Molecular Modeling 2020 -- lecture 18 … Tues Mar 31
Model building with restraints

Building a small molecule

Energy

Energy minimization

What is energy?

- **Energy** (G) is a measure of the **probability** of the **state of the system**. Energy is the **negative log of the probability ratio, times temperature**.

- $\Delta G = -RT \ln (A / \text{not } A)$
  or $-RT \ln(P / (1-P))$, where P = probability.

- The system = the atoms.

- State = where the atoms are.
  (This is a vague definition so we can be flexible about what the energy means.)

- Energy is always **relative** (see fig).

- Energy is measured between **two** states.

- Energy is expressed in J/mole, or kJ/mole.

- Energy breaks down into **enthalpy** (H) and **entropy** (S).
  $\Delta G = \Delta H - T\Delta S$.

- Energy also breaks down to **potential** energy and **kinetic** energy.
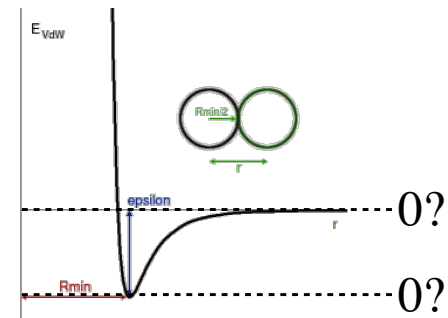
# The reference state is not a physically possible state for a protein.

- Reference state is the state at which energy is zero.

- For bonds, bond angle, dihedral angle, improper angle and planarity, the reference state is the ideal distance, angle, or planarity.

- For non-bonded interactions, the reference state is infinite distance.

- No protein can be drawn with ideal bond angles and infinite distance!

∴ the reference state for protein is not a *real* state.
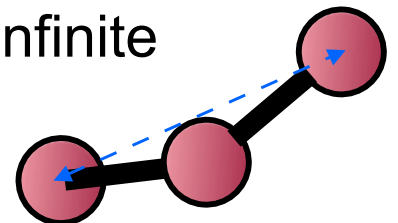
What good is the number if the reference state is not physical?

- Energy calculations should not be used "at face value". In other words, a negative number does not mean your molecule is stable. A positive number does not mean it is unstable.

- Instead, energy calculations should be used in a <u>relative sense.</u>
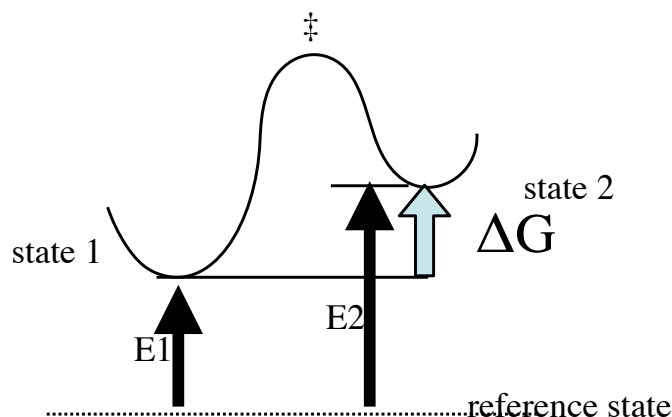
» If state 1 has energy E1 relative to the reference state, and state 2 has energy E2 relative to the reference state. Then the difference between state 1 and state 2 is independent of ht ereference state. ΔG = E2 - E1.

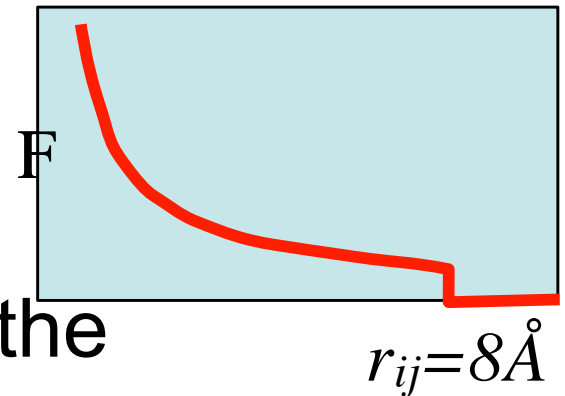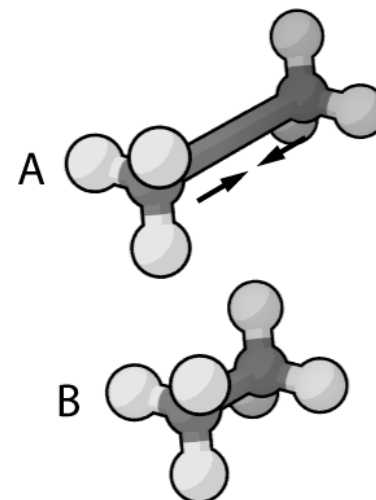# Electrostatics are truncated

but should it be?

- Pairwise calculations go up with the square of the number of atoms

$$F = k_e \frac{q_1 q_2}{r^2}$$

- So, we set an upper limit on the distances that will be calculated.

- After all, the Coulomb term goes to zero....

F

$r_{ij}=8\overset{\circ}{A}$

- Or does it?

- Electrostatic *force* goes **down** with the square of the distance, but *3D space* goes **up** with the square of the distance.
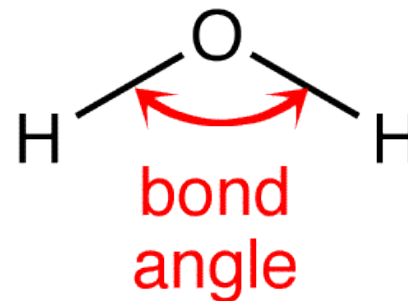
Review of force fields: molecular mechanics

- Molecular mechanics
  - bond length
  - bond angle
  - dihedral angle
  - improper angle (chirality)
  - planarity

A

B

Review of force fields: molecular mechanics

- Molecular mechanics
  - bond length
  - bond angle
  - dihedral angle
  - improper angle (chirality)
  - planarity

Review of force fields: molecular mechanics
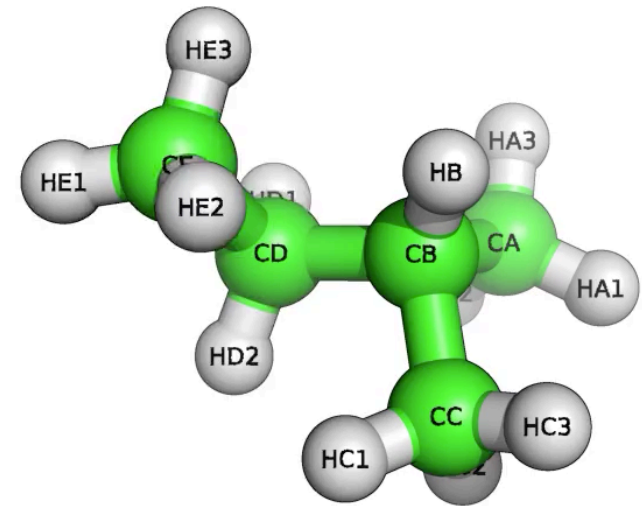
- Molecular mechanics
  - bond length
  - bond angle

  - dihedral angle
  - improper angle (chirality)
  - planarity

Review of force fields: molecular mechanics
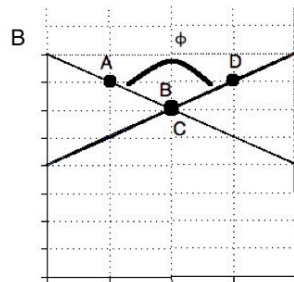
- Molecular mechanics
  - bond length
  - bond angle
  - dihedral angle

  - improper angle (chirality)
  - planarity

isoleucine beta carbon is chiral

Review of force fields: molecular mechanics

- Molecular mechanics
  - bond length
  - bond angle
  - dihedral angle
  - improper angle (chirality)
  - planarity

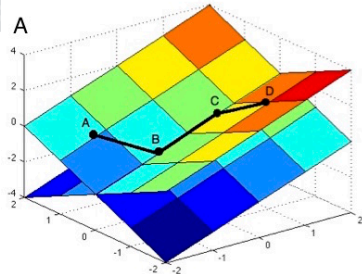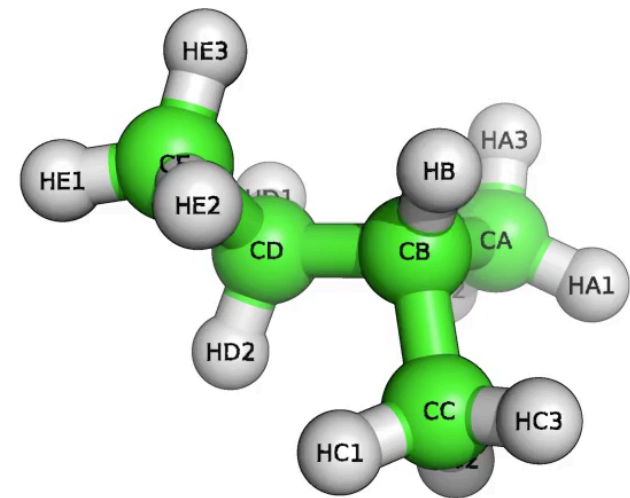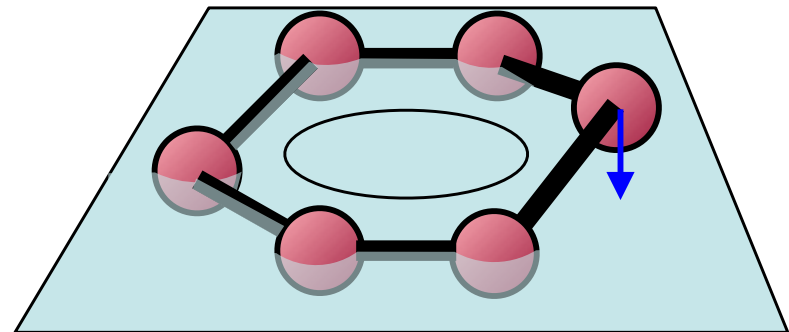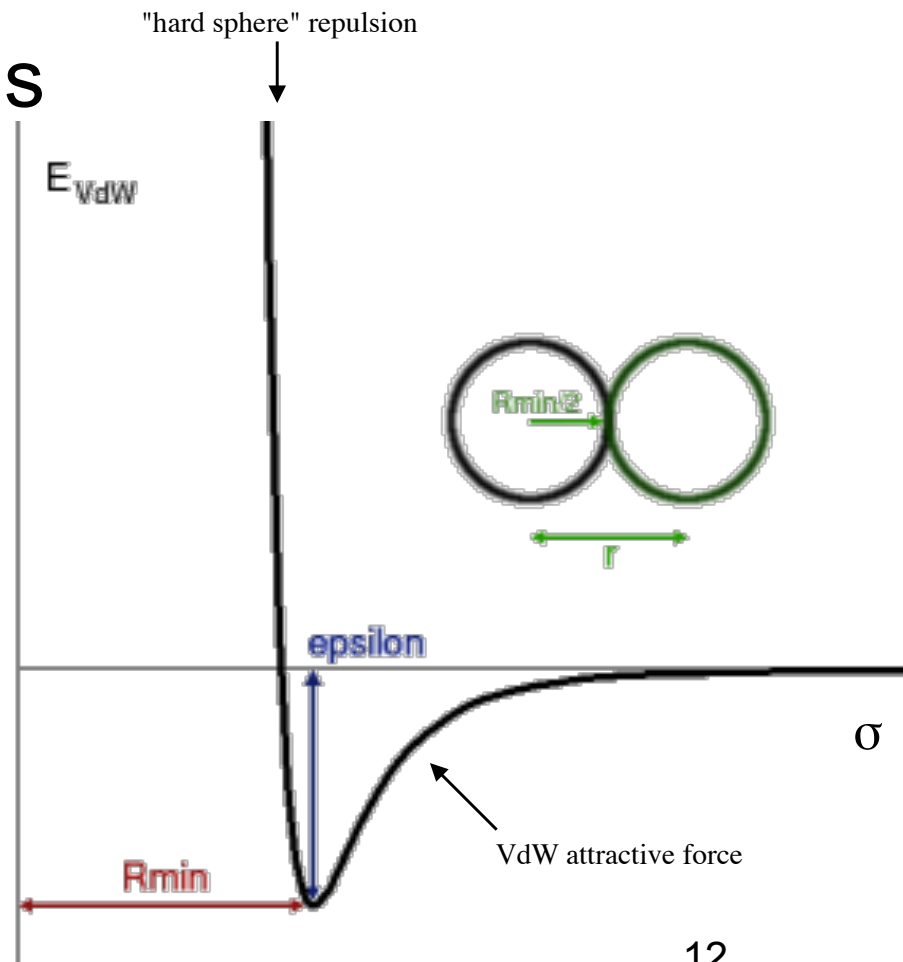# Review of force fields: molecular mechanics

- Molecular mechanics

- Non-bonded interactions
  - electrostatics
  - van der Waals

$$F = k_e \frac{q_1 q_2}{r^2}$$
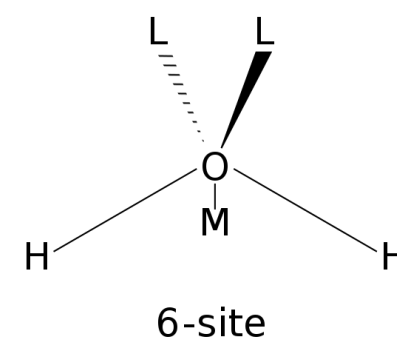
Review of force fields: molecular mechanics

- Molecular mechanics

- Non-bonded interactions
  - electrostatics

  - van der Waals

$$V_{\mathrm{LJ}} = 4\varepsilon \left[\left(\frac{\sigma}{r} \right)^{12} - \left(\frac{\sigma}{r} \right)^{6} \right]$$

"hard sphere" repulsion

E_{VdW}

Rmin/2

r

epsilon

σ

VdW attractive force

Rmin

12

Review of force fields: solvation

- Explicit solvation
  - water models
    - TIP3P
    - TIP4P
- Implicit solvation

3-site   4-site   5-site   6-site
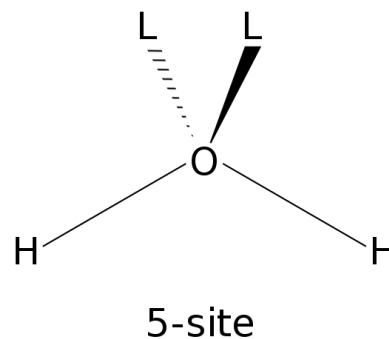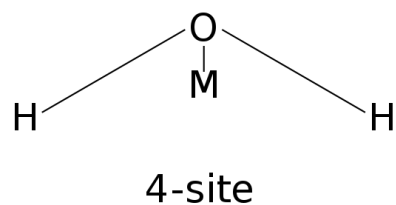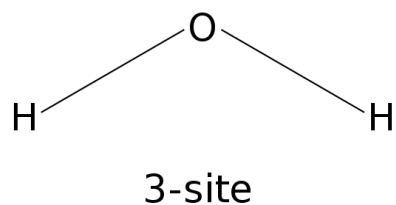
Review of force fields: solvation

- Explicit solvation

- Implicit solvation
  - Distance dependent dielectric
  - Poisson-Boltzmann
  - Generalized Born
  - Accessible surface area

water
Continuum solvent model

Hydrophobic effect is roughly proportional to surface area

torsion angle

Lennard Jones

Distance

bond length or 3-atom angle
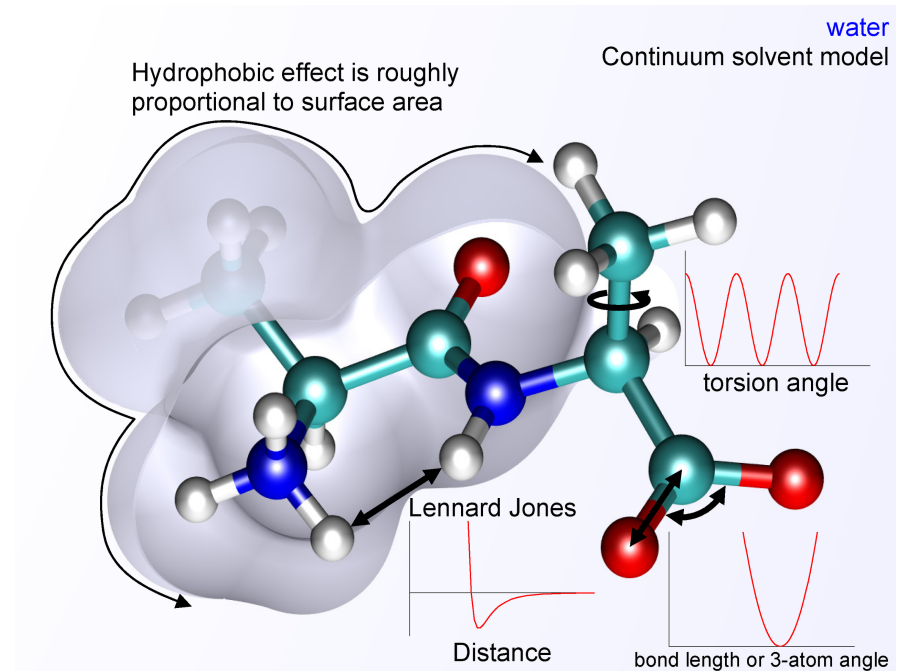
Three energetic terms that are <u>not properly calculated</u> in protein force fields.

- Electrostatics
- H-bonds
- The hydrophobic effect

Electrostatics are truncated
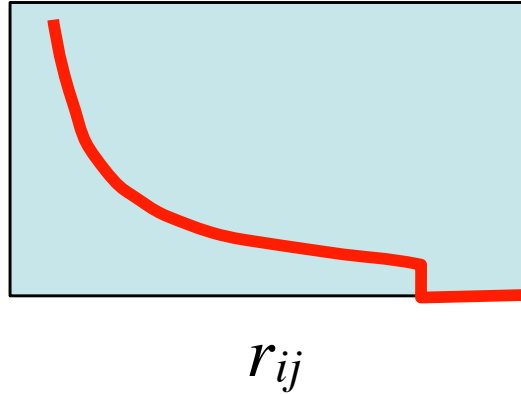
Calculated Atom pair force

$$F = k_e \frac{q_1 q_2}{r^2}$$

|F|

r_{ij}

Cutoff to zero force
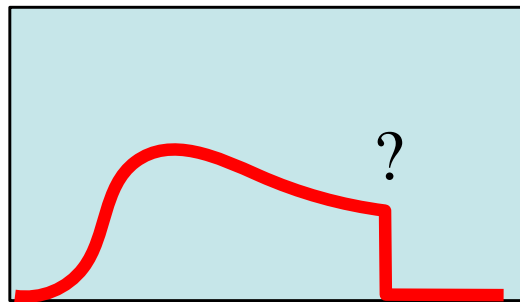at $r_{ij} = 8\overset{\circ}{A}$

Volume

r_{ij}
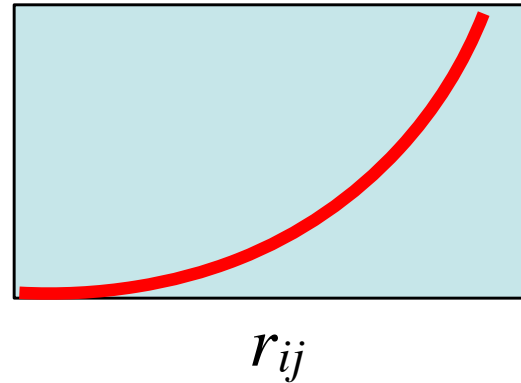
Number of atoms
goes up quadratically
with distance.

Calculated Total Σ |F|

$$\phi_i = \frac{1}{4\pi\varepsilon_0} \sum_{j=1(j\neq i)}^{N} \frac{Q_j}{r_{ij}}.$$

?

Although F is close to
zero at $> 8\overset{\circ}{A}$, total sum of
forces at $8\overset{\circ}{A}$ may be
significantly non-zero.

16

Electrostatics are truncated

$$\phi_i = \frac{1}{4\pi\varepsilon_0} \sum_{j=1(j\neq i)}^{N} \frac{Q_j}{r_{ij}}.$$

- Energy is the integral of the forces.
- The cumulative effect of many distant charges is **not negligible**.
- Case in point, highly charged proteins use electrostatics to attract ions from a distance (*e.g.* the enzyme *superoxide dismutase*).
- Other case in point. Highly charged proteins repel each other at long distances, increasing solubility. Decreasing aggregation.

come here.

ok

get away. far away.

not just 8Å. ok?

ok
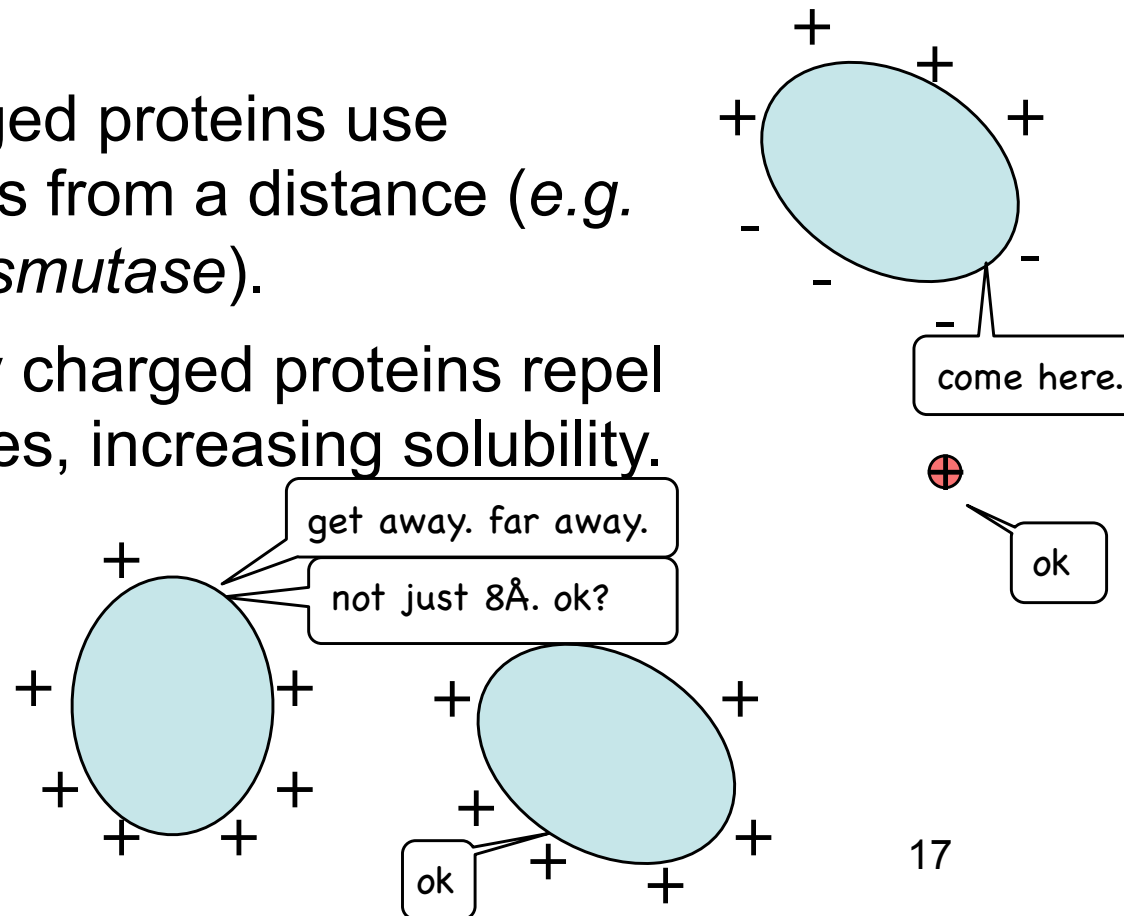
H-bond strength depends on environment

- H-bond donors have a polar H (usually N, or O)
- H-bond acceptors have lone-pairs (usually O, or N)
- Together they form a hydrogen bond.
- Part electrostatic, part covalent.
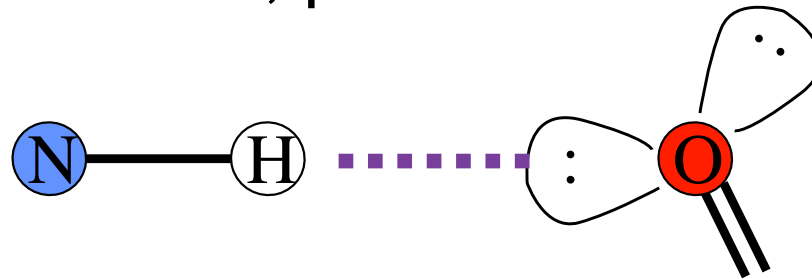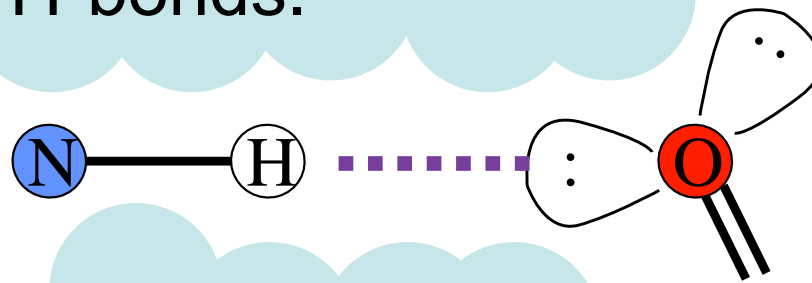
- Because charge-charge interactions are stronger when **buried**, buried h-bonds are **stronger** than solvated H-bonds.

Energy of unsatisfied H-bonds not calculated

- H-bond donors and acceptors do not want to be left **unsatisfied**.

- Force fields don't penalize unsatisfied H-bond donors/acceptors, unless a long MD simulation is carried out.

- If MD is not considered, then **buried** and **exposed** unpaired donors and acceptors are assigned the **same energy,** which is **wrong!**

buried, unsatisfied h-bond acceptor

not donor atoms
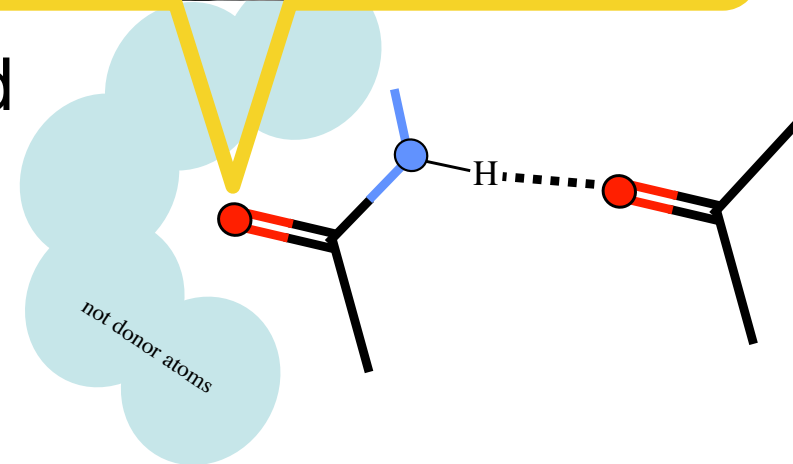
Energy of unpaired or paired buried charges not properly calculated

- The energy (ΔH) of paired positive and negative buried charges (a "salt bridge") is overestimated by forcefields. (It is really <u>more</u> negative)
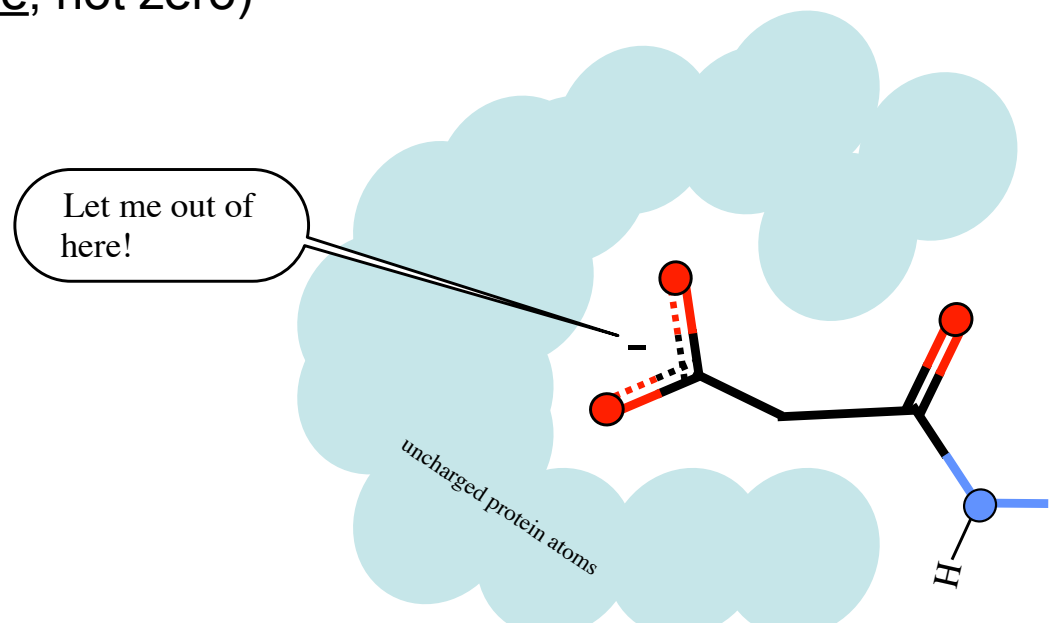
$$F = k_e \frac{q_1 q_2}{r^2}$$

  The energy (ΔH) of <u>un</u>paired positive or negative buried charges (see fig) is underestimated by forcefields. (It is really <u>positive</u>, not zero)

The Hydrophobic Effect : As hydration spheres coalesce, volume decreases, free energy decreases

Solvent accessible positions (dashed line) around non-polar atoms contain "high energy waters" because those waters lose some H-bonds.
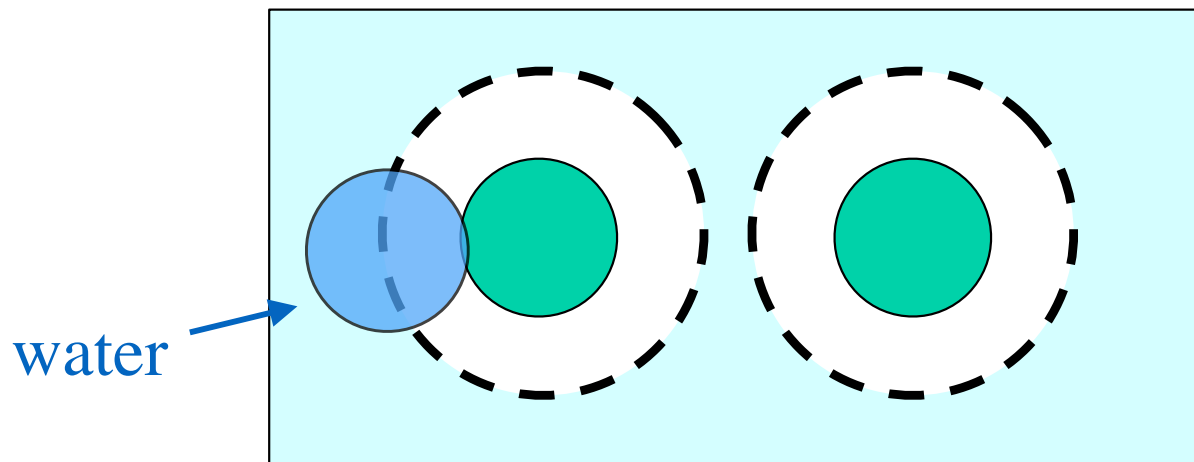
water

When non-polar atoms come together it decreases the number of high energy waters. (Even at the cost of creating some void space (brown).
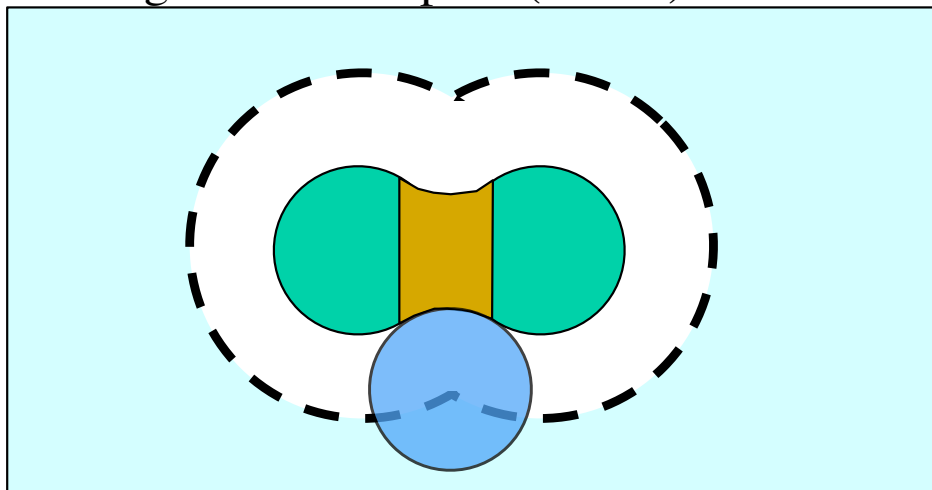
Solvent-excluded surface (SES) is a good estimator of hydration layer volume.

The Hydrophobic Effect : an emergent property of water.

- The hydrophobic effect is expressed if waters are modeled and a long simulation is done, because water is naturally attracted to water .
- Thus, the hydrophobic effect is an ***emergent*** property of a long simulation.
- It is **not a bug** that the hydrophobic effect is not in the force field explicitly.
- Don't expect energy minimization alone to bury hydrophobic surfaces. It won't. You have to do it manually.

Other imperfections in molecular force fields

- <u>Partial charges</u> are calculated, but are not allowed to change dynamically. They do change! But not much.
- <u>Dihedral angles</u> are poorly modeled by a cosine function. The true barrier to rotation depends on "1-4" interactions. This usually does not matter.
- Overpacking a protein core (*i.e.* when designing a protein) makes a protein unstable, but the calculated E_{VDW} (or E_{LJ}) energy is better! Why? Because dynamic movement is ignored. More movement means more <u>entropy</u>.

What do we do about it?

- We remain vigilant!

- If you find a buried, unsatisfied H-bond, <u>satisfy it</u> or understand that it is high-energy.

  - either move atoms or add a water.

- Minimize empty space between side chains in the core but don't overpack.

- Be aware that long-range electrostatics are not calculated. Visualize electrostatic surfaces to predict long-range behavior.

24

Restraints: energy minimization helpers.

constraints versus **restraints**

restraint = a function that approaches a minimum as the parameters approach ideal values.

For example, the bonded distance A-B is restrained to 1.52Å using the restraint $E(A,B) = (D_{AB} - 1.52)^2$

—- versus —-

Distance D_{AB} from atom A and atom B

constraint = a function that reduces the number of variable parameters in the system.

For example, atoms A,B,C and D are constrained to be in the same plane. Move atoms, then solve for the constrained atom position.

Stereochemistry energy functions are *restraints*.

Harmonic and non-harmonic restraints

Restraint forces are applied to move the atoms to their **ideal** distances/angles/positions/geometry.
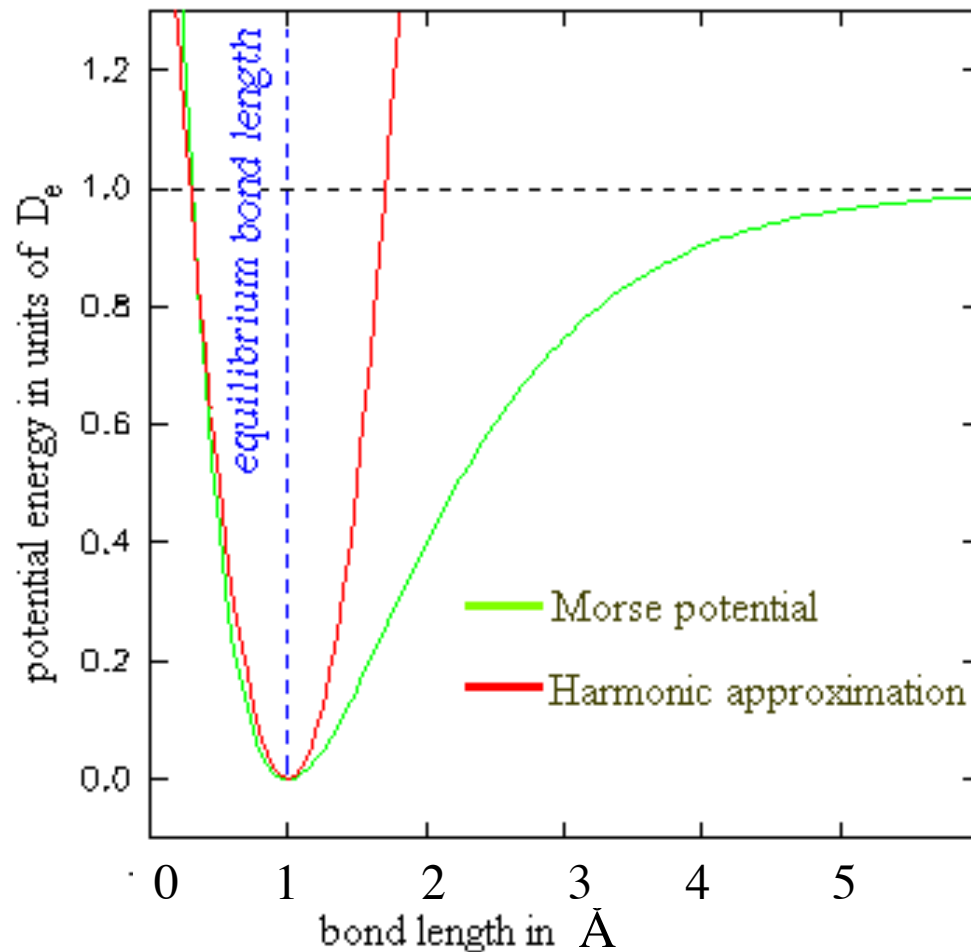

© O.S. Smart, 1995

Harmonic potential:

$$E(i,j) = \omega\left(x_{ij} - T\right)^2$$

where x_{ij} is the current distance between i and j, and T is the ideal distance between i and j.

How to force hydrogen bonds using restraints

- To add a restraint
  Edit | Potential | Restrain, distance,
     Target 1.8, 1.8, Weight 50
       Pick amide H and carbonyl O.
       Click **Create**.
  Cancel | Restrain (or esc) when done
- Energy minimize

  Compute | prepare | Structure preparation

  Checks for missing atoms, assigns energies.

  SVL: run 'gizmin.svl'

  When finished, be sure to **Cancel | GizMOE_Minimizer**
- To remove or modify restraints
  Potential setup (button at far lower left)
  Restraints tab
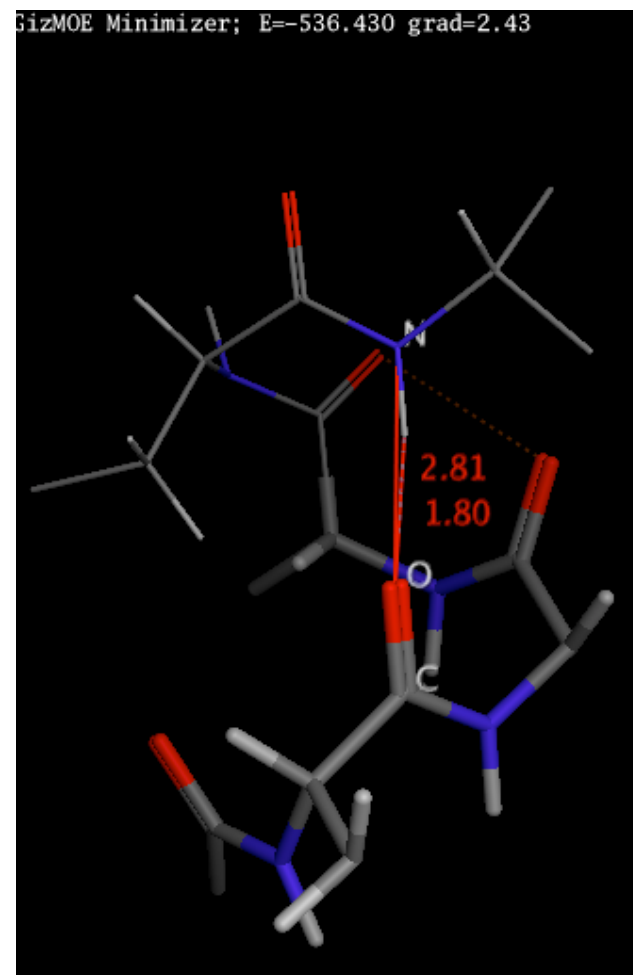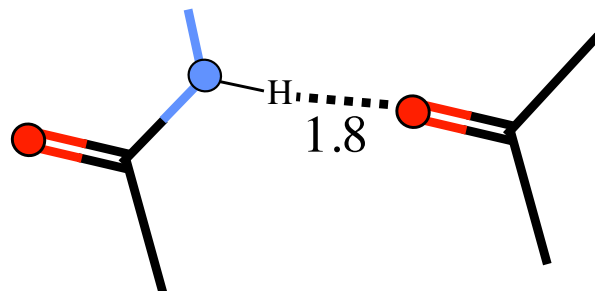
EPUSIEPF

- *Select the region you want energy minimize*

- Edit | Potential | Unfix

- Select | Invert

- Edit | Potential | Fix.

- Minimize.
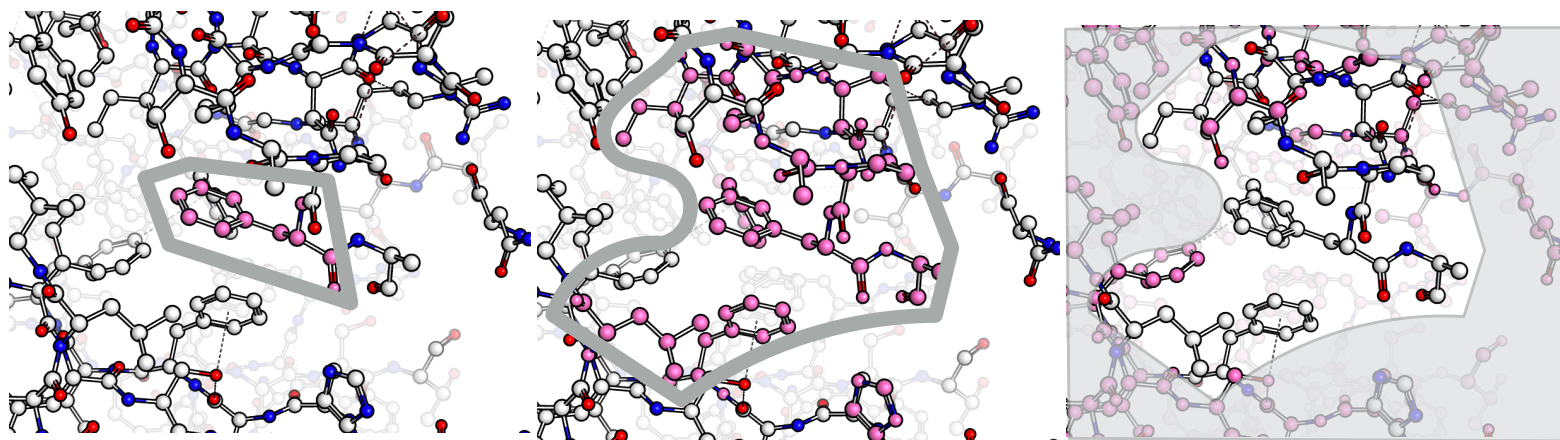
select   extend, unfix   invert, fix

review questions

- What does *sp2* hybridization mean?
- How is energy related to probability?
- What constitutes a "system"?
- Give an example of a state of a system.
- What changes when we minimize the energy? (besides the energy)
- Energy can be broken down into what two components?
- Name the molecular mechanics energy functions.
- What is a restraint, mathematically?
- The hydrophobic effect is an emergent property of what two properties of water?
- In what way are H-bonds not properly modeled?
- In what way are electrostatics not properly modeled?
- Is the high energy of a buried unsatisfied H-bond donors an emergent property?