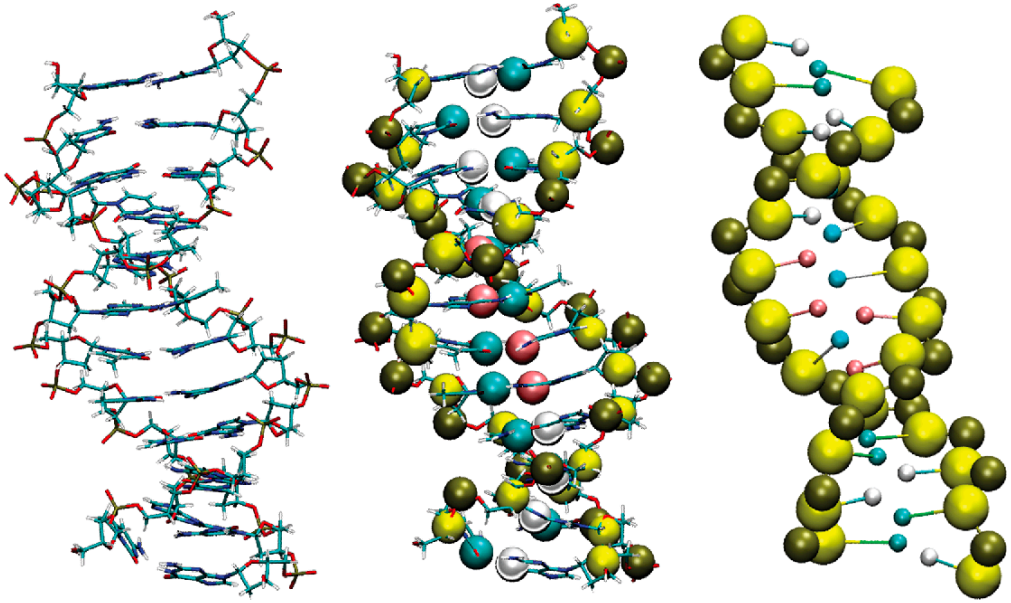
An introduction to force fields

Prof. Michael Shirts
University of Colorado Boulder
i-CoMSE Summer Workshop
Oklahoma State University, July 2022

What are we trying to simulate?

- We have some MODEL of our molecule.
- We have the hypothesis that the physical behavior only depends on SOME of the physics of the molecule
- We test that hypothesis with a simulation



- Can be atomistic, quantum mechanical, or coarse grained.
- Each particle isn't an atom, but a methane particle, or an amino acid.

i-CoSE

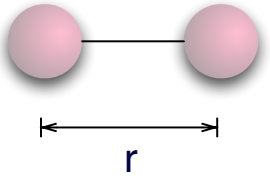
What is a force field?

- A force field is some function $U(x_1,x_2,x_3...x_N)$
  - Input: particle coordinates
  - Output: an energy
  - $-\nabla U$ (dU/dx_1, dU/x_2, $dU/x_3...dU/x_N$) is the **force** on each particle.
  - A force vector is assigned to each set of coordinates.
- Usually, a force field tries to capture the potential energy as a function of atomic nucleus coordinates.
- We AVERAGE OUT the quantum mechanics.
- Let's look at the mathematical and code description of some of the most common force fields (CHARMM/AMBER/GAFF/OPLS/OpenFF)

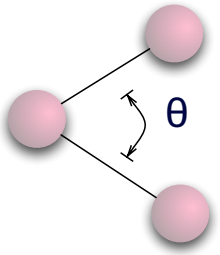
"Bonded Terms"

- We replace the electrons in orbitals with simple functional forms that are 10^6 times cheaper (but WAY more approximate).

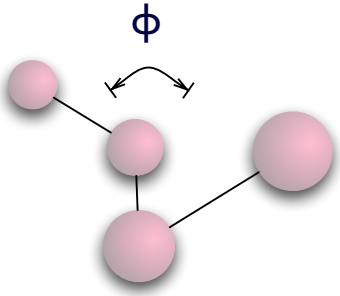
$$+ \sum_{\text{all bonds}} \frac{1}{2} K_b \left(r - r_0\right)^2$$



$$+ \sum_{\text{all angles}} \frac{1}{2} K_\theta \left(\theta - \theta_0\right)^2$$



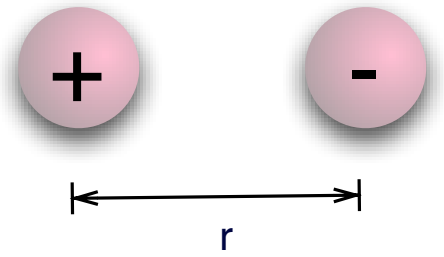
$$+ \sum_{\text{all torsions}} \sum_{n} K_\phi \left[1 - \cos\left(n\phi + \phi_0\right)\right]$$

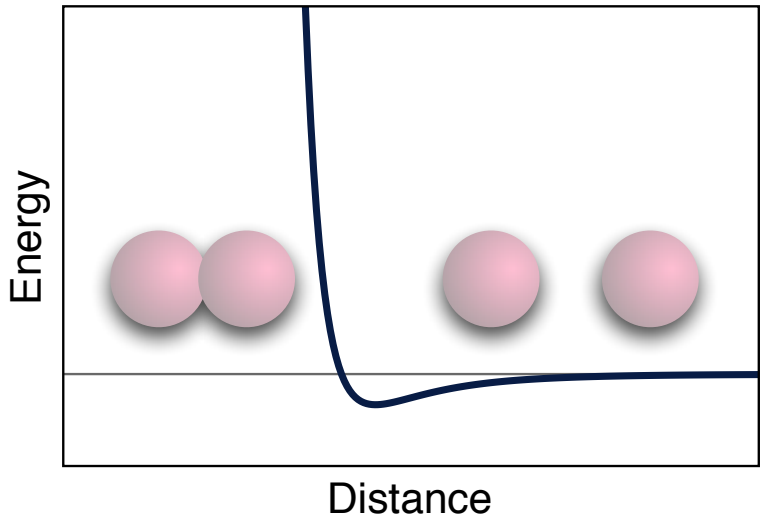


"Nonbonded terms"

- We replace the electrostatic / dipole / induced dipole interactions with physically motivated functional forms.

$$U(\mathbf{x}) = \sum_{\text{all pairs}} \frac{q_i q_j}{r} + \sum_{\text{all pairs}} \frac{A_{ij}}{r^{12}} - \frac{B_{ij}}{r^6}$$



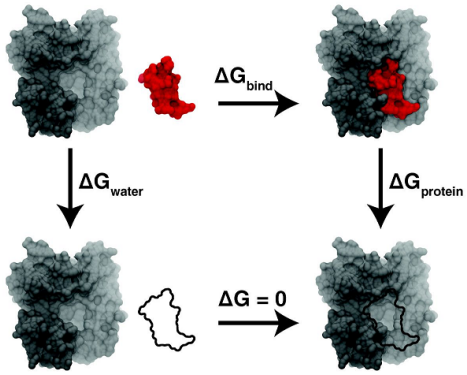


i-CoSE

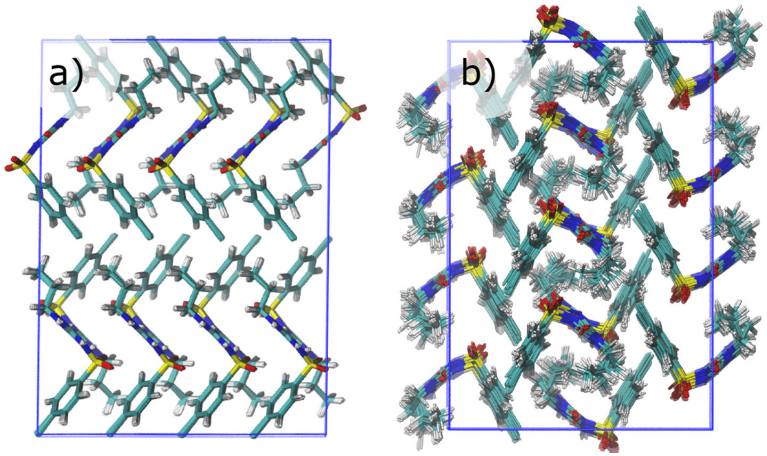
Small molecule force fields are widely-used tools in computational biophysics and soft matter systems

- Surprisingly, all-atom fixed charge force fields strike an appropriate balance between chemical accuracy and computation efficiency for many systems of interest.

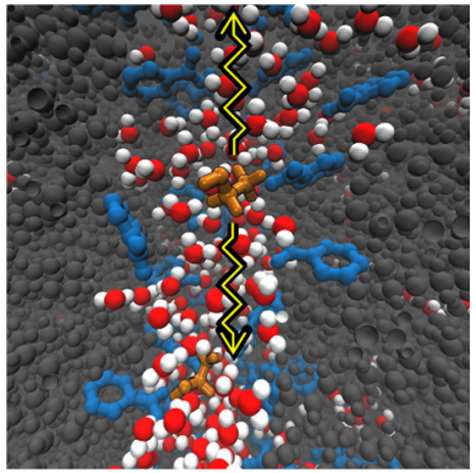
Protein-Ligand Binding



Drug molecule screening



Molecular transport

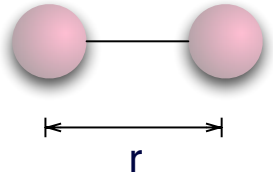


i-Co SE

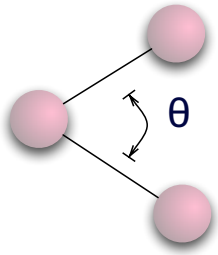
"Bonded Terms"

- We replace the electrons in orbitals with simple functional forms that are 10^6 times cheaper (but WAY more approximate).

$$+ \sum_{\text{all bonds}} \frac{1}{2} K_b \left(r - r_0\right)^2$$



$$+ \sum_{\text{all angles}} \frac{1}{2} K_\theta \left(\theta - \theta_0\right)^2$$

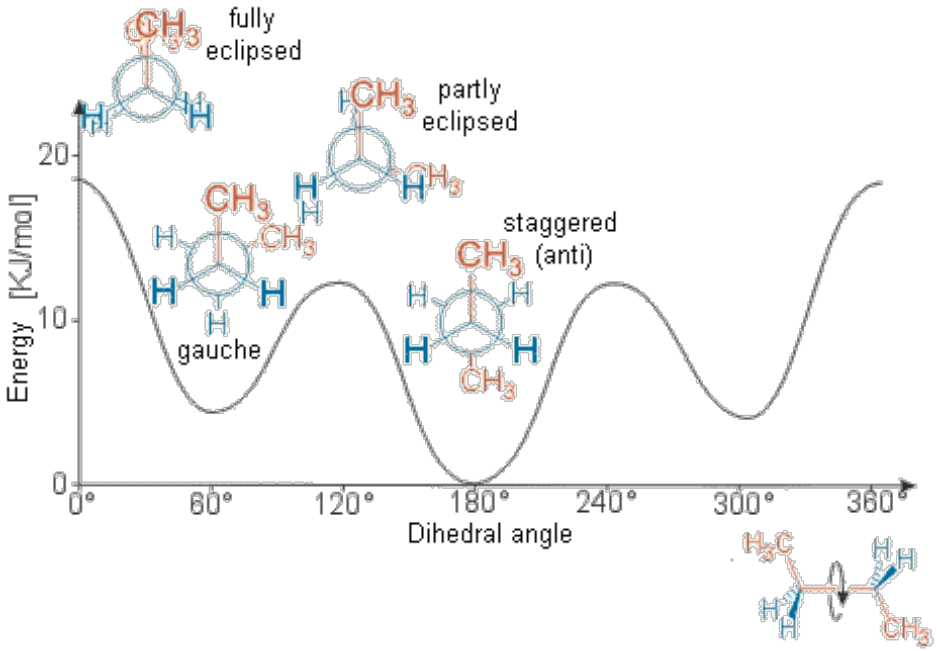


Harmonic bonds and angles
Pretty good approximation to quantum mechanics!

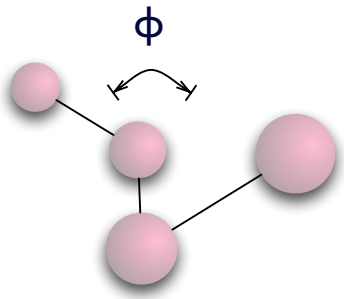
Go To Examples!

Torsions, or the junk bin of force fields

- Replace QM with a sum of periodic trigonometric functions

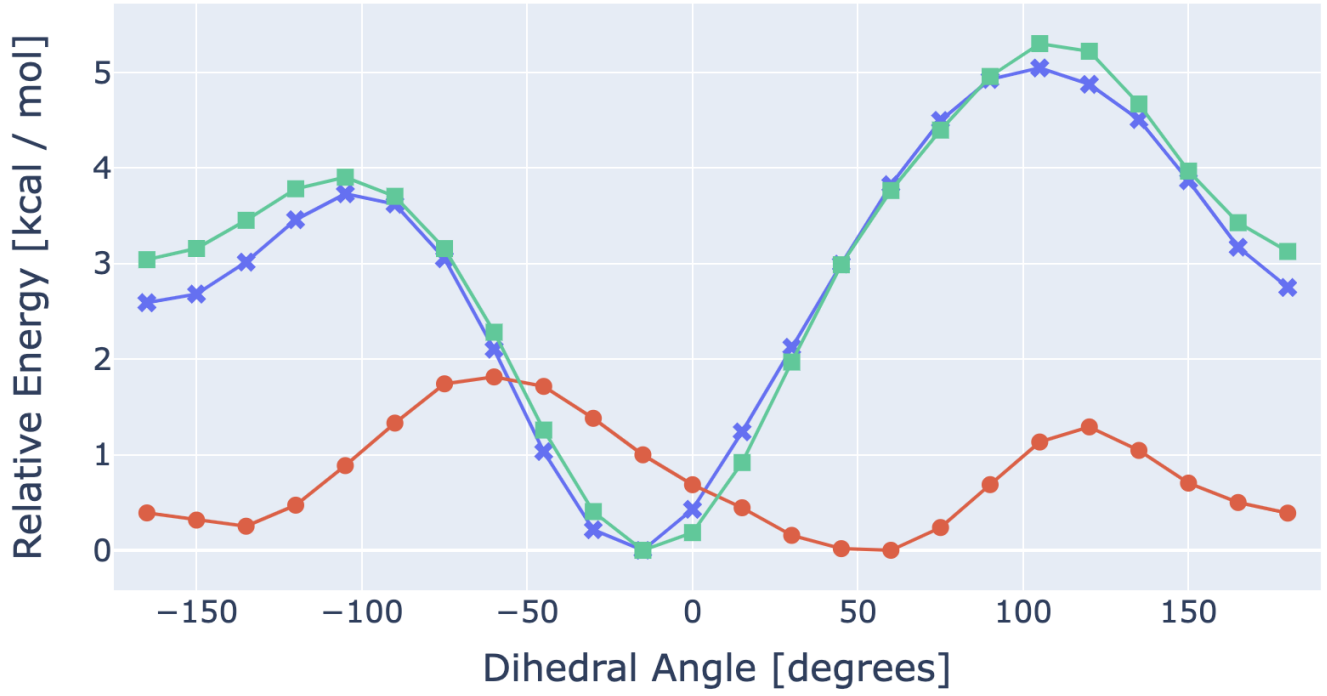


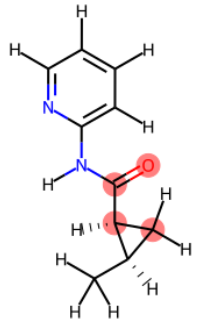
$$+ \sum_{\text{all torsions}} \sum_{n} K_\phi \left[1 - \cos\left(n\phi + \phi_0\right)\right]$$



i-CoSE

It can be difficult to get these right!

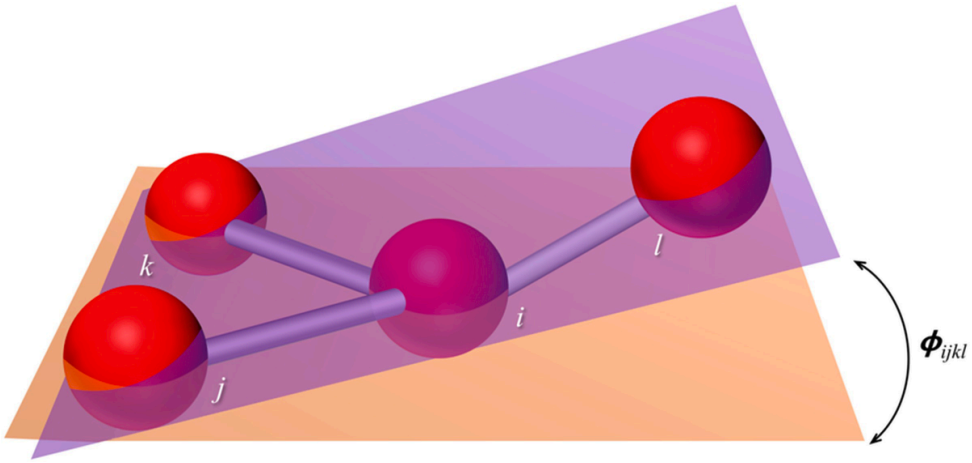


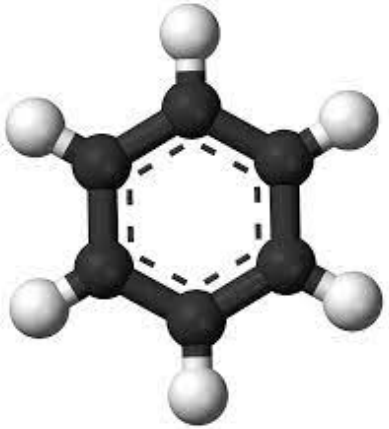


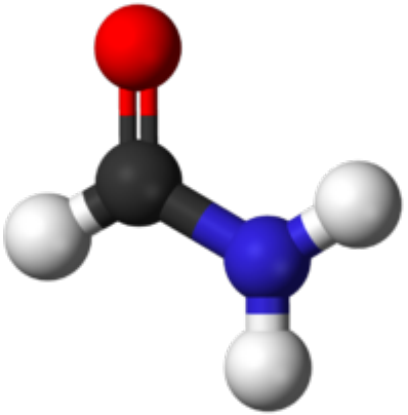
i-Cose

Improper torsions

- Ways to keep things with conjugated double bonds planar





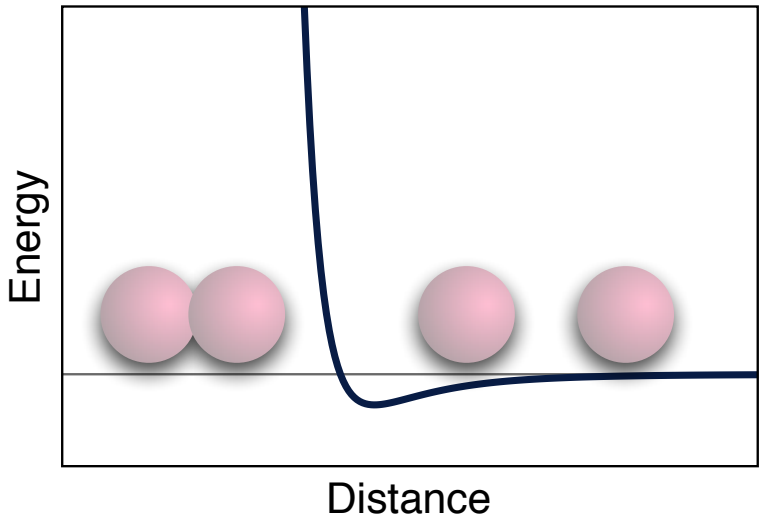


i-Co SE

Go To Examples!

Lennard-Jones terms

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



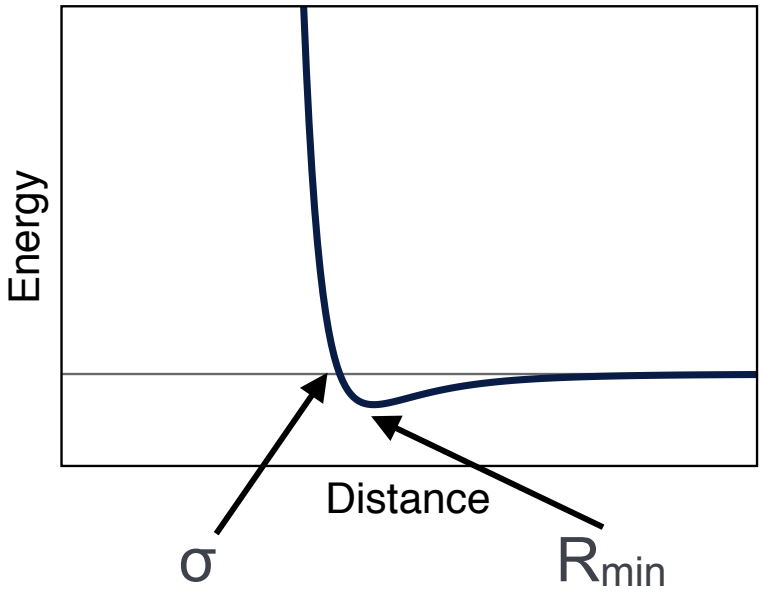

```
[ atomtypes ]
; atom type name type     mass                          sigma          epsilon
 opls_111       OW       15.99940      0.000       A    3.12171e-01  7.94960e-01
 opls_112       HW        1.00800      0.000       A    0.00000e-01  0.00000e-01
```

- This just has sigma and epsilon for O and H. What about the O-H interaction???

Lots of different, equivalent ways to write Lennard-Jones parameters!

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

$$U(r_{ij}) = \frac{A_{ij}}{r^{12}} - \frac{B_{ij}}{r^{6}}$$



$$R_{min} = 2^{\frac{1}{6}}\sigma$$

or

$R_{min/2}$

Combining rules

```
[ defaults ]
; nbfunc        comb-rule       gen-pairs       fudgeLJ fudgeQQ
1               3               yes             0.5     0.5
```

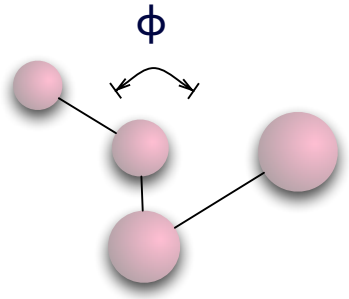
Geometric $\epsilon_{ij} = \sqrt{\epsilon_i \epsilon_j}$ $\sigma_{ij} = \sqrt{\sigma_i \sigma_j}$

Lorentz-Berthelot $\epsilon_{ij} = \sqrt{\epsilon_i \epsilon_j}$ $\sigma_{ij} = \frac{1}{2}\left(\sigma_i + \sigma_j\right)$

Works amazingly well! (but not perfect)

Exclusions?

- What about atoms separated by no bonds (1,2)?
- What about atoms separated by 1 bond (1,3)?
- What about atoms separated by 2 bonds (1,4)?




```
[ defaults ]
; nbfunc        comb-rule       gen-pairs       fudgeLJ fudgeQQ
1               3               yes             0.5     0.5
```

Go To Exercise!

Other terms you could put in:

- Hydrogen bonding terms
- Polarizability: The ability for charge distributions to rearrange.
- Other many body terms (charge penetration)
- Force field that break bonds (ReaxFF)
- Machine learning terms (Ani neural net potentials, others)

- BUT pairwise nonbonded potentials (plus some multibody bonded terms) are **much** faster than general multibody terms for both MD and MC.

- As long as you have a function $U(x_1,x_2,x_3...x_N)$, an energy output with an coordinate input, you can do molecular dynamics (or Monte Carlo!)

What sorts of force fields are out there?

- OPLS-AA / OPLS-UA
  - Small molecules, proteins
  - Evolved out of Bill Jorgensen's work
  - Now run by Schrodinger
  - Biomolecules and small molecule ligands

- TRAPPE
  - Coming mostly from Ilja Siepmann's group
  - Designed for phase equilibrium

Other Force Fields

- CHARMM
  - Evolved from the work in Martin Karplus's group at Harvard over ~50 years.
  - Alex MacKerrell (Maryland) and Benoit Roux (Chicago) are longest contributing developers but MANY more.
  - Proteins, lipids, nucleic acids small molecules

- AMBER
  - Evolved from the work in Peter Kollman's group at UCSF, and his academic descendants.
  - GAFF small molecule force field
  - Proteins and nucleic acids

Other Force Fields

- Open Force Field
  - http://www.openforcefield.org
  - New effort, focused on open software as well as open force fields
  - Small molecule based, in the process of adding biopolymers
  - Can build from SMILES or mol2
  - Difference is: it matches molecular parameters in a more general way than foyer

Other Force Fields

- What about polymers?
  - OPLS often used.
  - COMPASS
  - MMFF

- What about metals?
  - Lots of metal specific functional forms.

Force Field Parameterization Tools

- AmberTools/tleap
- CHARMM-GUI/CGENFF
- OpenFF toolkit
- foyer
- pysimm (polymers)
- OpenKIM (metals)

But What Force Field Should I Use?

- The one that allows you to test the hypothesis that you want to

OR

- The one that is validated as an effective screening mechanism to predict the property of interest to other researchers.

- Has the force field I want to use been validated on sufficiently similar systems?
- "All models are wrong. Some models are useful."
-