

7cUfgY; fUJbYX Molecular Dynamics

A tool in biophysical chemistry

Material Properties

Statistical Thermodynamics

Potentials & Distribution

Meaning of Time

Numerical Approximations

Material Properties

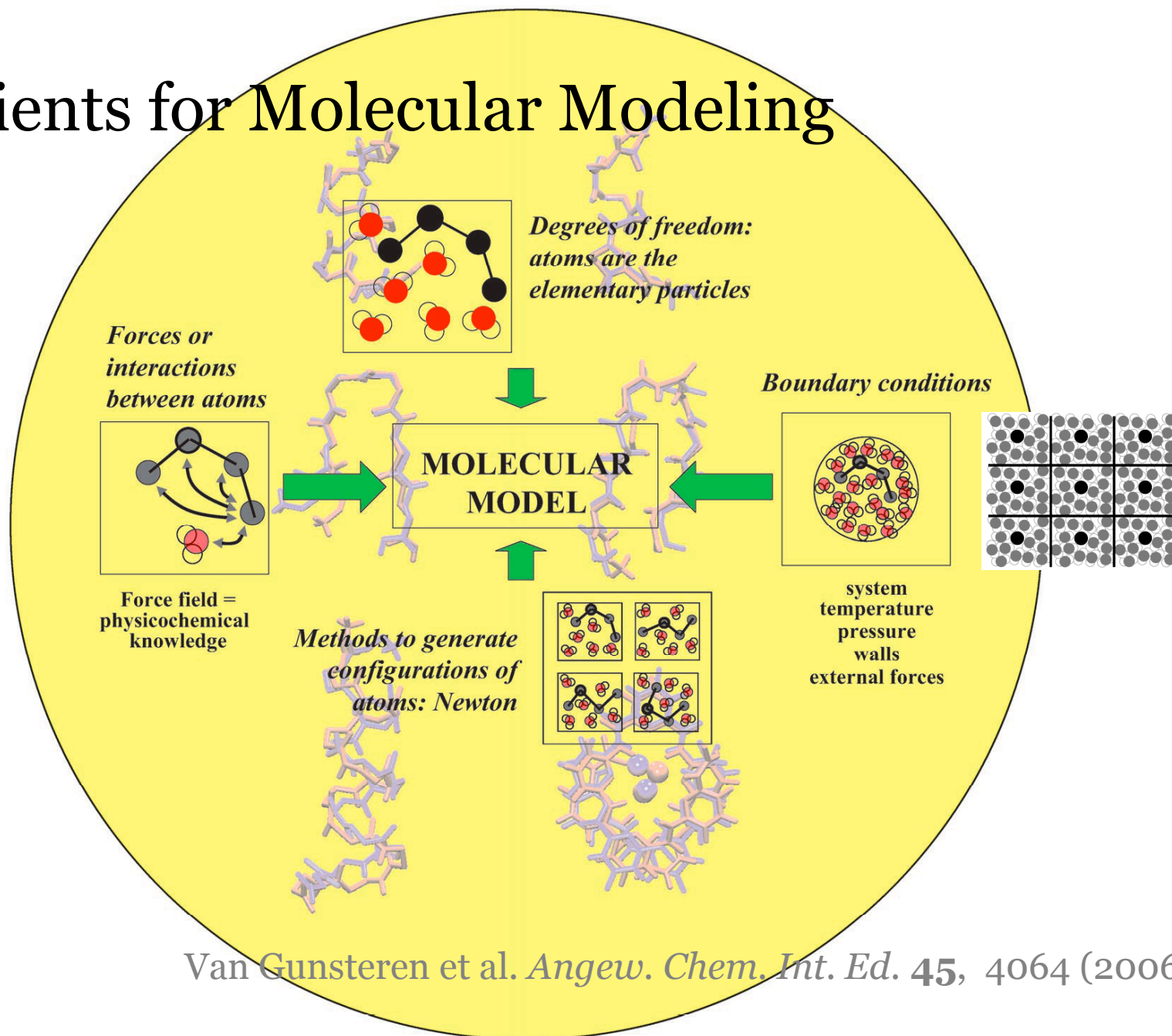
- › Matter is characterized by
 - Structure
 - Spatial relationships between building blocks
 - ORGANIZATION
 - Dynamics
 - Temporal behavior of structure
 - FLUCTUATIONS
 - RESPONSE

(Molecular) Driving Forces

- › Material properties are due to interactions between building blocks
 - ENERGY
 - Electrostatic, dispersion, repulsion
 - Chemical bond

- › and heat
 - Fundamental restlessness of nature
 - ENTROPY
 - Counts the accessible states of the system
 - Temperature

Ingredients for Molecular Modeling

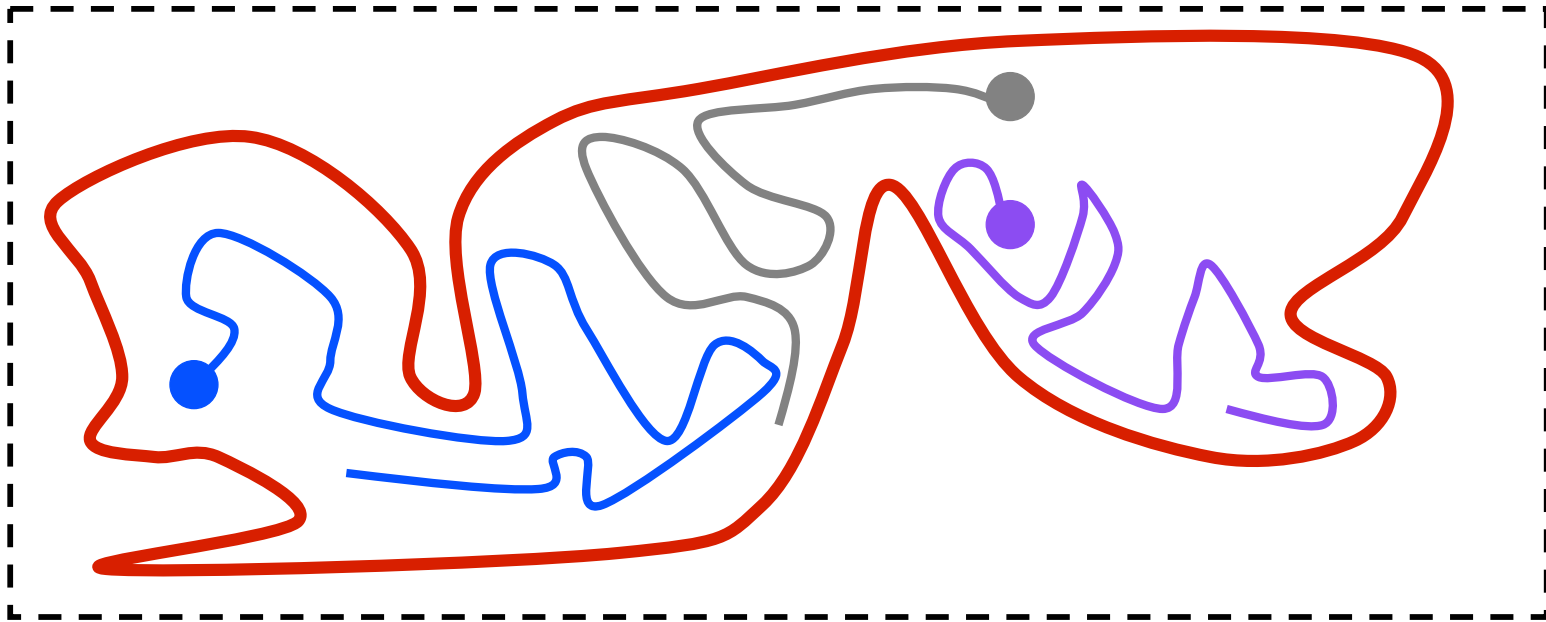


Statistical Thermodynamics

- › Enables computation of (macroscopic) properties
 - in terms of INTERACTIONS
 - between (microscopic) building blocks for systems containing many building blocks
 - Quantified in energy function with variables
 - positions (**R**) and momenta (**P**)
 - GIVEN CERTAIN BOUNDARY CONDITIONS
 - controlling the system
 - Constant number of particles OR chemical potential
 - Constant volume OR pressure
 - Constant energy OR temperature

Phase Space

- › Describes possible states of the system
 - Dimension $2 \times N_{\text{dof}}$ or $2 \times N_{\text{dof}}$ (\mathbf{R}, \mathbf{P}) or $N_{\text{dof}} \times N_{\text{dof}}$ (\mathbf{R})
 - A single system, in time, follows an allowed path through phase space visiting allowed states (\mathbf{R}, \mathbf{P})



Partition Function $Q = \sum w_i$

- › Describes appropriate probabilities of the possibilities of the system being realized
 - determined by boundary conditions
 - KNOWLEDGE OF PROBABILITIES
 - Enables calculation of properties

$$A = \langle A \rangle = \frac{\sum w_i A_i}{\sum w_i} = \frac{1}{Q} \sum w_i A_i$$

Partition Function of Macroscopic Systems

- › Depends on potential (internal) energy V
 - By building ensembles of coupled replicas of the system of interest it can be shown that

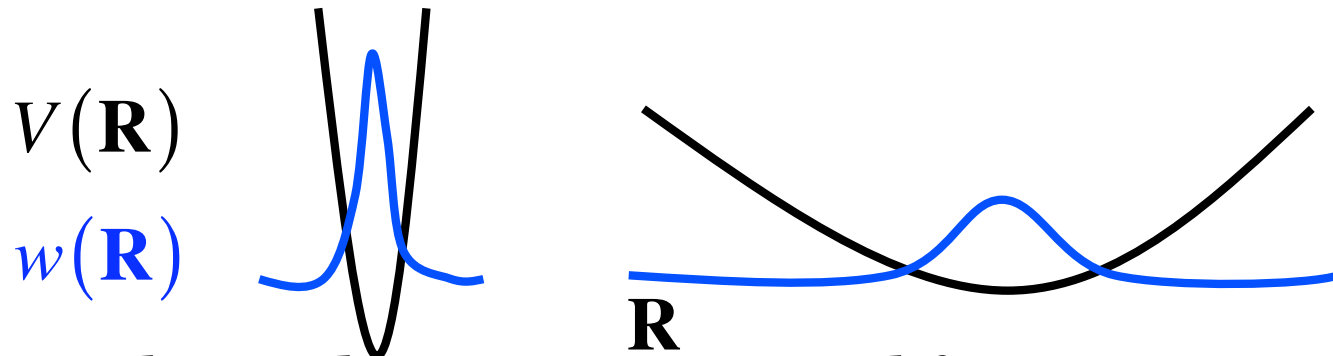
$$w(\mathbf{R}) = f(V(\mathbf{R}))$$

- Potential is of central interest!
- For constant Temperature

$$w(\mathbf{R}) = \frac{\exp(-V(\mathbf{R})/k_B T)}{Q}; \quad Q = \int d\mathbf{R} \exp(-V(\mathbf{R})/k_B T)$$

Potential and Distribution

- › Close link between potential and distribution



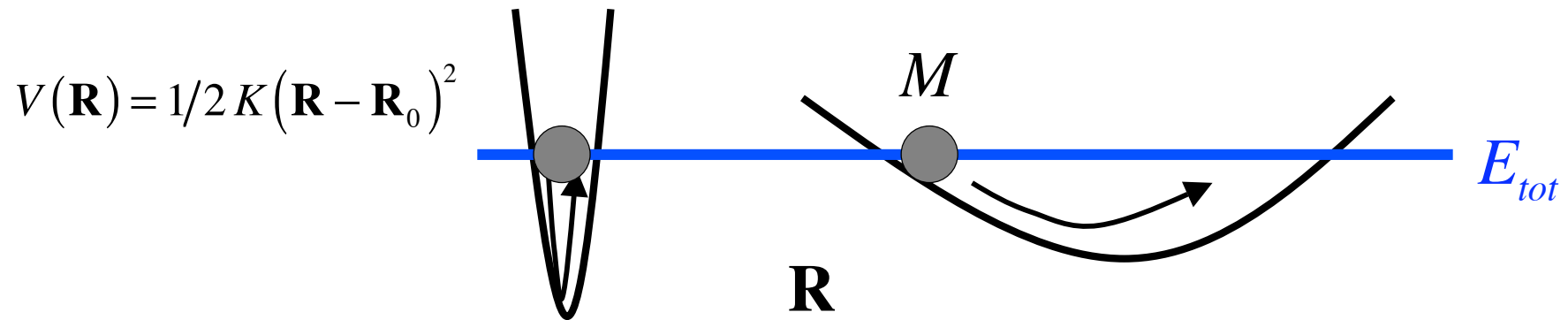
- Can be used to extract potential from structural data

$$w(\mathbf{R}) = \frac{\exp(-V(\mathbf{R})/k_B T)}{Q} \Leftrightarrow V(\mathbf{R}) = -k_B T \ln[w(\mathbf{R})Q]$$

- Inverse Boltzmann techniques
- In MARTINI for fine-tuning bonded interactions

Potential and Time

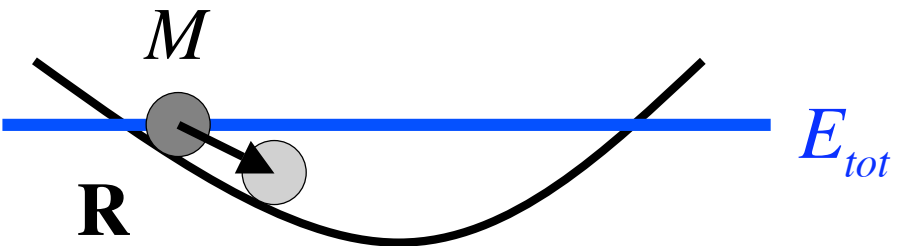
- › Close link between potential and frequency



- Time required to traverse potential depends on shape of the potential and the mass
- For harmonic oscillator, the characteristic frequency
$$\omega = (K/M)^{1/2}$$
- Can be used to gauge time

Equations of Motion

- › Distribution and characteristic frequency obtained by following the system in time

$$M\ddot{\mathbf{R}} = \mathbf{F} = -\partial V(\mathbf{R})/\partial \mathbf{R}$$


The diagram illustrates a particle of mass M moving in a potential well. A horizontal blue line represents the total energy E_{tot} . The particle is shown at a position \mathbf{R} on the potential curve, with an arrow indicating its direction of motion.

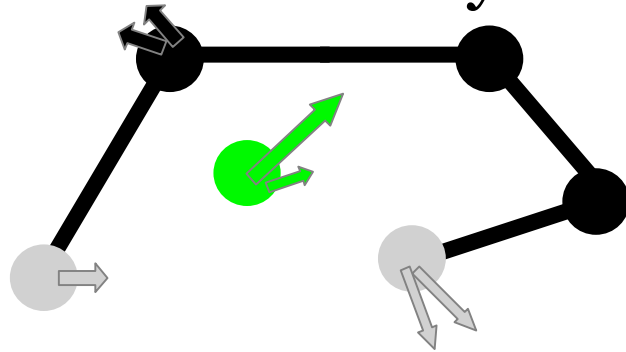
$$\mathbf{R}(\Delta t) \approx \mathbf{R}(0) + \dot{\mathbf{R}}(0)\Delta t + \mathbf{F}(0)(\Delta t)^2/2M$$

- Propagate in time from a starting point according to Newtonian equations of motion
- Conservation of energy is important because otherwise we get wrong distribution and frequency

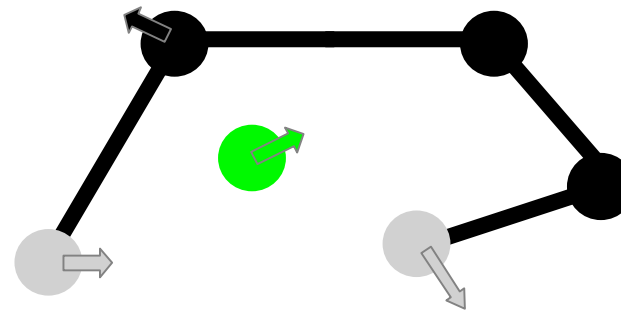
$$\Delta \mathbf{R} = \dot{\mathbf{R}} \Delta t + \mathbf{F} (\Delta t)^2 / 2M$$

Basic MD Algorithm

- > Numerically integrate Newton's Law: $\mathbf{F} = m \cdot \mathbf{a}$

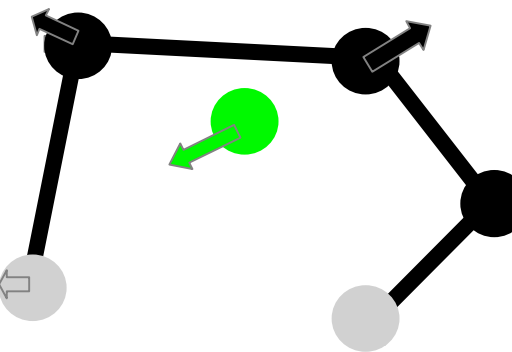
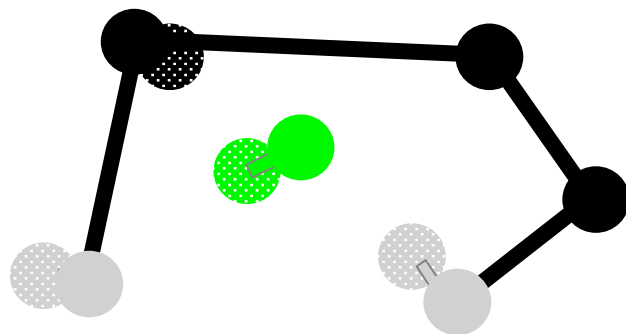


- > Forces & velocities



- > displacements

> t=0



- > new displacements, etc...

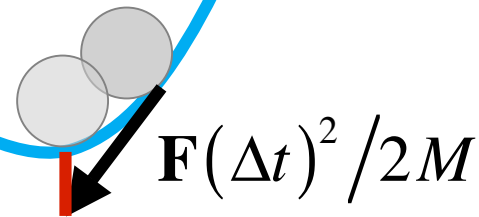
> t=1

Conservation of Energy

- › In numerical approaches, time-step should be such that PES* is followed

$$M\ddot{\mathbf{R}} = \mathbf{F} = -\partial V(\mathbf{R})/\partial \mathbf{R}$$

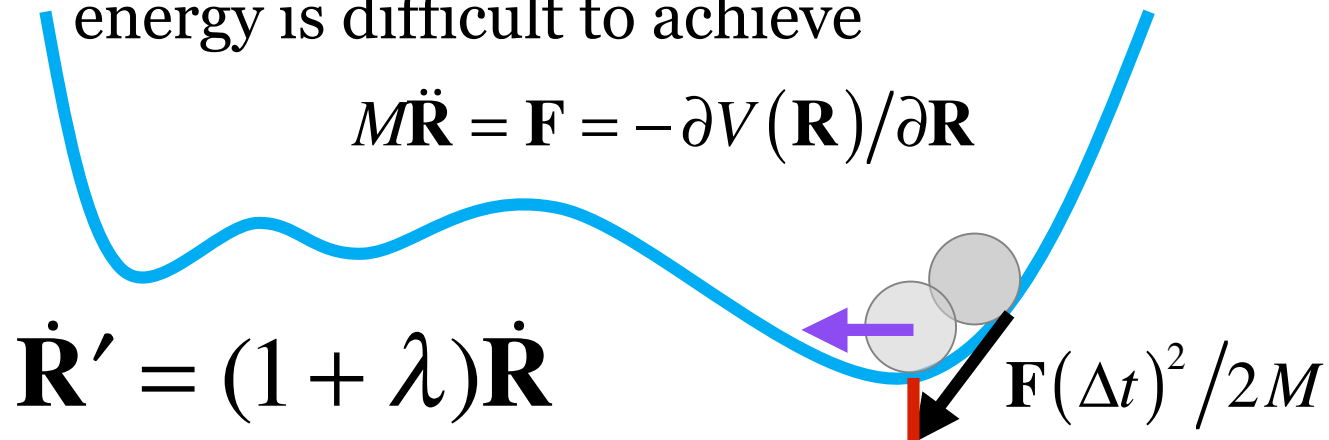
- › We should end up at the energy we expect or energy is lost from/added to the system



- › *Potential Energy Surface: multidimensional surface that gives the energy as a function of all co-ordinates

Conservation of Energy

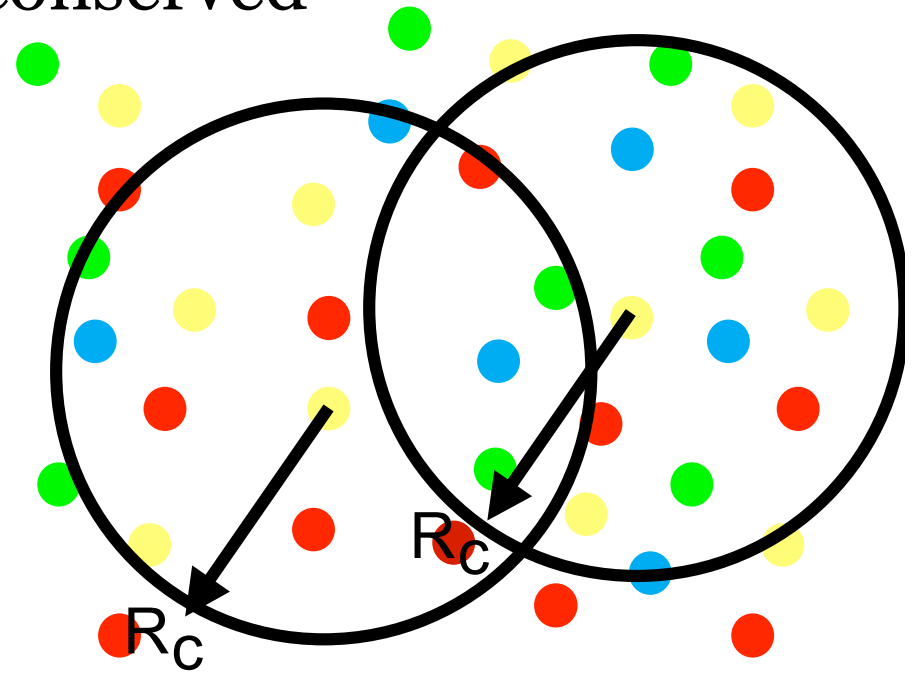
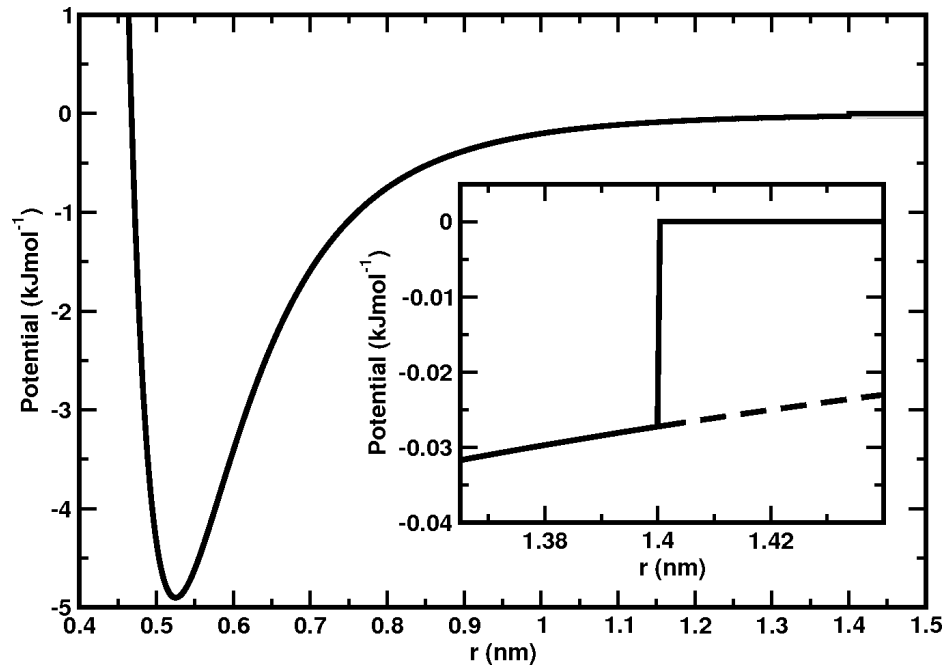
- › In practice, even with small time-steps, conservation of energy is difficult to achieve



- › Couple to heat bath to dissipate or gain heat
- › Popular methods are velocity rescaling (e.g. Berendsen) or extended ensembles (e.g. Nose-Hoover)
- › Strictly speaking, we are not doing Newtonian mechanics
- › Time step and bath coupling strengths are part of the parameter set!

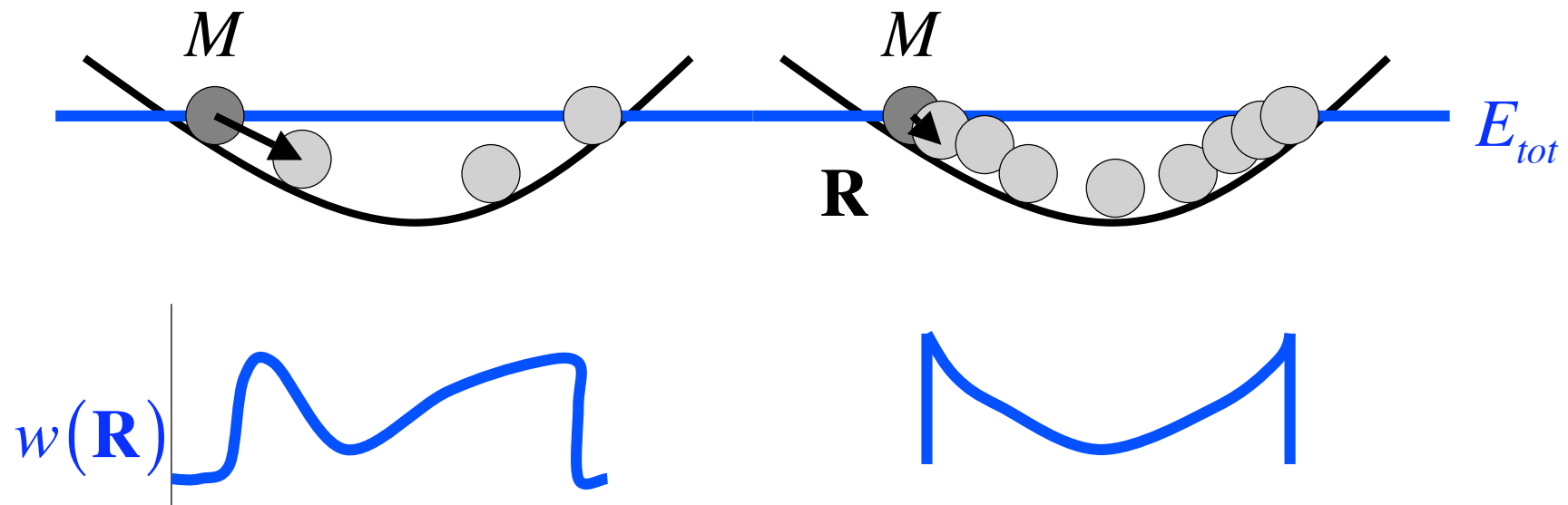
Cut-off noise

- › Particles start/stop interacting at certain distance: energy is not conserved



Sampling

- › Accuracy of distribution depends on sampling

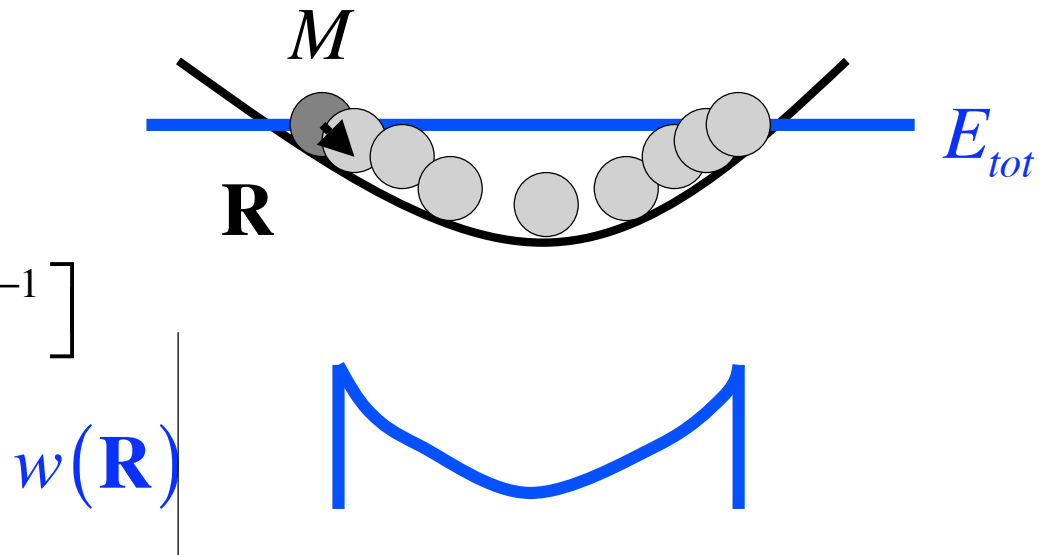


- Smaller time step leads to better sampling
- Rule of thumb: 5-10 points on a harmonic potential characterize it reasonably well

Appropriate time step for sampling

- › At atomistic level, time step for fastest vibrations

$$\begin{aligned} 1 / \Delta t_{vib} &= \nu_{IR} = c \tilde{\nu}_{IR} \approx \\ &3 * 10^8 * 3,000 [ms^{-1} cm^{-1}] \\ &\approx 1 * 10^{14} Hz \end{aligned}$$



- Time step ~ 1 fs

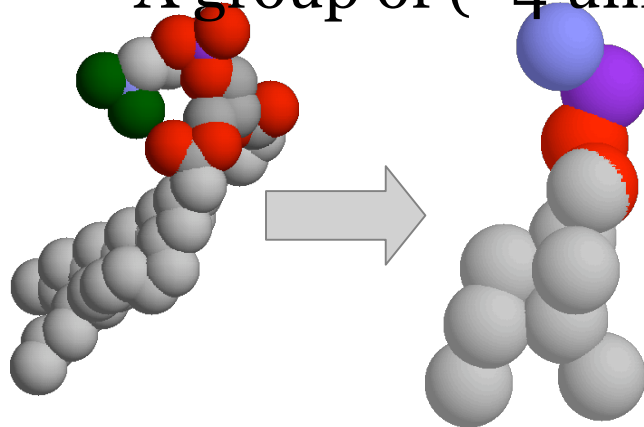
Aim and Scope of Simulation

- › For larger systems, PES is very complicated
 - Aim is to obtain a REPRESENTATIVE SAMPLE of the REALIZATIONS of the system
 - Structural and/or dynamic

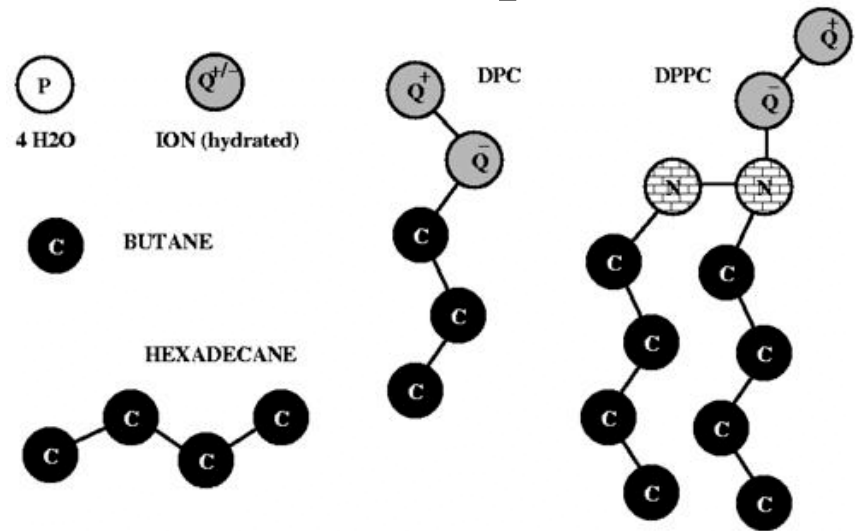
- › Phase space (\mathbf{R} , \mathbf{P}) to be sampled depends on
 - Degrees of freedom
 - WE DEFINE THE SYSTEM OF INTEREST AND CHOOSE A MODEL TO YIELD INFORMATION ON A CERTAIN LENGTH AND TIME SCALE
 - Realism depends on judiciousness of our choice

Molecular Models Considered

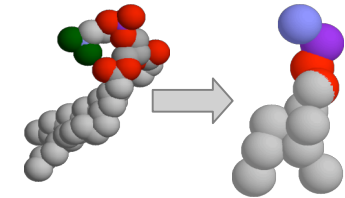
- > (United) Atoms
 - Each atom (or CH_n group) is treated as a particle
- > Coarse grained superatoms (MARTINI)
 - A group of (~4 united atoms) is treated as a particle



- > DPPC molecule (a lipid) in UA and CG (MARTINI) representations



COARSE GRAINING BENEFITS



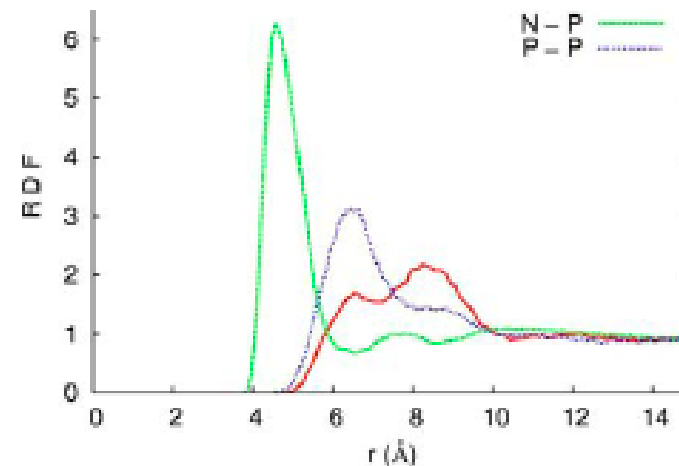
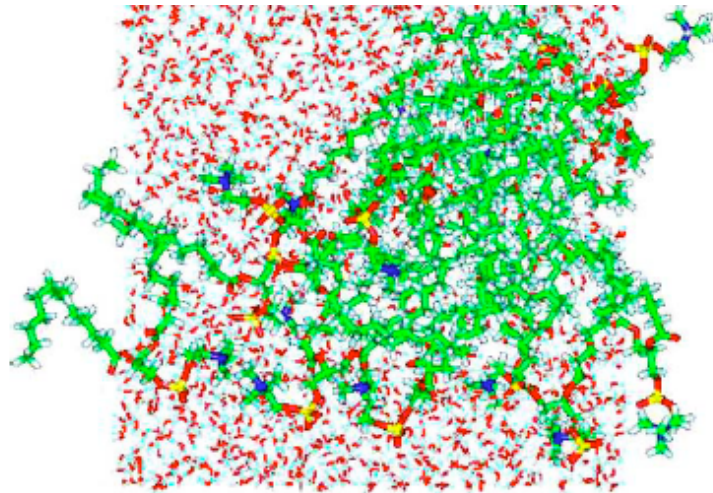
- › Reduced Complexity
 - Physics: not all detail is relevant for our question
- › Efficiency: increase length and time scales
 - Space: reduced density reduces number of interactions
 - 4 (number) x 4 (pairs)
 - 4 neighbor searching
 - Time: smoother energy landscape, increased time step
 - in algorithm: 10-20
 - effective time: 4
 - Total MARTINI: 2.5-5 10^3 speed-up compared to UA

SUMMARY

- › Molecular Dynamics
 - Generate representative ensemble of structures connected through representative dynamics
- › Molecular Model
 - Should reflect research question
 - Potentials determine distributions and dynamics
 - Complexity increases with level of detail
 - Rougher potential energy surface
 - Larger phase space
 - Slower sampling

Appropriate Interaction Potentials

- › Study structure through distribution functions
 - Find that in liquid state we need fairly short-ranged potentials



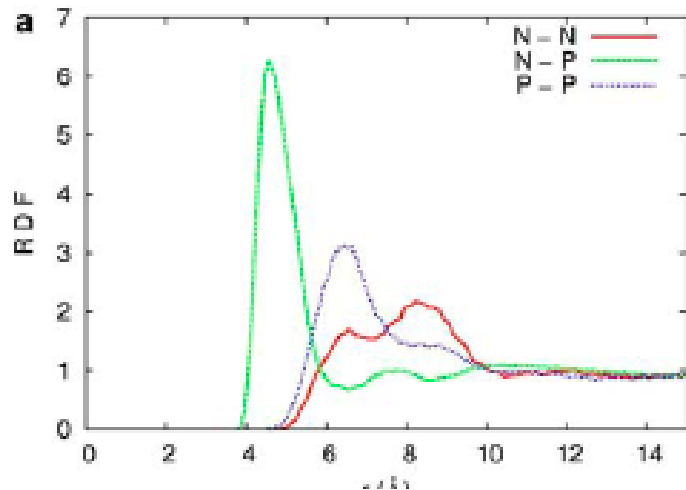
- Best correspondence by numerical potentials obtained via Boltzmann Inversion techniques
- For convenience, simple analytical forms are often used (variations of the Lennard-Jones potential)

Pictures: A. Lyubartsev *Eur Biophys. J.* **35**, 53-61 (2005)

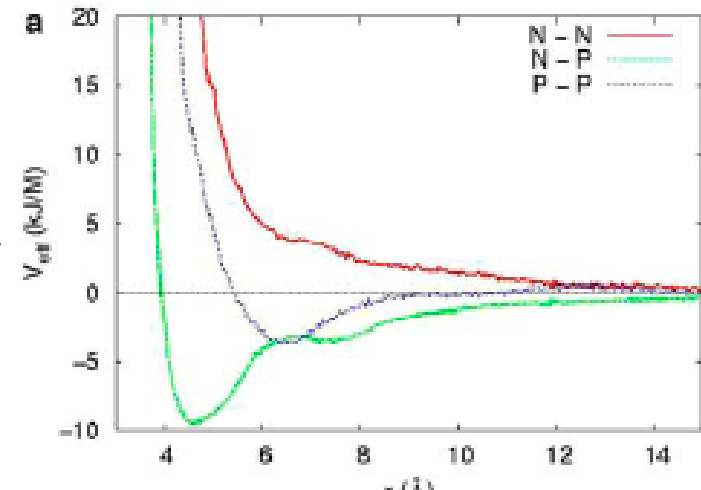
$$V(\mathbf{R}) = -k_B T \ln[w(\mathbf{R})Q]$$

Appropriate Interaction Potentials

- › Study structure through distribution functions
 - Find that in liquid state we need fairly short-ranged potentials



Boltzmann
inversion

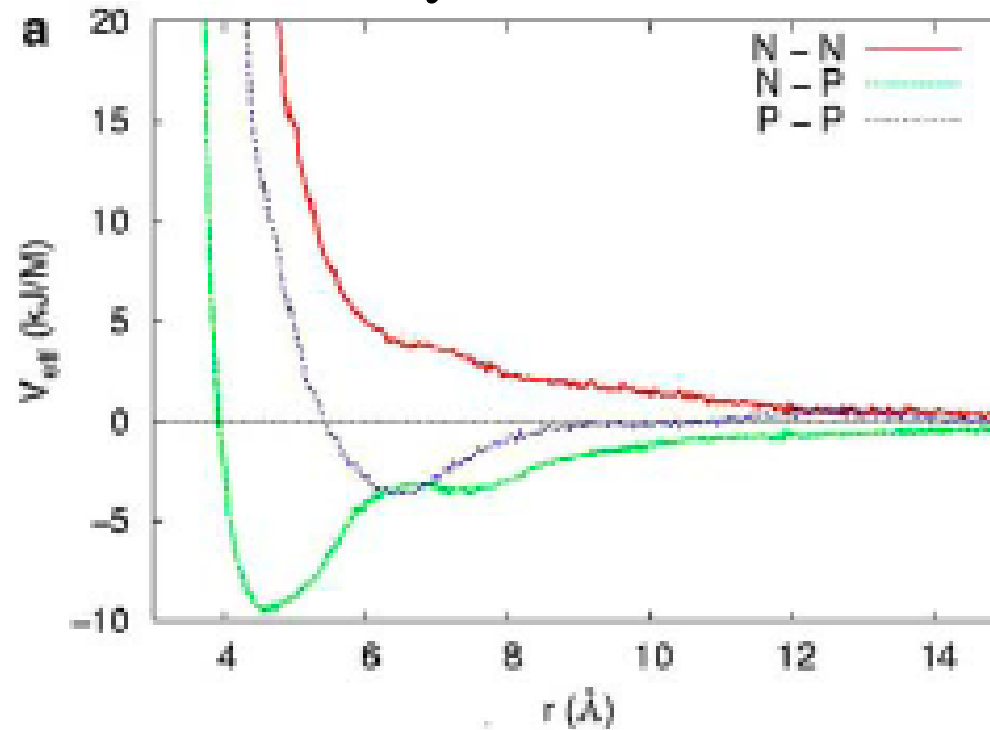


- No clear structure beyond ~ 1.0 nm
- Similar findings for many liquid-state systems from experiment

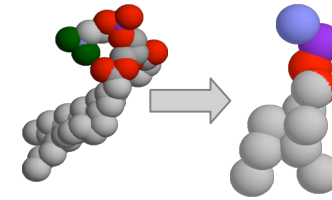
Pictures: A. Lyubartsev *Eur Biophys. J.* **35**, 53-61 (2005)

Appropriate Interaction Potentials

- › Approximate potentials
 - Compare IBP to common analytical potentials
 - Approximations may not be too bad



Coarse-graining Philosophies



› HIERARCHICAL MODELING

- Part of a wider scheme
 - Interactions at less detailed level are the result of the collective interactions at more detailed level
 - General method applicable to any system (like an algorithm)

› EFFICIENT MODEL AT CERTAIN SCALE

- Reproduce faithfully certain chosen properties
 - Developed with certain application in mind
 - Nevertheless aiming at wide use through considering the physics of the problems in mind

Hierarchical modeling: pros and cons

› PRO

- **UNBIASED**
 - Physics follows through the hierarchy of models
- **STRAIGHTFORWARD MULTISCALING**
 - Enables reliable combination of levels of modeling
 - Entirely general approach

› CON

- **REQUIRES LARGE WORK LOAD**
 - Need detailed level simulations to derive CG potentials
 - Complicated numerical potentials
- **LIMITED VALIDITY**
 - Strictly valid for one state point only (new system, new potentials)

Multiscaling

- › Compatible models at different length and time scales
 - Combine them in one system
 - Space resolution (e.g. QM-MM)
 - Resolution exchange (cf. Replica exchange)
 - Switch between resolutions
 - Propagate at coarse level to generate an ensemble
 - Zoom in on snapshots and add detail

- › What is compatible?
 - Phase space at less detailed level matches phase space at detailed level mapped to the less detailed level

Semi-empirical modeling: pros and cons

› PRO

- CHEAP

- Parameterize on empirical data available
- Simple analytical potentials

- TRANSFERABLE

- After parameterizing building blocks, many similar systems can be treated straightforwardly

› CON

- BIASED

- Toward parameterized properties

- PROBLEMATIC MULTISCALING

- Different levels do not need to correspond closely
- Extensive validation required

Semi-empirical modeling: partitioning

> MARTINI and GROMOS 53A6 Philosophy

- Partitioning properties are important in biomolecular modeling

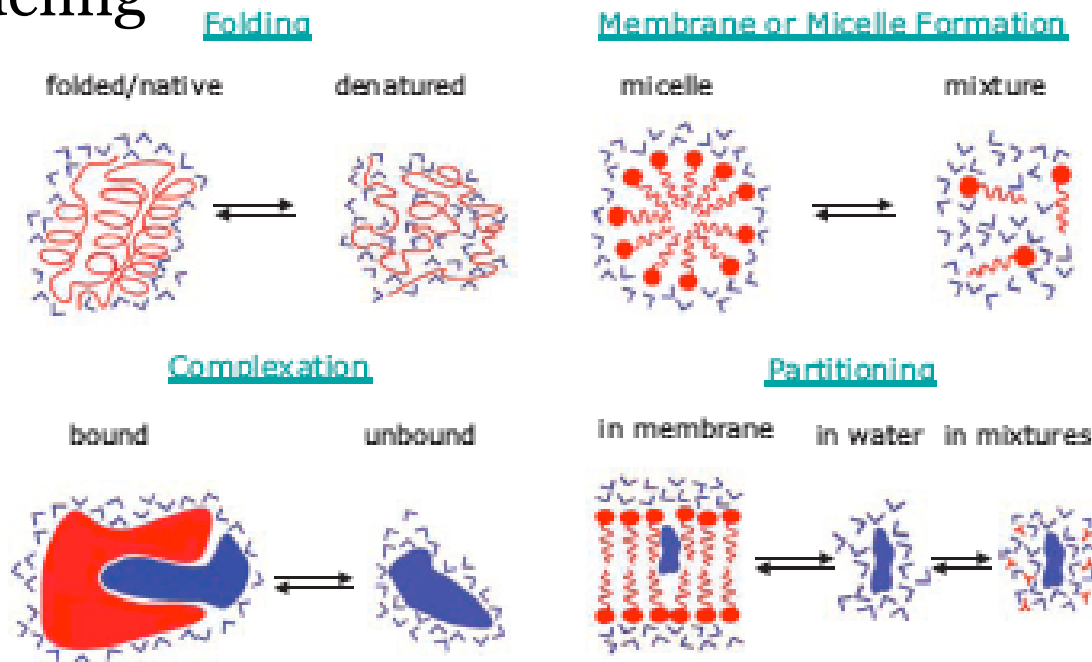


Figure 1. Four biomolecular processes that are governed by thermodynamic equilibria.

Semi-empirical modeling: parameterization

> SIZE

- From density
 - Partial volumes

> COHESION

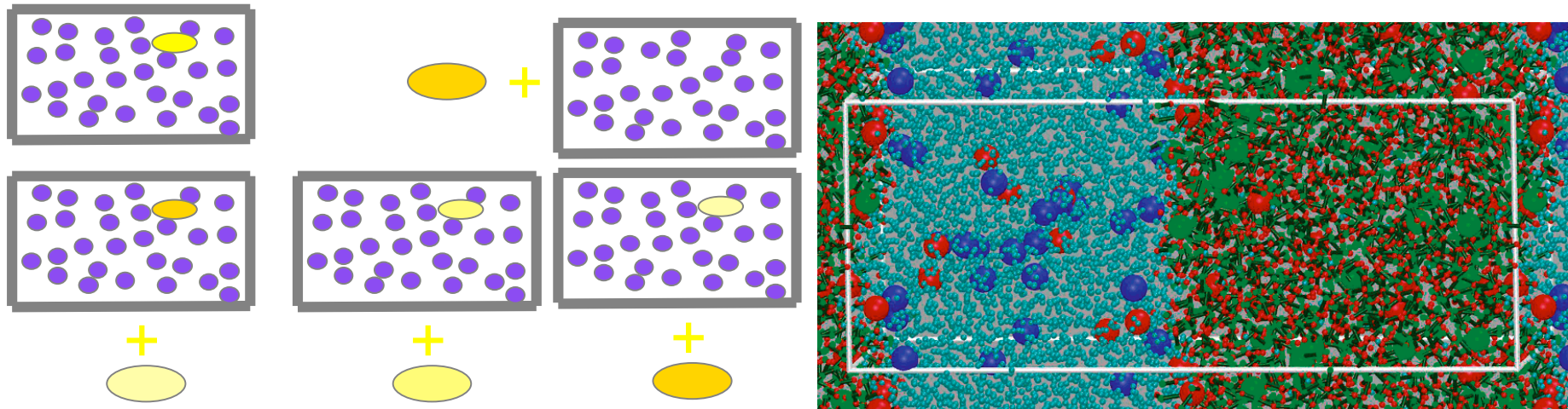
- From heat of vaporization

> PARTITIONING

- From free energy of solvation
 - In water
 - In alkane
 - Everything in everything

Partitioning from Simulations

- › Indirect by restricted sampling
- › Direct by counting



$$\Delta G_{12} = \int_1^2 d\lambda \frac{\partial H(\lambda)}{\partial \lambda} \quad K_{12} = \frac{c_2}{c_1} = \frac{N_2}{N_1} \frac{V_1}{V_2} \hat{=} \frac{p_2}{p_1}; \Delta G_{12} = RT \ln K_{12}$$

- › Thermodynamic integration

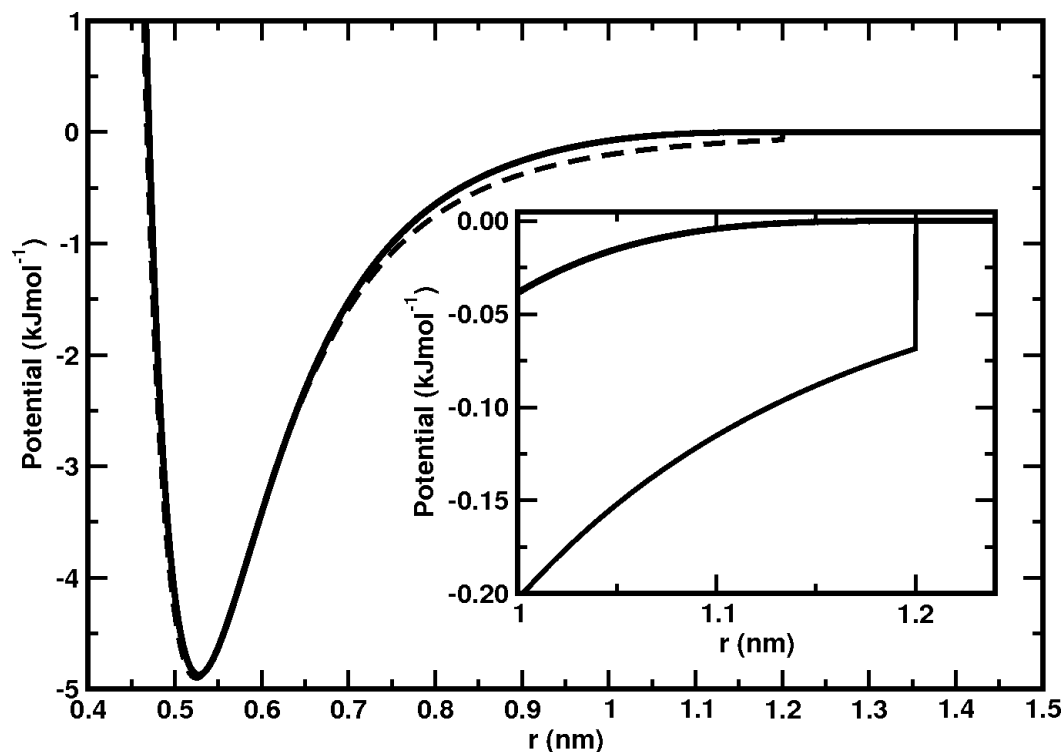
MARTINI modeling: potentials

› NON-BONDED

- Modified Lennard-Jones and Coulomb

Smooth transition at cut-off to reduce cut-off noise can be achieved in different ways.

Potential and Force vanish at cut-off!

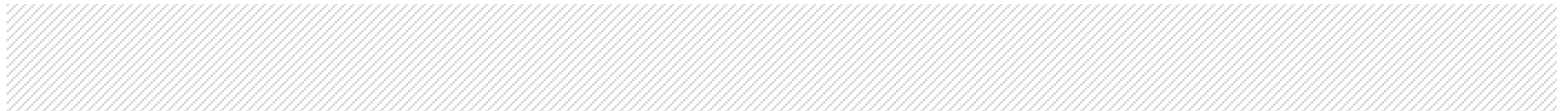


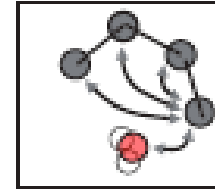
S.J. Marrink et al. *J. Phys. Chem. B.* **108**, 750-760 (2004)
R. Baron et al. *Phys. Chem. Chem. Phys.* **8**, 452-461 (2007)

Appropriate Interaction Potentials

› Bonds and angles

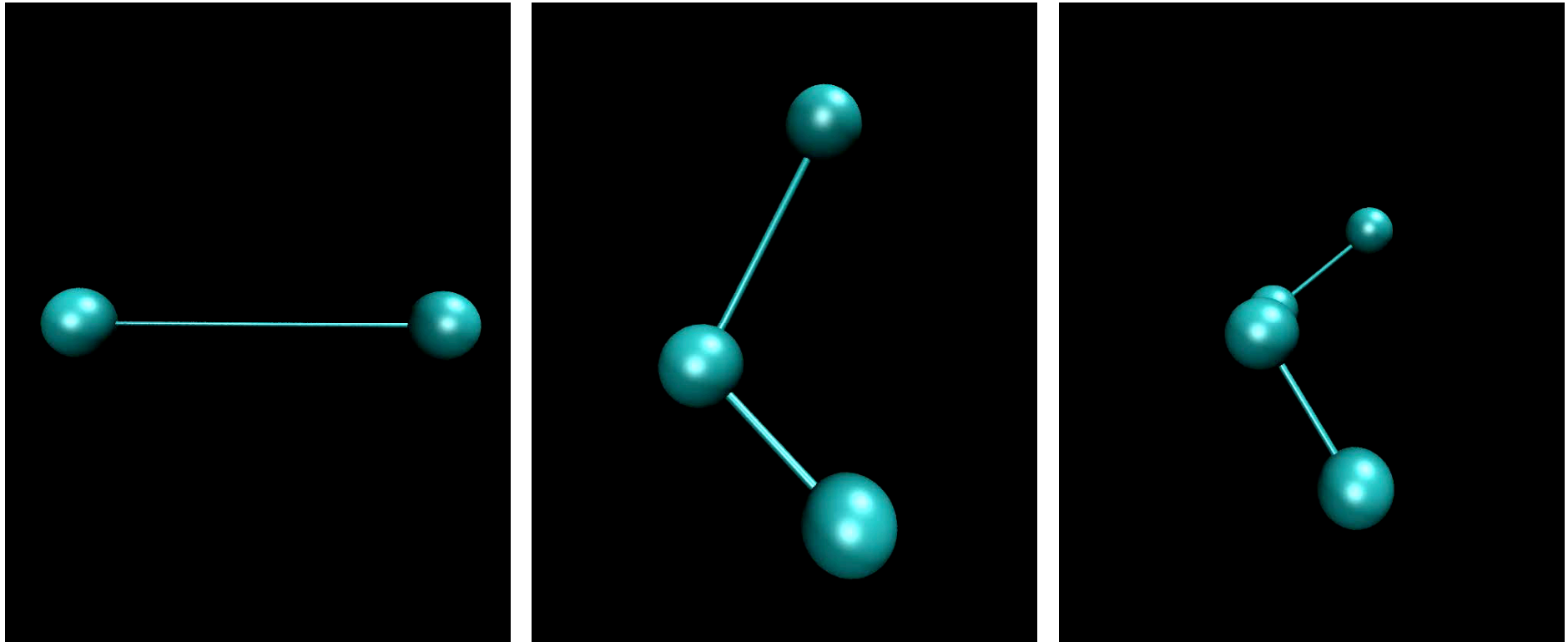
- At UA level, bonds and angles aptly described by harmonic potentials (narrow Gaussian distributions, e.g. from X-ray scattering experiments)
- MARTINI model also uses these, but with much lower force constants





MARTINI Potentials

- › BONDED: Simple harmonic/cosine bonds & angles



- › Torsional potentials are not used in standard MARTINI
 - Can be added, but long harmonic bonds more stable

COARSE GRAINING CAVEATS

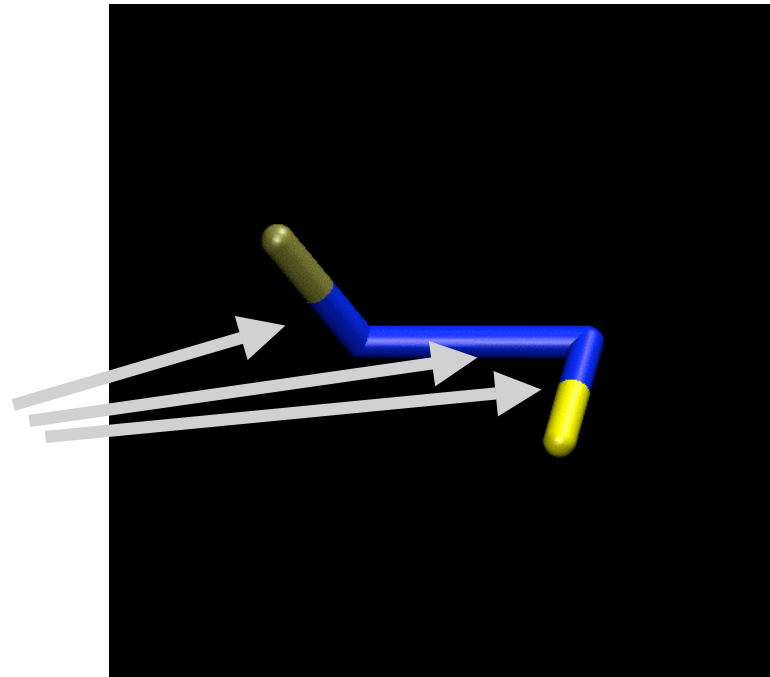
> A SIMPLE EXAMPLE

- 2-to-1 mapping scheme of harmonic oscillator
- Weak coupling through collisions (gas)

> Weak LJ potential

> Small mass

> Harmonic springs

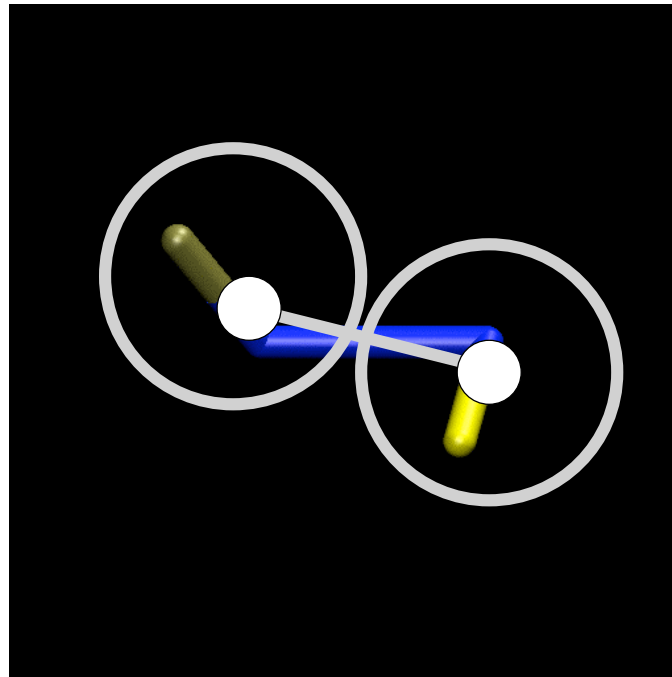


> Large mass

COARSE GRAINING CAVEATS

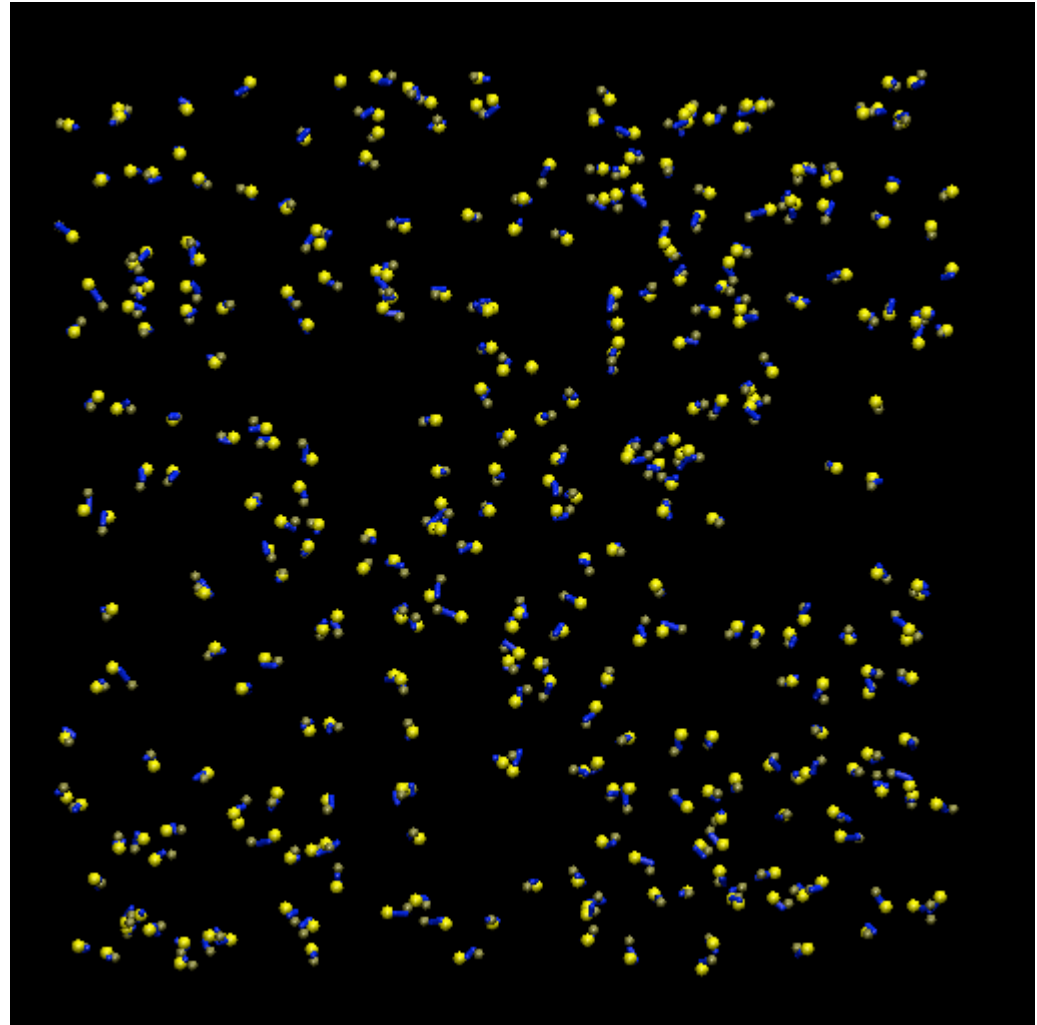
> A SIMPLE EXAMPLE

- 2-to-1 mapping scheme of harmonic oscillator
- Coarse grain on Center of Mass



COARSE GRAINING CAVEATS

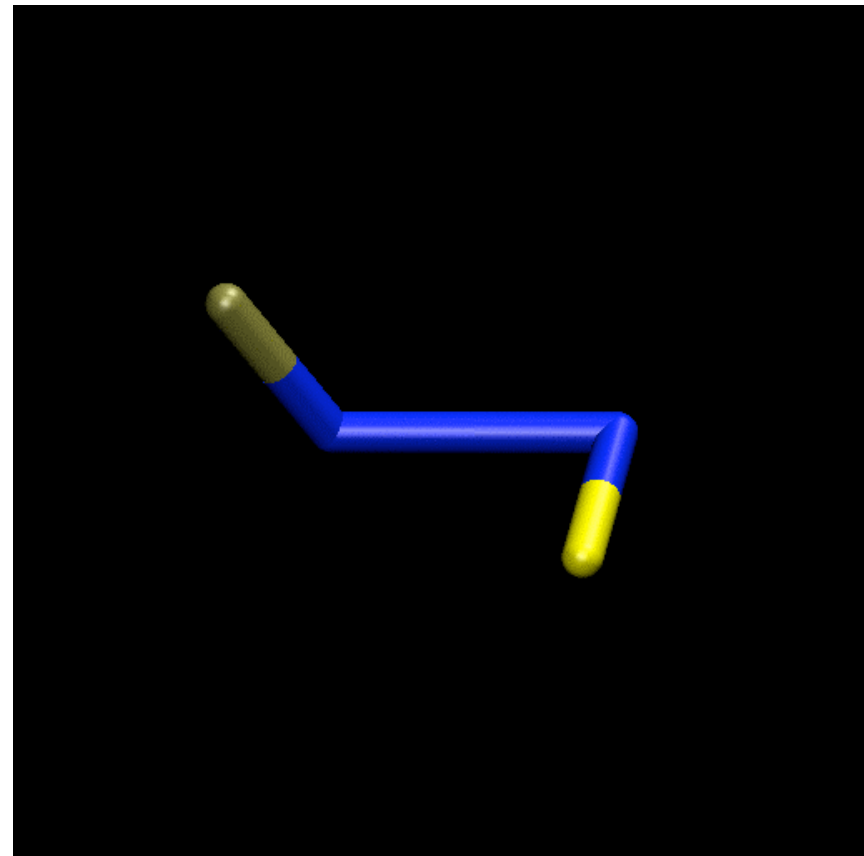
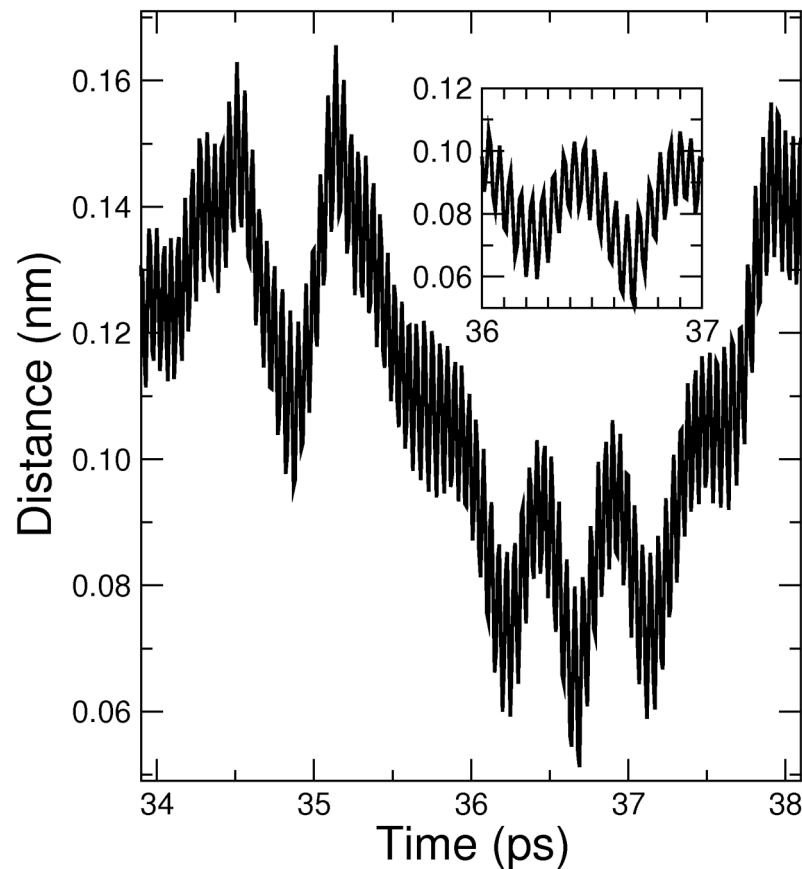
- › A SIMPLE EXAMPLE
 - A look at 100 ps trajectory for the system
 - Energy exchange through collisions



COARSE GRAINING CAVEATS

> A SIMPLE EXAMPLE

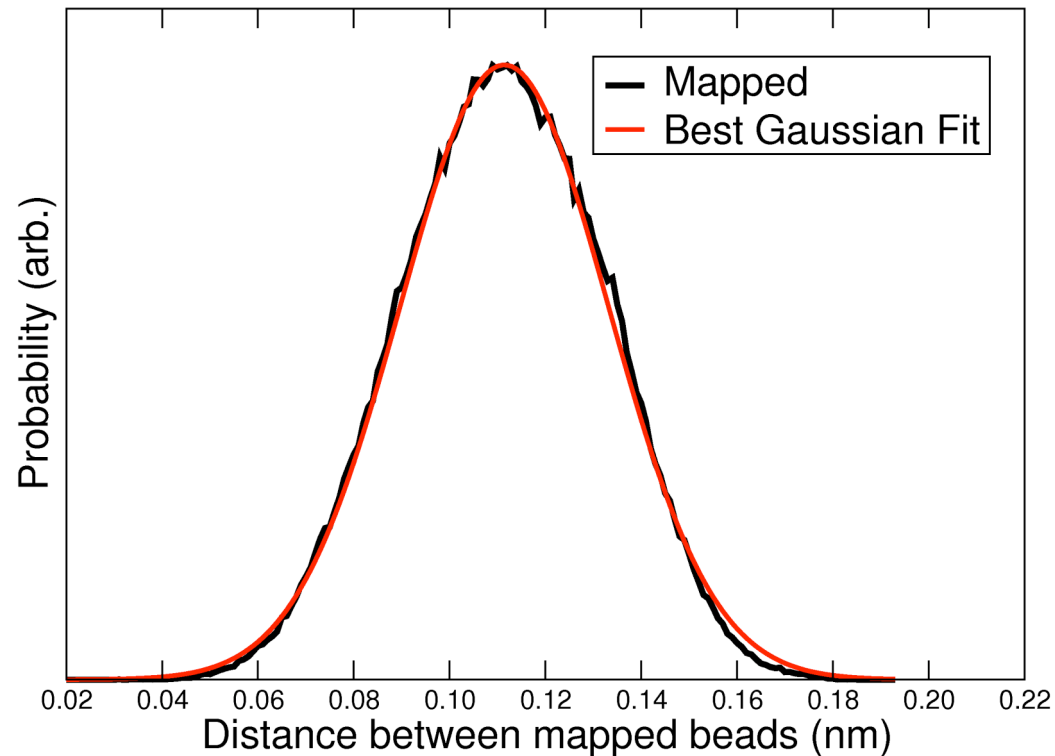
- A look at 10 ps trajectory for one oscillator



COARSE GRAINING CAVEATS

> DETAIL IS LOST

- Distribution reflects effective interaction
 - Can possibly be achieved by a simple potential



$$V(\mathbf{R}) = K(\mathbf{R} - \mathbf{R}_0)^2 / 2$$

DERIVING POTENTIAL

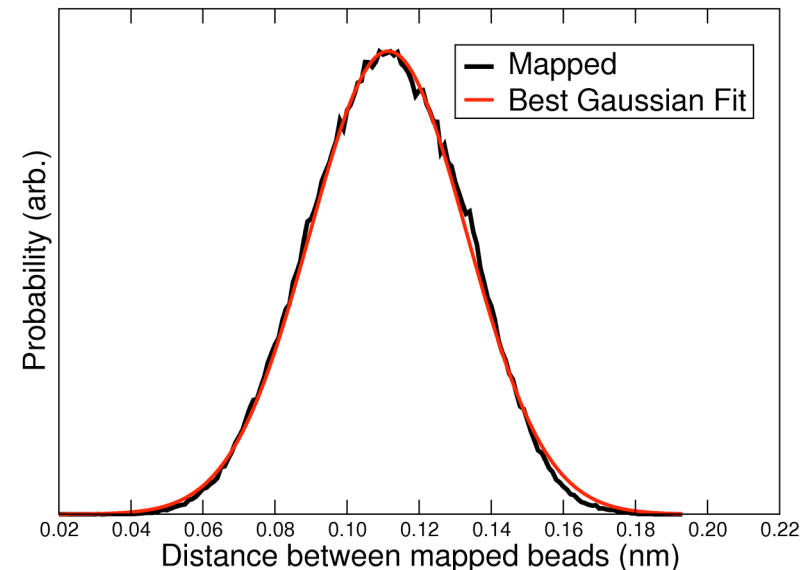
- › Harmonic force constant from Gaussian distribution

$$w(\mathbf{R}) \propto \exp\left[-K(\mathbf{R} - \mathbf{R}_0)^2 / 2k_B T\right]$$

$$\int dx \frac{1}{\sigma\sqrt{\pi}} \exp\left[-(x - x_0)^2 / \sigma^2\right] = 1$$

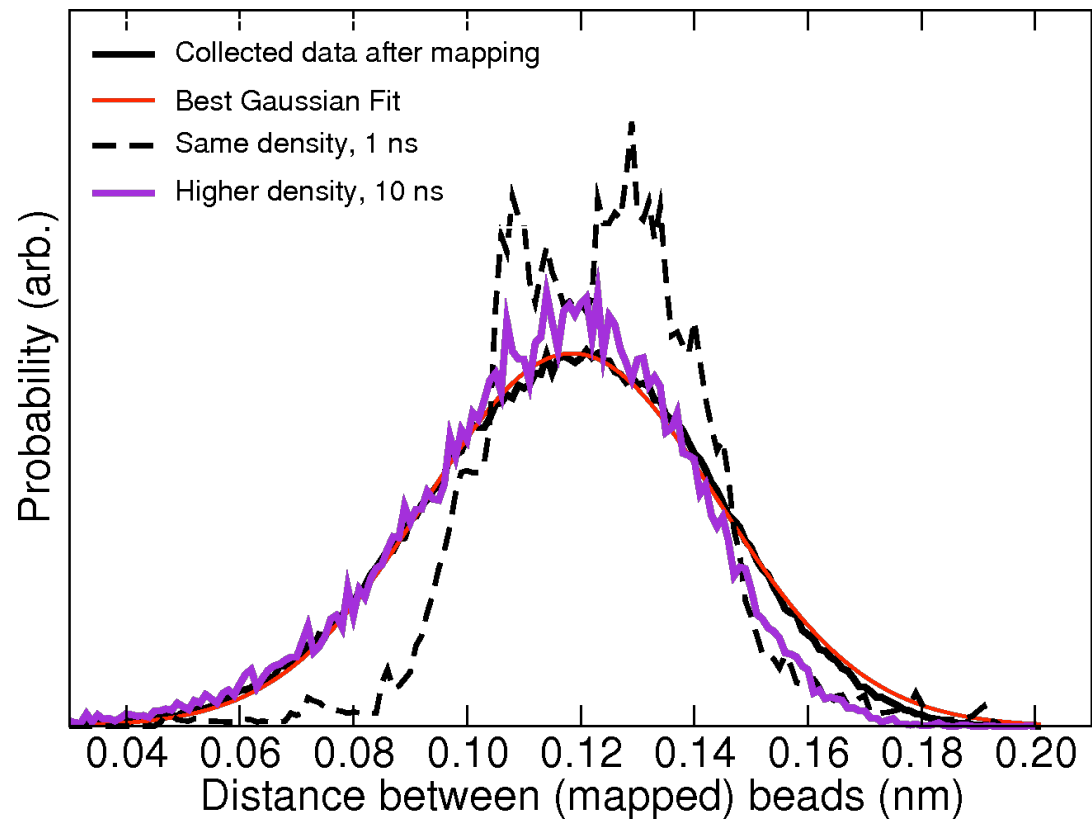
- › Fit for best x_0 and σ :

$$K = 2k_B T / \sigma^2$$



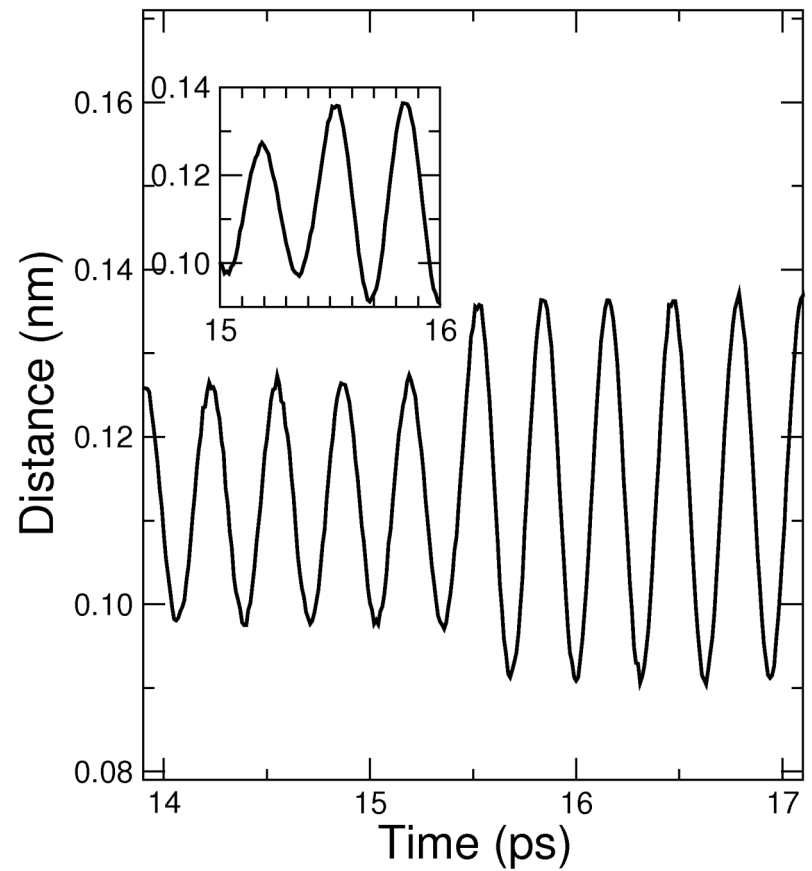
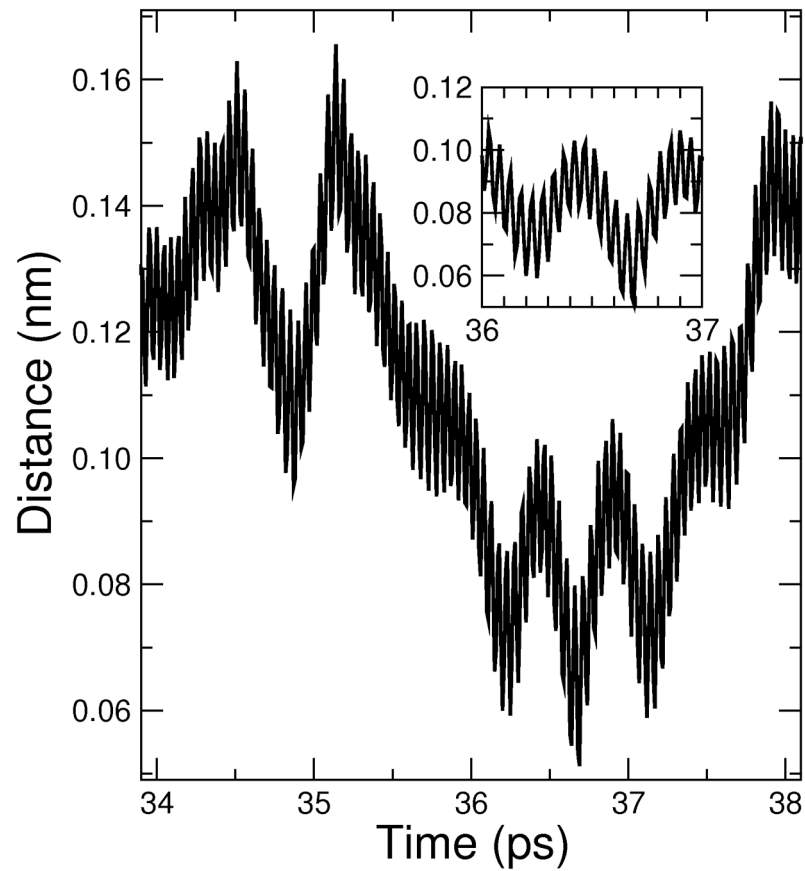
COARSE GRAINING CAVEATS

- › Compare distributions at CG level
 - In this simple(?) example, there are some complications
 - Frequent collisions required to get good statistics



COARSE GRAINING CAVEATS

- › Smoother interaction, smoother motion



COARSE GRAINING CAVEATS

› THE MEANING OF TIME

- Smoother interaction, smoother motion
 - Enables larger time steps
 - Friction is lower, sampling speeds up
 - Barriers are more easily overcome

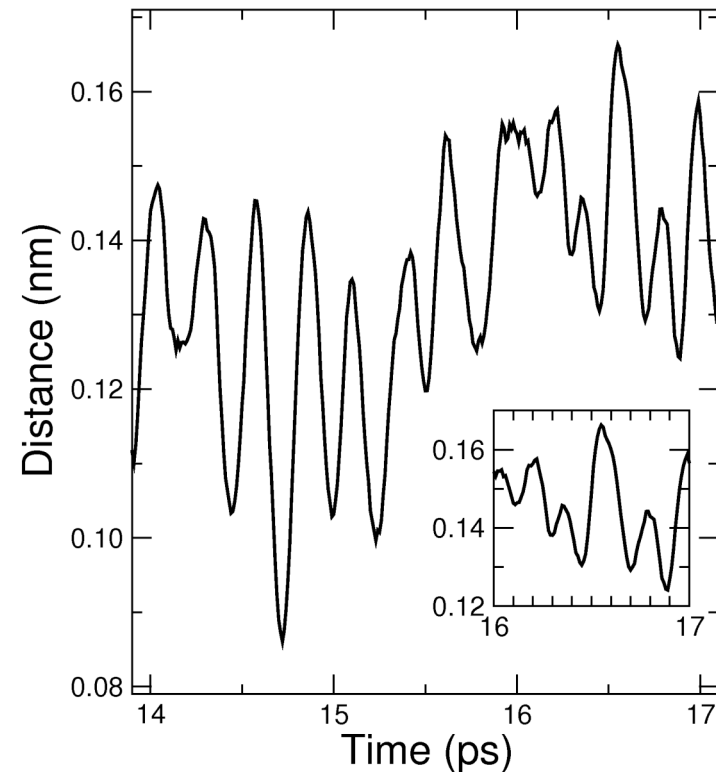
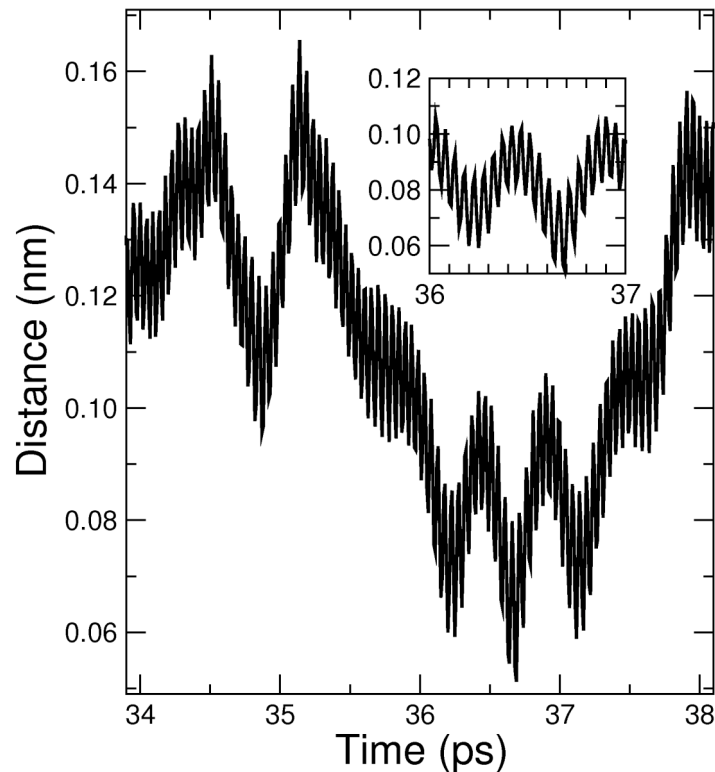
› DETAIL IS LOST

- Physics may be disturbed
- Exchange of energy between modes less efficient

COARSE GRAINING CAVEATS

› THE MEANING OF TIME

- Re-introduce friction through stochastic term
 - Part of benefit is gone, but dynamics are more realistic



COARSE GRAINING CAVEATS

› THE MEANING OF TIME

- Effective time scale can be investigated by comparing to more detailed-level simulations
 - Diffusion
 - Build-up of configurational entropy
- MARTINI model for lipids: effective time is approximately 4 times simulation time

COARSE GRAINING SUMMARY

> SPEEDING UP SEARCH IN PHASE SPACE

- Use an effective interaction from distribution
 - Smooths Potential Energy Surface
 - Reduces number of interactions
 - Increases effective time

> DETAIL IS LOST

- Beneficial: detail may not be required
- Detrimental: some system characteristics are gone