This semester

Models – mostly proteins

- from detailed to more abstract models

Some simulation methods

Books

None necessary – for my group and Prof Rarey

- “Molecular Modelling: Principles and Applications” Leach, Andrew (good, not cheap)
- Essentials of Computational Chemistry : Theories and Models / Christopher J. Cramer (ebook, in library) bits of chap 2 & 3

- Folien sind immer in openolat (nicht stine)

Grand Plan

- Models for proteins (mostly)
  - from detailed to less details
- Energies
- Dynamics

How to work with them

- simulations, calculations
  - protein folding, evolution

- What kind of model is appropriate for different problems ?

Themes

- atomistiche Energie-Modelle
- Energie / Freie Energie
- Monte Carlo (MC)
- Molecular Dynamics Simulations (MD)
- Applications and Extensions of MC and MD
- Advanced Monte Carlo
- Water
- Coarse Grain / Mesoscopic / low resolution methods
- Lattice Methods
- Evolution
- Protein Folding

Organization

Im Prinzip

- 2 × Vorlesung
- 1 × Vorlesung
- 1 × Übung

In der Praxis

- 2 × Vorlesung + 2 × Vorlesung or
- 2 × Vorlesung + 2 × Übung

People

- Nils Petersen

Zeitplan

- Mo 10:15
- Di 12:15

Sprache ?

Organization

Klausurtermine in Stine (15 Juli, 25 Sep)

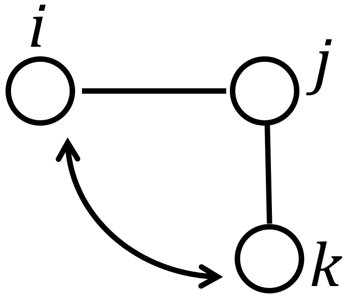
Andrew der Pedant - bei Klausuren

- Energie/potenzielle Energie/freie Energie
- absolute Werte / relative Werte
- größere / kleinere, mehr positiv / mehr negative
- Vektoren, Skalar
- Vorzeichnungen…

Some questions

3 bonded atoms – push atom *k* left

- which atoms will feel a force ?



I have a multiple sequence alignment

- the conserved sites are most important
- is this true ?

Can you compare the free energy of two conformations of

- a drug ?
- a protein ?

To be answered during the semester..

Do I have to memorise all the formula ?

Very few

- Coulombs law
- example form of energy for bonds and angles
- Boltzmann distribution
- definition of entropy
- relationship of free and potential energy

Most other examples will be derived

Why are we here ?

- observables
- distributions
- two ways to get at distributions
  1. from Boltzmann relation (later)
  2. from dynamics simulations

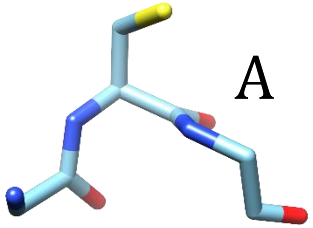
- dynamics simulations
  - need forces
- forces need energies

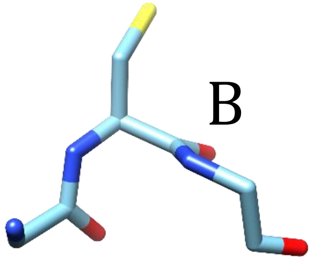
Observables

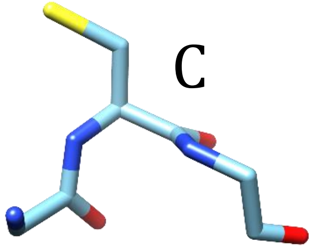
- $\mathcal{A}$ is density, coordinates, something you can measure
- what you see is $\langle\mathcal{A}\rangle$ the average over time and molecules $\langle\mathcal{A}\rangle = \frac{1}{\tau}\int_0^\tau \mathcal{A}_t \, dt$

- simulating gives us a sampling over time / molecules

Can we just average over values ?


A


B

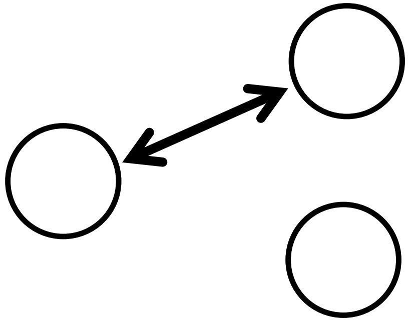

C

- 3 rotamers in a protein
- they are not equally likely
  - A is most common, C is high energy and not seen (no contribution)

How to get the correct averages ?

- simulate the system and do the averaging over every $\mathcal{A}$ in the simulation
- how would you simulate ?

Simulations need forces and energies



Distributions depend on energies

- we need energies

Dynamics simulations need forces

- forces come from energies $F = \frac{-dU}{dr}$

So we spend weeks discussing energies

Atomistic Energy Models

Why do we need models ?

- Previous lectures – need for low energy configurations
  - really needs definition of energy

Can we define energy ?

- for very simple systems yes
- for more complex systems
  - only approximations

Need to know when approximations are small and when bad

- are charges on atoms like fixed charges ?
- are bonds like springs ?

Is energy sufficient ?

Does the world care about potential energy ?

- No. Really cares about free energy

$G = U - TS$ (usually speak of ΔG)

Approach to free energy

- calculate potential energy "U" with a model
- get entropy "S" from some sampling method (often implicit)

Need good models for energy

Definitions

This topic is classical / atomistic

- often referred to as “molecular mechanics”
- quantum effects are not reproduced

Different levels of models

QM	ab initio	very detailed
	semi-empirical	
atomistic		
coarse grain		
continuous		
lattice		
big molecule as single point		very coarse (colloids, diffusion)

Why we like atomistic models

Intuitive

- how do we draw structures ? store coordinates ?

Atoms sometimes correspond to measurable properties

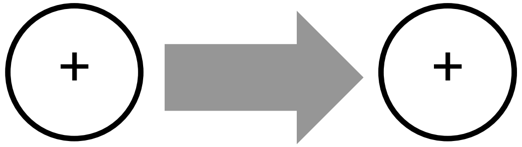
- x-ray crystallography, NMR
- predicting some dynamics
- interactions between proteins, proteins + ligands

What is a force field ?

Set of equations / formulae that tell us about the force acting on a particle

Classic example

- I have charge
- bring another charge near, it feels a force due to the first
- what is the equation telling me about the energy ?



$$U\left(r_{ij}\right) = \left(\frac{1}{4\pi\epsilon_0}\right)\frac{q_i q_j}{r_{ij}}$$

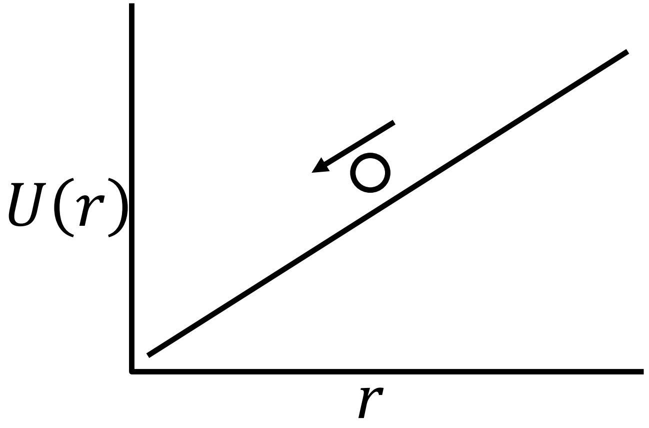
and the force…

Force fields, energies, derivatives

Important rule $\vec{F} = \frac{-dU}{d\vec{r}}$

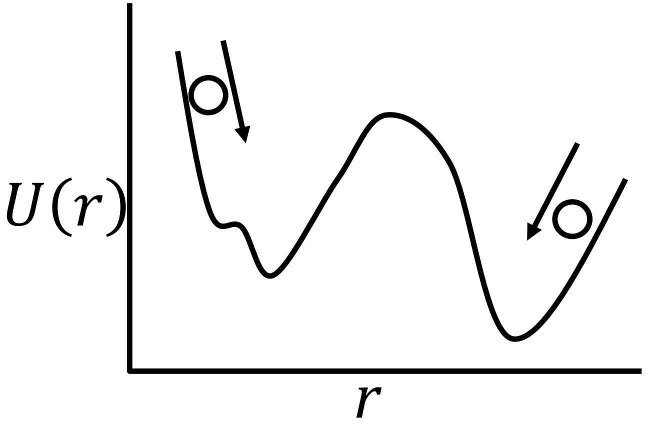
A toy example, one dimension $U(\vec{r}) = k\vec{r}$

- force is $\vec{F} = \frac{-dU}{d\vec{r}} = -k$



How do we want force ?

- in x, y, z terms $\vec{r}$



really
$$F_x = \frac{-\partial U}{\partial x}$$
$$F_y = \frac{-\partial U}{\partial y}$$
$$F_z = \frac{-\partial U}{\partial z}$$

A simple force

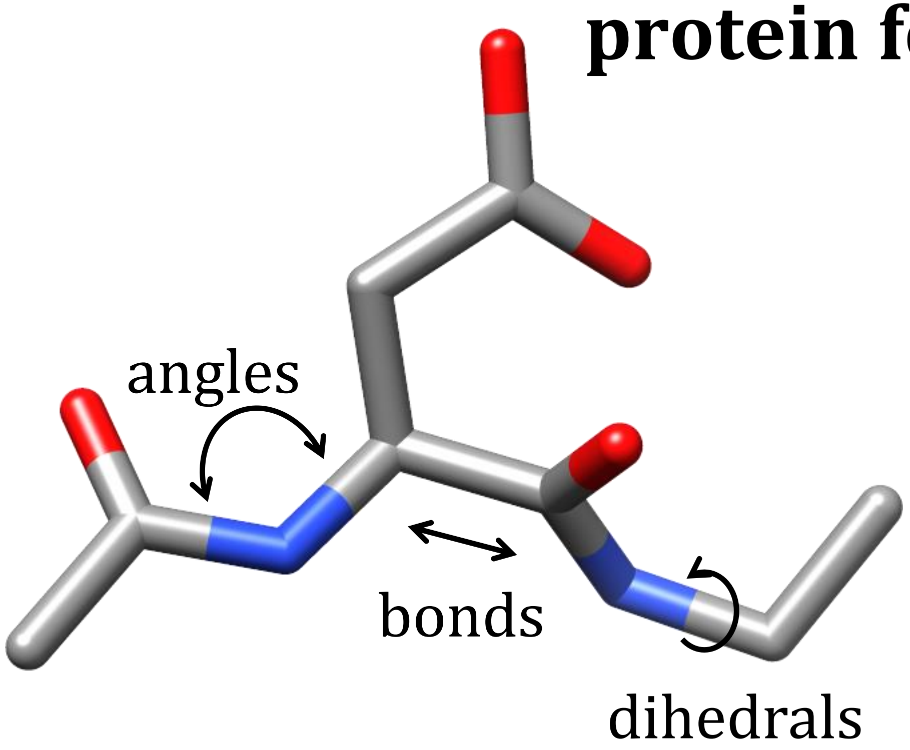
With an energy $U(r_{ij}) = \frac{1}{4\pi\epsilon_0}\frac{q_i q_j}{r_{ij}}$

force is $\vec{F} = \frac{-dU}{d\vec{r}} = \frac{1}{4\pi\epsilon_0}\frac{q_i q_j}{r_{ij}^2}\ \hat{r}_{ij}$

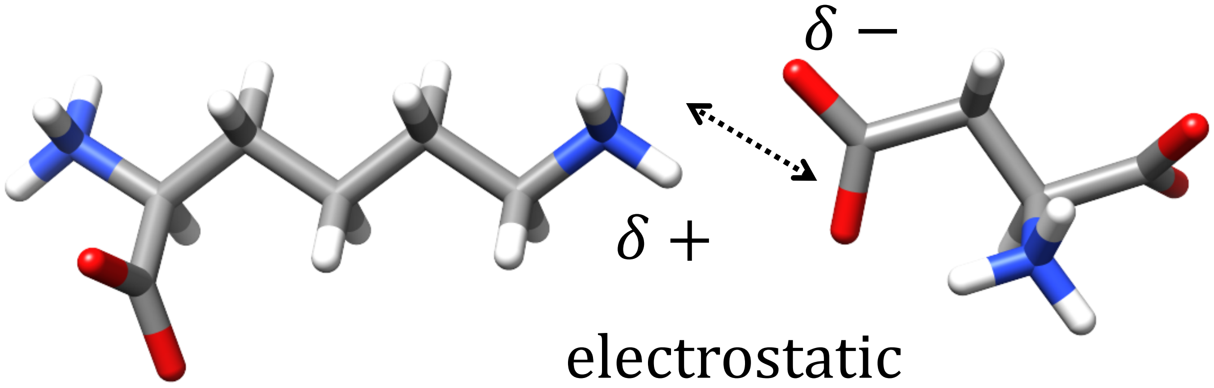
Rule

- if the derivative of an energy is non-zero
  - there is a force (remember for exams)

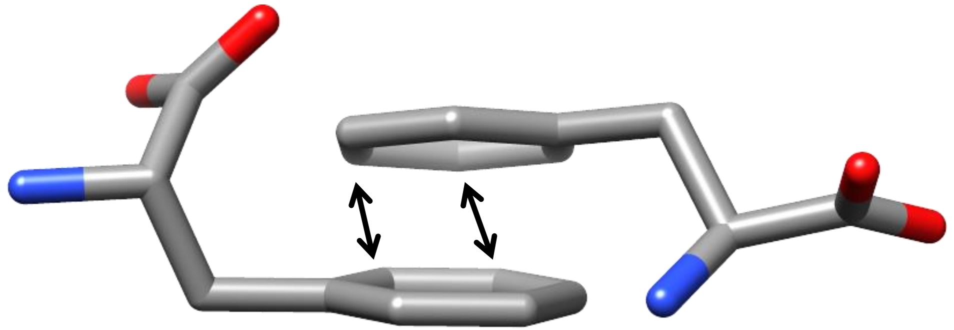
protein force field (pictures)



3 types of
bonded energies



electrostatic



van der Waals

2 types of
non-bonded energies

Protein force fields as equation

$$U(\{\vec{r}\}) =$$

$$\sum_{i=1}^{N_{bond}} \frac{k_i}{2}\left(r_i - r_{i,0}\right)^2$$ bonds

$$+ \sum_{i=1}^{N_{angles}} \frac{k_i}{2}\left(\cos\theta_i - \cos\theta_{i,0}\right)^2$$ angles

$$+ \sum_{i=1}^{N_{dihedral}} k_i\left(1 + \cos\left(n\varphi_i - \nu_{i,0}\right)\right)$$ dihedrals

$$+ \sum_{i=1}^{N_{atom}} \sum_{j=i+1}^{N_{atom}} \left(4\epsilon_{ij}\left(\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right) + \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}}\right)$$ non−bonded

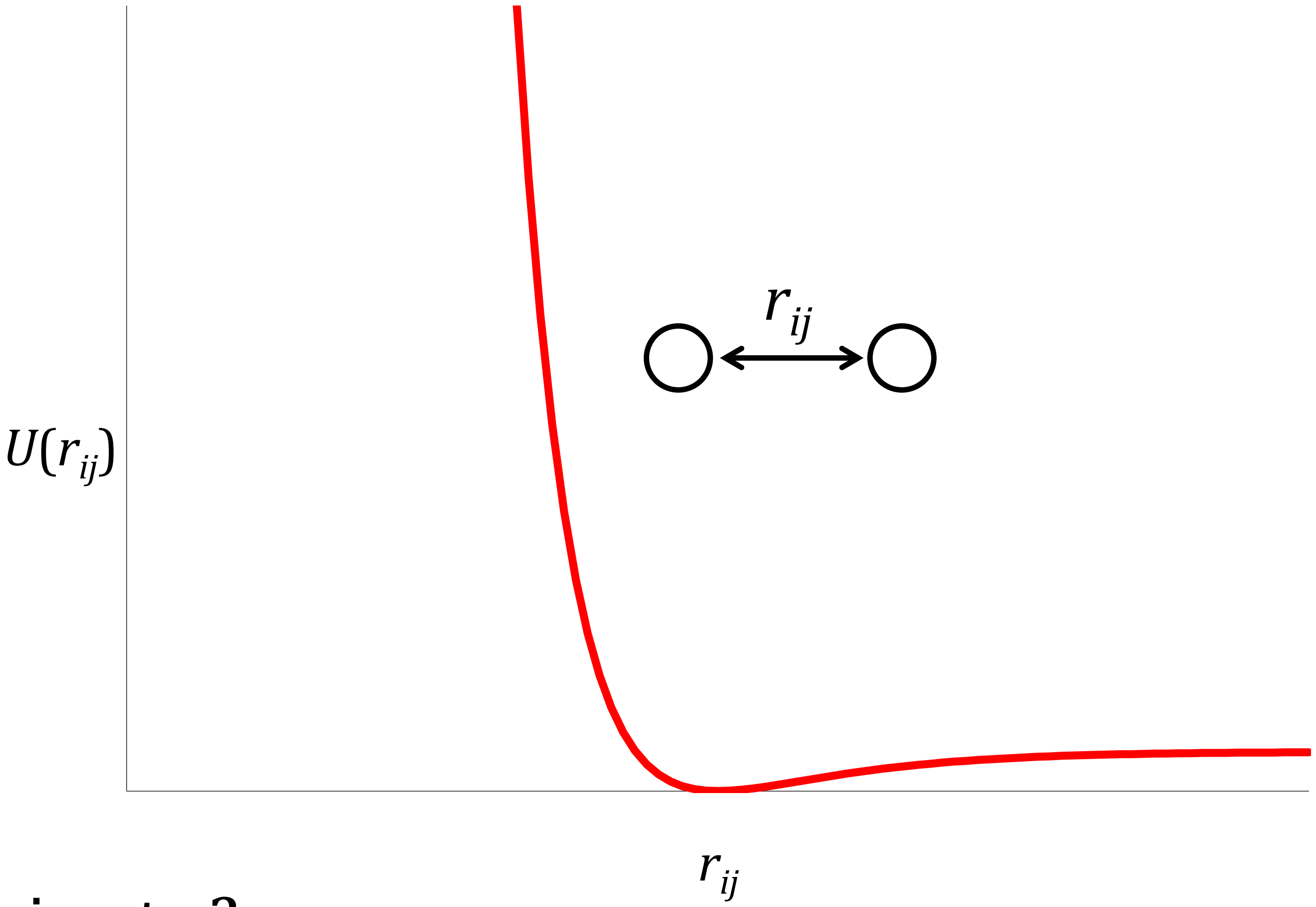
Not a bad approximation

x_i is a value in the coordinates, $x_{i,0}$ is a literature reference value

Bonds

What are bonds really ?

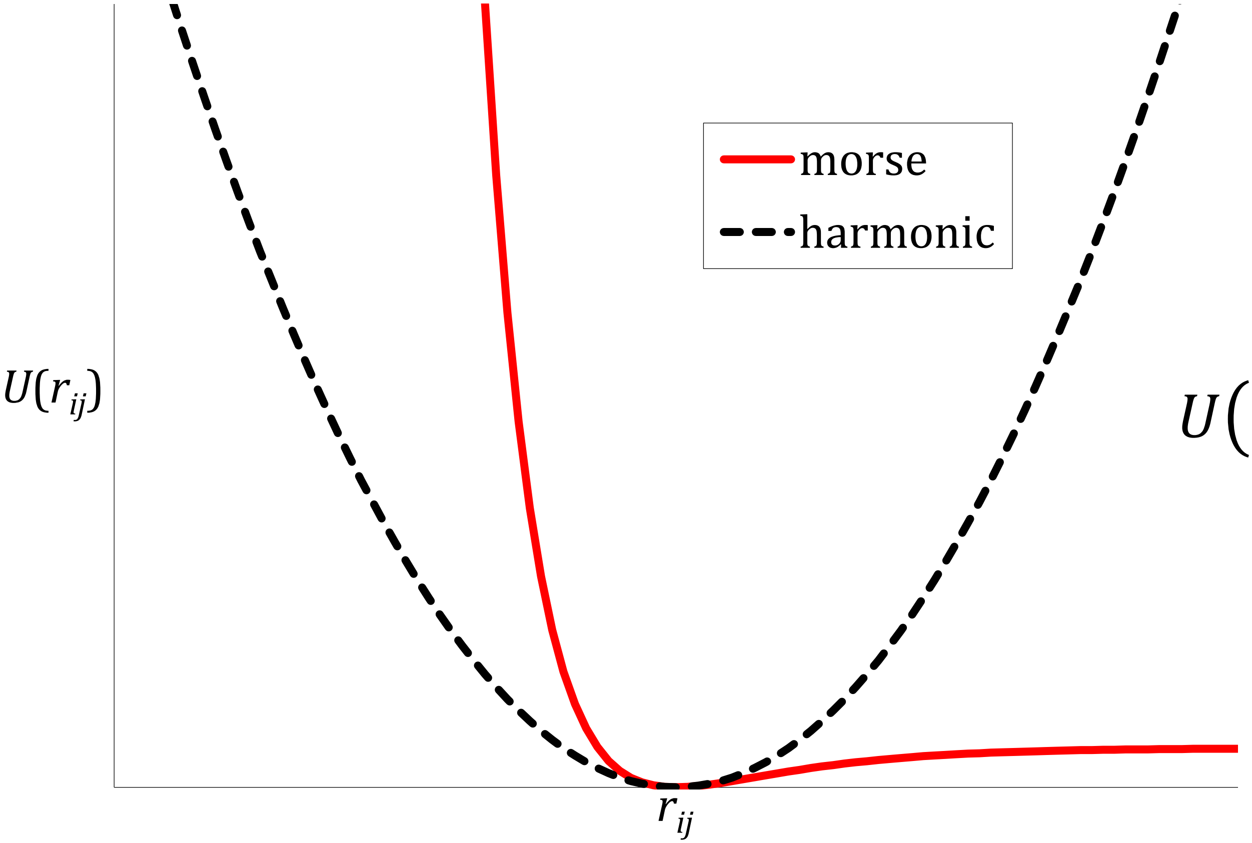
- as we pull two particles apart, what happens ?



- can we approximate ?

Bonds - approximate

Is the black line a good approximation ?

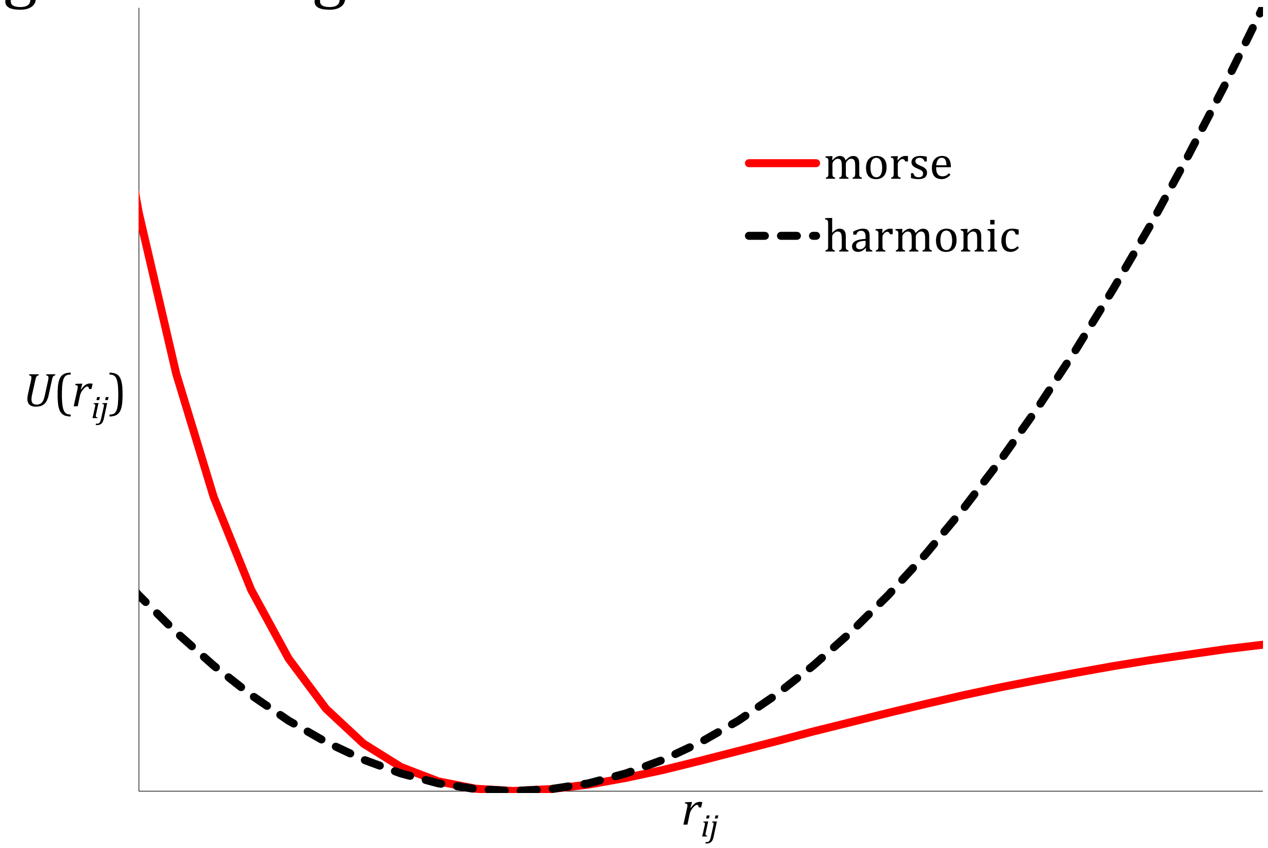


$$U(r_{ij}) = \frac{k}{2}(r_{ij} - r_0)^2$$

Bonds – good enough

Bonds do not stretch much

- harmonic is often good enough



Why do we write $\frac{k}{2}\left(r_{ij} - r_0\right)^2$? (not just k)

Bonds – the end

We normally write $U\left(r_{ij}\right) = \frac{k}{2}\left(r_{ij} - r_0\right)^2$

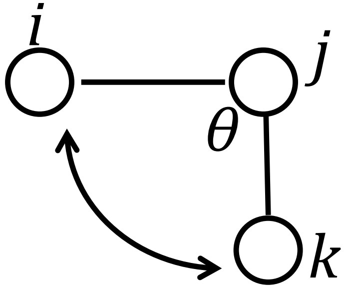
as scalar.. $F = \frac{-dU}{dr} = -k(r - r_0)$ which looks like Hookes law

Do bonds matter ?

- fluctuations very small at room temp (< 0.1 Å)
- our structures are not so accurate / we simulate to look at coarse features
- often treated as rigid joints (maybe more in MD lectures)

Angles

Angles are not as boring as bonds



$$U_{angle(\vec{r}_i,\vec{r}_j,\vec{r}_k)} = \frac{k}{2}\left(\cos\theta_{ijk} - \cos\theta_0\right)^2$$

or

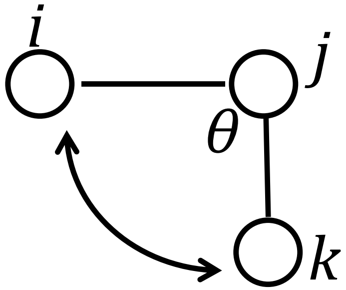
$$U_{angle(\vec{r}_i,\vec{r}_j,\vec{r}_k)} = \frac{k}{2}\left(\theta_{ijk} - \theta_0\right)^2$$

but to get forces is messy (use chain rule) … why ?

Why are forces difficult ?

We use physical models like

$$U_{angle}\left(\vec{r}_i, \vec{r}_j, \vec{r}_k\right) = \frac{k}{2}\left(\cos\theta_{ijk} - \cos\theta_0\right)^2$$



Force is obvious

- looking at picture
- in terms of $\cos\theta$
- in terms of θ *?*
- we want x, y, z

$$\vec{F}_{angle}(\vec{r}_i\ ) = \frac{-\partial U(\vec{r}_i)}{\partial(\vec{r}_i)}$$

$$= \frac{-\partial U(\vec{r}_i)}{\partial \cos\theta}\frac{\partial \cos\theta}{\partial(\vec{r}_i)}$$

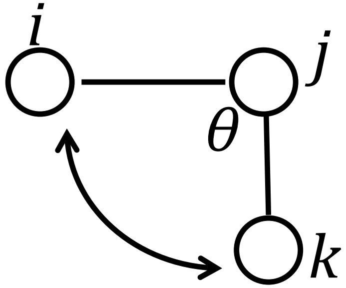
$$= \frac{-\partial U(\vec{r}_i)}{\partial \cos\theta}\frac{\partial \cos\theta}{\partial\theta}\frac{\partial\theta}{\partial(\vec{r}_i)}$$

Angle forces

$$= \frac{-\partial U_{angle}(\vec{r}_i)}{\partial \vec{r}_i}$$

$$F_{angle_i} = \frac{-\partial U_{angle}(\vec{r}_i)}{\partial \cos\theta_{ijk}} \frac{\partial \cos\theta_{ijk}}{\partial \vec{r}_i}$$

$$= k\left(\cos\theta_{ijk} - \cos\theta_0\right)\left(\frac{\vec{r}_{kj}}{r_{jk}} - \frac{\vec{r}_{ij}}{r_{ij}}\cos\theta_{ijk}\right)\frac{1}{r_{ij}}$$



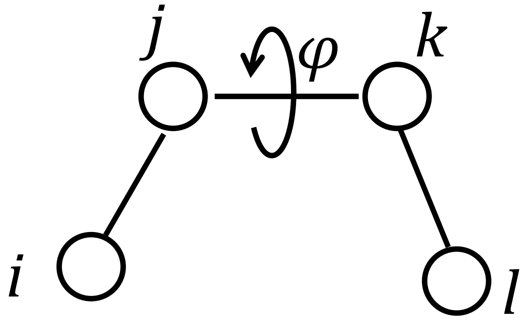
The other atoms ?

similar expression for F_k

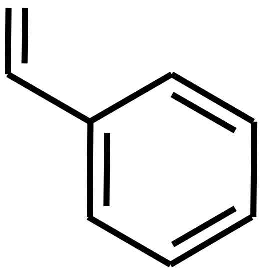
$F_j = -(F_i + F_k)$

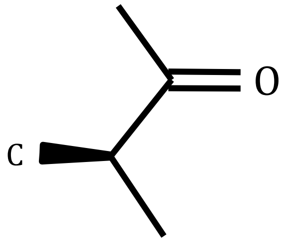
dihedral / torsion angles

basically…



but details will vary

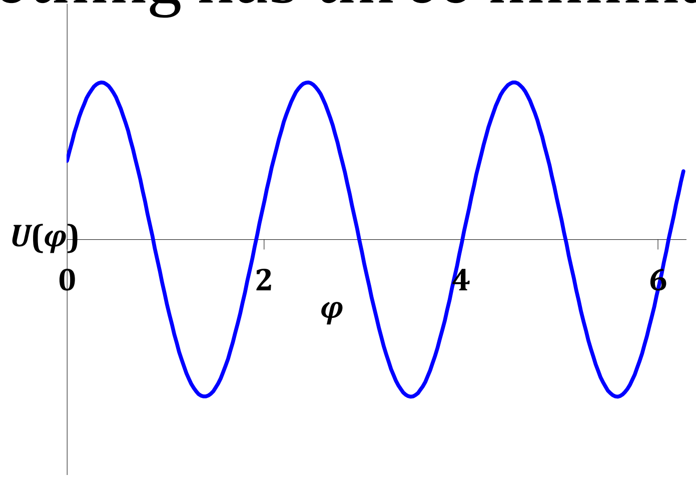


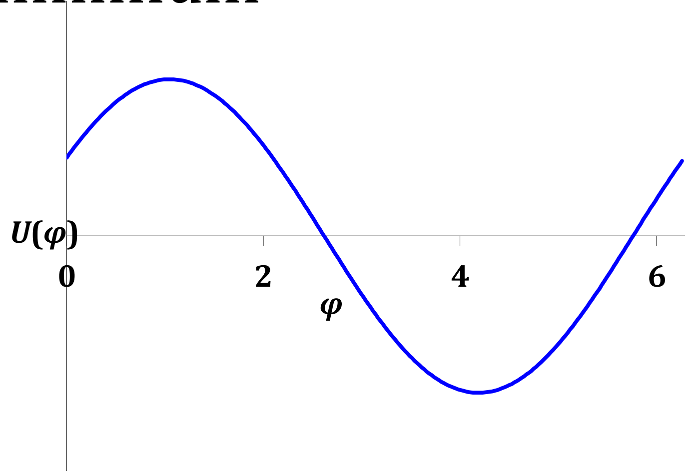


- how large are energy barriers ?
- how many minima are there ?

Form of dihedral term

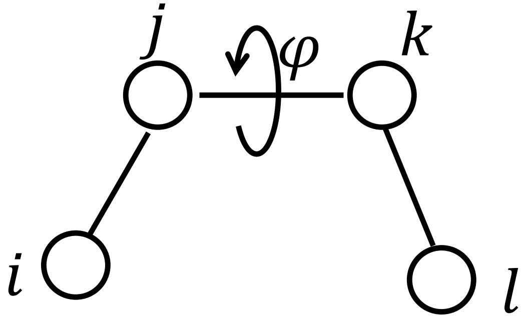
Maybe something has three minima or one minimum



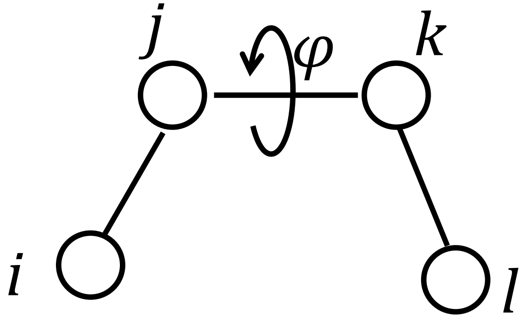


What are the causes ?

- do i and l interact ?
- electron clouds from j and k ?
- model $\quad k_i(1+\cos(n\varphi_i - \nu_0))$



model for dihedral angles



$U(\vec{r}) = k_i(1 + \cos(n\varphi_i - \nu_0))$

n controls multiplicity

- n = 3 butane
- n = 2 peptide bond

At room temperature, barriers are similar to kT so

- atoms do rotate, but there are preferences (from spectroscopy)

How good is the model ?   Not perfect

- some rotamers are preferred (need other terms)

What do forces look like ? scary (much trigonometry)

- intuitively easy, maths messy- think of j and k

Non bonded forces

- van der Waals / Lennard-Jones / dispersion + attraction
- electrostatic

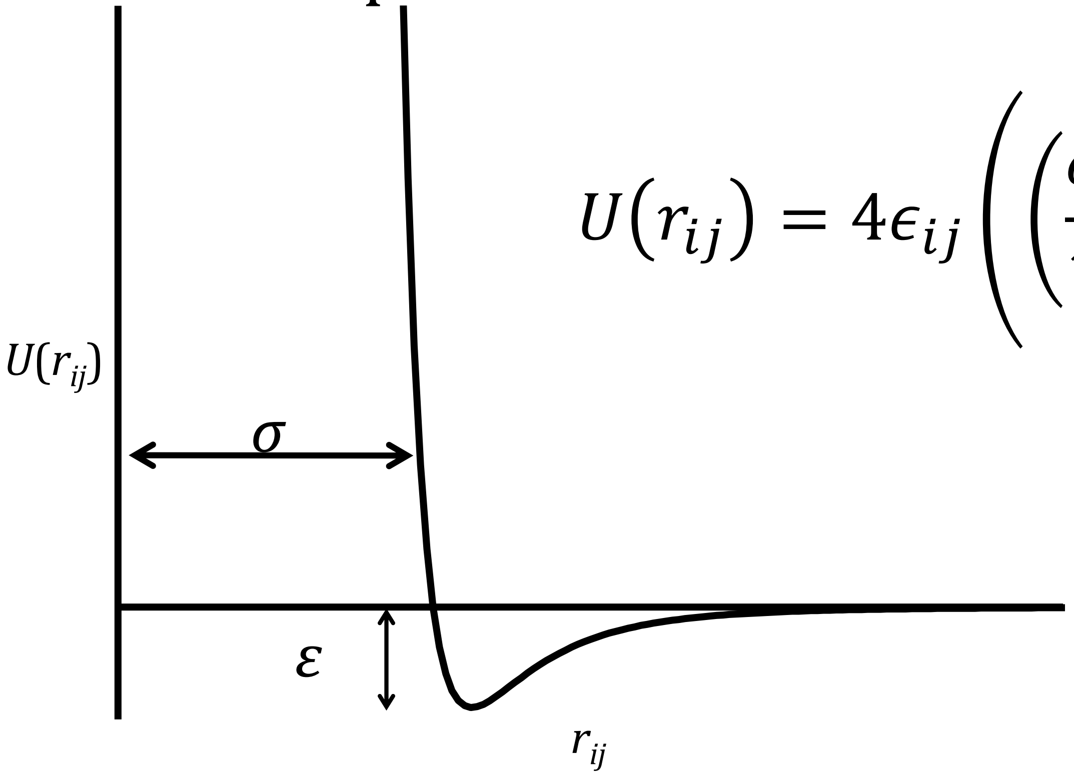
Why are they separated from others ?

- bonds, angles and dihedrals
    - you know the participants in advance
- non-bonded
    - atoms can move to and from each other

van der Waals

What do we know in advance ?

- “inert” gases do form liquids (atoms like each other)
- atoms do not sit on top of each other



$$U(r_{ij}) = 4\epsilon_{ij}\left(\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right)$$

- distance of minimum energy $= 2^{1/6}\,\sigma$

van der Waals – how good is it ?

Repulsive and attractive look similar

- quite different

$$U(r_{ij}) = 4\epsilon_{ij}\left(\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right)$$

Attractive part ?

- electrons wobble and talk to neighbours – induce charges
  - r^{-6} pretty good

Repulsive part ?

- what do atoms look like ?
- really exponential, r^{-12} is just convenient

Lennard-Jones terms – how real ?

How real is it ?

- good for liquid argon
  - diffusion, transport…
- Lennard-Jones fluids

Nice features

- ε, σ = “well depth” and size
- ε, σ specific for atom pairs
  - σ for H is tiny, for C is much bigger

Electrostatics

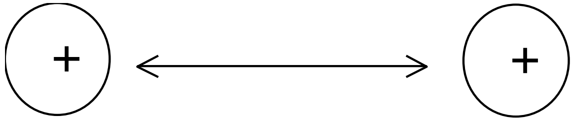
Coulombs law

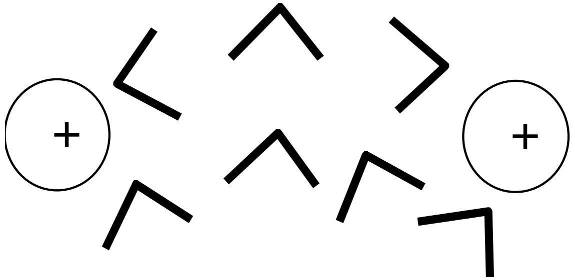
Sounds easy $U(r_{ij}) = \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}}$ or $U(r_{ij}) = \frac{q_i q_j}{D\, r_{ij}}$

D is dielectric constant

Why is it difficult ?

- what is *D* ?
- with and without water ?
- intervening protein ?





- more on solvent models later

Model can be made better / worse

Is this model good or bad ? fast or slow ?

What will it be used for ?

- molecular dynamics simulations, energy minimising
  - rarely a simple energy evaluation
- CPU intensive

Conflicting goals

- make model cheaper, but maintain quality
- add details to make model better

Cost of model

Cost of model ?

- bonds, angles, dihedrals $O(n)$
- non-bonded n particles interact with n particles $O(n^2)$
- probably ≈ 90% of time spent on non-bonded interactions

Many simplifications

- some are popular and important

Cheaper models - United atoms / heavy atoms

When does a proton matter ?

- charge interactions, H bonds

When is a proton not interesting ?

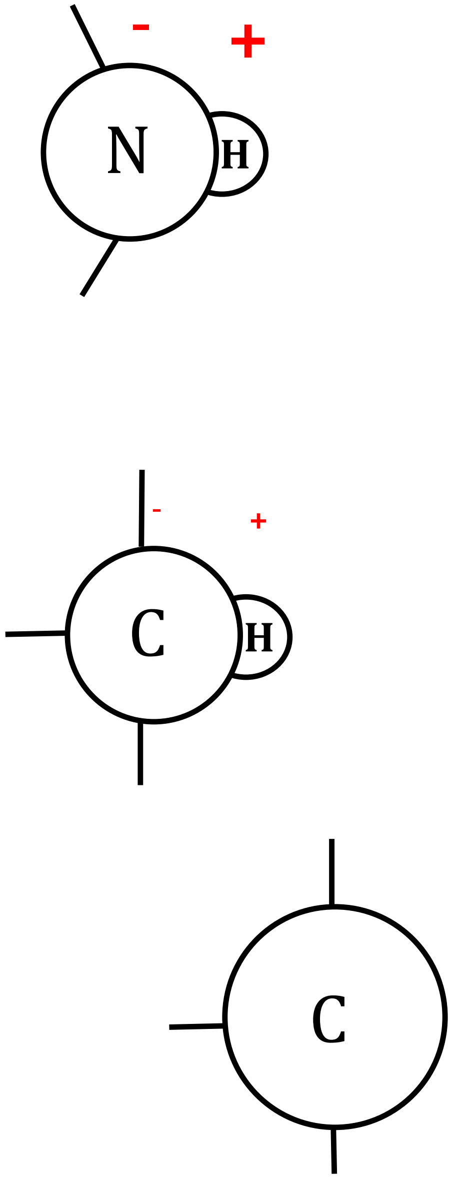
- most aliphatic and aromatic
- hardly a charge / tiny radius

Do we even need the hydrogen ?

- Use a "united atom"
  - mass – easy 12 + 1
  - charge ? nothing
  - radius slightly larger

Rule

- use explicit H in polar groups
- absorb everywhere else
- CH, CH_2, CH_3



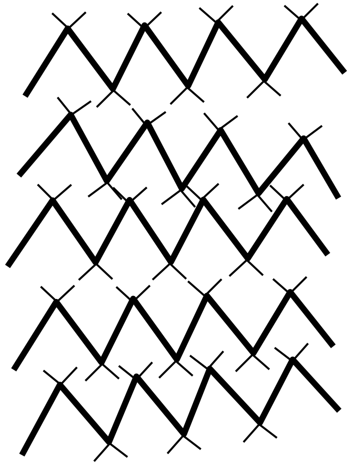
United atoms how bad ?

Gain of united atoms ?

- roughly halves number of atoms
- Costs ?
  - dynamics ? no problem
  - structure ? not too bad

Problems ?

- some small effects can be seen in certain systems
  - lipids
- maybe some effects in proteins

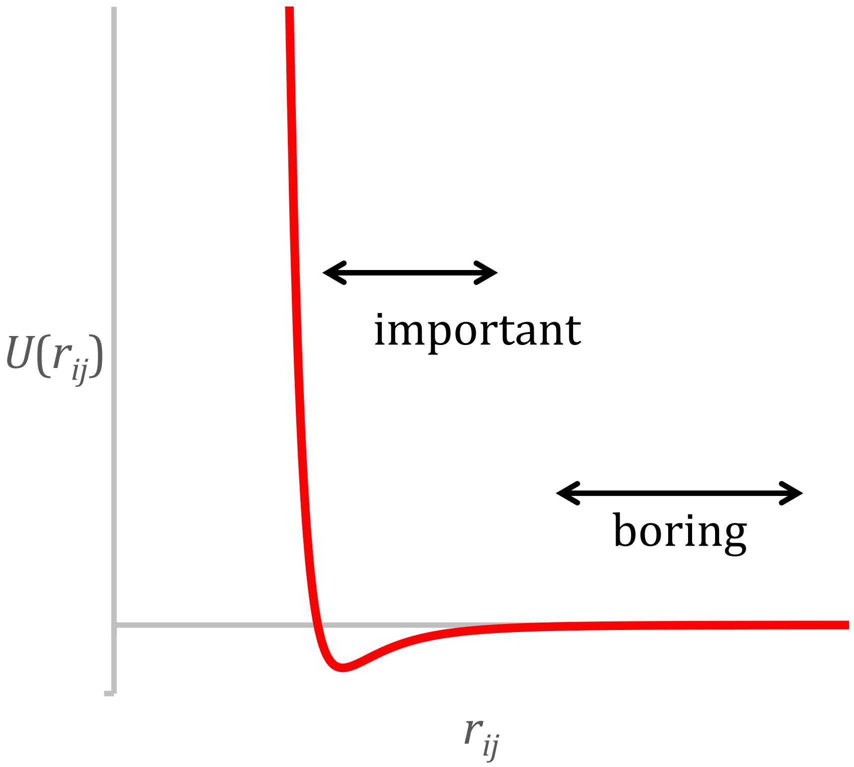


Cutoffs (cheaper models)

How important is an interaction ?

- some are not interesting (bonds)
- some depend where you are

$$U(r_{ij}) = 4\epsilon_{ij}\left(\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right)$$



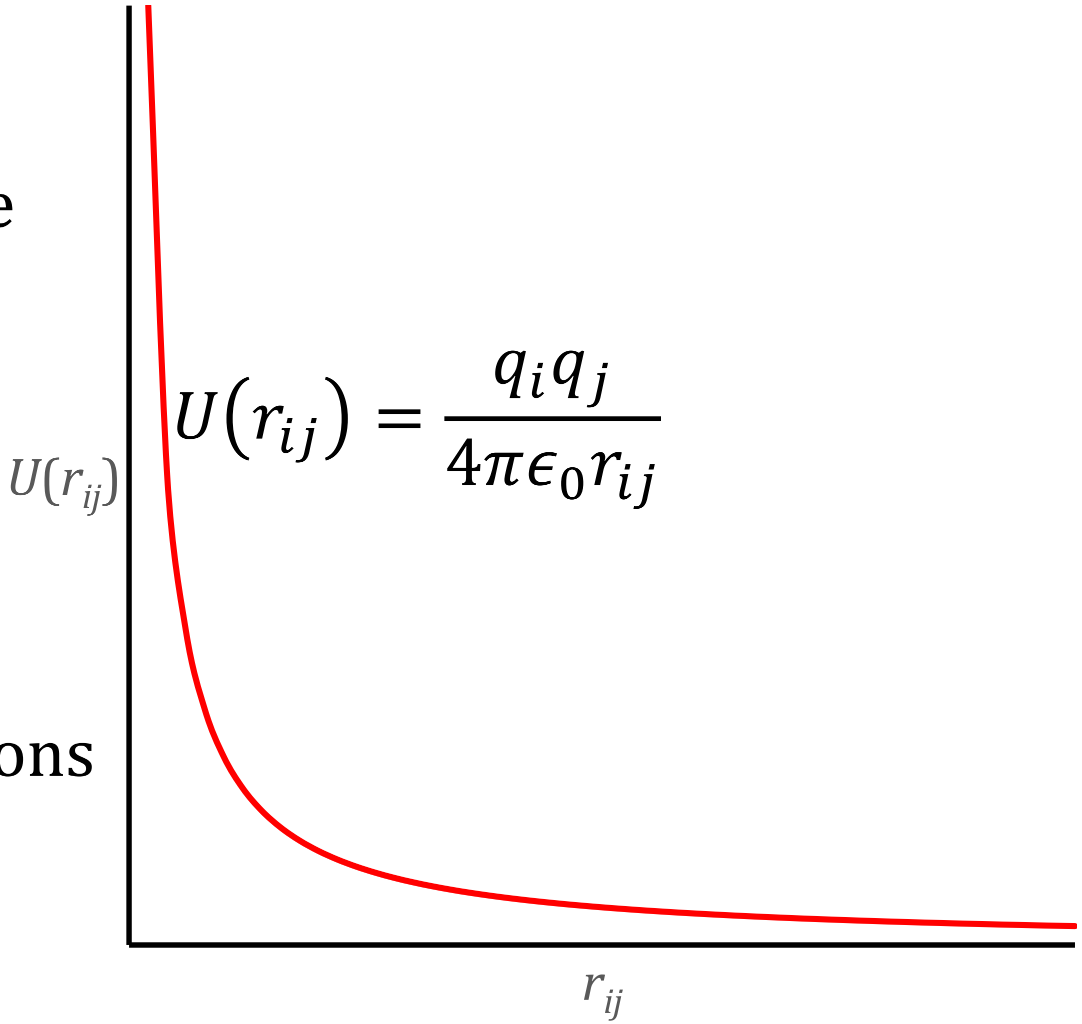
  - r^{-6} and r^{-12} become small quickly
  - if $r_{ij} > 6$ to 8 Å, $U_{LJ}(r_{ij}) = 0$
    - not a problem

- more difficult.. for other terms

cutoffs - problems

Electrostatics

- r^{-1} shrinks slowly
- but eventually, we should be able to ignore
- bigger cutoff (10 – 15 Å)



$$U(r_{ij}) = \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}}$$

In practice

- use one cutoff for all non-bonded calculations

Problems

  - subtle – look at derivative

Possible improvements

Better bonds ?

- more sophisticated than harmonic $(r - r_0)^2$
- can do - not very interesting

L-Jones better than r^{-12} ?

- can do - not worth worrying about

Electrostatics

- polarisation ?
  - completely lacking from model so far (fixed partial charges)
  - we know it is important
  - popular, difficult

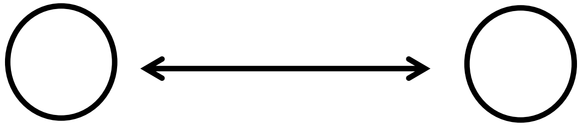
Water

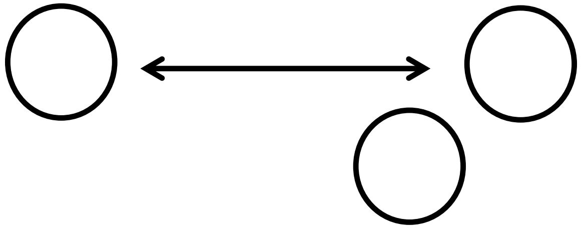
- model so far is *in vacuo* ... more later

Problems and fixes in model

Model is fundamentally wrong

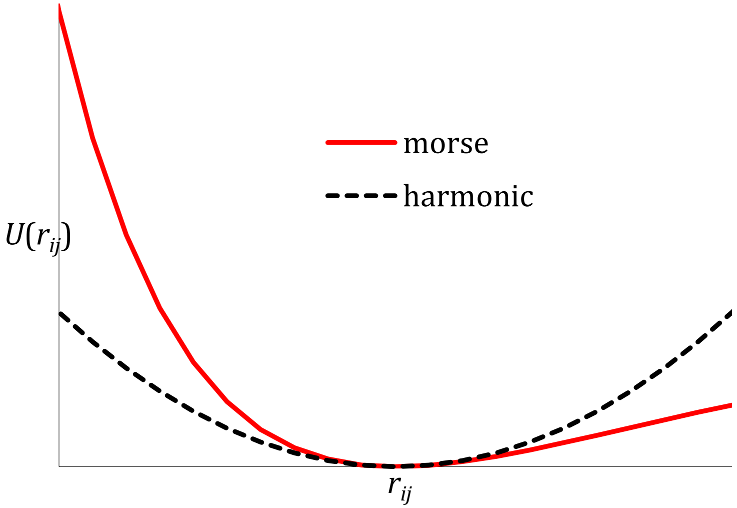
- think about forms of functions
- all two-body based, compare...





Can we represent interactions with two body form ?

- can be good for a small range
- remember earlier picture...
- what works well at 300K may not work at 600K



Why problems may not be seen

Model is not perfect, why do simulations work ?

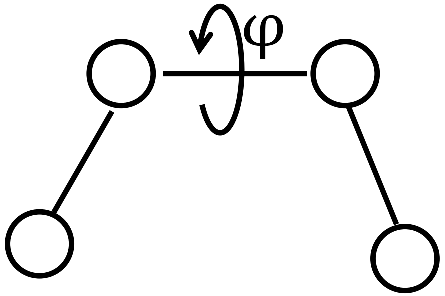
- lots of parameters, not independent

Simple example

- what controls density of a polar fluid
  - charges ?
  - Lennard-Jones terms ?   Both

Complicated example

- rate of rotation..
  - torsional term (k in torsion term)
  - size of atoms and barriers (ε, σ) (less important)
  - angles ? (less important)



Errors in one part of force field compensated elsewhere

Testing force fields

What should a force field do ?

- structure
  - simulate a protein
    - it should not blow up (necessary / not sufficient)
  - density
  - anything you can calculate and measure (distances, chemical shifts..)
- reproduce energy changes (free energies)
- global minimum (free) energy should agree with experiment
- dynamics properties
  - which torsion angles rotate at room temperature ?
  - order parameters from NMR

Transferability

Perfect model of physics would work in all cases

- atomic parameters same from protein to protein
- from protein to organic molecule
- across temperature ranges ?

Disappointments /difficulties

Special systems / special problems

- highly charged systems
  - DNA
    - solvent and charges
- lipids
  - repetitive nature emphasises some problems

Meaning of disappointment

- simulate a protein and it falls apart
- it implodes
- density of a system is wrong
- energetic predictions are wrong
- dynamic predictions are wrong
- smaller structural predictions are wrong

Parameters

- Force field / model has lots of parameters
- charge, mass, ε, σ,
- bond lengths, angles, … for each type of bond / angle /dihedral
- Sources
  - literature
    - mass
  - partial charges ?
    - high level calculations
  - measurements on small molecules (crystallography)
    - bond lengths, geometry
  - trial and error (example)
    - simulate a liquid
      - reduce σ to increase density
      - decrease ε to make it boil more easily

Parameters are a compromise

Model is not perfect

- internal compensation
- compromise example
    - partial charges are not really fixed
    - depend on environment + geometry
- make a decision and adjust others to work in important area

Atomistic force field summary

Model for potential energy
- proteins, organic molecules ..

Main model
- ignore water
- 3 kinds of bonded interactions
- 2 non-bonded

Conservative force field
  - $U(r_{ij})$energy depends on coordinates only
    - no time component

Energy continuous
  - useful
  - derivative always defined
    - very useful

more summary

- main model
- speed-ups – cutoffs, united atoms
- where do parameters come from ?
- what should parameters do ?
- why may you not see errors in parameters ?
- transferability of parameters

Where next ?

- Better models (not much)
- Fixing worst aspects
- Simpler models

- Simulation
- first some rules from statistical mechanics

Simulations / statistical mechanics / …

Topics

- very simple stat mechanics
- Monte Carlo
- molecular dynamics

Favourite books

- "Understanding Molecular Simulation", Frenkel and Smit, Academic Press, 2002
- "Computer Simulation of Liquids", Allen & Tildesley, Oxford Science Publications, 1990 or 2017 (both are good)

Concepts

- potential energy
- work
- entropy
- free energy

Goal

- There are models for potential energy
- The world is driven by free energy changes
- for free energies we need simulations

Nomenclature

- T temperature
- N all kinds of things, usually number of particles
- k and k_B Boltzmann's constant
- R gas constant = $k\,N_a$
- S entropy
- F Helmholtz (NVT)
- G (Gibbs) free energy (NPT)
- U internal energy

- H enthalpy = $U + pV$
- E energy, E_{kin}, E_{pot}, E_{bond}, …
- Q heat
- W work
- V volume
- Ω (x) number of states of system with property (x)

more nomenclature

For homework/seminars/happy life/ lecture

- lower case italics for scalars (n, r)
    - historic exceptions like $pV = nRT$
- other fields
    - statisticians like $P(a|b)$
    - Σ is a summation, but not in Stellingen (alphabet)
- maths is maths, πr^2 and never $\pi * r^2$ and never ever pi * radius^2
- $\vec{r}$ or $\vec{r}$ or $\mathbf{r}$ vectors
- $\{\mathbf{r}\}$ collection of vectors
- if x is a scalar, $\dot{x}$ is the time

Potential energy

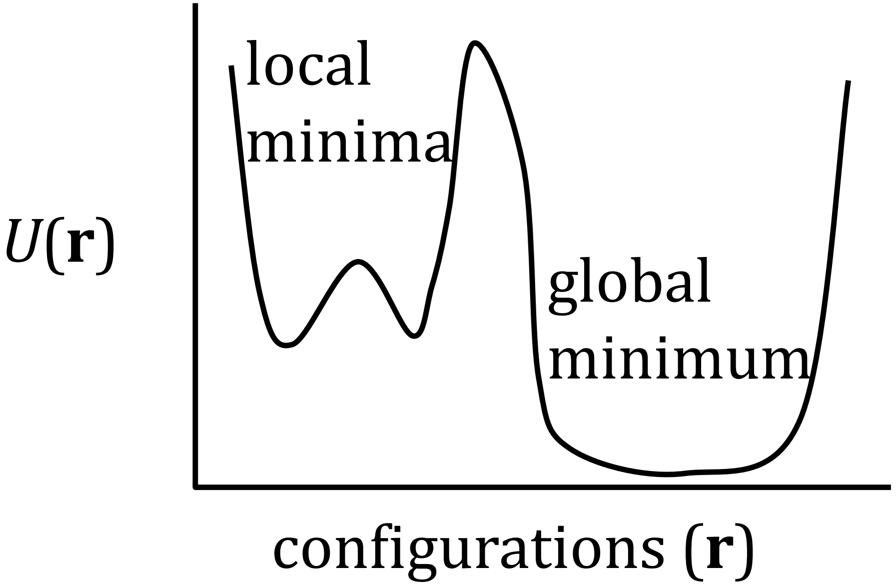
electrostatic $U(r) = \frac{q_1 q_2}{4\pi\epsilon_0 r}$

gravity $U(r) = \frac{G m_1 m_2}{r}$

elastic, …

potential energy of a protein…

- where would the protein like to be ?
- will be more interesting in ten minutes



Conserving energy

For this course:

- first law
- relevant energies
- is it obvious ?
- when do we not conserve energy ?

First law

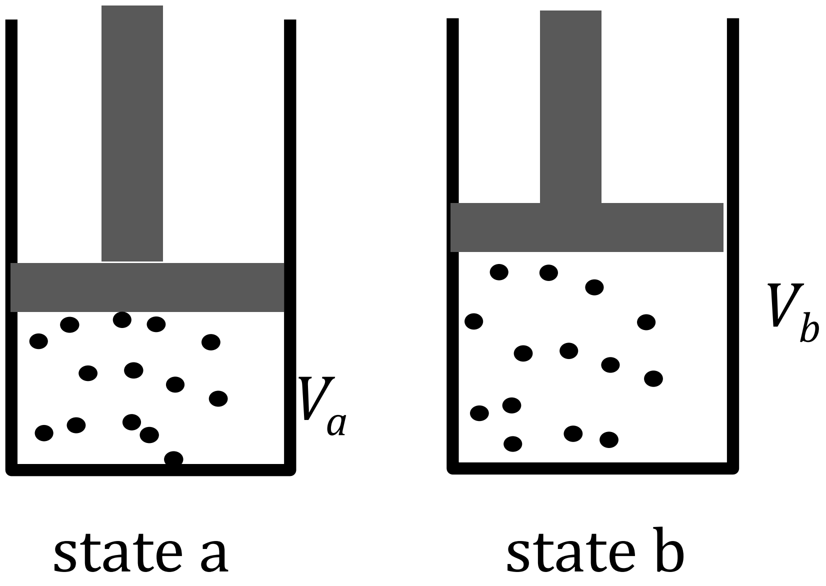
- $dU = \Delta Q + \Delta W$
- dU change in internal energy
- ΔQ heat given to system
- ΔW work done on system

- example of work…

Work on a gas

$$\Delta W = -P\Delta V$$

$$\begin{aligned} dU &= \Delta Q + \Delta W \\ &= \Delta Q - P\Delta V \end{aligned}$$



Others

- charges in electric fields, surface tension / area, ….

- we do not often do this in this course
- in one lecture: W will be important
  … how hard is it to move a ligand from solvent into a protein ?

Conserving energy

Newtonian dynamics coming soon

- what is the energy of a particle ?

$$E_{tot} = E_{pot} + E_{kin}$$

- energy of a system ? (box of gas, protein in water)

$$E_{tot} = \sum_{i=1}^{N_{particles}} \left(E_{pot_i} + E_{kin_i}\right)$$

- total energy is conserved – potential + kinetic

Do we always have kinetic energy ?

- Monte Carlo – no
- Wirkstoffentwurf-Vorlesungen - no

Conserving energy – always obvious ?

In the real world

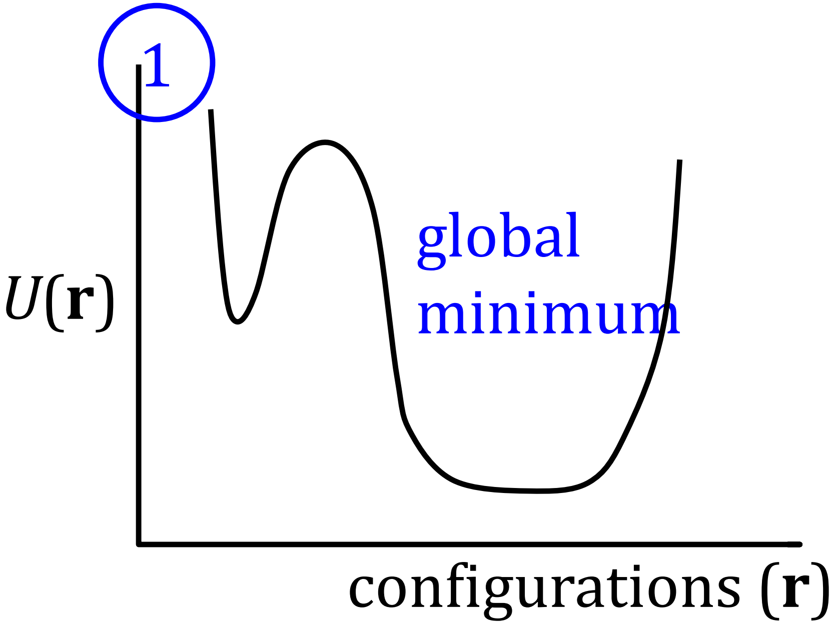
- closed systems ? yes
- almost everything in this course – yes

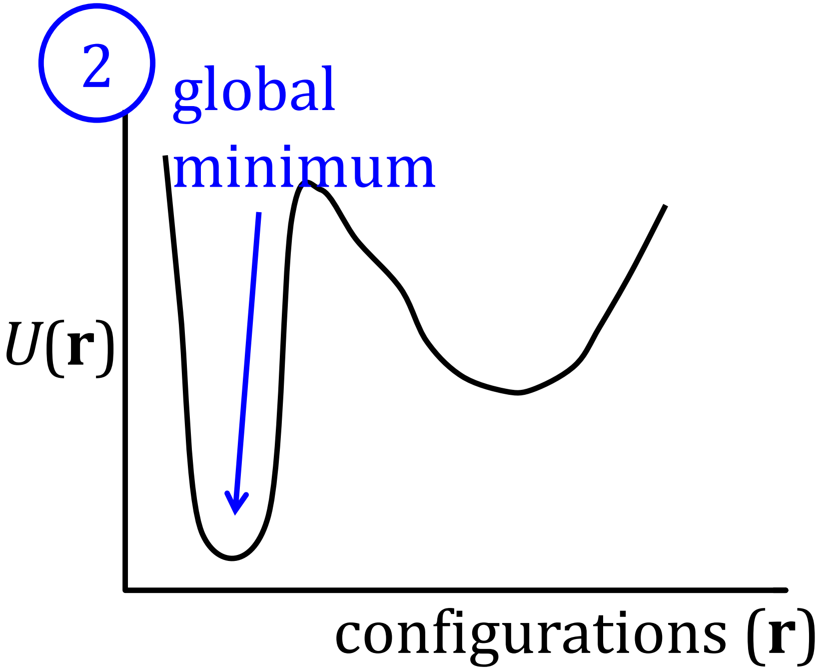
Simulations

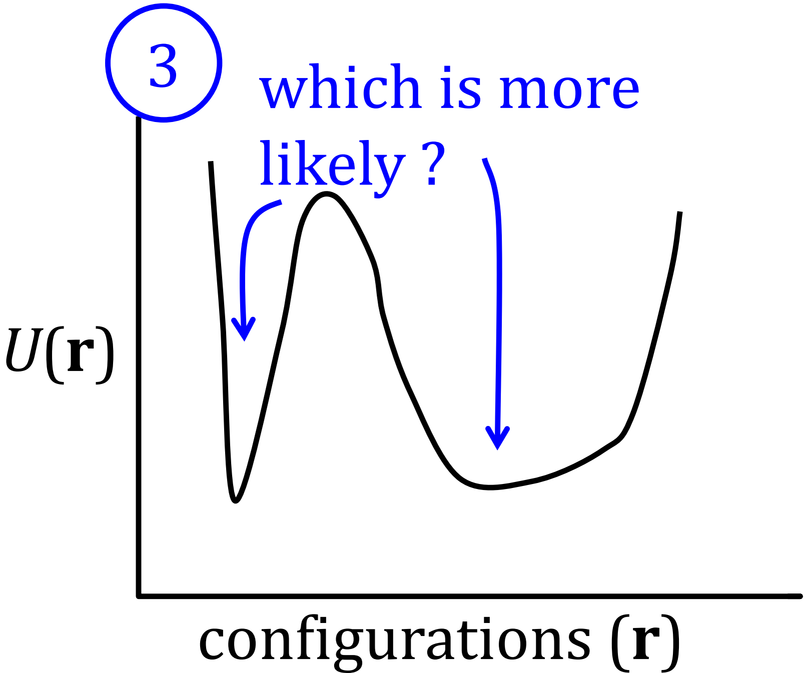
- if I want to use statistical mechanics, standard theory
  - no heating
  - no changing the force field (sounds obvious ?)
    - I want to simulate protein (un)folding
      - pull on the C-terminus – illegal – does work on system
      - heat system – illegal
  - push a drug (Medikament) into binding pocket – illegal

There is more than kinetic and potential energy..

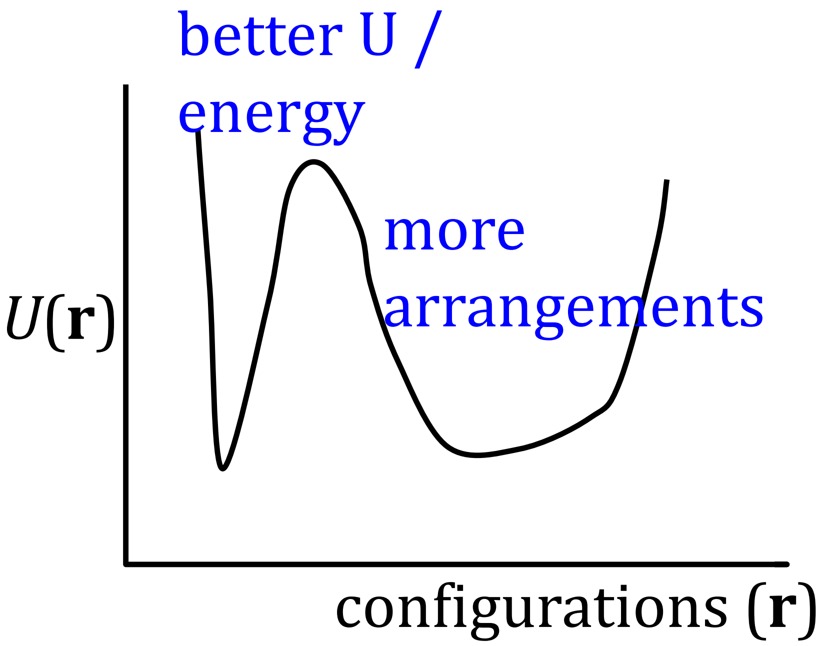
More than potential energy







Simulation and formal statistical mechanics



Quantified ?

- entropy and free energy

Dumb simulation, preferring to go downhill

- should show how probabilities (entropy) are balanced against energy

History

Statistical mechanics

- derived by summing up properties of individual particles

Thermodynamics

- less emphasis on individual particles

Lots of formulae which cannot be completely applied to proteins

- sums over infinite volumes, numbers of particles, time

Rules and Limitations

Always at equilibrium

Entropy

- you tend to become disordered (2nd law)
- role in free energy
- formula for simple systems

Disorder, how many ways can the system arrange itself..

- depends on energy (and volume and number of particles)
- two formulae coming… simple, more useful

How many states can the system have (specified E) ? $\Omega(E)$

Entropy

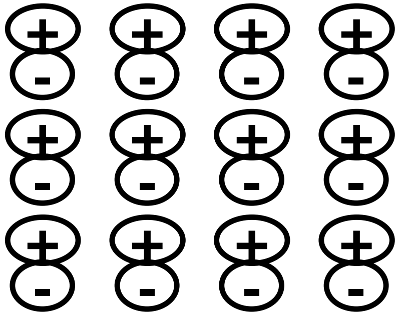
$S = k \ln \Omega(E)$ (very simple form)

Number of states $\Omega(X)$

$T = 0$

$\Omega(E) = 1$

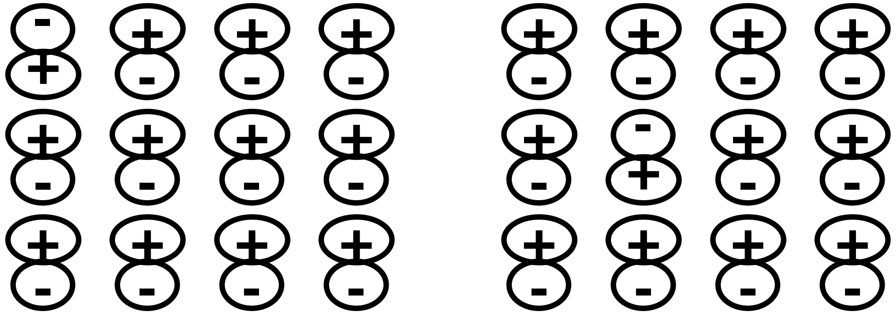
$S = k \ln 1 = 0$



$T > 0$ small

$\Omega(E) = n_{mol}$

$S = k \ln n_{mol}$



A bit more energy ? more states

- more.. solid →liquid .. many many more

Gibbs

What if states are not equally likely ?

N_{state} states with distribution

- Let $p_1 = 0.999, p_2 = 0.00001, p_3 = 0.00001$, ….

  just as if the system had one state - low entropy
- Let $p_1 = 0.5, p_2 = 0.1, p_3 = 0.1$, ….

  a bit more entropy, but still very much dominated by p_1
- Let $p_1 = 0.01, p_2 = 0.01, p_3 = 0.01$, ….

- lots of states, all equally likely
  - lots of entropy

In general

$$S = -k \sum_{i=1}^{N_{state}} p_i \ln p_i$$

Gibbs entropy !!

Applicability

Can one really estimate ? $S = -k \sum_{i=1}^{N_{state}} p_i \ln p_i$

- liquid.. no
- protein ..no

With a simulation ?

- too many states
- approximations to ΔS later

Simple system (grid / lattice)

- yes .. later in semester

Units

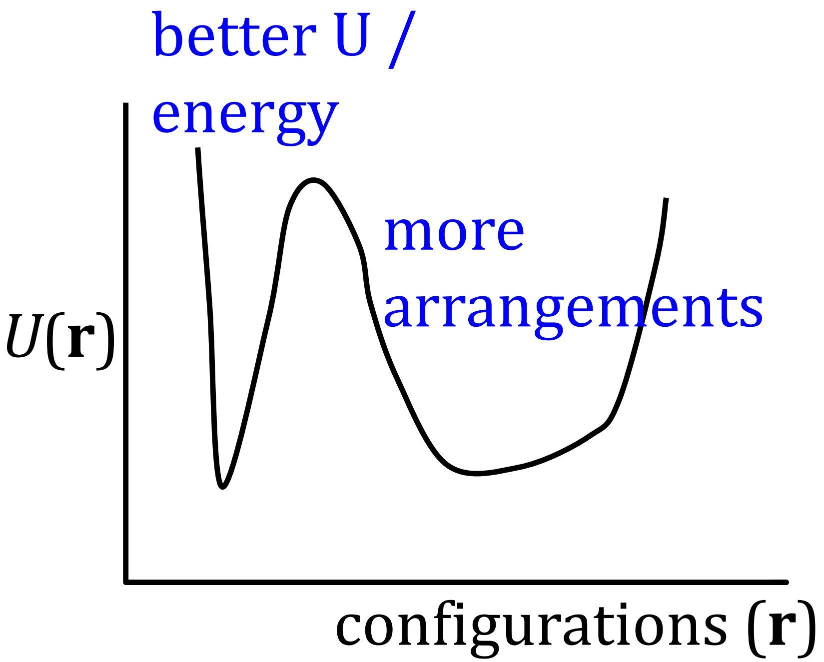
More arrangements = more entropy

Connected to energy

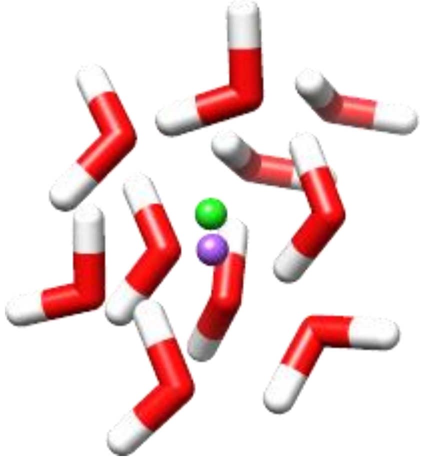
- balance of energy and entropy

Entropy units

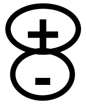
- energy / temp (JK^{-1})
- what formula helps me remember this ?



Where can we use this ?



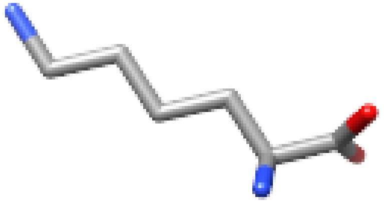
What is the entropy of a molecule ?



- sounds easy
- really…

Entropy depends on solvent (system)

Entropy of a conformation ?



- has no meaning

$$S = -k \sum_{i=1}^{N_{state}} p_i \ln p_i$$

Why are entropy approximations so bad ?

Is S additive ? Only rarely – not for interesting systems

Adding entropy

Ω_A Ω_B $\Omega_\mathrm{AB} = \Omega_\mathrm{A}\,\Omega_\mathrm{B}$

Entropy is "extensive"

$$S_{AB} = -\,k\ln(\Omega_A\,\Omega_B) = -\,(k\ln(\Omega_A) + k\ln(\Omega_B)) = S_A + S_B$$

Assumption

- for my new system A and B do not interact

What if they interact ?

- putting A in state 1 changes probability of B in state 1

$$S = -k\sum_{i=1}^{N_{state}} p_i \ln p_i$$

$$\Omega_{AB} \neq \Omega_A\,\Omega_B$$

Adding entropy

I cannot usually add the entropy of two systems to get the entropy of new total system

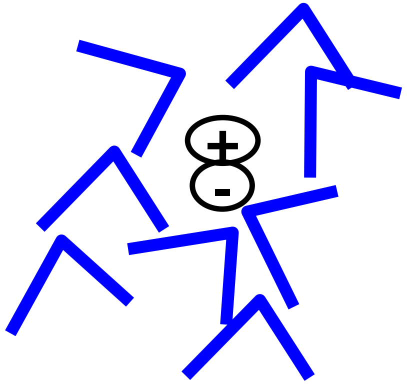
Two systems

- could be boxes (previous slide)
- could be parts of a molecule (coming)

Do not add entropy terms ! Exam questions (promise)

Can I decompose entropy ?

May we decompose entropy ?



Remember energy in proteins

$$E_{tot} = E_{bonds} + E_{vdw} + E_{\dots}$$

$$S_{tot} = S_{bonds} + S_{vdw} + \dots ?$$

No.

- Makes no sense unless bonds are decoupled from atoms and angles…

$$S_{ligand} + S_{solvent} + \dots$$

- obviously do interact

Free energies

$G_{solv}, G_{protein}, G_{ligand}, G_{bonds}$ not really legal

- imagine bonds are system A, angles are system B, ..

Decomposing Free Energy

Would you see this in the literature ?

Example

- in protein asp → asn mutation (COO$^-$ to CON)
- protein becomes more stable

Claim:

- the protein is more stable due to a change in electrostatic free energy

What has been implicitly assumed ?

There is no such thing as ΔG_{el}

More examples – particle interactions

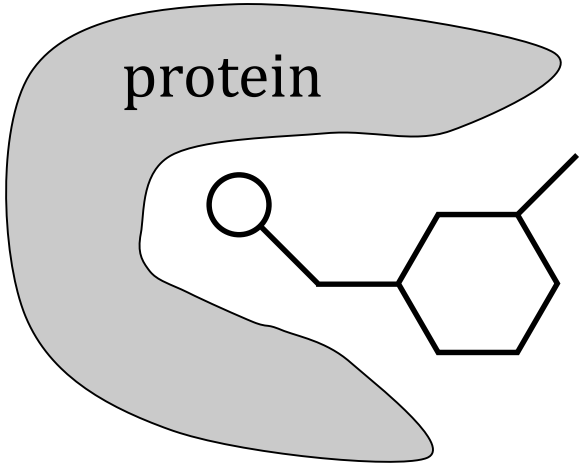
You have a scoring function for interactions

- usually called ΔG

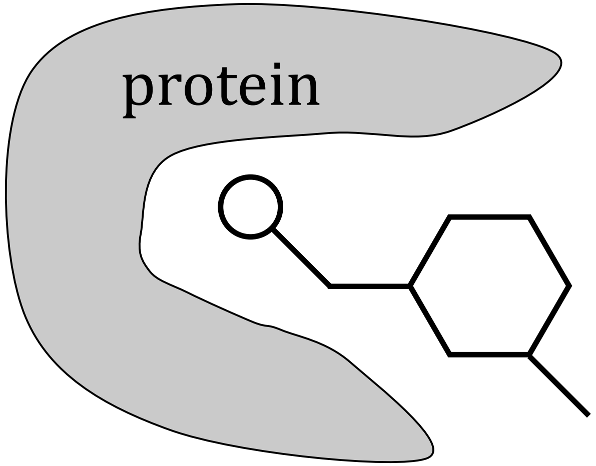
$\Delta G_{tot} =$	ΔG_{HB} +	H-bonds
	ΔG_{ionic}+ ….	ionic

- is this legal ?

Entropy of a conformation ?



binding 1



binding 2

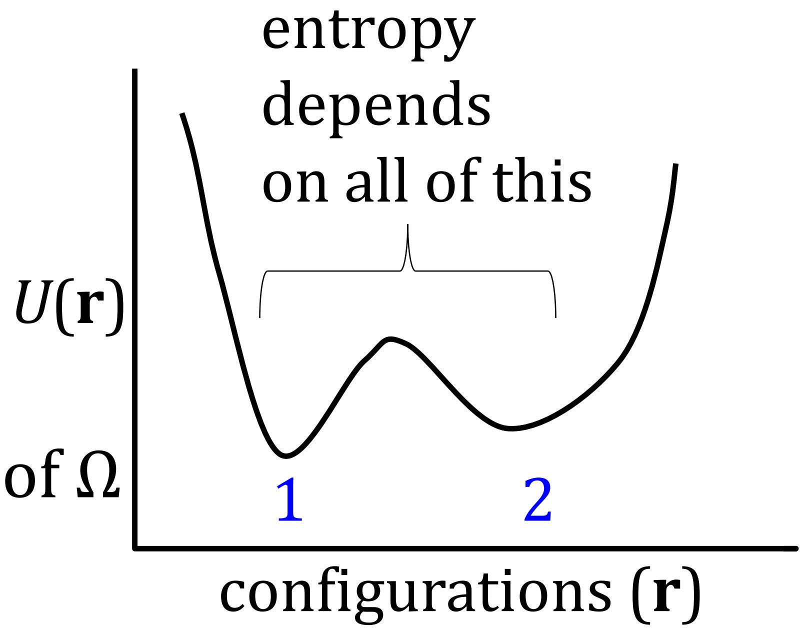
Can I talk about ΔG_1 vs ΔG_2 ?

$$S = -k \ln \Omega$$

but Ω depends on all accessible states

If binding 1 can change to binding 2, they are both part of Ω

- I do believe we can talk about U_1 and U_2



Meaning of labelled free energies

If we have "free energy contributions"

- we have entropy contributions

ΔG_{HB} implies $H_{HB} + T\Delta S_{HB}$

ΔG_{ionic} implies $H_{ionic} + T\Delta S_{ionic}$

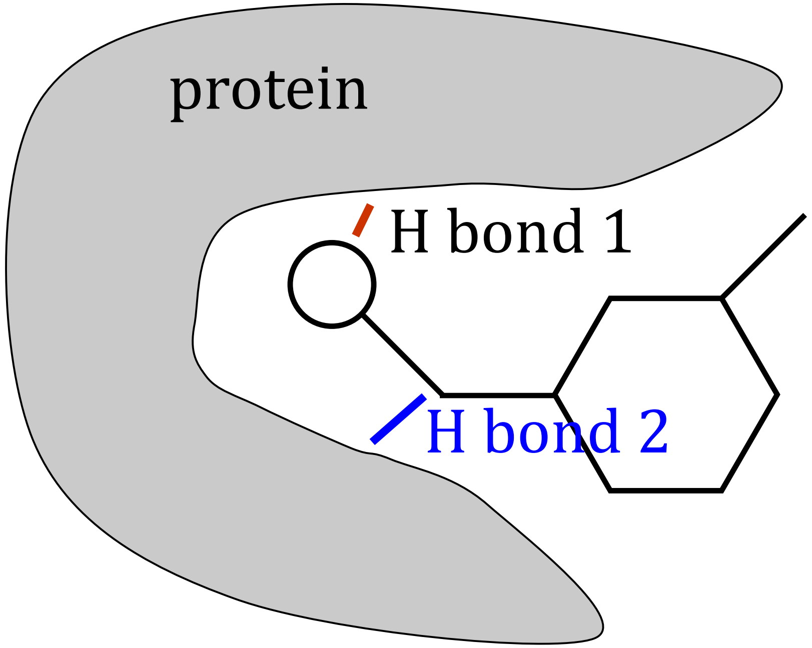
what is S_{ionic}? $S = -k \ln \Omega_{ionic}$?

- no meaning

One cannot have "free energy contributions"

independence of terms

Even without labelled entropies



Can we add free energy due to Hbond 1 and 2 ?

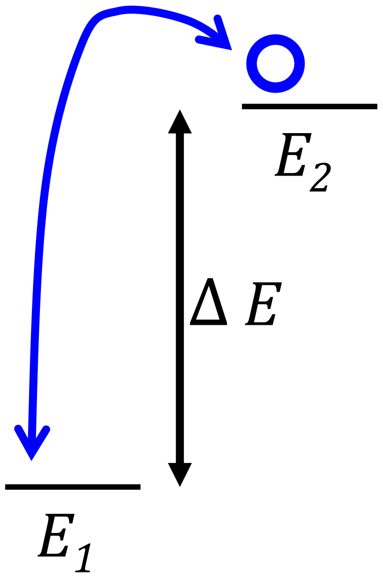
- implies $\Delta S = \Delta S_1 + \Delta S_2$
- $-k \ln \Omega = -\ (k \ln \Omega_1 + \ k \ln \Omega_2)$
- $-k \ln \Omega = -\ k \ln(\Omega_1\, \Omega_2)$

- only possible if there is no interaction between 1 and 2
- additivity of free energies is widely used (usually wrong)

Stop and summarise

- Internal energy U – things like springs
- Enthalpy H – includes pressure
- Gibbs free energy G what we usually use
  - no discussion of different ensembles in this course
- Free energy lets us incorporate the concept of what is most likely
- Entropy is not additive between systems (parts of systems) that interact

Boltzmann distribution – two states



System with two energy levels

- I know the temperature and energy difference
- How likely is system to sit in E_1 ? E_2 ?
  - System is very cold,
    - ΔE seems big
  - System is very very hot T→∞
    - ΔE does not matter, seems small
- Relative populations (probabilities) p_1, p_2
- $\Delta E = E_2 - E_1$

$$\frac{p_1}{p_2} = e^{-\Delta E / kT}$$

- sometimes $\frac{p_1}{p_2} = e^{-\beta \Delta E}$

Sign conventions

$\Delta E = E_1 - E_2$ or $\Delta E = E_2 - E_1$?

For exams

- the lower energy state is always more populated

- there are always questions about probabilities / energies
- state with most negative energy is most populated

Boltzmann distribution

What is the probability of a certain energy level ?

- depends on all available levels

$$p_i = \frac{e^{-\beta E_i}}{\sum_j e^{-\beta E_j}}$$

- name of bottom … partition function, $Z = \sum_i e^{{-E_i}/{kT}}$

- does this agree with previous slide ?

Going from two to n states

Start with

$$\frac{p_1}{p_2} \begin{aligned} &= e^{-\beta\Delta E} \\ &= e^{-\beta(E_1 - E_2)} \\ &= \frac{e^{-\beta E_1}}{e^{-\beta E_2}} \end{aligned}$$

but this is for the ratio of two states. If we have n states, we are interested in $\frac{p_1}{p_1+p_2+p_3}$

Let us turn this upside down and start with $\frac{p_1+p_2+p_3+\cdots}{p_1}$

let us consider some more states and get the absolute fraction for p_1

$$\frac{p_1 + p_2 + p_3 + \cdots}{p_1} = \frac{p_1}{p_1} + \frac{p_2}{p_1} + \frac{p_3}{p_1} + \cdots$$

$$= \frac{e^{-\beta E_1}}{e^{-\beta E_1}} + \frac{e^{-\beta E_2}}{e^{-\beta E_1}} + \frac{e^{-\beta E_3}}{e^{-\beta E_1}} + \cdots$$

$$= \frac{e^{-\beta E_1} + e^{-\beta E_2} + e^{-\beta E_3} + \cdots}{e^{-\beta E_1}}$$

$$= \frac{\sum_j e^{-\beta E_j}}{e^{-\beta E_1}}$$

turn upside-down,

$$p_i = \frac{e^{-\beta E_i}}{\sum_j e^{-\beta E_j}}$$

where the summation runs over all states of the system

Consequences of Boltzmann distribution

At absolute zero

- only lowest energy state is populated

At low temperatures

- low energy states favoured

High temperature

- system can visit high energy regions

Infinite temperature

- all states equally likely

For two states (bound / unbound)

- exponential term means populations quickly become big/small

$$\frac{p_1}{p_2} = e^{-\Delta E/_{kT}}$$

How big are energy differences ?

Simplest case, equal populations $\frac{p_1}{p_2} = e^{\frac{\Delta E}{kT}}$

$p_1 = p_2$

$$1 = e^{\Delta E / kT}$$
$$\ln 1 = \Delta E / kT$$
$$\Delta E = 0$$

k =1.38 × 10^{-23} J K^{-1}, but use

$$R = k\, N_A$$
$$= k\;\; 6.02 \times 10^{23} \text{ J K}^{-1}\text{mol}^{-1}$$
$$= 8.314 \text{ J K}^{-1} \text{ mol}^{-1}$$

Examples of populations

$$\frac{p_1}{p_2} = e^{\frac{\Delta E}{kT}}$$

$$\ln\frac{p_1}{p_2} = \frac{\Delta E}{kT}$$

$$\Delta E = kT\ln\frac{p_1}{p_2}$$

For 99:1 at 300 K

- ΔE = 11 kJ mol^{-1}

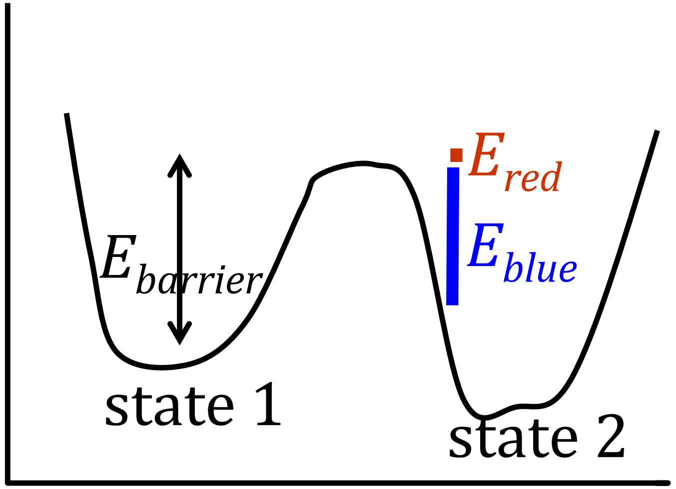
Drugs

- Dissociation constant k_d of drug D to protein P $k_d = \frac{[D][P]}{[DP]}$
- nanomolar drugs are 10^{-9},
- difference between 1 and 10 nm binding drug
  - 46 vs 52 kJ mol^{-1}
- topic will return later

Barrier crossing

How likely are you to cross a barrier ?

- p_{red} vs p_{blue}



- all the blue copies of system will not make it over $E_{barrier}$
- red population is small, even for $E_{barrier} \sim kT$
- but explains why $\ln(rate) \propto T$

where next ?

- how to see some of these properties by simulating
- how simulating depends on these properties