Building a small molecule

Energy

Energy minimization

Make a beta hairpin using restraints
Exercise 18.1 Building a small molecule
Building aspartame
Building aspartame

Starting with an empty Moe window:
• Edit/Build/Molecule, or use Builder button
• Create backbone using atoms buttons: N,C,C,N,C,C,O,C (Notice the chain is made in the fully reduced state.)

• Add carbonyl oxygens: Select an H, hit O in Builder. the H becomes an O.
Building aspartame

- Select carbonyl groups. Click double bonds (=)

- Add sidechains:
  - Select the \textit{back} H on the first alpha-carbon. Click C, then -COO-
  - Select the \textit{front} H on the second alpha-carbon. Click C, then benzene.
Building aspartame

• Fix ionization of NH$_3$
Select N. In Builder, click "+1" (a proton is added)

• Fix hybridization of NH.
• Double-click second N. Choose Geometry: "sp2". Click "Apply"
• Click "Minimize".
Real-time energy minimization using GizMOE

SVL: run ‘gizmin.svl’

Scientific Vector Language window. You can write programs for MOE to run.

Runs Minimizer without stopping.

To stop Gizmin, Cancel/MOE:Giz_minimizer
..or, make *gizmin* a function key
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.

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

- 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.
Three energetic terms that are not properly calculated in protein force fields.

- Electrostatics
- H-bonds
- The hydrophobic effect
Electrostatics are truncated

• Pairwise calculations go up with the square of the number of atoms
• So, we set an upper limit on the distances that will be calculated.
• After all, the Coulomb term goes to zero....
• 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
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

iso leucine 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

\[
V_{\text{LJ}} = 4\varepsilon \left[ \left( \frac{\sigma}{r} \right)^{12} - \left( \frac{\sigma}{r} \right)^{6} \right]
\]
Review of force fields: solvation

- Explicit solvation
  - water models
    - TIP3P
    - TIP4P
- Implicit solvation
Review of force fields: solvation

• Explicit solvation

• Implicit solvation
  – Distance dependent dielectric
  – Poisson-Boltzmann
  – Generalized Born
  – Accessible surface area
Electrostatics are truncated

Calculated Atom pair force

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

Volume

\[ \text{Number of atoms goes up quadratically with distance.} \]

Calculated Total \( \sum |F| \)

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

Cutoff to zero force at \( r_{ij}=8\text{Å} \)

Although \( F \) is close to zero at > 8Å, total sum of forces at 8Å may be significantly non-zero.
Electrostatics are truncated

- 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.

\[ \phi_i = \frac{1}{4\pi\varepsilon_0} \sum_{j=1\ (j \neq i)}^{N} \frac{Q_j}{r_{ij}}. \]
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!
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 more negative)

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

The energy (ΔH) of unpaired positive or negative buried charges (see fig) is underestimated by forcefields. (It is really positive, 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.

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

- **Partial charges** are calculated, but are not allowed to change dynamically. They do change! But not much.
- **Dihedral angles** 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 entropy.
What do we do about it?
(relevant when we start designing)

• We remain vigilant!
• If you find a buried, unsatisfied H-bond, satisfy it 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.
Restraints: energy minimization helpers.
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 ---

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.
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
Exercise 18.2

Make a beta hairpin

anti-parallel sheet with valine side chains all on the same side of the sheet.

Edit | Build | Protein,  Geometry: anti-strand. Residue: ADVDKVSPNGVEVKVRA

Zoom out.

Select the second half of the chain starting with NG.

Rotate and translate it (shift-alt-middlemouse) so that the first three valines (3,5,7) are lined up with other three valines (12,14,16), and the valine backbone H-bonding groups (NH and CO) are close to the H-bonding distance (1.8Å from H to O)

Hide side chains to help see the backbone atoms better.

Edit | Potential | Restrain.  Set Target 1.8, 1.8, Weight 50. Select H and O atoms. Create.

When done you have 2 restraints for each of the three paired valines for a total of 6 restraints.

Compute | Prepare | Structure preparation.  Hit Correct if necessary.

SVL: run 'gizmin.svl'.

If there are errors in the restraints, Cancel/GizMOE, open Potential Setup (extreme lower left of the MOE window).

Restraints. Click on restraints to delete or modify them.

Restart SVL: run 'gizmin.svl'.

Look at out the structure.
It should have beta pleating when viewed from the edge of the sheet. Sidechains should alternate up and down in that view. Residues SPNG form a beta-turn.

Cancel/Gizmin . Remove the restraints. SVL: run 'gizmin.svl'.

Does the structure hold together or fall apart? It should hold together.
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?
• What is a constraint?
• The hydrophobic effect is an emergent property of what 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?
Supplementary: How to make a new peptide bond

Zoom in on splice points.

Delete extra oxygen, if present.
Select C, N. **Build | single bond.**
Select N. **Window | Atom manager** (or cntrl-a)
Select the atom. Set geometry to sp2. **Apply.**