Water models / solvation

Andrew Torda, May 2012, strukt & sim

Biggest effects of water

- electrostatic
- dynamic

Model types

- explicit
- implicit

Dynamic effects of water

Dynamic effects of water

one lonely moving particle

- initial velocity $\dot{x}_t$
- future velocity easy $\dot{x}_{t+\delta t} = \dot{x}_t$
- energy ? constant $\frac{m\,\dot{x}^2}{2}$

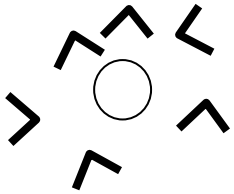
two particles ? interacting ?

- future velocity a bit more difficult
- easily bounded – cannot be more than $\frac{m_1\dot{x}_1^2+m_2\dot{x}_2^2}{2}$

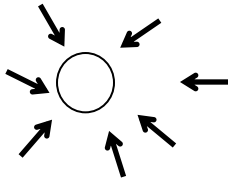
one particle in water...

Velocities of particles in water

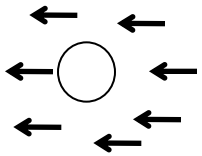
- lots of random interactions



- a small velocity ?



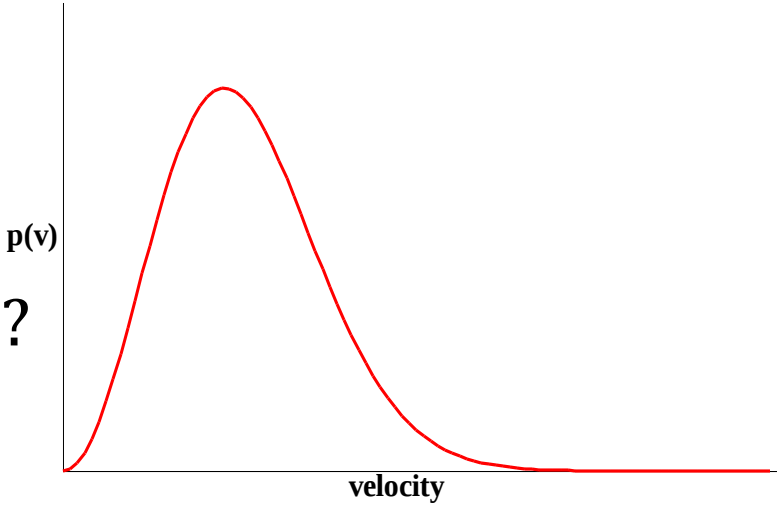
- a big velocity ?



- a probability distribution

+

- how does $\dot{x}_t$ tell us about $\dot{x}_{t+\delta t}$?
  - much less



Modelling dynamic effects

Summary

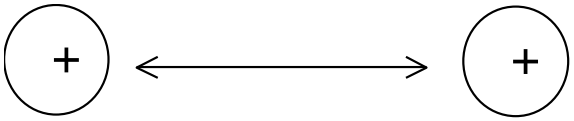
- solvent will add fluctuations
- makes us forget velocity faster

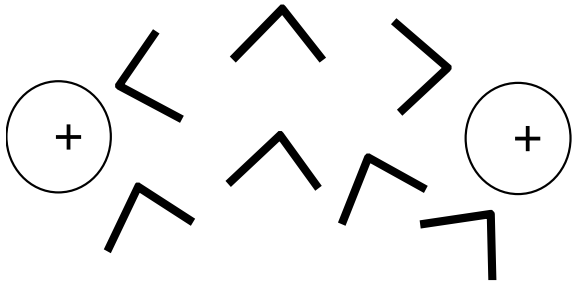
Can this be modelled ?

- yes (in molecular dynamics simulations)
- not really a force field / energy topic
- add random fluctuations to velocities
- can be made to look like water

Electrostatic effects of water

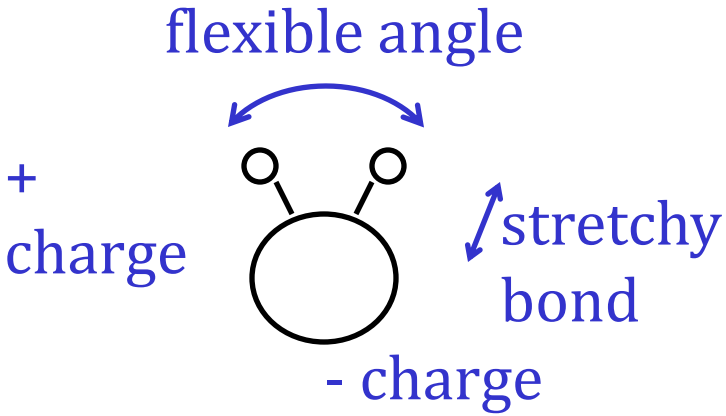
- water molecules
  - not charged
  - polar
- interaction between charges very different if water in between
- details soon



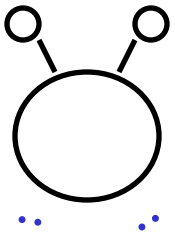


Explicit water

- earlier descriptions of proteins
  - a set of connected atoms
- extend to include water
- what does water look like ?



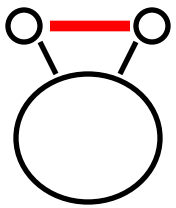
- what else has it got ?
  - think about electron pairs on “O”
- what is really important ?



Important features of a water model

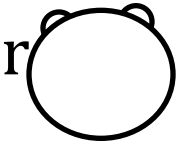
Do we care about water internal dynamics ?

- usually not
- make bonds rigid
- make angle rigid
  - treat as a bond



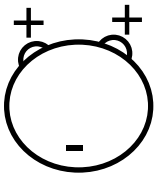
Dimensions

- protons are really small
- does water geometry matter
  - usually not



Charge

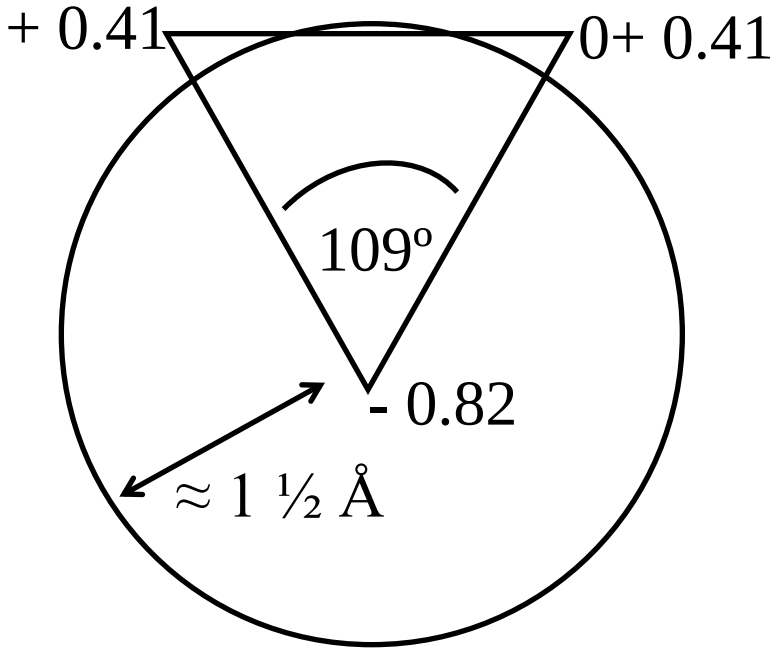
- most important



Final result..

A useful explicit water model

- 3 charges
- 1 Lennard-Jones radius
- 3 masses - why ?
  - only for molecular dynamics
- 3 bonds (completely rigid)
- has a name – “SPC”
- what can it do ?
  - diffusion, density, compressibility, heat capacity
  - dielectric constant
  - solvation energies ?
- perfect ?
  - no
    - add polarisation, bonds, bond angles, offset charge from mass



Explicit water + protein

Protein water interactions

- via charge
- via Lennard-Jones term (r^{-12} and r^{-6})

Only average properties are interesting

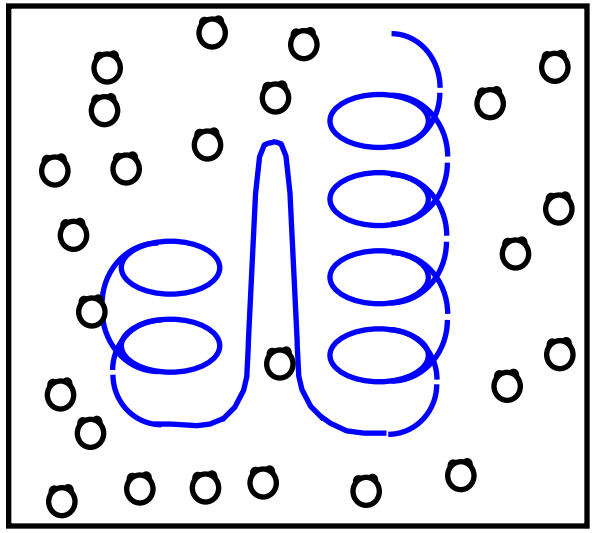
- useful only in simulations (MC, MD)

What is neat

- automatically incorporates
  - dynamic effects
  - electrostatics

Problems

- very very expensive
- typical simulation 10^3 protein atoms
- 10^4 solvent atoms



worst case for proteins + water

Imagine a world with no cutoffs for interactions

- scales as $O(n^2)$
- adding water takes 5 or 10 times as many atoms
- takes 25 or 100 times as much CPU time

Even worse

- proteins move more slowly in water (viscosity)

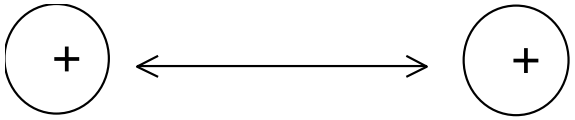
What to do ?

- look for cheaper model

Cheaper water models

- Do we really need dynamic effects of water ?
  - maybe not
    - only want energies
    - only care about structures
  - or
    - model with a random force
- then look for model which gets most essential aspects of water
  - electrostatics
    - distance dependent dielectric
    - reaction field
    - surface area methods

Distance dependent idea

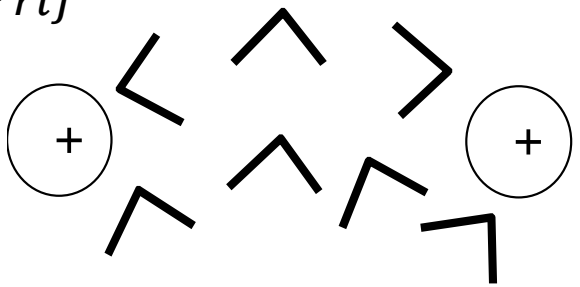


$$U(r_{ij}) = \left(\frac{1}{4\pi\epsilon_0}\right)\frac{q_i q_j}{r_{ij}}$$

Bare charges

$$\equiv \frac{q_i q_j}{D r_{ij}}$$

With solvent $U(r_{ij})$ changes less than $\frac{q_i q_j}{D_{rij}}$



Net effect ?

- water is very polar and tends to orient itself around charges
- as if the water "screened" the charges (makes them smaller)

Distance dependent dielectric implementation

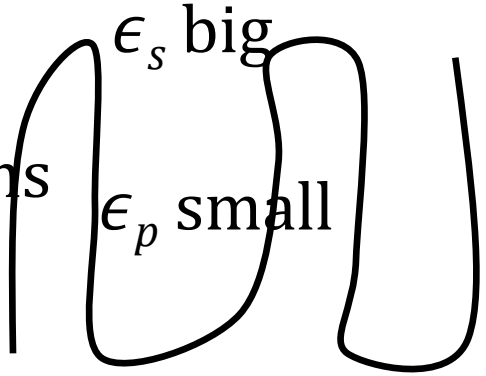
- invent approximation $D_{eff} = r_{ij}$ then

$$U(r_{ij}) \approx \frac{q_i q_j}{D_{eff} r_{ij}} \approx \frac{q_i q_j}{r_{ij}^2}$$

- is this physics ?
  - no
- does it work ?
  - a bit (ugly)
  - little real physical basis
- water does not behave so simply
- fundamental problem...

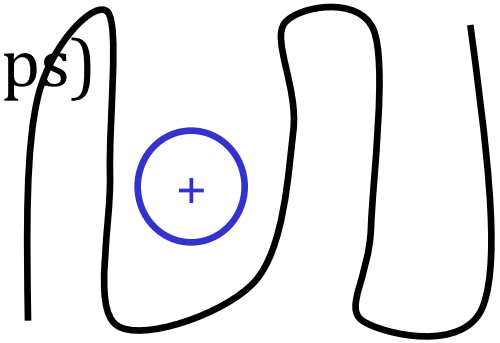
Fundamental problem with distance dependent D

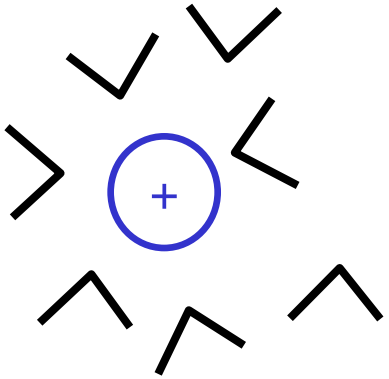
- if we rely on distance dependent dielectric constant
  - assume one 'fix' works everywhere (not true)
- think of formula $U(r_{ij}) = \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}}$
- model will differ on big and small proteins



Reaction field idea

- different problem to before
- charge in a protein (lots of neutral CH groups)
  - not much happens
- particle in water
  - what does the water do ?
  - tends to orient
  - lots of q^+q^- interactions
    - much better energy
  - is this like a force ?
    - yes, think $-dU / dr$
  - can this be modelled ?

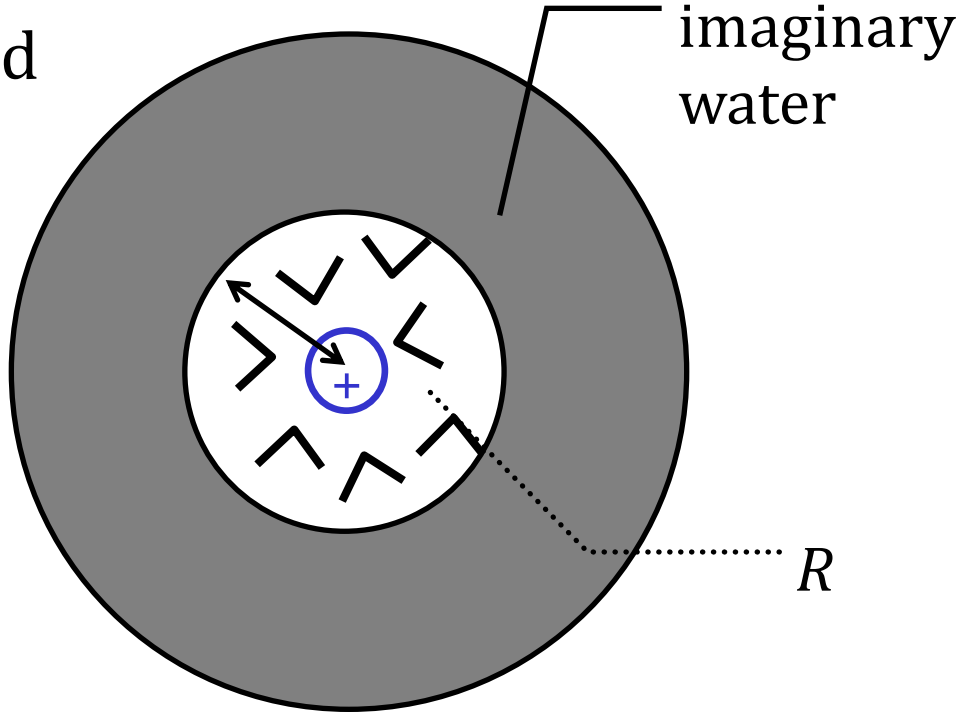




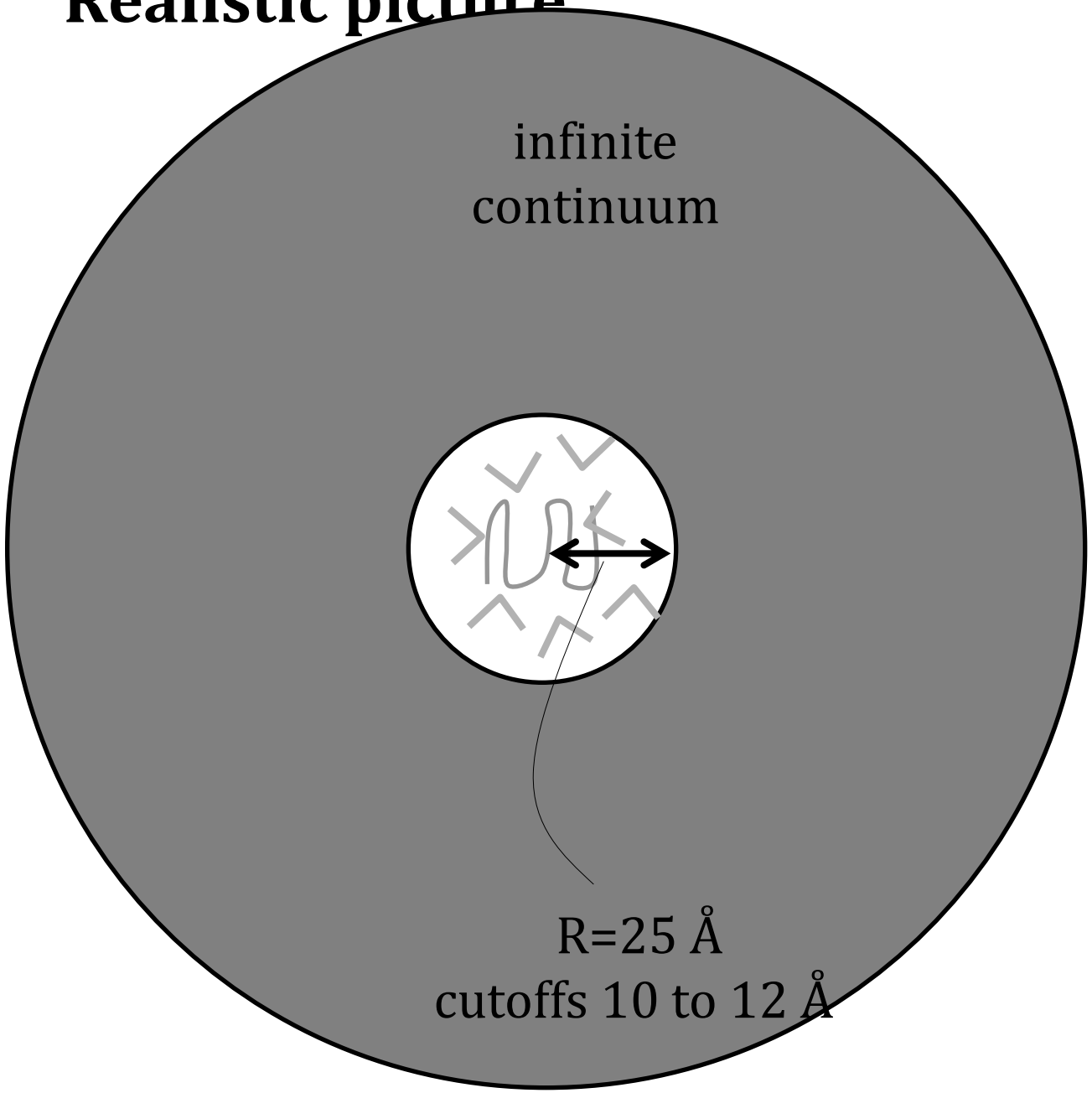
Friedman, H.L., Mol. Phys. 29, 1533-1543 (1975) Image approximation to the reaction field

interaction with imaginary solvent

- think of particle interacting with distant water molecules
- our charge interacts with them all but
  - if they are far away (big R) less important
  - depends on dielectric constant
  - inside white region ε_r and
  - grey region ε_s
- within white region
  - treat atoms with a correction
- grey region
  - treat as continuum

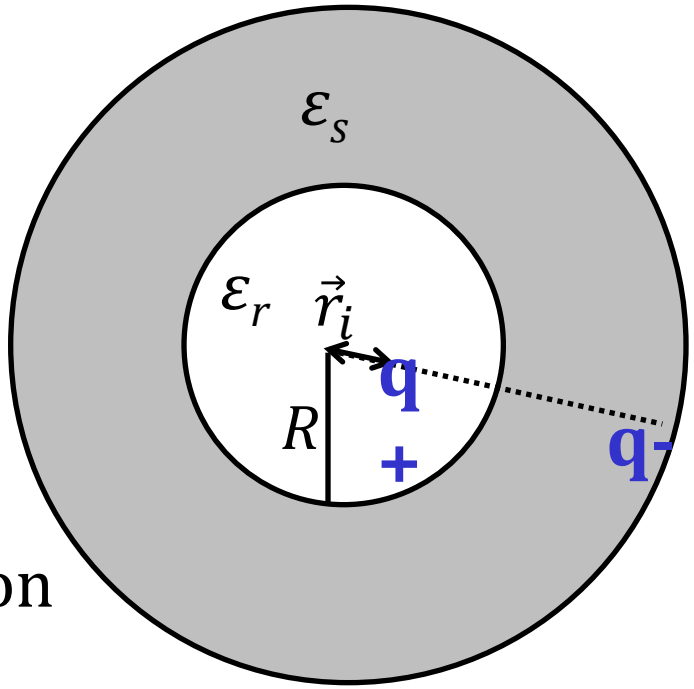


Realistic picture



Reaction field / image charge formula

- as if we interact with an "image" charge
- size $q_{im} = -\frac{\epsilon_s - \epsilon_r}{\epsilon_s + \epsilon_r} \frac{q_i R}{r_i}$
- location $\left(\frac{R}{r_i}\right)^2 \vec{r}_i$
- near middle
  - $R \gg r_i$
  - image far away
- near boundary
  - imaginary solvent important
  - strong (favourable) interaction
- important result
  - we have modelled the happiness of a charge in solution
  - charges happiest on outside of protein



Reaction fields and pairs

- charge q_i interacts with water
- water responds
- q_j feels effect of water

- no longer $U(r_{ij}) = \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}}$

- instead, $U(r_{ij}) = \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}} \cdot f(q_i, q_j, \text{distances to centre}, \ldots)$

Simpler ways to model solvent

Problem with real physics

- if you use this model, you are obliged to use
  - real charges, real coordinates...
- parameters not perfect
- hard to rationalise repairs

Many effects simultaneously

- charges interacting with water dipoles
- loss of water – water interactions
- change of solvent entropy
- change of solute entropy ?

Different approach

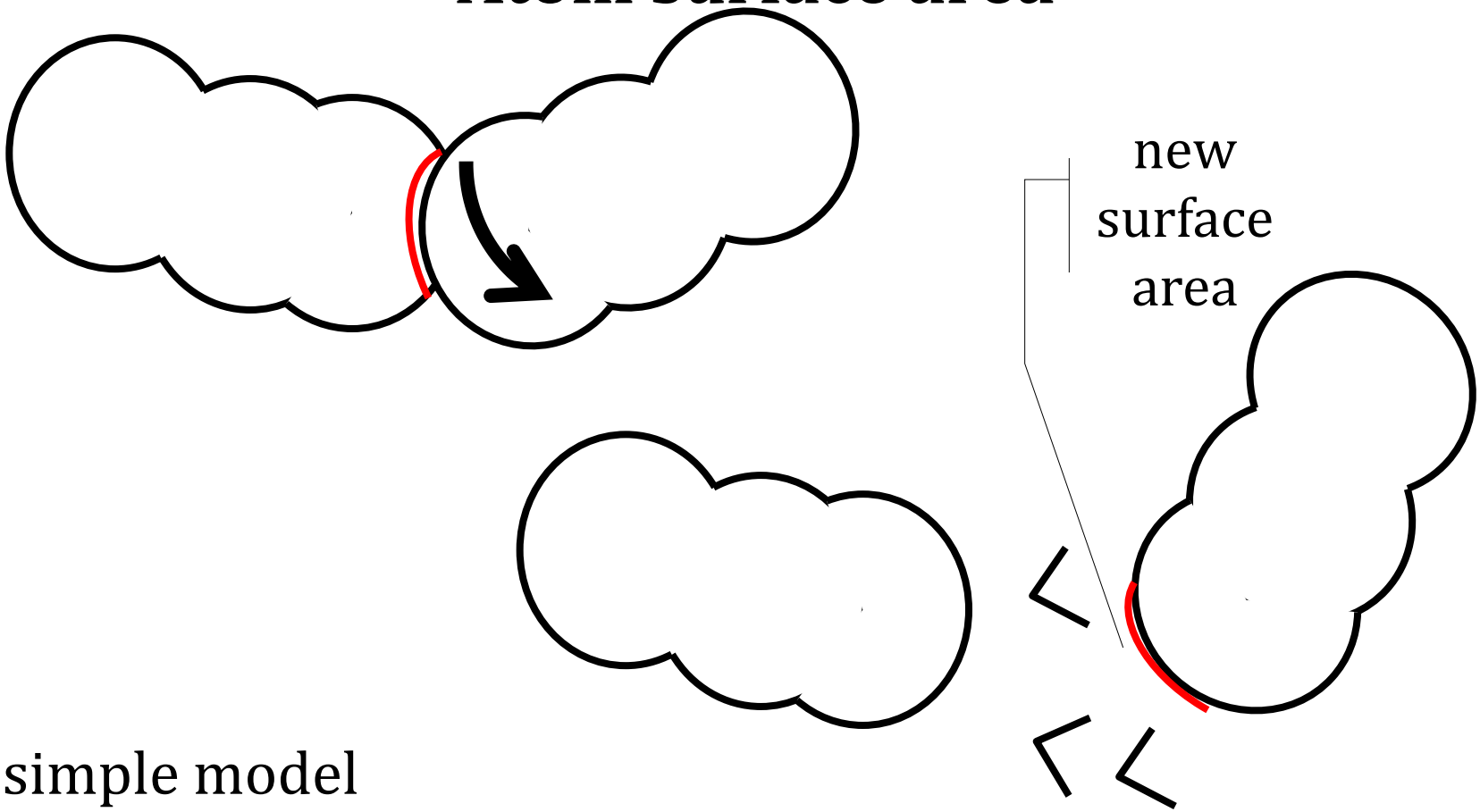
- less rigorous models

Basis of quick water model

Philosophy

- I can not model water properly
- find a very general way to incorporate effects
- Water makes some atoms happy
- Others do not care too much
- Find some very general way to include water effects
  - whether they are favourable / unfavourable
- what is easiest way to think about water influence ?

Atom surface area



- simple model
  - for each atom, energy depends on surface area

Formalising SASA model

- Solvent accessible surface area (SASA)
- for every atom, i $\quad G_i^{solv}(\vec{r}_i) = \gamma_i A_i(\vec{r}_i)$
- G because we no longer have a pure potential energy
- $G_i^{solv}(\vec{r}_i)$ because the energy term depends on coordinates
- γ_i is a specific parameter for each kind of atom
  - for O, N will be negative
  - for CH, CH_2, CH_3 will be positive or near zero
- area, A_i, has to be calculated

Problems

- A_i is difficult to calculate
  - use approximation
- γ_i not easy to estimate

Example SASA calculation[1]

- classical atomistic force field
- distance dependent dielectric
- two γ_i parameters, $\gamma_{C,S}$ =0.012 and $\gamma_{O,N}$= – 0.060 kcal mol^{-1}

Results

- better than *in vacuo*
  - deviation from known structure during simulation
  - not too many H-bonds formed
  - radius of gyration ? (how big is protein)
- why do they appear OK ? why only two γ_i ?
  - not tested in detail
  - worst problems fixed

[1]Ferrara, P, Apostolakis, J and Caflisch, A. 2002, Proteins, 46, 24-33

summary

- Explicit water is best, but expensive
- We have not discussed dynamic effects
- distance dependent dielectric +
  - SASA style models
  - complementary
- many variations
  - surface accessible volume
  - more γ_i parameters
  - add in reaction field for better long range electrostatics
- changes and flaws in one parameter are hidden by others