

Molecular Biophysics –  
Part 3 –  
Lecture 3  
Relating Molecular Dynamics to  
Macroscopic Quantities

David Mathews

# Contact Info:

- [David\\_Mathews@urmc.rochester.edu](mailto:David_Mathews@urmc.rochester.edu)
- Phone: x56313
- Office: 1.11132 (MRBX)
- Web: <http://rna.urmc.rochester.edu>
- Office Hours:
  - Wednesday April 6, 10 AM – noon
  - Wednesday April 13, 10 AM – noon
  - Or by appointment

# Review of Lecture 1:

- Dynamic programming refers to calculations in which the cost/score is built progressively from smaller solutions.
- Dynamic programming algorithms guarantee the global minimum solution.

# Review of Lecture 2:

- Molecular mechanics forcefields approximate the energetics of a molecule using classical mechanics.
  - Bonded terms
    - Bond stretch
    - Bond angles
    - Dihedrals (including improper dihedrals)
  - Nonbonded terms
    - Electrostatics
    - Van der Waals
- Geometry optimizers
  - Steepest Descent (local minima)
  - Conjugate Gradient (local minima)
  - Simulated annealing (not limited to local minima)

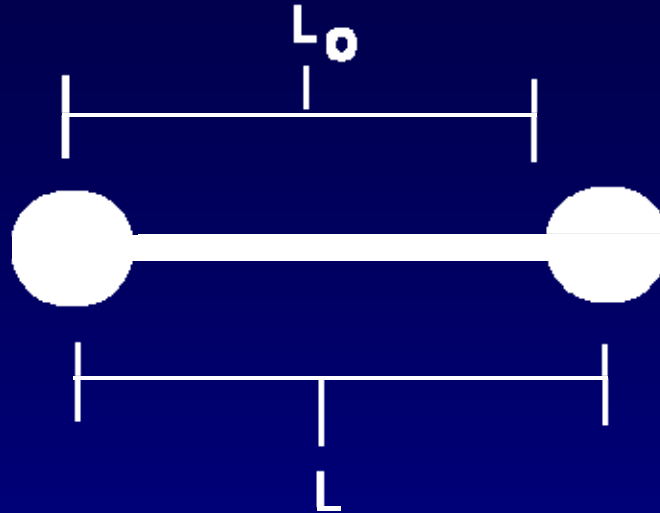
## Review of Lecture 2:

- Molecular dynamics simulates the motion of a molecule as a function of time.
- The equations of motion are chosen to minimize propagation errors.
  - Verlet algorithm
  - Leapfrog algorithm

# Overview of Lecture 3:

- Sample systems for understanding the equations of motion.
  - Basis for next homework assignment
- Using simulations to predict macroscopic quantities.
  - Ergodic Hypothesis
  - Free energy methods
  - Entropy, Enthalpy, temperature
- Implicit Solvent

# Recall Bond Stretch:



- Bond stretch is modeled using Hooke's Law, i.e. a spring with spring constant  $K_L$  and equilibrium length  $L_0$ .

$$V(\text{bond}) = K_L (L - L_0)^2$$

# Harmonic Oscillator:

- $V(x) = k(x - x_0)^2$ , set  $x_0 = 0$

$$V(x) = k(x)^2$$

$$F = -(d/dx) V(x) = -2kx = ma = m (d^2/dt^2) x(t)$$

$$\frac{d^2x}{dt^2} + \frac{2kx}{m} = 0$$

$$x(t) = A \sin\left(\sqrt{\frac{2k}{m}}t - \delta\right)$$

$$= A \cos\left(\sqrt{\frac{2k}{m}}t - \delta\right)$$

$$x(t) = A \cos\left(\sqrt{\frac{2k}{m}}t\right)$$

$$\frac{dx(t)}{dt} = v(t) = -A \sqrt{\frac{2k}{m}} \sin\left(\sqrt{\frac{2k}{m}}t\right)$$

$$\tau \sqrt{\frac{2k}{m}} = 2\pi$$

$$\tau = \pi \sqrt{\frac{m}{k}}$$

# Period of H Vibration in Methyl Group:

- (Approximate C position as fixed.)
- $Sp^3$  C – methyl H
- $k=340$  kcal/(mol  $\text{\AA}^2$ )

$$\tau = \sqrt{2}\pi \sqrt{\frac{1.008 \text{ AMU}}{340 \frac{\text{kcal}}{\text{mol } \text{\AA}^2}}}$$

$$\tau = \sqrt{2}\pi \sqrt{\frac{(1.008 \text{ AMU})(1.66054 \times 10^{-27} \text{ kg/AMU})}{\left(340 \frac{\text{kcal}}{\text{mol } \text{\AA}^2}\right) \left(\frac{4184 \text{ J/kcal}}{6.02 \times 10^{23} \times (1 \times 10^{-10} \text{ m/\text{\AA}})^2}\right)}}$$

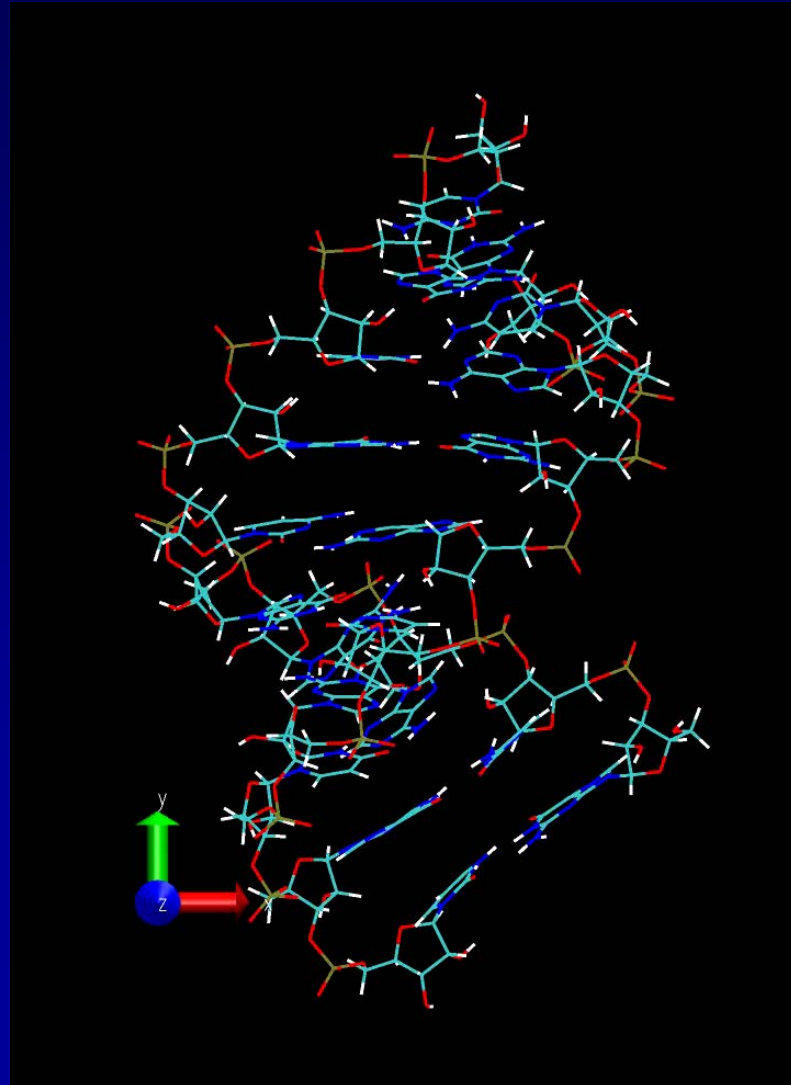
$$\tau = 1.2 \times 10^{-14} \text{ s}$$

$$\tau = 12 \text{ fs}$$

# Homework #2:

- For the harmonic oscillator approximation of a carbon-hydrogen bond, calculate the position, velocity, and acceleration for the hydrogen atom bonded to a carbon with fixed position from 0 to 10 fs, at 1 fs timesteps.
- Start the calculation with  $v(0) = 0$ ,  $x(0) = 0.5 \text{ \AA}$
- First, calculate the position, velocity, and acceleration analytically.
- Next, use the poor equations of motion from lecture on 3/31 to calculate position, velocity, and acceleration:
  - $x(t+\Delta t) = x(t) + v(t) \times \Delta t + (1/2) \times a(t) \times (\Delta t)^2$
  - $v(t+\Delta t) = [x(t+\Delta t) - x(t)]/\Delta t$
- Finally, calculate position, velocity, and acceleration using the Verlet algorithm:
  - $x(t+\Delta t) = 2x(t) - x(t-\Delta t) + a(t) \times (\Delta t)^2$
  - $v(t) = [x(t+\Delta t) - x(t-\Delta t)]/2\Delta t$
- Use the analytic solution of  $x(-\Delta t)$  to seed the first step.
- Note that the Verlet algorithm is a better approximation of the analytic solution (as advertised).

# Connecting Simulation to Thermodynamics:



# Connecting Simulation to Thermodynamics:

- Thermodynamics uses statistical methods to understand the behavior of molecules.
- How can we predict thermodynamic behavior using MD?
- Bad Approach:
  - Start a simulation with a large number of molecules.
  - Memory use is  $O(N)$  and time scales  $O(N^2)$  (or  $O(N \log N)$ ) where  $N$  is the number of atoms.

# Ergodic Hypothesis: (Ergodicity):

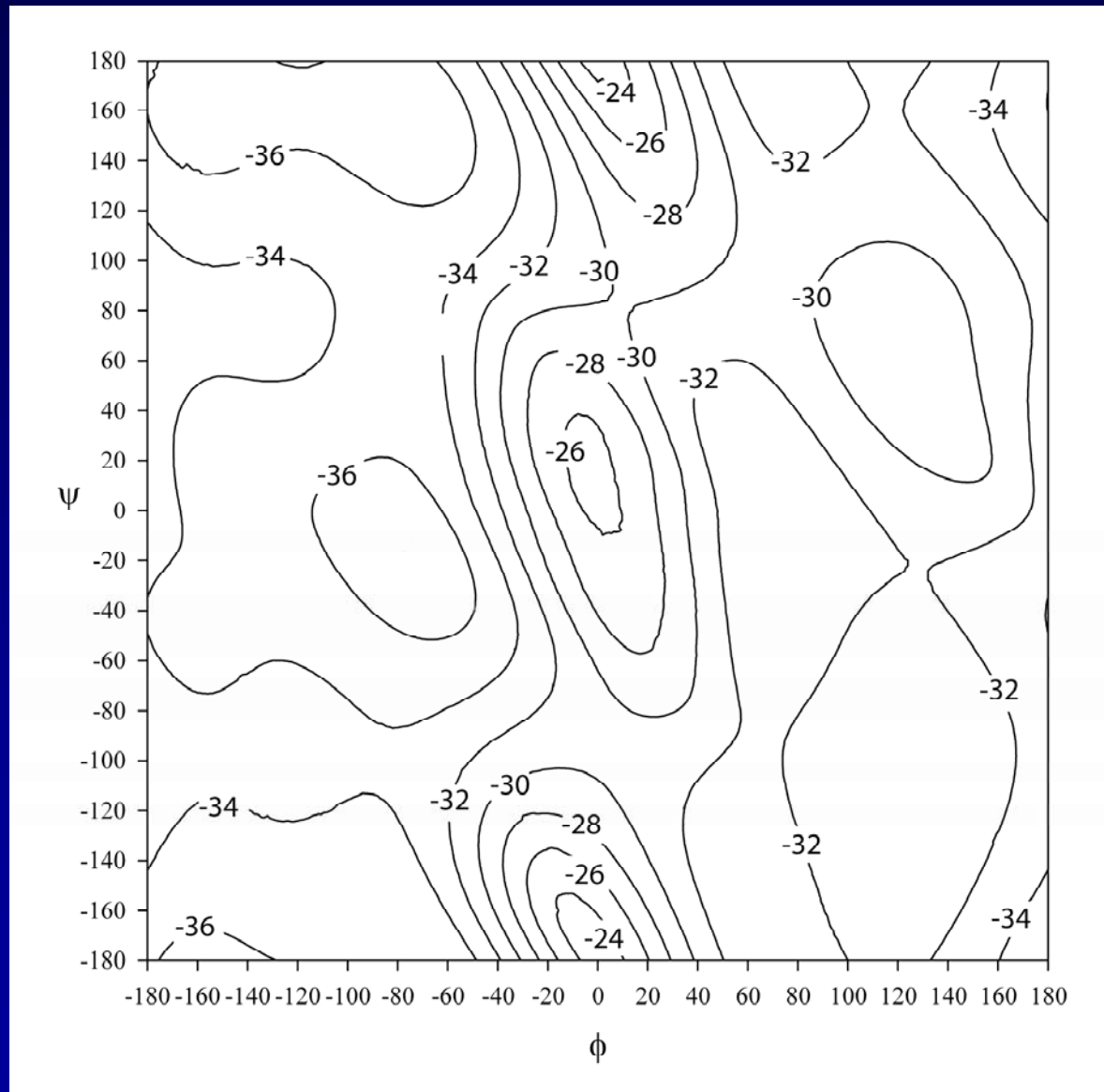
$$\bar{A} = \frac{1}{t} \int_{t_0}^{t_0+t} A(t) dt$$

This will change depending on the starting conditions.

$$\bar{A} = \lim_{t \rightarrow \infty} \frac{1}{t} \int_{t_0}^{t_0+t} A(t) dt = \frac{1}{M} \sum_{i=1}^M A_i = \langle A \rangle$$

This says that an average quantity measured over long enough time will approach the ensemble average.

# Is the Ergodic Hypothesis True in General?



# Convergence:

- What if we want to know the ensemble average for  $\Phi$ ?
- We would run the simulation for some period of time (1 ns).
- We would calculate  $\Phi$  ( $\Phi_1$ ).
- Then we would run another 1 ns and recalculate  $\Phi$  ( $\Phi_2$ ) using the whole time.
  - If  $\Phi_1 = \Phi_2$ , there is evidence that the simulation converged.
  - If  $\Phi_1 \neq \Phi_2$ , the calculation has not converged and should be run longer.
- Alternatively, we could start a new simulation from a different starting conformation and calculate  $\Phi$ .

# Free Energy Methods:

In a canonical ensemble :

$$Q = \int \int e^{-V/kt} dqdp$$

$$A = -kt \ln Q$$

$$= kt \ln(1/Q)$$

$$= ktC \ln \frac{\int \int e^{V/kt} e^{-V/kt} dqdp}{\int \int e^{-V/kt} dqdp}$$

$$= kt \ln \langle e^{E/kt} \rangle$$

- This is self-defeating because the probable states ( $e^{-V/kt}$ ) with low energy are the states that contribute little to the integral ( $e^{V/kt}$ ) and *vice versa*.

# About Free Energy:

- Free energies in the NVT ensemble are Helmholtz Free Energies (A).
- Free energies in the NPT ensemble are Gibb's Free Energies (G).

# Free Energy Differences are Tractable:

- Calculate the free energy difference between A and B.
- $V_B = V_A + V_{BA}$

$$\Delta A = A_B - A_A = -kt \ln Q_B/Q_A$$

$$\begin{aligned}\Delta A &= -kt \ln \frac{\int \int e^{-V_B/kt} dqdp}{\int \int e^{-V_A/kt} dqdp} \\ &= -kt \ln \frac{\int \int (e^{-V_A/kt} e^{+V_A/kt}) e^{-V_B/kt} dqdp}{\int \int e^{-V_A/kt} dqdp} \\ &= -kt \ln \frac{\int \int e^{-V_A/kt} e^{-V_{BA}/kt} dqdp}{\int \int e^{-V_A/kt} dqdp} \\ &= -kt \ln \langle e^{-V_{BA}/kt} \rangle_A\end{aligned}$$

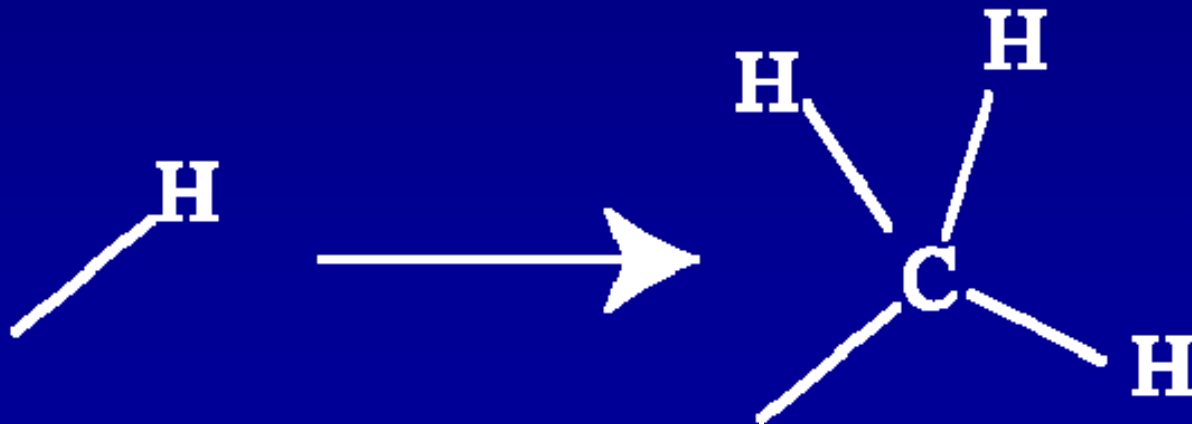
# Free Energy Perturbation:

$$\Delta A = -kt \ln \langle e^{V_{BA}/kt} \rangle_A$$

- The free energy difference is a function of the potential difference between state A and state B as an ensemble average over states in A.
- Unlike perturbation theory (Quantum mechanics) this is an exact equation.
- Examples:
  - Predict the effect on stability of changes in amino acid sequence.
  - Predict pKa's for titratable side chains

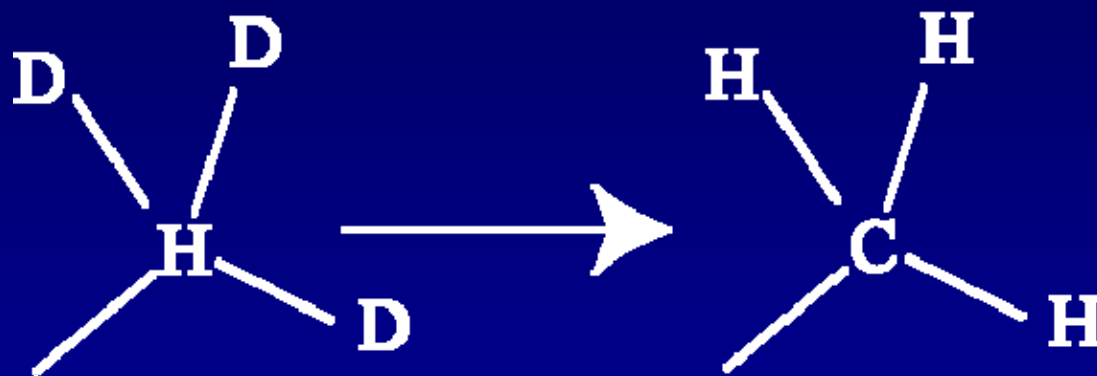
# What About Atom Numbers:

- We stated that we are using the NVT ensemble.
- Interesting perturbations require changes in atom number.
- Consider the change from glycine to alanine:



# Dummy Atoms:

- Add dummy atoms that do not interact with the system:



# Practical Scheme:

$$\Delta A = -kt \ln \langle e^{V_{BA}/kt} \rangle_A$$

- Errors are smaller if the perturbation term,  $V_{BA}$  is smaller.

$$\begin{aligned} V(\lambda) &= (1-\lambda)V_A + \lambda V_B \\ \Delta A &= -kt \int_0^1 \frac{\partial \ln \langle e^{-V(\lambda)} \rangle_\lambda}{\partial \lambda} \partial \lambda \\ &= -kt \int_0^1 \frac{\partial \ln \langle 1 + \frac{-V(\lambda)}{kt} \rangle_\lambda}{\partial \lambda} \partial \lambda \\ &= -kt \int_0^1 \frac{\partial \langle \frac{-V(\lambda)}{kt} \rangle_\lambda}{\partial \lambda} \partial \lambda \\ &= \int_0^1 \frac{\partial \langle -V(\lambda) \rangle_\lambda}{\partial \lambda} \partial \lambda \end{aligned}$$

- Calculation can be divided into a number of calculations from  $\lambda = 0$  to  $\lambda = 1$ .

# Solving the Integral:

- Use a Gaussian Quadrature to solve the integral:

$$V(\lambda) = (1-\lambda)V_A + \lambda V_B$$

$$\Delta A = \int_0^1 \frac{\partial \langle -V(\lambda) \rangle_\lambda}{\partial \lambda} d\lambda = \sum_{i=1}^n w_i \frac{\partial \langle -V(\lambda) \rangle_{\lambda_i}}{\partial \lambda}$$

- Press *et al.* *Numerical Recipes in {C, Fortran, C++}*. Cambridge University Press.

| n | $\lambda_i$ | $w_i$  |
|---|-------------|--------|
| 5 | .04691      | .11846 |
|   | .23076      | .23931 |
|   | .5          | .28444 |
|   | .76923      | .23931 |
|   | .95308      | .11846 |
| 9 | .01592      | .04064 |
|   | .08198      | .09032 |
|   | .19331      | .13031 |
|   | .33787      | .15617 |
|   | .5          | .16512 |
|   | .66213      | .15617 |
|   | .80669      | .13031 |
|   | .91802      | .09032 |
|   | .98408      | .04064 |