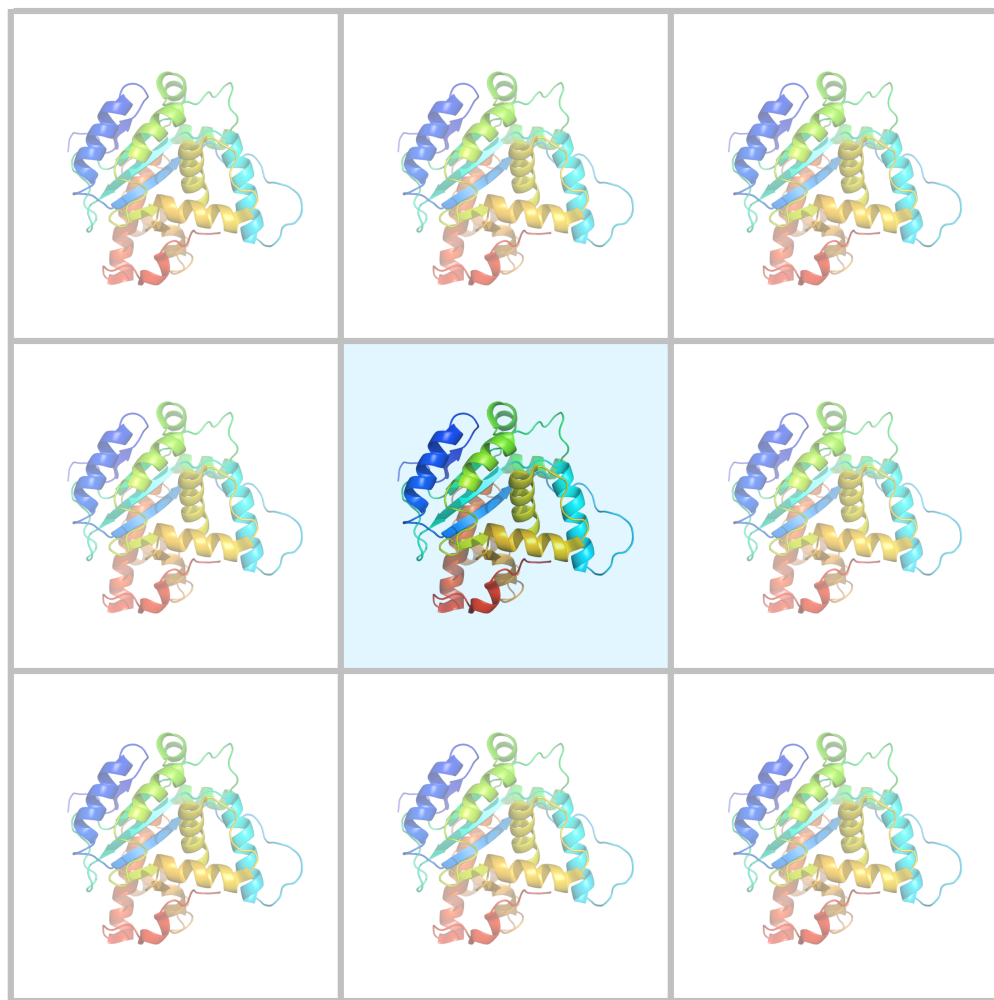






Periodic boundary conditions

(with electrostatic energy calculated using an Ewald summation)



Lattice size should be large enough to accommodate tumbling motions of the solute and to preserve a layer of solvent of **at least 2 Debye lengths** between the images.

Numerical integration

$$m_i \mathbf{a}_i(t) = \mathbf{F}_i(t)$$

Non-integrable system of
differential equations

Finite-difference
approximations :

$$\mathbf{a}_i(t) \simeq \frac{\mathbf{v}_i(t + \frac{1}{2}\delta t) - \mathbf{v}_i(t - \frac{1}{2}\delta t)}{\delta t}$$

$$\mathbf{v}_i(t + \frac{1}{2}\delta t) \simeq \frac{\mathbf{r}_i(t + \delta t) - \mathbf{r}_i(t)}{\delta t}$$

$$\mathbf{v}_i(t - \frac{1}{2}\delta t) \simeq \frac{\mathbf{r}_i(t) - \mathbf{r}_i(t - \delta t)}{\delta t}$$

Finite-difference version
of Newton's equations :

$$m_i \frac{\mathbf{r}_i(t + \delta t) - 2\mathbf{r}_i(t) + \mathbf{r}_i(t - \delta t)}{\delta t^2} = \mathbf{F}_i(t)$$

We can calculate $\mathbf{r}_i(t+\delta t)$ from
 $\mathbf{r}_i(t)$ and $\mathbf{r}_i(t-\delta t)$:

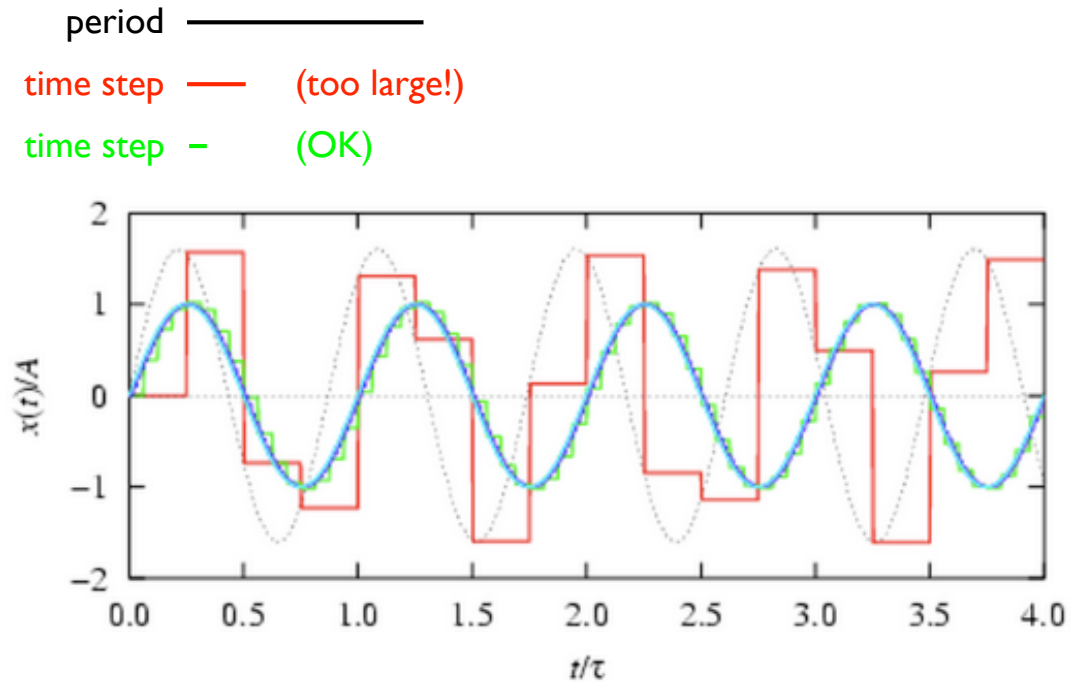
$$\mathbf{r}_i(t + \delta t) = 2\mathbf{r}_i(t) - \mathbf{r}_i(t - \delta t) + \delta t^2 \frac{\mathbf{F}_i(t)}{m_i}$$

(Verlet formula, 1967)

Choice of time step (δt)

Rule of thumb :

δt should be **at least one tenth** of the shortest period of oscillation of the system, but **ideally one twentieth**.



O–H stretch : $\sim 3500 \text{ cm}^{-1} \Rightarrow \tau \sim 10 \text{ fs} \Rightarrow \delta t \sim 0.5 \text{ fs}$

C=O stretch : $\sim 1700 \text{ cm}^{-1}$

N–O stretch : $\sim 1600 \text{ cm}^{-1}$

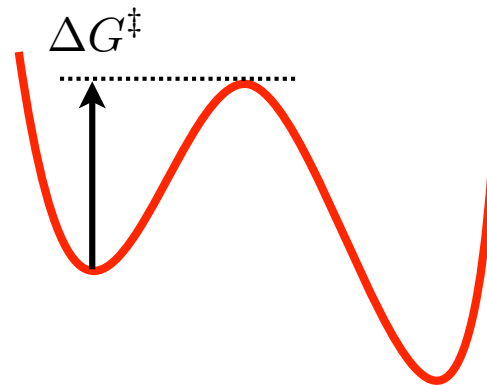
C=C stretch : $\sim 1600 \text{ cm}^{-1} \Rightarrow \delta t \sim 1.0 \text{ fs}$

To avoid using a 0.5 fs time step, the X–H bonds can be kept rigid (using the SHAKE algorithm).

How long should the simulation be ?

Typical time for crossing an energy barrier :

$$\tau = \tau_o e^{\Delta G^\ddagger / k_B T}$$



Using $\tau_0 = 1$ ps :

$$1 \text{ kcal mol}^{-1} \Rightarrow \mathbf{1.2 \text{ ps}^{-1}}$$
 (“1.2 crossings per ps”)

$$5 \text{ kcal mol}^{-1} \Rightarrow \mathbf{1.4 \text{ ns}^{-1}}$$

$$10 \text{ kcal mol}^{-1} \Rightarrow \mathbf{< 1 \text{ ms}^{-1}}$$

The problem gets worse if the system has multiple barriers...

“Thermostatted” Newton’s equations (Nosé–Hoover equations)

$$m_i \ddot{\mathbf{r}}_i = \mathbf{F}_i - m_i \dot{\mathbf{r}}_i \dot{\eta}$$
$$Q \ddot{\eta} = 2K - N_f k_B T$$

variable friction coefficient

inertia factor

number of degrees of freedom

Boltzmann constant

$$N_f = 3N - N_c$$

with

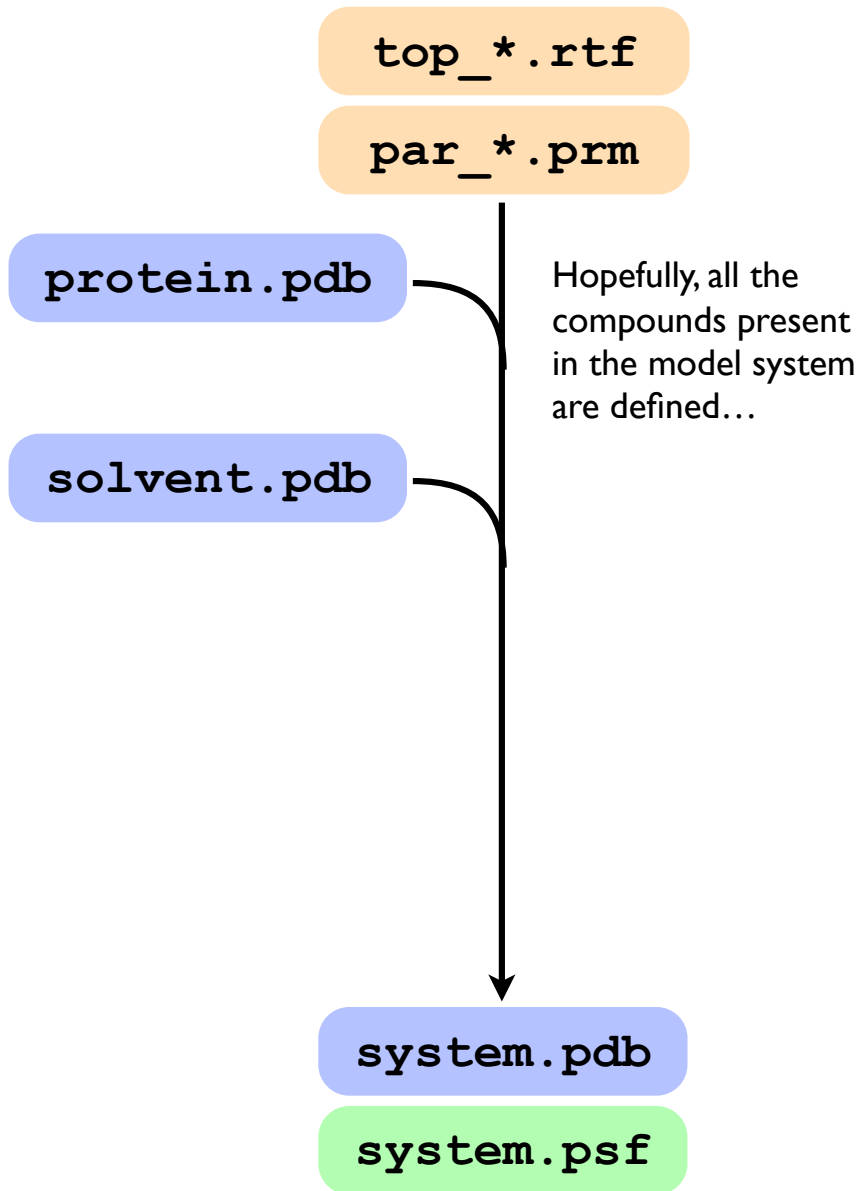
$$2K = \sum_j m_j \dot{\mathbf{r}}_j^2$$

2 times the total instantaneous kinetic energy of the system

According to the equipartition theorem, each degree of freedom of the system should have a kinetic energy of $\frac{1}{2} k_B T$.

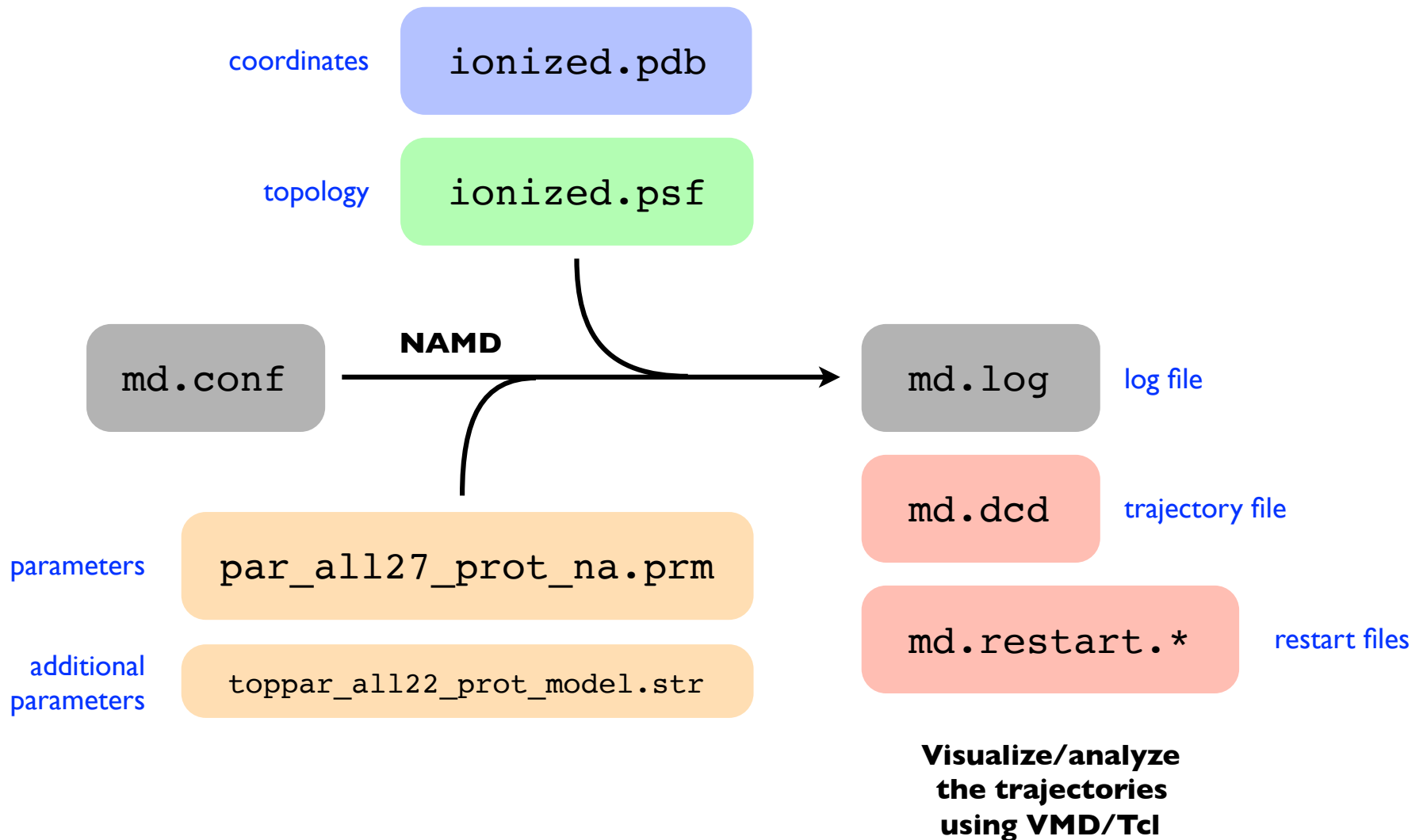
For that reason, K is related to the instantaneous temperature of the system.

General procedure for system setup



- **Read topology and parameter files**
- **Read PDB file**
(protein coordinates)
- **Add solvent / membrane**
- **Specify details of energy function**
(cutoff scheme, periodic boundary conditions, additional restraints, etc.)
- **Minimize energy**
(to remove potential bad contacts)
- **Write PDB and PSF**
(coordinates and topology)

MD simulation in NAMD



Molecular dynamics

“Putting the thermal fluctuations
back into a protein structure”

Basic techniques :

Simulation of a system
undergoing thermal fluctuations

- Conformational sampling / searching
- Exploration of dynamics
- Debye–Waller factors (x-ray)
- Diffusion coefficients
- IR spectra
- NMR observables
- Raman spectra

Advanced techniques :

Free energy calculations

- Reaction profiles (conformational change, substrate permeation, etc.)
- Binding free energies of ligands
- Partition constants (“log P ”)
- Thermodynamic effects of a mutation
- pK_a calculations
- Rate constants / kinetics

Debye–Waller factors (a.k.a. B-factors)

To compare with experimental B-factors, we have to account for crystal disorder, which is not present in the simulations.

$$B_i = \frac{8\pi}{3} \langle |\Delta\vec{r}_i|^2 \rangle + B_W$$

$$\Delta\vec{r}_i = \vec{r}_i - \langle \vec{r}_i \rangle$$

Example: KcsA

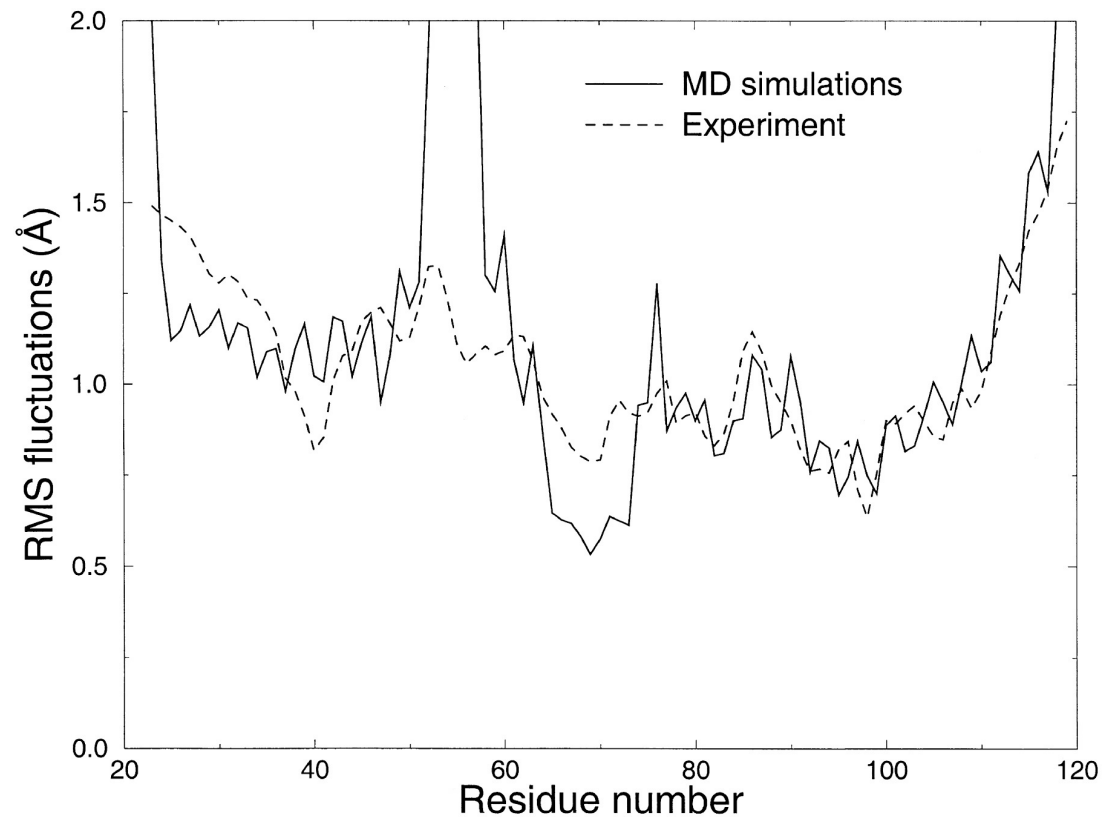


Figure from :

Bernèche and Roux, *Biophys. J.* **78**, 2900–2917 (2000)

Diffusion coefficients

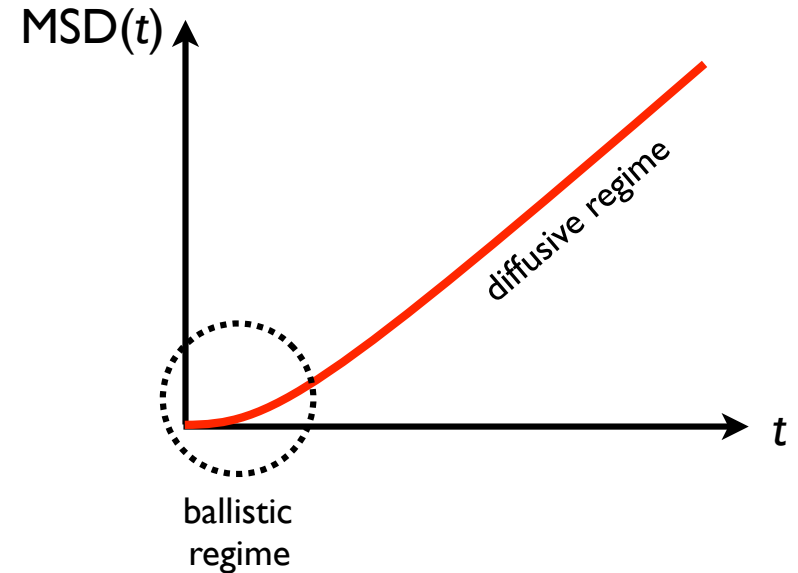
(for a molecule in a homogeneous environment)

$$\text{MSD}(t) = \left\langle |\mathbf{r}(t) - \mathbf{r}(0)|^2 \right\rangle$$

$$D = \lim_{t \rightarrow \infty} \frac{1}{6t} \text{MSD}(t)$$

We usually calculate the MSD by assuming that the fluctuations are *stationary* :

$$\begin{aligned} \text{MSD}(t) &= \left\langle |\mathbf{r}(t + \tau) - \mathbf{r}(\tau)|^2 \right\rangle \\ &= \frac{1}{N_\tau} \sum_{\tau}^{N_\tau} |\mathbf{r}(t + \tau) - \mathbf{r}(\tau)|^2 \end{aligned}$$



Radial distribution functions (RDFs)

(between pairs of atoms)

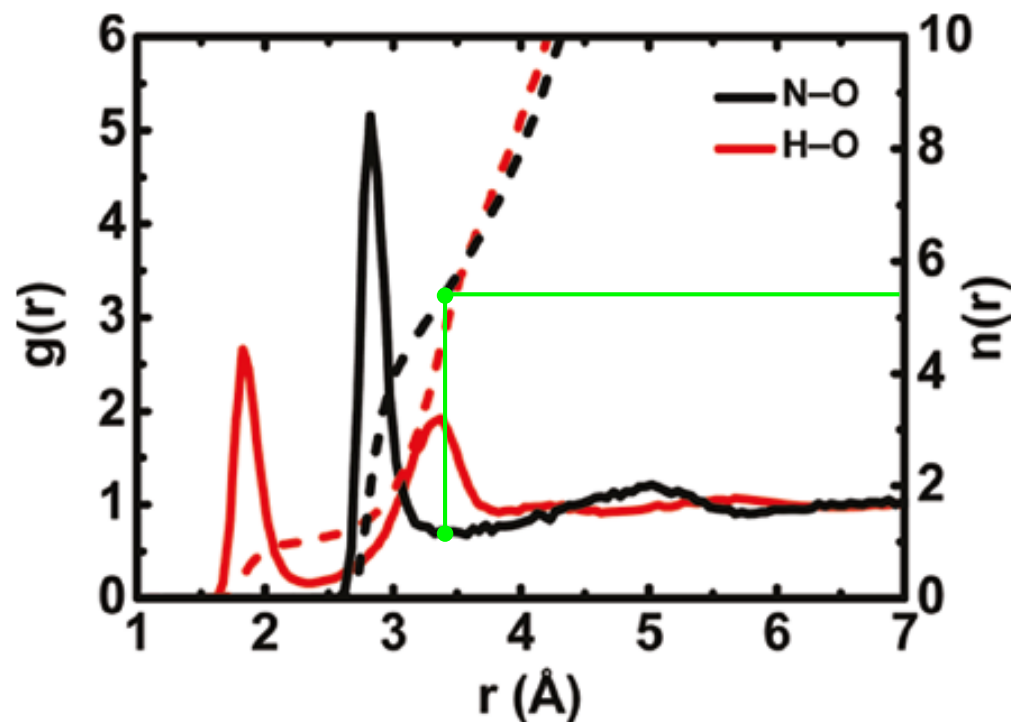
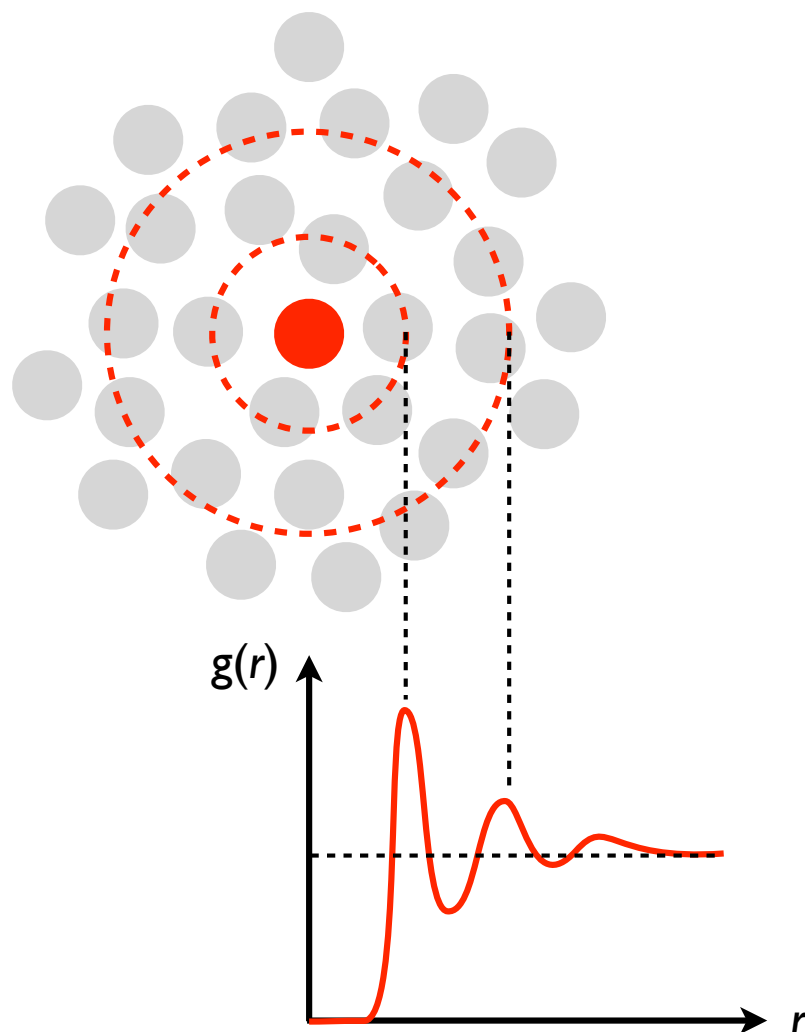


Figure 4. N–O and H–O radial distribution functions (solid lines, scale on left) and running integration numbers (dashed lines, scale on right) of NH_4^+ in water at 298.15 K.

Figure from :

Orabi and Lamoureux, *J. Chem. Theory Comput.* **8**, 182–193 (2012)

Gramicidin A channel

15 residues, alternating L and D amino acids :

Formyl-Val-Gly-Ala-Leu-Ala-Val-Val-Val-Trp-Leu-Trp-Leu-Trp-Ethanolamine

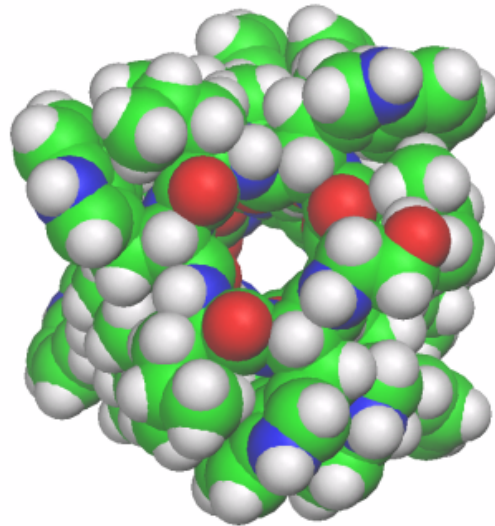
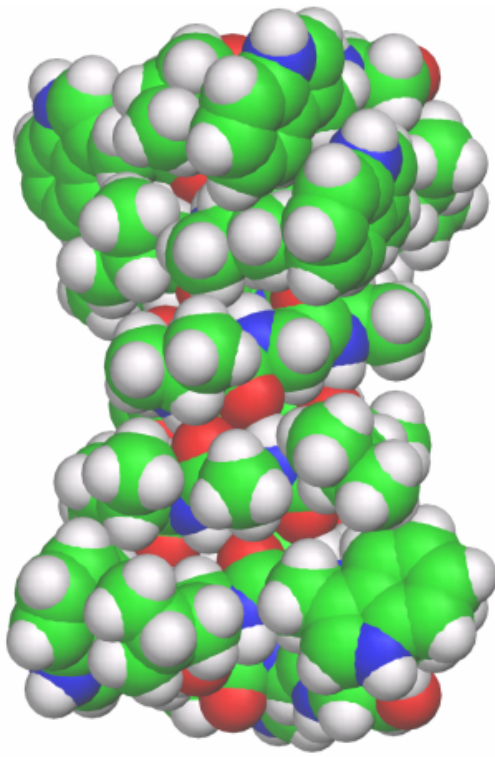
Head-to-head β -helical dimer (26 Å long)

Urry. *PNAS* **68**, 672–676 (1971)

Arseniev, Barsukov, Bystrov, Lomize, and Ovchinnikov. *FEBS Lett.* **186**, 168–174 (1985)

Ketchum, Hu, and Cross. *Science* **261**, 1457–1460 (1993)

Townsley, Tucker, Sham, and Hinton. *Biochemistry* **40**, 11676–11686 (2001)



Conducts H^+ , Cs^+ , Rb^+ ,
 K^+ , Na^+ , Li^+

Impermeable to anions

Figure from :

Roux, *Acc. Chem. Res.* **35**, 366–375 (2002)

Solid-state NMR observables

Chemical shift anisotropy (CSA)

$$\Delta\sigma = \sigma_{||} - \sigma_{\perp}$$

$$\sigma_{||} = \left\langle \hat{\mathbf{Z}} \cdot \sigma(t) \cdot \hat{\mathbf{Z}} \right\rangle$$

direction of the
magnetic field

$$\sigma_{\perp} = \frac{1}{2} \left\langle \hat{\mathbf{X}} \cdot \sigma(t) \cdot \hat{\mathbf{X}} + \hat{\mathbf{Y}} \cdot \sigma(t) \cdot \hat{\mathbf{Y}} \right\rangle$$

directions perpendicular
to the magnetic field

$$\sigma(t) = \sum_{i=1}^3 \hat{\mathbf{e}}_i(t) \sigma_{ii} \hat{\mathbf{e}}_i(t)$$

instantaneous unit vector
along the i -th principal axis of
the static chemical shift tensor

corresponding
tensor element
magnitudes

Tensor oriented relative
to dipolar couplings
(C=O, N-C_α, N-H, etc.)

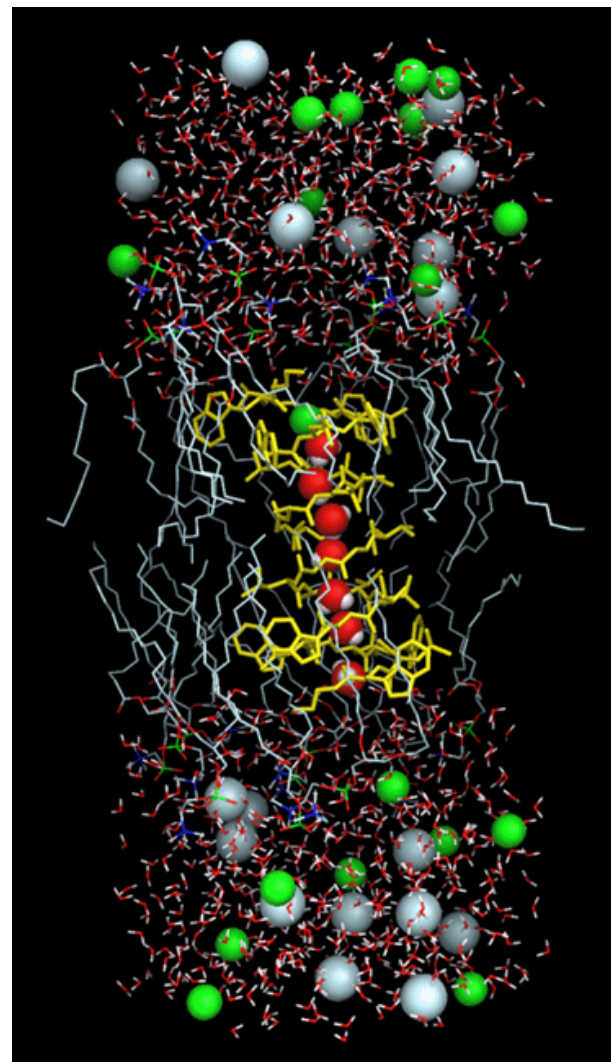


Figure from :
Allen, Andersen, and Roux,
JACS **125**, 9868–9877 (2003)

Solid-state NMR observables

Examples :

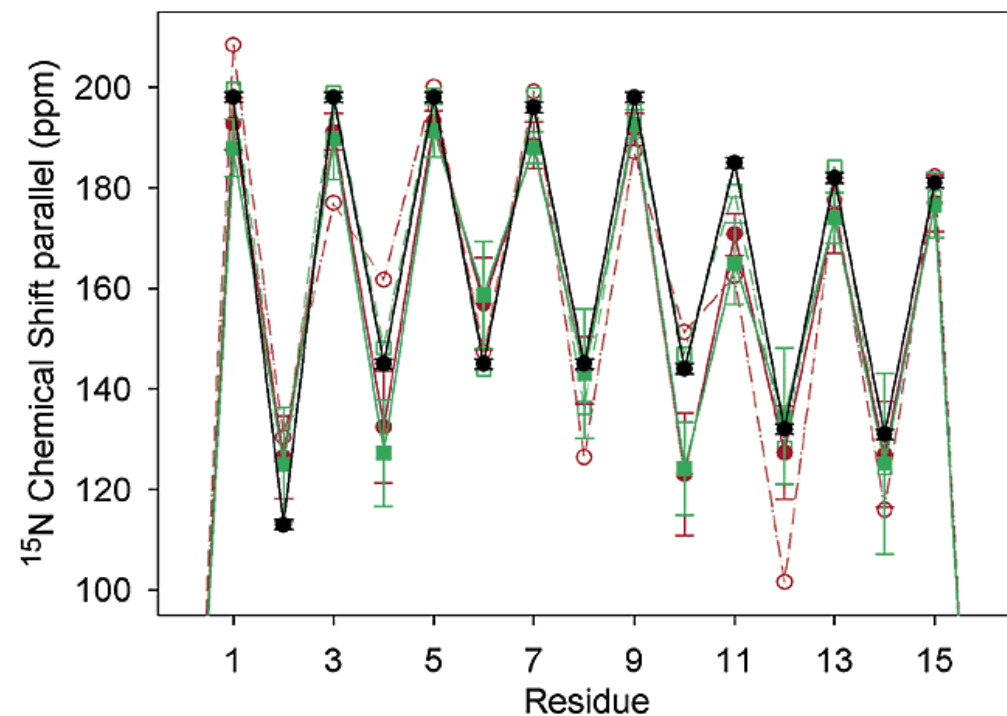
^{15}N – ^1H dipolar coupling (DC)

C_α – ^2H quadrupolar splitting (DQS)

$$\langle \Delta\nu \rangle = \Delta\nu_0 \left\langle \frac{3 \cos^2 \theta(t) - 1}{2} \right\rangle$$

angle between the two atoms
and the direction of the field

standard
frequency
splitting



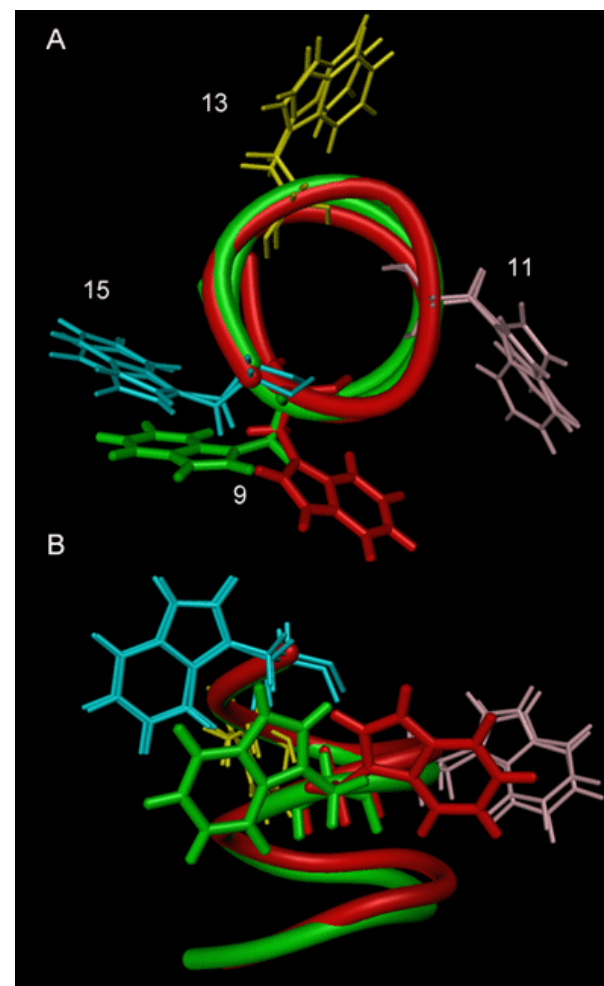
PDB: 1MAG

Ketchum, Lee, Hu, and Cross.
J. Biomol. NMR **8**, 1–14 (1996)
 (solid-state NMR in DMPC bilayers)

PDB: 1JNO

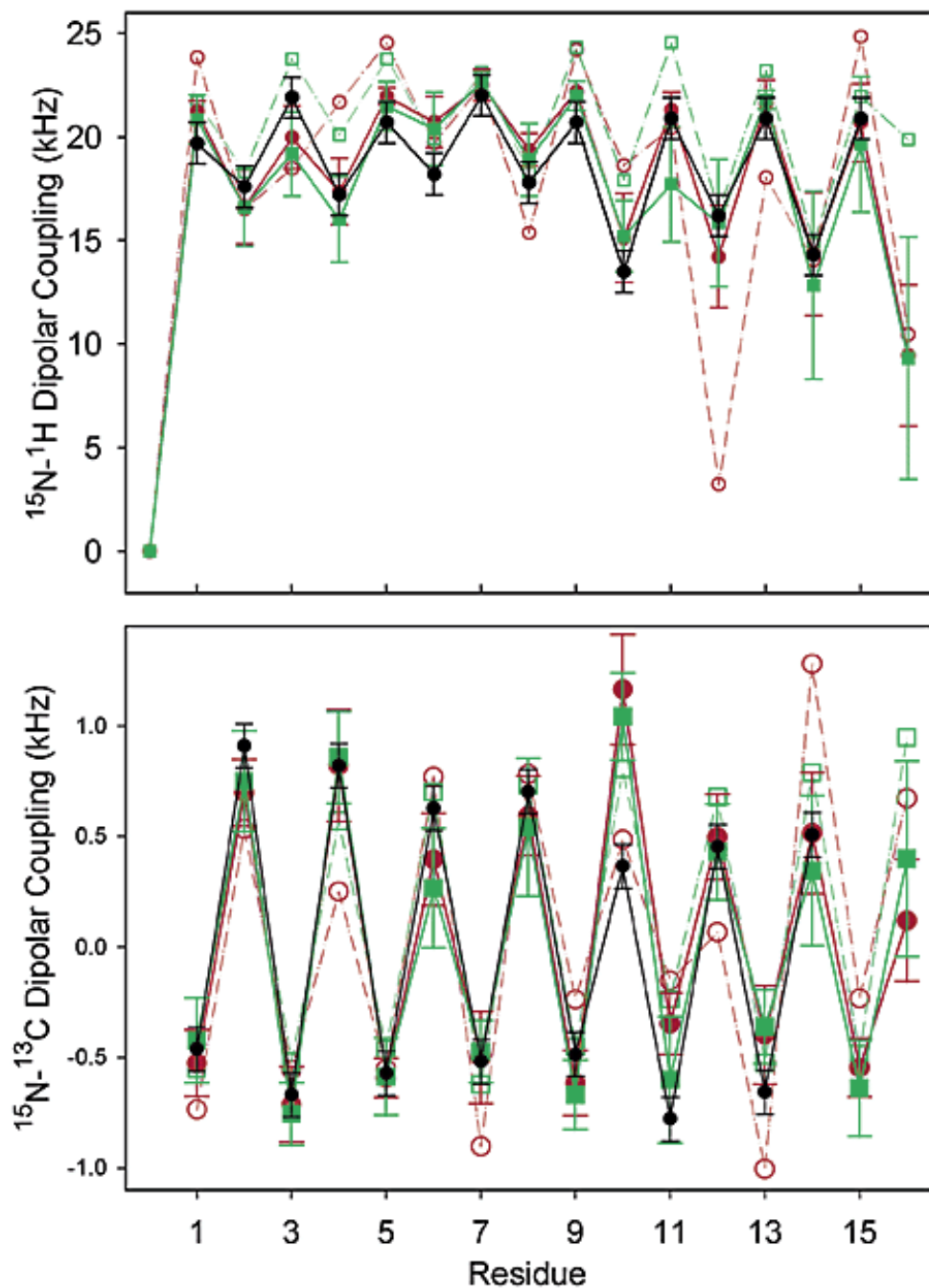
Townsley, Tucker, Sham, and Hinton.
Biochemistry **40**, 11676–11686 (2001)
 (solution NMR in SDS micelles)

The structures have different helical pitches (6.5 residues per turn for **1MAG** and 6.3 residues per turn for **1JNO**) and different rotameric states for residue Trp9 (stacking with Trp15 in the **1MAG** structure, and away from Trp15 in the **1JNO** structure)

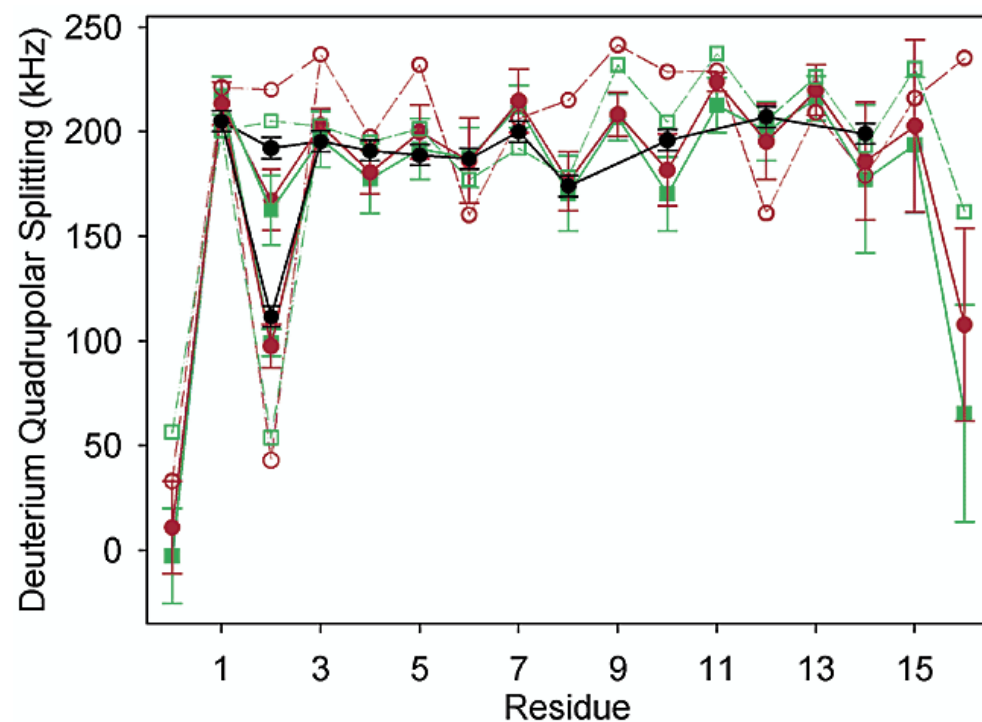


Figures from :

Allen, Andersen, and Roux,
JACS **125**, 9868–9877 (2003)



Figures from :
 Allen, Andersen, and Roux,
JACS **125**, 9868–9877 (2003)



Backbone :

“Despite the differing helical pitches and NMR observables of the initial PDB:1JNO and PDB:1MAG structures, the two backbones become nearly indistinguishable after dynamic relaxation in the bilayer, attaining helical pitches consistent with independent X-ray determination and backbone NMR observables in close agreement with experiment.”

Trp9 :

“The observed isomerizations of the Trp9 side chain suggest that a mixture of rotameric states may better reproduce experiment. In fact, empirical fitting, supported by umbrella sampling calculations, suggests that a mixture of states consisting of 80% T and 20% K rotamers best reproduces side chain NMR observables.”

Molecular dynamics

“Putting the thermal fluctuations
back into a protein structure”

Basic techniques :

Simulation of a system
undergoing thermal fluctuations

- Conformational sampling / searching
- Exploration of dynamics
- Debye–Waller factors (x-ray)
- Diffusion coefficients
- IR spectra
- NMR observables
- Raman spectra

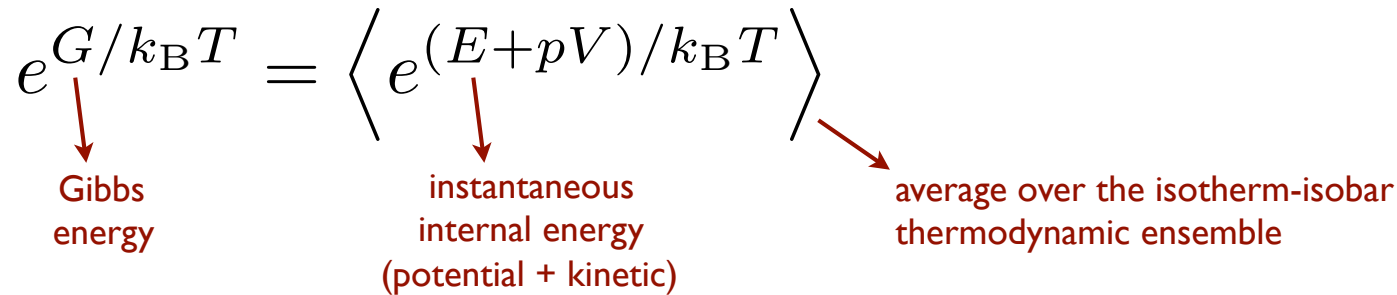
Advanced techniques :

Free energy calculations

- Reaction profiles (conformational change, substrate permeation, etc.)
- Binding free energies of ligands
- Partition constants (“log P ”)
- Thermodynamic effects of a mutation
- pK_a calculations
- Rate constants / kinetics

Free energy (Gibbs)

$$e^{G/k_B T} = \left\langle e^{(E+pV)/k_B T} \right\rangle$$



Gibbs energy

instantaneous internal energy (potential + kinetic)

average over the isotherm-isobar thermodynamic ensemble

It seems that computing the Gibbs energy of a system only amounts to calculating the exponential of the instantaneous enthalpy ($H = E + pV$) in an MD simulation at constant temperature and pressure.

Problem !

The higher the enthalpy of a region of phase space, the more it contributes to the average but the less often it is “visited” during the MD simulation.

Fortunately...

We care about free energy *differences* :

$$\Delta G = G_{\text{after}} - G_{\text{before}}$$

Still !

The free energy difference may be too large to be efficiently sampled from a free simulation.

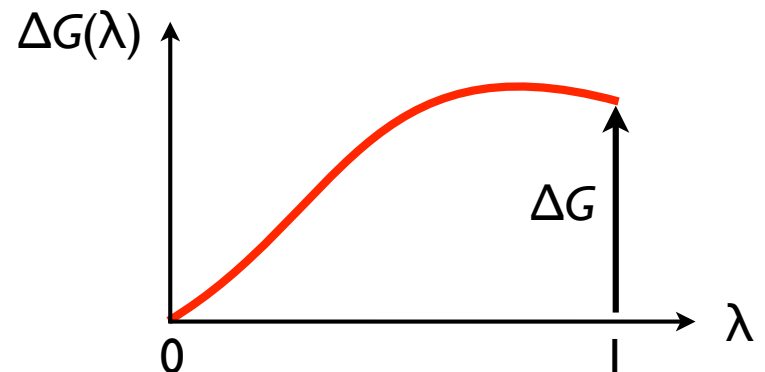
Thermodynamic integration (TI)

We introduce a parameter λ that controls the simulation of the system in such a way that :

$$G(\lambda) = G_{\text{before}} \quad \text{when } \lambda = 0 \quad \text{and} \quad G(\lambda) = G_{\text{after}} \quad \text{when } \lambda = 1.$$

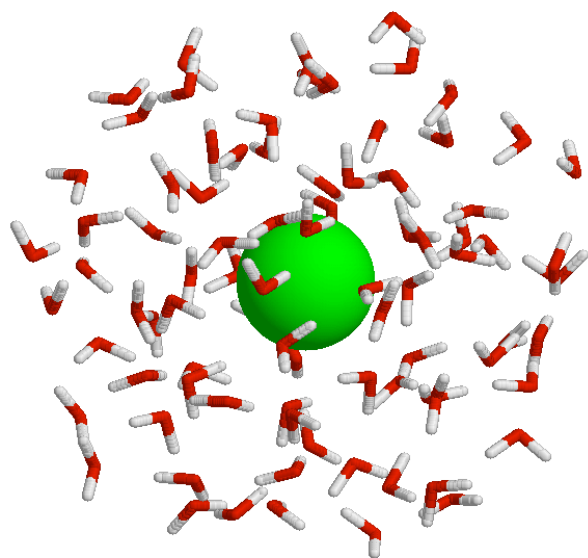
$$\begin{aligned} \Delta G &= \int_0^1 d\lambda \frac{\partial G}{\partial \lambda} \\ &= \int_0^1 d\lambda \left\langle \frac{\partial U}{\partial \lambda} \right\rangle_{\lambda} \end{aligned}$$

We can calculate a free energy difference by integrating the “profile” of the derivative of U (the potential energy) with respect to λ , averaged over the isotherm-isobar ensemble corresponding to a specific value of λ .

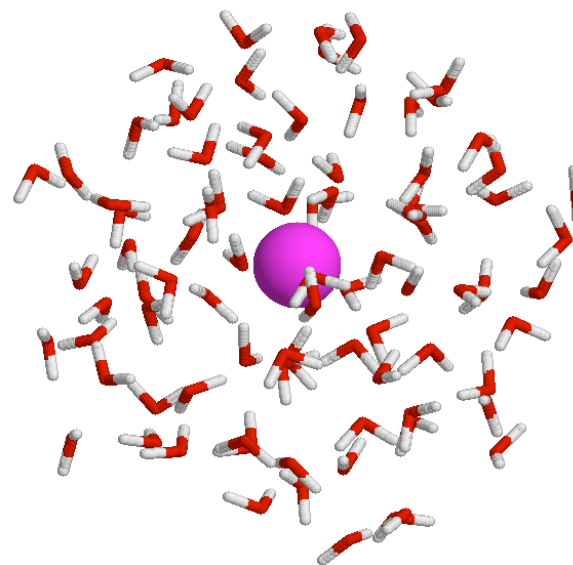


Application of TI :
Relative hydration free energies

$$U(\lambda) = (1 - \lambda) U_{\text{K}^+} + \lambda U_{\text{Na}^+}$$



K^+
 $(\lambda = 0)$

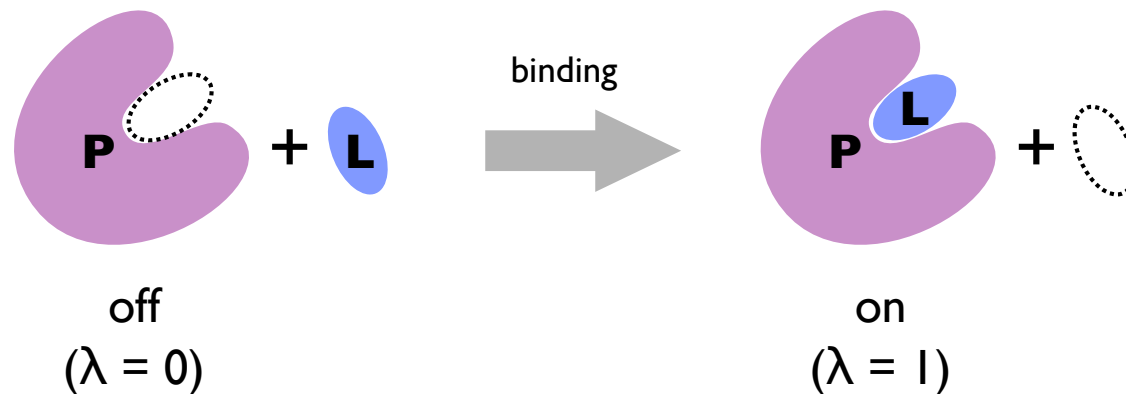


Na^+
 $(\lambda = 1)$

$$\Delta G = -17.2 \text{ kcal/mol}$$

Application of TI : Binding free energies

Simple approach : $U(\lambda) = (1 - \lambda) U_{\text{off}} + \lambda U_{\text{on}}$



“Well-behaved” approach :

We perform the calculation in stages.

First stage: λ is used to “turn on” the shape of the molecule

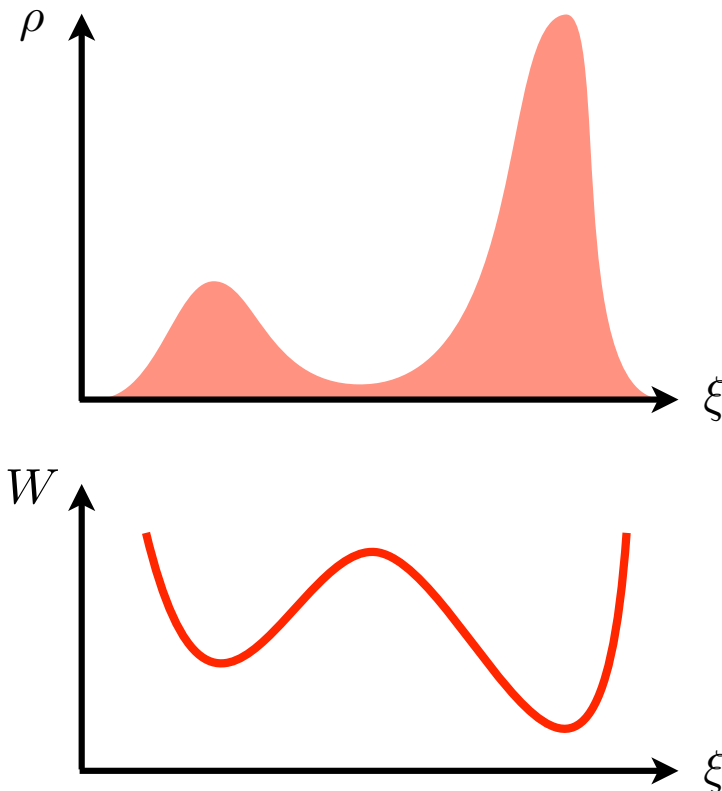
Second stage: λ is used to “turn on” the charges on the molecule

Free energy profile

$$e^{-W(\xi)/k_B T} \propto \rho(\xi)$$

↓
distribution of the
reaction variable
over the simulation

$$W(\xi) = -k_B T \ln \rho(\xi) + \text{constant}$$



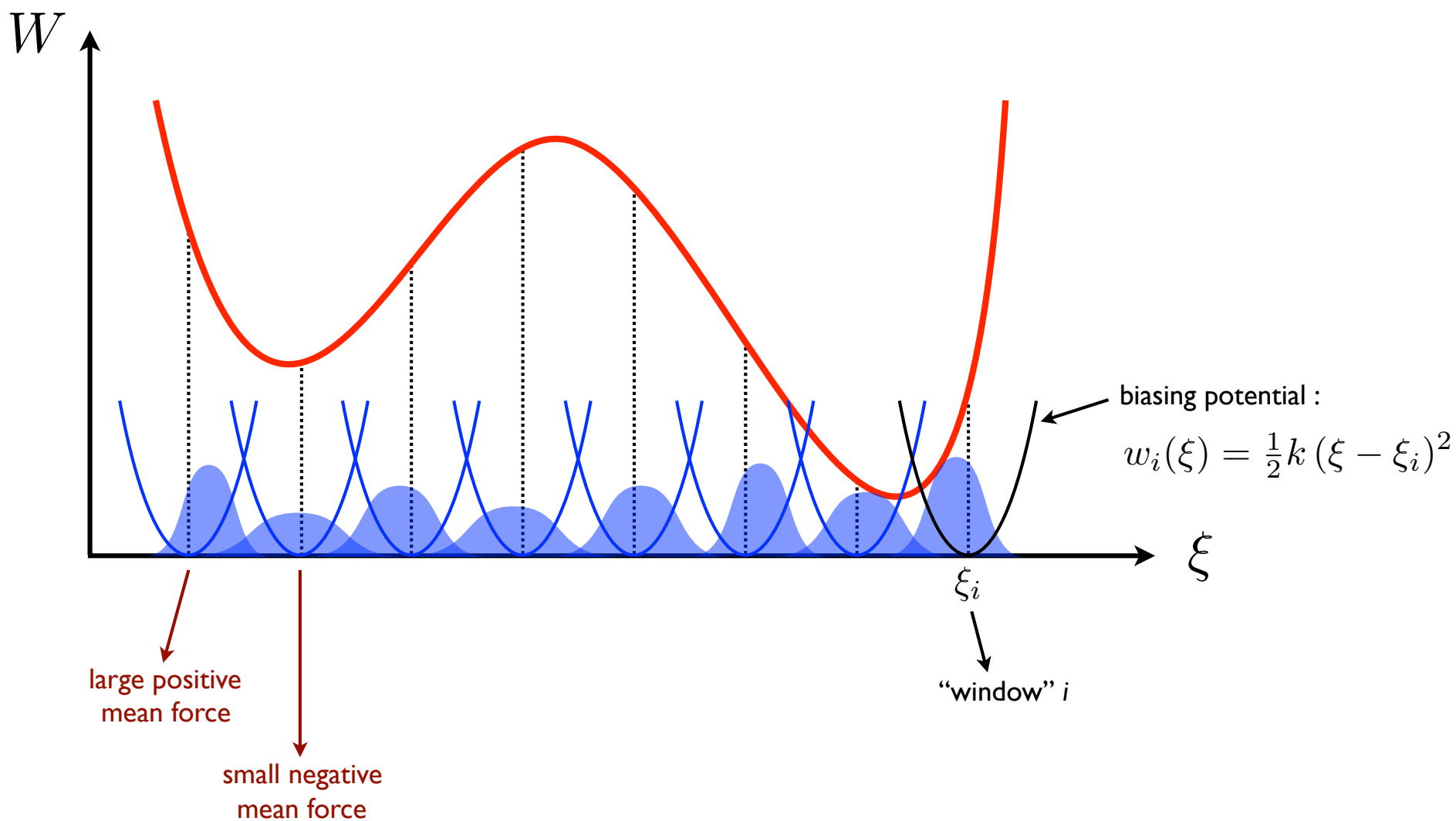
Potential of mean force

$$\langle F(\xi) \rangle = -\frac{\partial W}{\partial \xi}$$

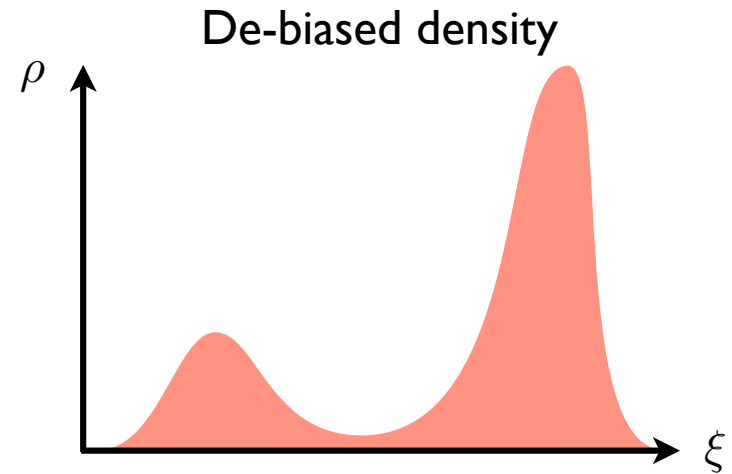
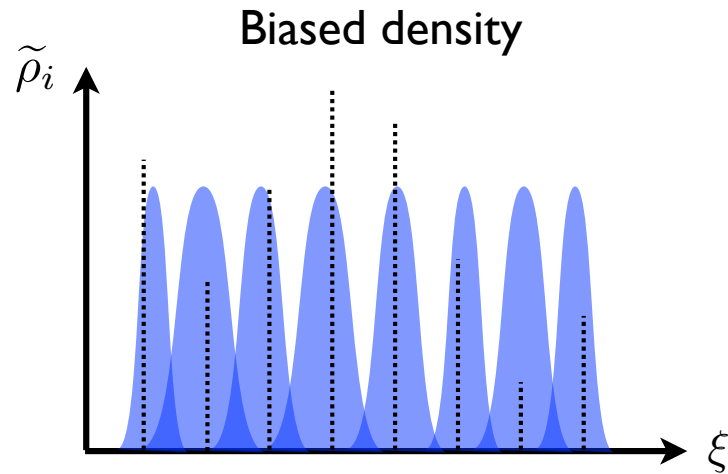
$$W(\xi) = W(\xi_0) - \int_{\xi_0}^{\xi} d\xi' \langle F(\xi') \rangle$$

We can calculate the PMF if we know the average force “felt” along the reaction coordinate.

Umbrella Sampling



Weighted Histogram Analysis Method (WHAM)



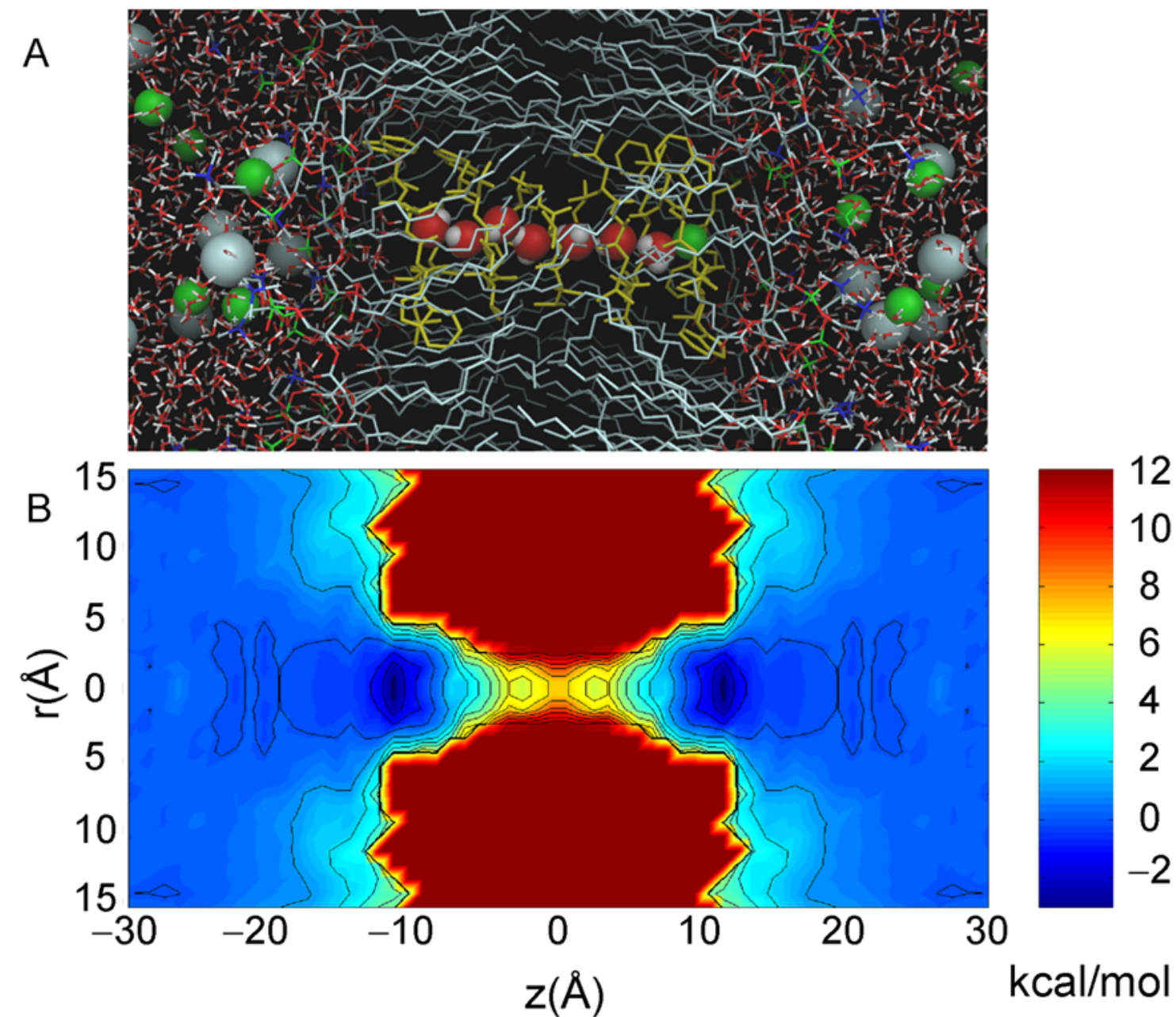
free energy added
to the system by
the introduction of
the biasing potential

$$\rho(\xi) = \sum_{i=1}^n p_i(\xi) \rho_i(\xi)$$

$$\rho_i(\xi) = \tilde{\rho}_i(\xi) e^{[w_i(\xi) - f_i]/k_B T}$$

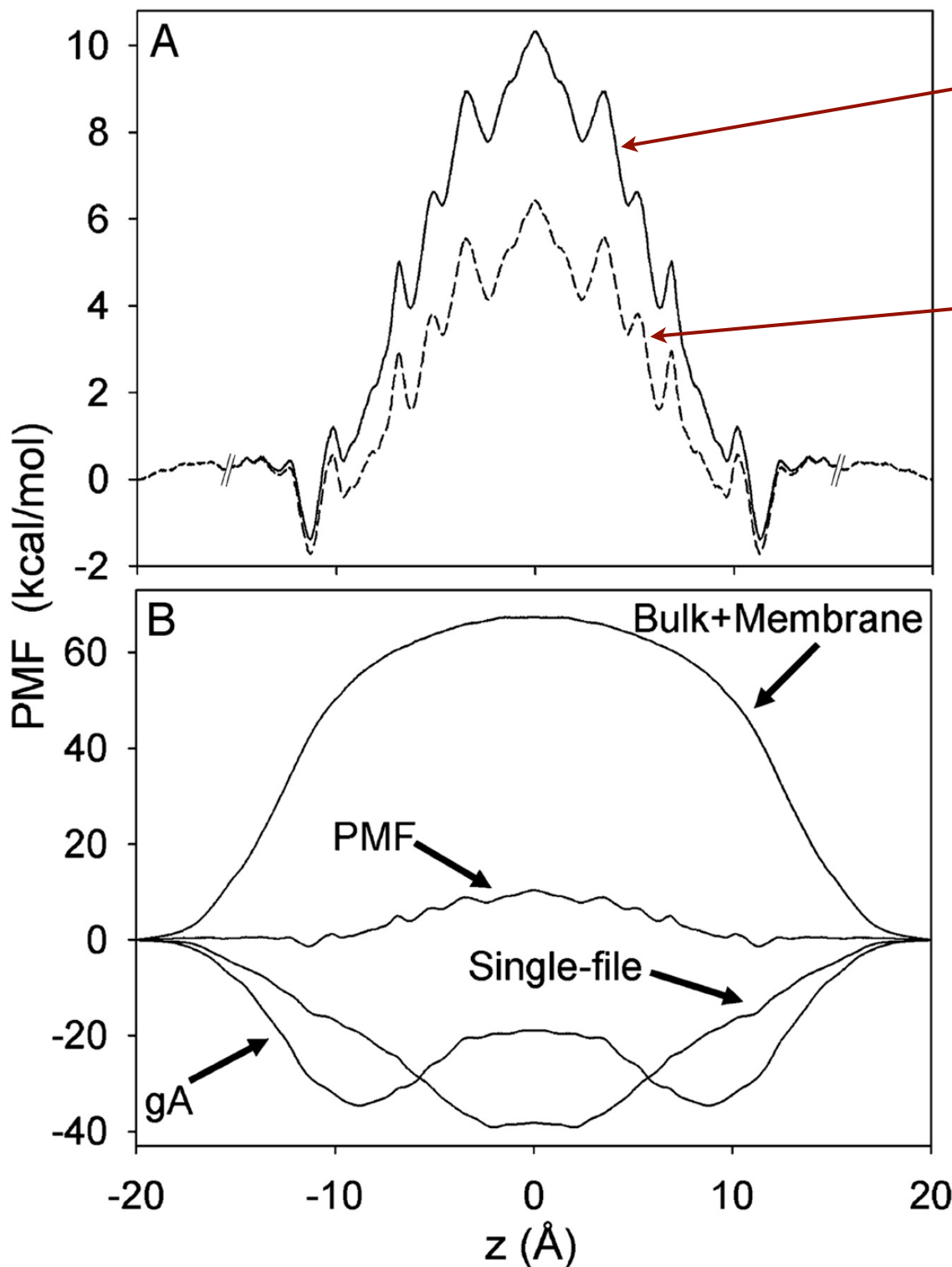
$$p_i(\xi) = \frac{n_i e^{-[w_i(\xi) - f_i]/k_B T}}{\sum_{j=1}^n n_j e^{-[w_j(\xi) - f_j]/k_B T}}$$

Free energy profile of K^+ around Gramicidin A



Reference :

Allen, Andersen, and Roux,
PNAS **101**, 117–122 (2004)



Unidimensional PMF

Unidimensional PMF corrected for simulation artifacts (lack of electronic polarizability of the lipids & periodic boundary conditions)

Free energy decomposition :

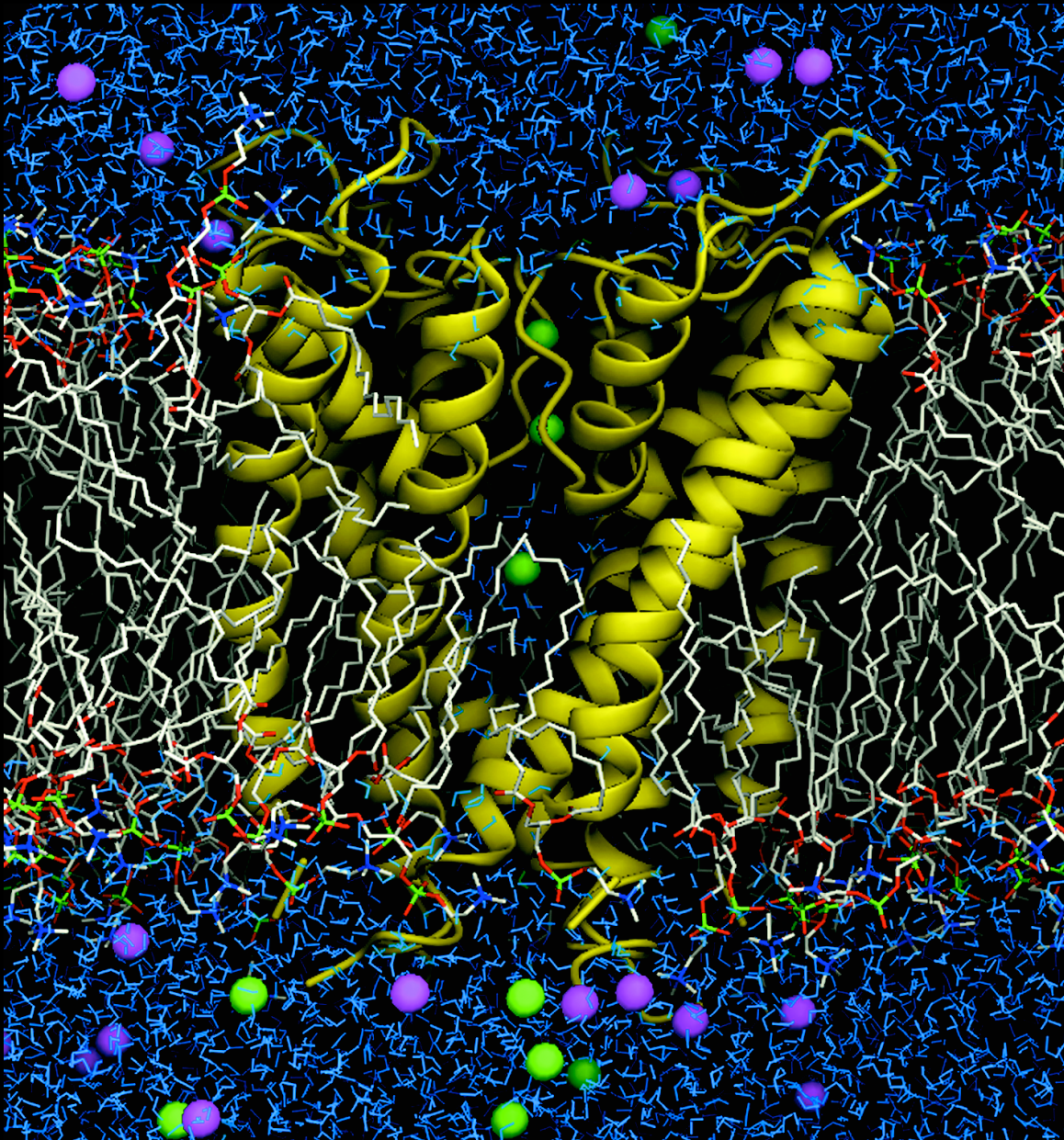
$$W_{\alpha}(z) = W_{\alpha}(z_0) - \int_{z_0}^z dz' \langle F_{\alpha}(z') \rangle$$

force on reaction coordinate z due only to atoms from subsystem α

Reference :

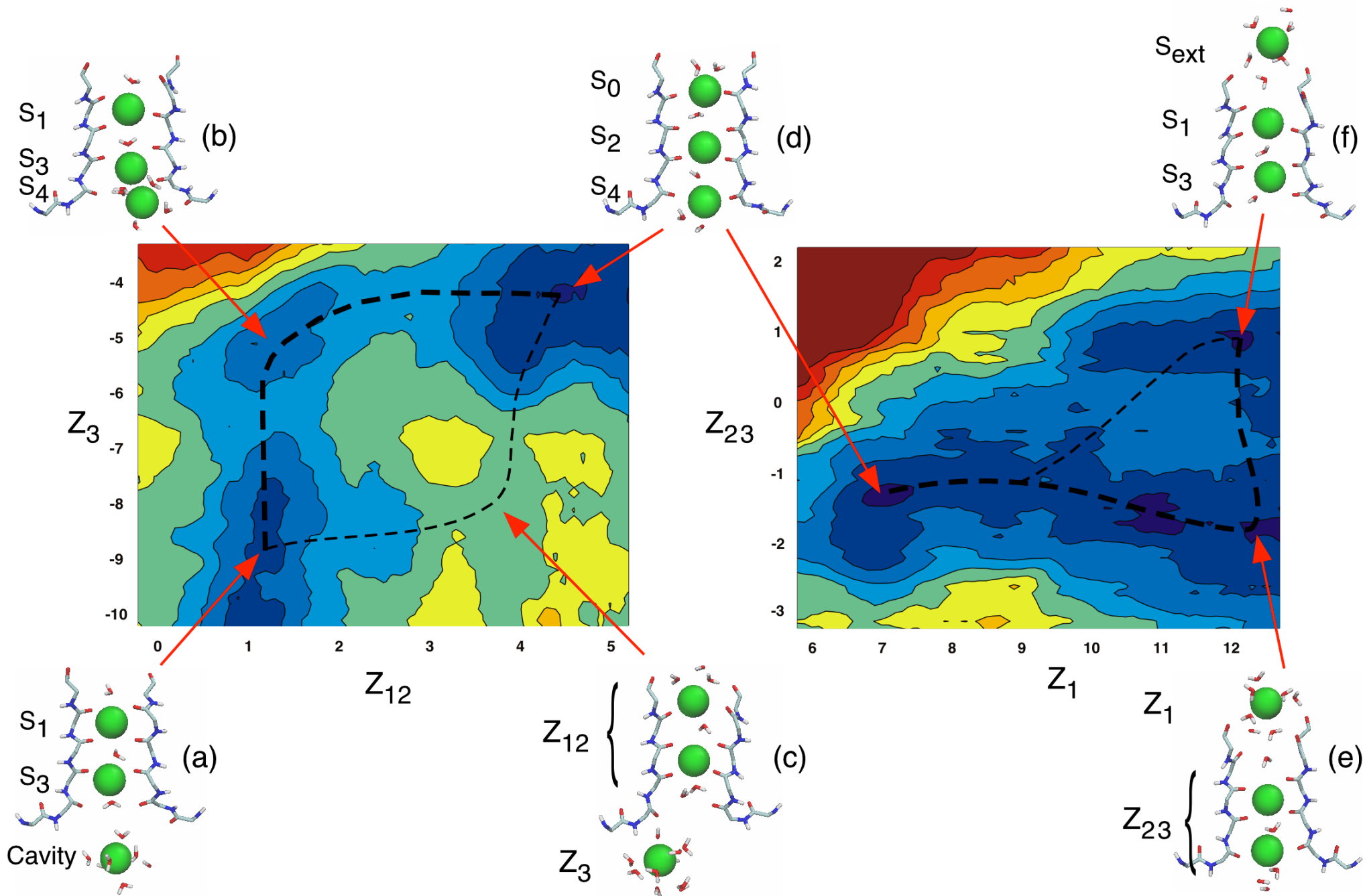
Allen, Andersen, and Roux, *PNAS* **101**, 117–122 (2004)

Molecular Dynamics simulations of the KcsA K⁺ Channel



Reference :

Bernèche and Roux, *Biophys. J.* **78**,
2900–2917 (2000)



Reference :
 Bernèche and Roux, *Nature* **414**, 73–77 (2001)

What to do with a PMF ?

Rate constants

Transition state theory :

(low-friction regime)

$$k_{\text{TST}} = \nu_{\text{w}} e^{-\Delta W^{\ddagger} / k_{\text{B}} T}$$

frequency of oscillation at
the bottom of the **well**

$$\nu_{\text{w}} = \frac{\sqrt{W''(\xi_{\text{w}}) / m_{\xi}}}{2\pi}$$

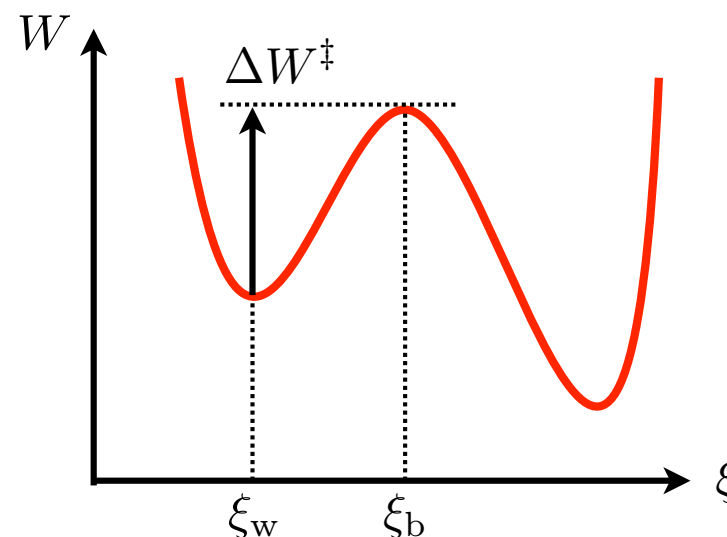
Kramers theory :

(high-friction regime)

$$k = \frac{D(\xi_{\text{b}})}{2\pi k_{\text{B}} T} \sqrt{W''(\xi_{\text{w}}) |W''(\xi_{\text{b}})|} e^{-\Delta W^{\ddagger} / k_{\text{B}} T}$$

diffusion coefficient at
the top of the **barrier**

curvature of the PMF at the bottom of
the **well** and the top of the **barrier**



This formula gets more complicated if the reaction coordinate is not a Cartesian coordinate.

Reference :

Berne and Borkovec, *J. Chem. Soc., Faraday Trans.* **94**, 2717–2723 (1998)

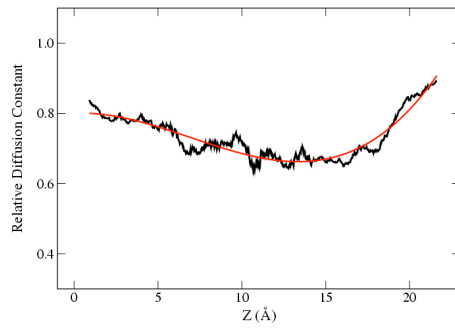
Simulate Ion Flow

**What to do
with a PMF ?**

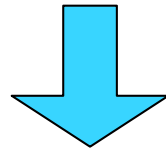
Molecular Dynamics (MD)

$$\ddot{r}_i(t) = -\frac{1}{M_i} \nabla_i U$$

Input: U



- all atoms, channel, ions, water, lipids
- ns simulations
- equilibrium properties
- dynamical diffusion constant D_i



W_{eq}

Free energy potential
of mean force from
MD simulations

+

V_{mp}

Transmembrane
potential profile
from PB-V continuum
electrostatics

=

W_{tot}

Total free energy
governing the
movements of
permeation ions

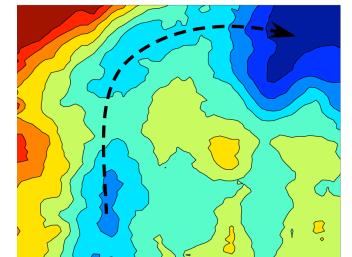
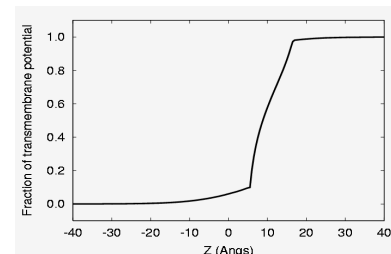
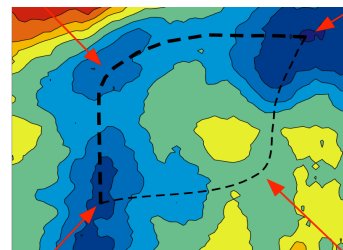
Brownian Dynamics (BD)

$$\dot{r}_i(t) = -\frac{D_i}{k_B T} \nabla_i W_{tot} + \xi_i(t)$$

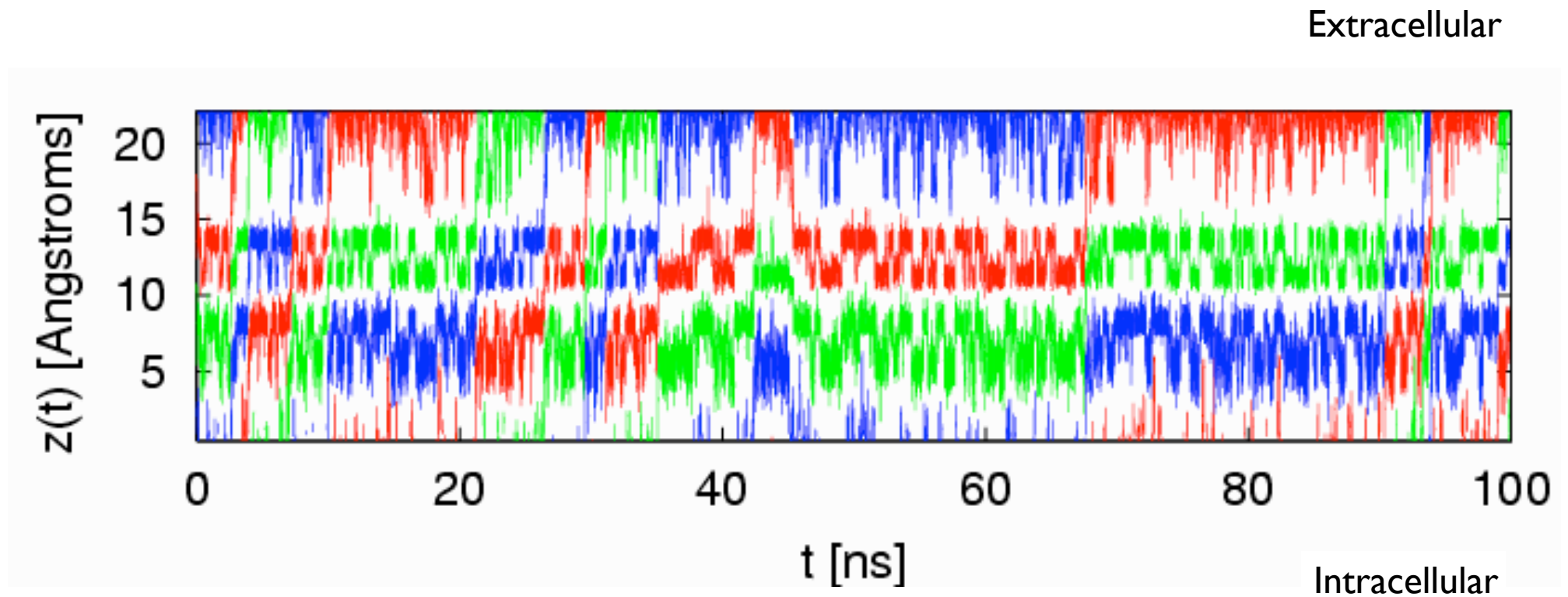
Inputs: D_i , W_{tot}

- ionic current
- μ s simulations
- just ions

Source :
Benoît Roux, Course on
ion permeation.



Brownian Dynamics Simulations of K^+ in KcsA



Source :

Benoît Roux, Course on
ion permeation.