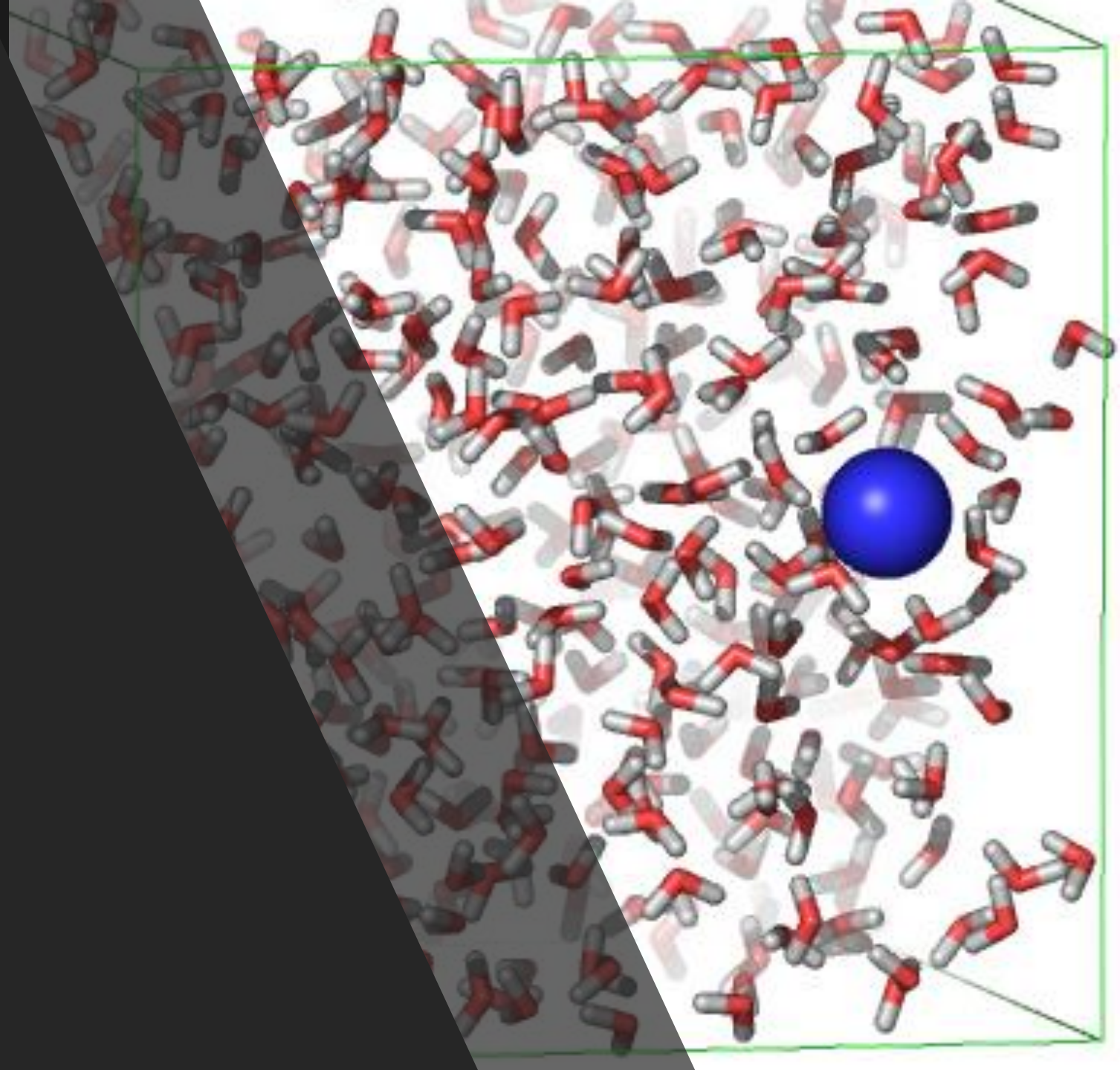


Introduction to Free Energy Simulations



Molecular Modeling Softwares & Databases:

- Molecular Mechanical Tools (Classical Minimization and MD):
 - https://en.wikipedia.org/wiki/Comparison_of_software_for_molecular_mechanics_modeling
- Quantum Mechanical Tools (QM Geom Opt and QM-MD):
 - https://en.wikipedia.org/wiki/List_of_quantum_chemistry_and_solid-state_physics_software
- Drug Design Tools:
 - https://www.click2drug.org/directory_Docking.html
- Molecular Databases:
 - https://en.wikipedia.org/wiki/List_of_biological_databases
 - (PDB, Chempyder, PubChem → some good ones)
- Building Molecules & Basic Minimization:
 - Avogadro (<https://sourceforge.net/projects/avogadro/>), Iqmol (<http://iqmol.org>)
- Visualization Softwares:
 - VMD (<http://www.ks.uiuc.edu/Development/Download/download.cgi?PackageName=VMD>)
 - Pymol (<https://sourceforge.net/projects/pymol/>)
- CHARMM Tutorial:
 - https://www.charmmtutorial.org/index.php/CHARMM_Tutorial



Preface: The “Ergodic Hypothesis” and its implications!

Consider
the
following





If you have 500 mL solution of 0.1 M *acetic acid (AA)*, how many molecules of acetic acid are in the solution??





If you have *500 mL* solution of *0.1 M acetic acid (AA)*, how many molecules of acetic acid are in the solution??

$$500\text{mL} \times \frac{10^{-3}\text{L}}{\text{mL}} \times \frac{0.1\text{mol AA}}{\text{L}} \times \frac{6.022\text{e}23 \text{ molecules AA}}{\text{mol}}$$





If you have *500 mL* solution of *0.1 M acetic acid (AA)*, how many molecules of acetic acid are in the solution??



$$500\text{mL} \times \frac{10^{-3}\text{L}}{\text{mL}} \times \frac{0.1\text{mol AA}}{\text{L}} \times \frac{6.022\text{e}23 \text{ molecules AA}}{\text{mol}}$$

$$= 3.011\text{e}20 \text{ molecules AA}$$



$3.011e20$ molecules AA



Energetic properties of this solution will be dependent on:

- **Intra**molecular interactions (dependent on molecular CONFORMATIONS!!)
- **Inter**molecular interactions (charge-charge/electrostatics, and vdW)
- Temperature (temperature makes the world go-round)



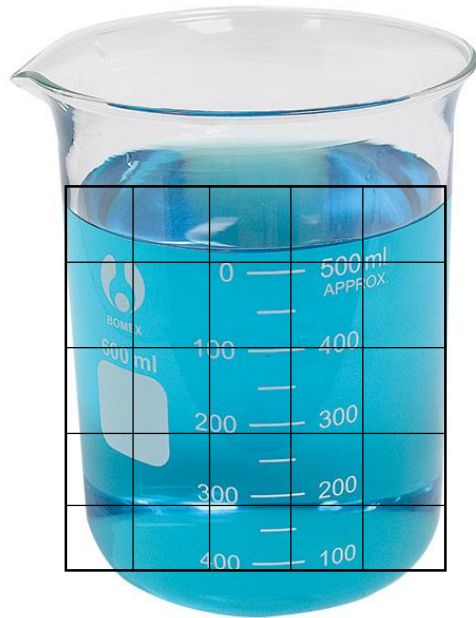
3.011e20 molecules AA

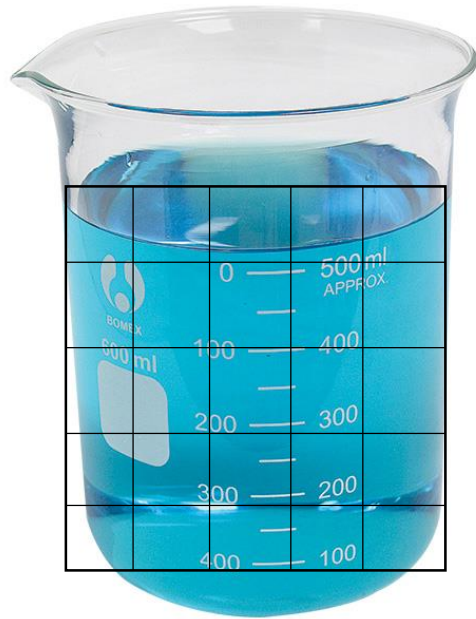


Energetic properties of this solution will be dependent on:

- ***Intra***molecular interactions (dependent on molecular CONFORMATIONS!!)
- ***Inter***molecular interactions (charge-charge/electrostatics, and vdW)
- Temperature (temperature makes the world go-round)

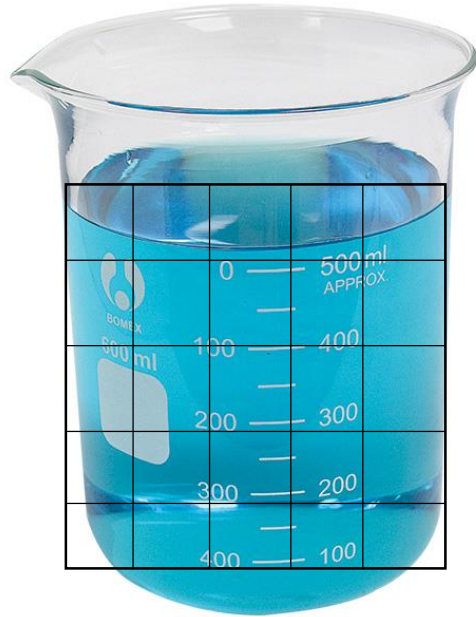
How in the HALIBUT do we simulate 3.011e20 molecules of anything?!?!





1. Try to generate each of the AA configurations randomly and separately?!



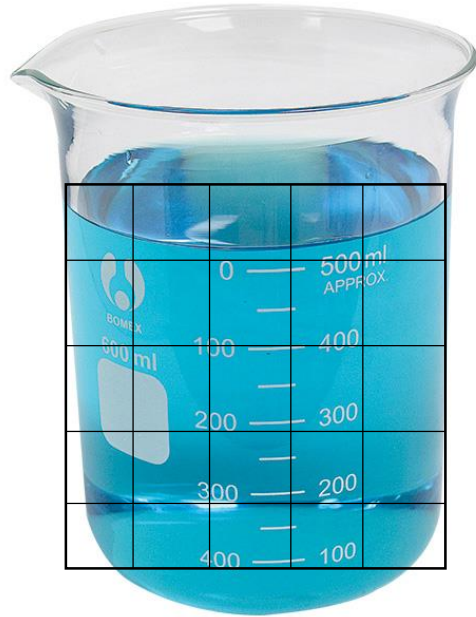


1. Try to generate each of the AA configurations randomly and separately?!



2. Generate one configuration, then allow it to *move* (dynamics) in space and time and “hope” that it samples all configurations!!

“Ergodicity”



1. Try to generate each of the AA configurations randomly and separately?!

=



2. Generate one configuration, then allow it to *move* (dynamics) in space and time and “hope” that it samples all configurations!!

Take Home Message on Ergodicity:

Ergodic Hypothesis states, for many natural systems of interest, averages of the system over time (2) are equal to an "ensemble average" (1).

Dynamical Simulations → Sample molecular "shapes" (conformations) by applying force over time which determine energetic profiles → better understanding of all $3.011e20$ AA molecules in solution!!



$$F = m \cdot a$$



$$F = m \cdot a$$



$$F = m \cdot a$$



Take Home Message on Ergodicity:

Ergodic Hypothesis states, for many natural systems of interest, averages of the system over time (2) are equal to an "ensemble average" (1).

Dynamical Simulations → Sample molecular "shapes" (conformations) by applying force over time which determine energetic profiles → better understanding of all 3.011×10^{20} AA molecules in solution!!

So long simulations may be needed to properly describe a system!!



$$F = m \cdot a$$



$$F = m \cdot a$$



$$F = m \cdot a$$





Final Thoughts on Ergodicity:

Do we really *need* $3.011e20$ frames of AA trajectory to sufficiently describe the system?!

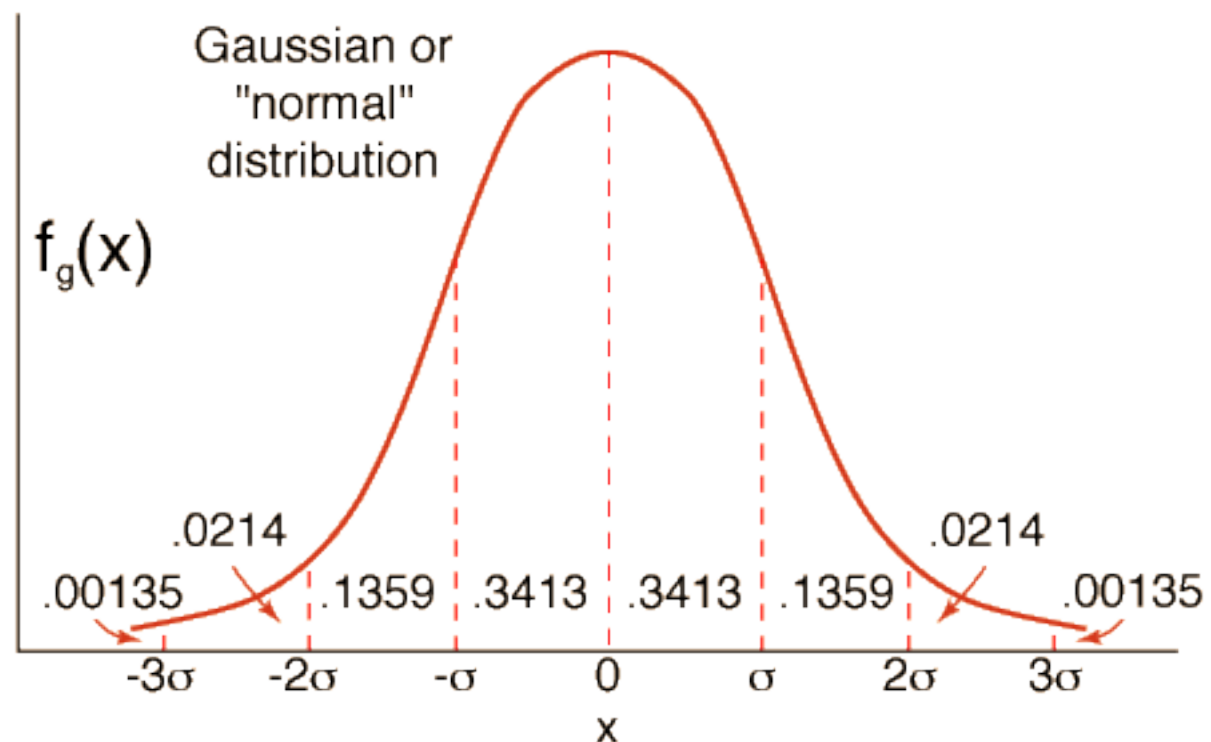


Final Thoughts on Ergodicity:

Do we really *need* $3.011e20$ frames of AA trajectory to sufficiently describe the system?!

Not really...

- Have $3.011e20$ molecules but many have similar conformations
- Populations of most degrees of freedom take a Gaussian distribution
- $\sim 10^5 - 10^7$ energy evaluations (frames) needed for $\Delta A!$ \rightarrow still a LOT!!



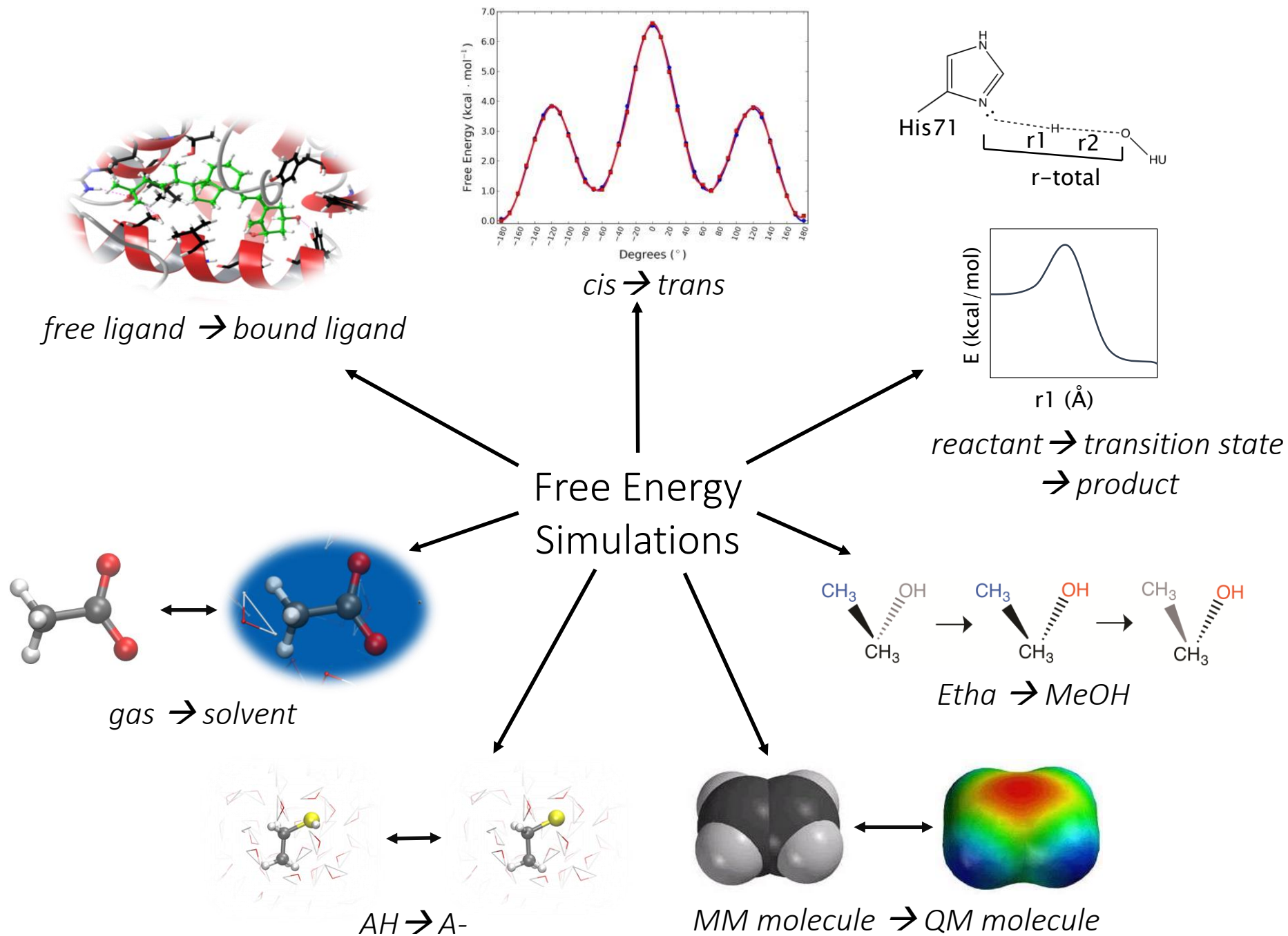


Puppy Break!!





Background to Classical Free Energy Simulations!



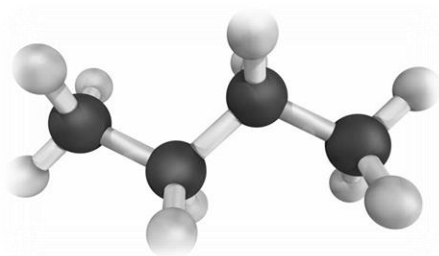
How to calculate free energies from simulation?

Zwanzig's Equation (aka Free Energy Perturbation):

0 = gas phase butane
1 = implicit solvent butane

$$\Delta A(0 \rightarrow 1) = -k_b T \ln \left\langle \exp \left(-\frac{U_1 - U_0}{k_b T} \right) \right\rangle_0$$

$10^5 - 10^7$ points in simulation



for every frame: calc U_0 and U_1

$$U(\vec{R}) = \sum_{\text{bonds}} K_b (b - b_0)^2 + \sum_{\text{UB}} K_{UB} (S - S_0)^2 + \sum_{\text{angle}} K_\theta (\theta - \theta_0)^2$$

$$+ \sum_{\text{dihedrals}} K_\chi (1 + \cos(n\chi - \delta)) + \sum_{\text{impropers}} K_{\text{imp}} (\phi - \phi_0)^2$$

$$+ \sum_{\text{nonbond}} \epsilon \left[\left(\frac{R_{\text{min}_{ij}}}{r} \right)^{12} - \left(\frac{R_{\text{min}_{ij}}}{r} \right)^6 \right] + \frac{q_i q_j}{\epsilon_1 r_{ij}}$$

How to calculate free energies from simulation?

Bennett's Acceptance Ratio (BAR)

0 = gas phase butane
1 = implicit solvent butane

$$\Delta A(0 \rightarrow 1) = k_b T \ln \left\{ \frac{\langle f(U_0 - U_1 + C) \rangle_1}{\langle f(U_1 - U_0 - C) \rangle_0} \right\} + C \quad C = k_b T \ln \frac{Q_0 N_0}{Q_1 N_1}$$

$$f(x) = (1 + \exp(-\beta x))^{-1}$$

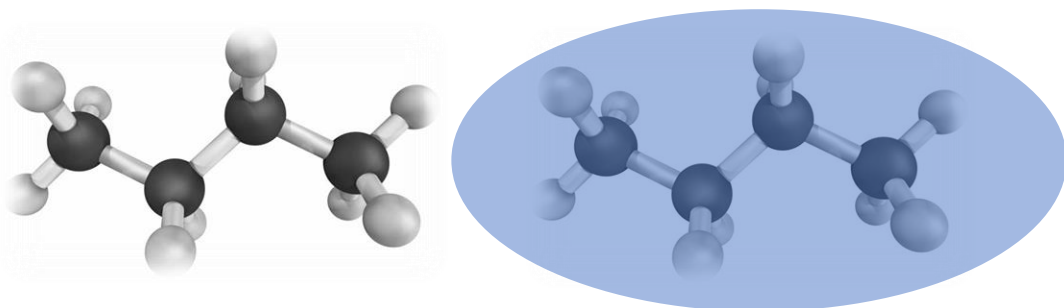
$10^5 - 10^7$ points in simulation

for every frame: calc U_0 and U_1

$$U(\vec{R}) = \sum_{\text{bonds}} K_b (b - b_0)^2 + \sum_{\text{UB}} K_{UB} (S - S_0)^2 + \sum_{\text{angle}} K_\theta (\theta - \theta_0)^2$$

$$+ \sum_{\text{dihedrals}} K_\chi (1 + \cos(n\chi - \delta)) + \sum_{\text{impropers}} K_{\text{imp}} (\phi - \phi_0)^2$$

$$+ \sum_{\text{nonbond}} \epsilon \left[\left(\frac{R_{\text{min}_{ij}}}{r} \right)^{12} - \left(\frac{R_{\text{min}_{ij}}}{r} \right)^6 \right] + \frac{q_i q_j}{\epsilon_1 r_{ij}}$$





“No plan survives first contact
with implementation.”
-Andy Weir, “The Martian”

(based on a quote in Sun Tzu’s “The Art of War”)

But wait! It's not always that easy!

CHALLENGES

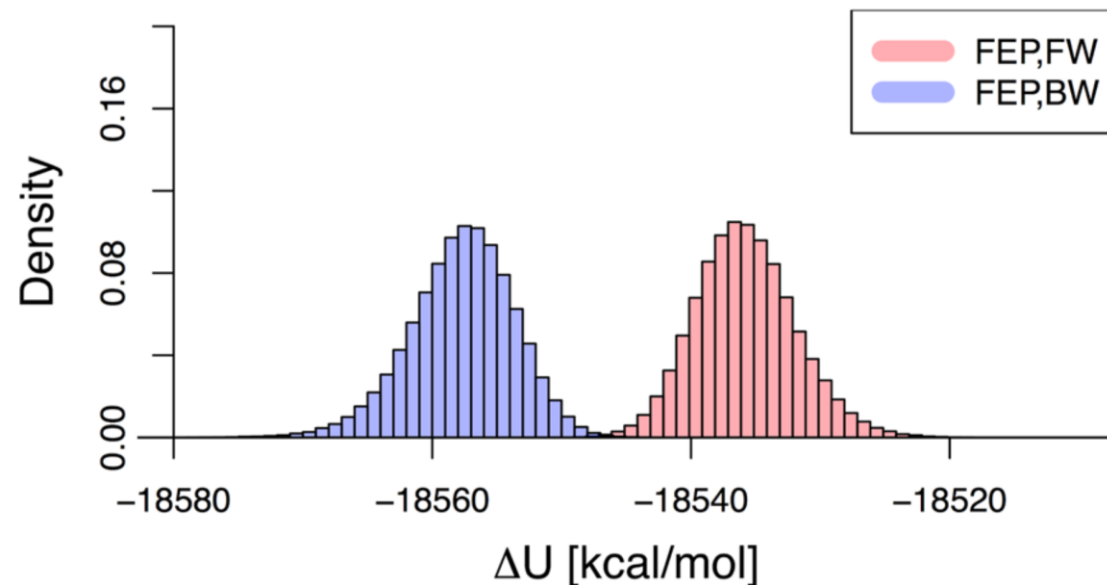
Gas \rightarrow Explicit Solvent: overlaps with waters
 \rightarrow catastrophic clashes

- Gas \rightarrow Implicit, likely done in 1 step, not explicit solvent

Complicated transformations: likely low "overlap" between endstates 0 and 1

- "overlap" = conformation found in state 0 is likely to exist in state 1, and vice versa
- when there overlap is low, potential energy surfaces are too dissimilar and the central assumption of FEP is violated

MM treatment not appropriate for reaction paths, or some other transformations



(a) ΔU Histograms



But wait! It's not always that easy!

CHALLENGES

Gas \rightarrow Explicit Solvent: overlaps with waters
 \rightarrow catastrophic clashes

- Gas \rightarrow Implicit, likely done in 1 step, not explicit solvent

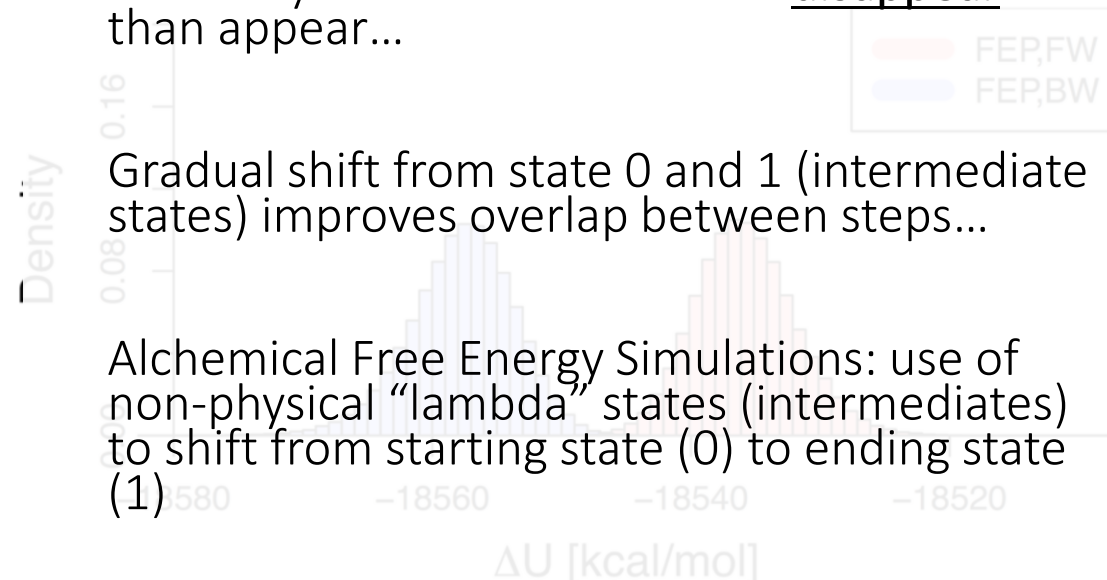
Complicated transformations: likely low "overlap" between endstates 0 and 1

- "overlap" = conformation found in state 0 is likely to exist in state 1, and vice versa
- when there overlap is low, potential energy surfaces are too dissimilar and the central assumption of FEP is violated

MM treatment not appropriate for reaction paths, or some other transformations

SOLUTIONS:

Generally easier to make atoms disappear than appear...



Alchemical Free Energy Simulations: use of non-physical "lambda" states (intermediates) to shift from starting state (0) to ending state (1)

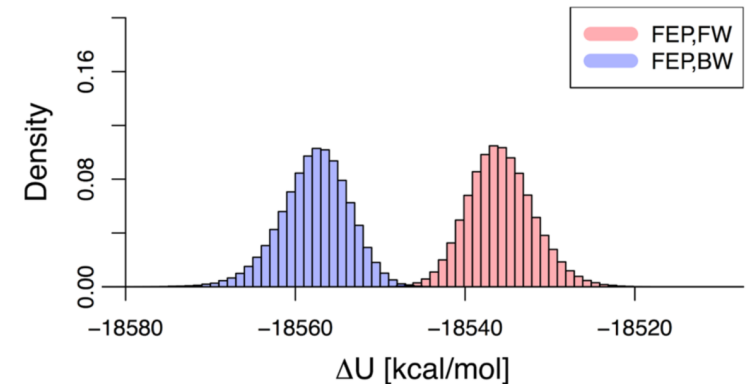
QM/MM Free Energy Simulations!

Alchemical Free Energy Simulations:



$$\Delta A(0 \rightarrow 1) = -k_b T \ln \left\langle \exp \left(-\frac{U_1 - U_0}{k_b T} \right) \right\rangle_0$$

$$\Delta A(0 \rightarrow 1) = k_b T \ln \left\{ \frac{\langle f(U_0 - U_1 + C) \rangle_1}{\langle f(U_1 - U_0 - C) \rangle_0} \right\} + C$$



(a) ΔU Histograms



Alchemical Free Energy Simulations:

0 → 0.1 → 0.2 → ... → 0.8 → 0.9 → 1

$$\Delta A(0 \rightarrow 1) = -k_b T \ln \left\langle \exp \left(-\frac{U_1 - U_0}{k_b T} \right) \right\rangle_0$$
$$\Delta A(0 \rightarrow 1) = k_b T \ln \left\{ \frac{\langle f(U_0 - U_1 + C) \rangle_1}{\langle f(U_1 - U_0 - C) \rangle_0} \right\} + C$$
$$\Delta A(0 \rightarrow 1) = \int_0^1 \left\langle \frac{\partial U}{\partial \lambda} \right\rangle d\lambda$$

Multiscale Free Energy Simulations: An Efficient Method for Connecting Classical MD Simulations to QM or QM/MM Free Energies Using Non-Boltzmann Bennett Reweighting Schemes

Gerhard König,[†] Phillip S. Hudson,[‡] Stefan Boresch,^{*,§} and H. Lee Woodcock^{*,‡}

[†]Laboratory of Computational Biology, National Heart Lung and Blood Institute, National Institutes of Health, Bethesda, Maryland 20892, United States

[‡]Department of Chemistry, University of South Florida, 4202 E. Fowler Avenue, CHE205, Tampa, Florida 33620-5250, United States

[§]Department of Computational Biological Chemistry, Faculty of Chemistry, University of Vienna, Währingerstraße 17, A-1090 Vienna, Austria



Contents lists available at ScienceDirect

Biochimica et Biophysica Acta

journal homepage: www.elsevier.com/locate/bbagen



Efficiently computing pathway free energies: New approaches based on chain-of-replica and Non-Boltzmann Bennett reweighting schemes[☆]



Phillip S. Hudson^{a,1}, Justin K. White^{a,1}, Fiona L. Kearns^a, Milan Hodoscek^b, Stefan Boresch^c, H. Lee Woodcock^{a,*}

^a Department of Chemistry, University of South Florida, 4202 E. Fowler Ave., CHE205, Tampa, FL 33620-5250, USA

^b Center for Molecular Modeling, National Institute of Chemistry, Hajdrihova 19, SI-1000 Ljubljana, Slovenia

^c Department of Computational Biological Chemistry, Faculty of Chemistry, University of Vienna, Währingerstraße 17, A-1090 Vienna, Austria

Use of Nonequilibrium Work Methods to Compute Free Energy Differences Between Molecular Mechanical and Quantum Mechanical Representations of Molecular Systems

Phillip S. Hudson,[†] H. Lee Woodcock,^{*,†} and Stefan Boresch^{*,‡}

[†]Department of Chemistry, University of South Florida, 4202 East Fowler Avenue, CHE205, Tampa, Florida 33620-5250, United States

[‡]Department of Computational Biological Chemistry, Faculty of Chemistry, University of Vienna, Währingerstraße 17, A-1090 Vienna, Austria

Methods for Efficiently and Accurately Computing Quantum Mechanical Free Energies for Enzyme Catalysis

F.L. Kearns^{*}, P.S. Hudson^{*}, S. Boresch^{†,1}, H.L. Woodcock^{*,1}

^{*}University of South Florida, Tampa, FL, United States

[†]Faculty of Chemistry, University of Vienna, Vienna, Austria

¹Corresponding authors: e-mail address: stefan@mdy.univie.ac.at; hlw@usf.edu

Convergence of single-step free energy perturbation

Stefan Boresch & H. Lee Woodcock

To cite this article: Stefan Boresch & H. Lee Woodcock (2016): Convergence of single-step free energy perturbation, Molecular Physics, DOI: [10.1080/00268976.2016.1269960](https://doi.org/10.1080/00268976.2016.1269960)

To link to this article: <http://dx.doi.org/10.1080/00268976.2016.1269960>

Computing Converged Free Energy Differences between Levels of Theory via Nonequilibrium Work Methods: Challenges and Opportunities

Fiona L. Kearns,^[a] Phillip S. Hudson,^[a] Henry L. Woodcock,^{*,[a]} and Stefan Boresch^{*,[b]}



Now you all are free energy experts!



Thank you! Questions?



