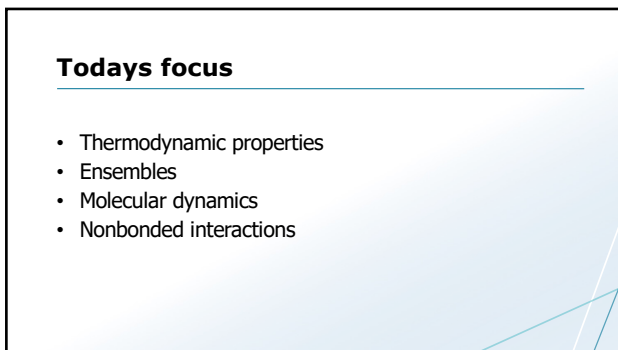


Molecular simulation methods

Bjorn Olav Brandsdal

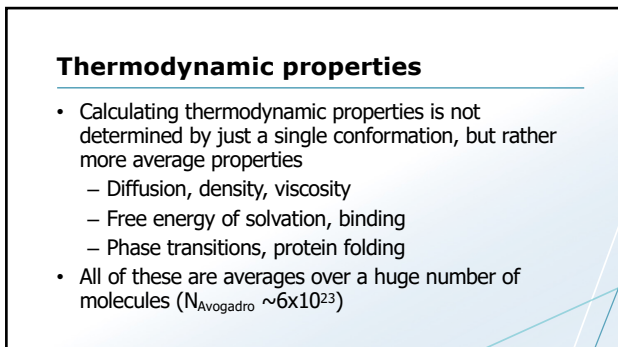
1



Todays focus

- Thermodynamic properties
- Ensembles
- Molecular dynamics
- Nonbonded interactions

2



Thermodynamic properties

- Calculating thermodynamic properties is not determined by just a single conformation, but rather more average properties
 - Diffusion, density, viscosity
 - Free energy of solvation, binding
 - Phase transitions, protein folding
- All of these are averages over a huge number of molecules ($N_{\text{Avogadro}} \sim 6 \times 10^{23}$)

3

How do we calculate them?

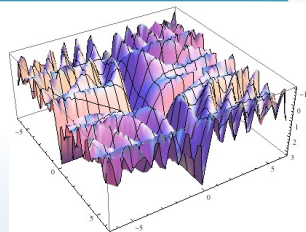
- Average over time?
 - Simulate the motions of the system for "infinity"
- Average over space? What space?
 - All possible combinations of coordinates and velocities, i.e. every state the molecule can be in (phase space)
 - Unlikely to be in high energy states, so weight it with Boltzmann factor:

$$e^{-\frac{\epsilon_i}{kT}}$$

4

Sampling a landscape

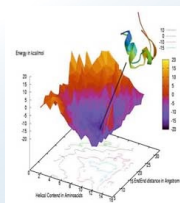
- 3N-dimensional space
- Native structure is the free energy minimum
- Ideally, we would sample all of phase space exhaustively
- In practice, we are limited to the most populated parts



5

Ensemble properties

- Focus on average thermodynamic properties - not just single configurations of the system
- Move around in "representative parts" of the energy landscape - how we do it is less important!
- If know the entire energy landscape we then know everything!



6

Sampling methods

- Monte Carlo
 - No time-dependence
 - Require smart moves
- Newton's equations of motion
 - Deterministic
 - Dynamic events
- Stochastic or Brownian dynamics

7

Molecular dynamics

Tries to calculate paths or trajectories of the individual particles

Gives you:

- Time dependence
- Conformational properties of molecules
- How is motion calculated?

8

Molecular dynamics

- Start with a set of coordinates
- Calculate the potential and forces using a force field
- From the force acting on each particle we can calculate the acceleration working on it:

$$a_i = \frac{F_i}{m_i} \quad \text{Newton's second law}$$

9

Molecular dynamics

- From the acceleration we can calculate how the atomic velocities change

$$v_i(t + \Delta t) = v_i(t) + a_i \Delta t$$

- From the velocities we can calculate new atomic positions

$$x_i(t + \Delta t) = x_i(t) + v_i \Delta t$$

- Repeat force calculation with new coordinates

10

More formal definition

- Solve Newton's equations of motion:

$$m_i \frac{\partial^2 r_i}{\partial t^2} = F_i \quad i = 1 \dots N$$

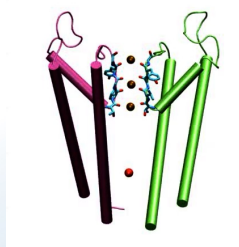
$$F_i = - \frac{\partial V(r)}{\partial r_i}$$

- Time step must be small (typically femtosecond)
- Forces depend on all particles in the system - expensive to calculate
- Generates a system trajectory over time

11

Trajectory

- K⁺ permeation through the Kv1.2 channel



12

Time step (Δt)

- The choice of time step is crucial; too short and phase space is sampled inefficiently, too long and the energy will fluctuate widely
- Instabilities are caused by motion of atoms being extrapolated into regions where the potential energy is prohibitively high (overlap)
- For flexible molecules, the time step should be an order of magnitude less than the fastest motions in the system (C-H vibrations 10 fs $\Rightarrow \Delta t=1$ fs)

13

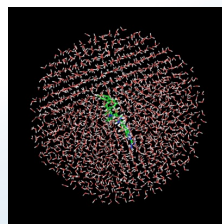
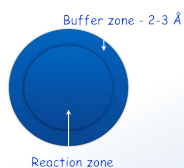
Boundary conditions

- Vacuum - no solvent
- Implicit model for solvent (not as good as explicit)
- Droplet: Spherical layer of water around a molecule or a protein, solvent molecules are restrained with random forces
- Periodic boundary conditions: Simulation box repeated in all directions, resembles bulk fluid

14

Example

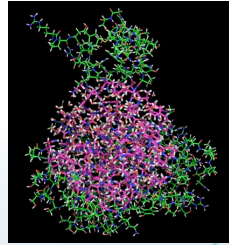
- Spherical - entire molecule immersed in a sphere
- Outermost molecules are restrained



15

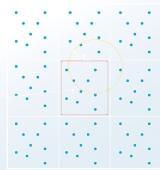
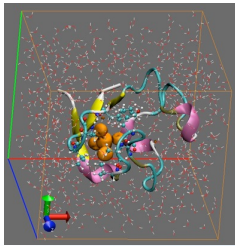
Example

- Spherical: Active site immersed in solvent
- "Green region" frozen in the simulations



16

Example: Periodic conditions



17

Periodic cell shapes

- Cubic or rectangular is by far the most common shape
- Truncated octahedron
- Rhombic dodecahedron



Truncated octahedron



18

Thermodynamic ensembles

- What are you simulating?
 - Energy is conserved for an isolated system, unless sharp cutoffs are used
 - If potential energy goes down, kinetic goes up (temperature increases)
 - Fixed temperature may be preferred?
 - Pressure, volume, number of particles?
- What are the physical "rules" for the simulation?

19

Some common ensembles

- Microcanonical ensemble - **NVE**
 - Most "pure" physical variant
 - Constant number of particles (N)
 - Constant volume (V)
 - Constant energy (E) - temperature varies

Mostly used in physics, can be problematic for proteins, what happens when the structure improves and we find lower parts in the energy landscape?

20

Some common ensembles

- Canonical ensemble - **NVT**
 - Constant number of particles (N)
 - Constant volume (V)
 - Constant temperature (T)
- NPT is common too!!

This ensemble uses special techniques to maintain the temperature e.g. at experimental conditions 298 K

21

Some common ensembles

- Isothermal-isobaric – NPT
 - Constant number of particles (N)
 - Constant pressure (P)
 - Constant temperature (T)
- Chemical reactions are usually carried under these conditions

22

Nonbonded interactions

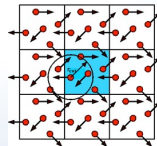
- We get lots of interactions when water is added
 - Most time-consuming part of simulations
- What happens at the boundaries?
 - Interact with atoms in neighbor cells
- Impossible to interact with all periodic copies
 - => Introduce cutoffs and neighbor lists

23

Cutoffs and minimum image

Atoms interact with all other atoms that are within the cutoff distance

Atoms interact only with the closest periodic copy of each atom - minimum image convention

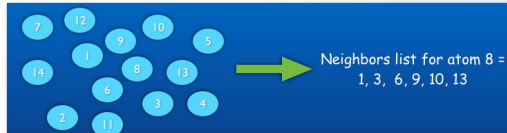


Cutoffs does not lead to dramatically speed reductions - we still need to calculate all the $N(N-1)$ distances

24

Nonbonded neighbor lists

- Atom's neighbors do not change must during 20-50 simulation steps
- Make a list to keep track of the neighbors within the cutoff distance



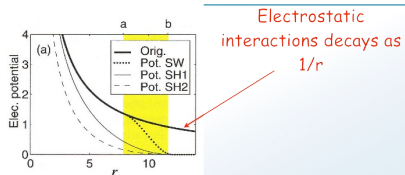
25

Nonbonded neighbor lists

- Also known as Verlet's neighbors list
- Update interval depends on your system
 - e.g. SHAKE/LINCS/RATTLE
 - Time step for integration
- Too often updating => no speed gain
- Too seldom updating => instabilities in the simulations

26

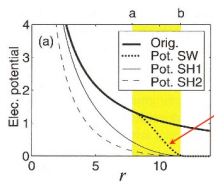
Cutoff interactions



Bad things can happen if we just truncate the these interactions - why?
What will happen to the derivate (force) at the cutoff?

27

Switched interactions



Called a switch-function - works for Lennard-Jones too

28

Switched interactions

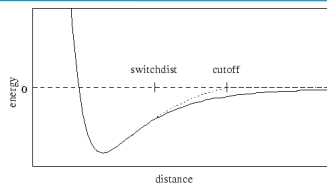
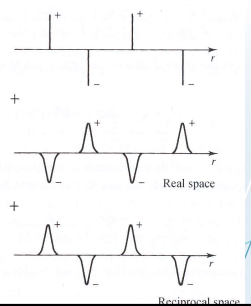


Figure 1: Graph of van der Waals potential with and without the application of the switching function. With the switching function active, the potential is smoothly reduced to 0 at the cutoff distance. Without the switching function, there is a discontinuity where the potential is truncated.

29

Ewald summation

- It is possible to solve the full electrostatic problem
- Add a screening function around charges to make their interactions decay fast
- Solve the rest in reciprocal space by using Fourier transforms



30

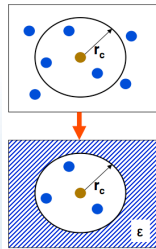
PME - Particle Mesh Ewald

- Particle-Mesh-Ewald
- Modern, fast way to perform Ewald summation
- Charge spreading on 3D grid
- Solve Poisson equation on grid
- Transform back, interpolate to get forces

31

Reaction field methods

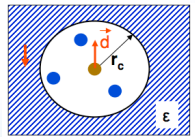
- A sphere is constructed around the atom/molecule with a radius equal to the cutoff
=> interactions within are computed directly
- Outside the cavity the system is treated as a dielectric continuum



32

Reaction field methods

- Any net dipole in the cavity induces a polarization in the dielectric, which in turn polarizes the charge distribution in the cavity
- Treating this mutual polarization in a self consistent manner leads to the reaction field in the cavity:



$$E_i = \frac{2(\epsilon - 1)}{(\epsilon + 1)r_c^3} \sum_j \mu_j$$

33

Problems with reaction field methods

- Permittivity - ad hoc
- Cavity radius
 - adjustable
 - problematic to determine
- Poor energy conservation
 - Discontinuities when a molecule leaves cavity

34

Temperature control

- How do we control the temperature?
- Temperature is just kinetic energy
- If we change velocities we change temperature
- Several alternative, for example:
 - Calculate actual temperature at each step
 - Too low? Scale up velocities
 - Too high? Scale down velocities

35

It's statistical mechanics!!!

- Remember:
Just because you see something in a simulation does NOT mean it is real
- Everything is about statistics
- When you see it 10 times it's significant - a single event is not!
- You should always try to calculate error estimates for predicted properties

36

MD programs

- GROMACS, Amber, Charmm, NAMD, ESPRESSO, Encad, BOSS, LAMMPS...
- Programs are closely tied to force fields (Amber99, Charmm19, OPLS)
- Most are freely available to academia
- We will use Q and different force fields...

37

A typical simulation

- Get structure (typically from PDB)
- Fix missing residue, segments, protonation states, His/Gln/Asn etc.
- Prepare topology and parameters
- Add solvent (normally water)
- Energy minimization
- Heating and equilibration
- Data production
- Analyze the trajectories produced

38

Summary

- **Think first then simulate!!!**
- Time scales & limitations
- Empirical force fields
- Phase space, dynamics, Monte Carlo
- Sampling distributions
- Algorithms, approximations, quality
- Boundary conditions, interactions, etc

39
