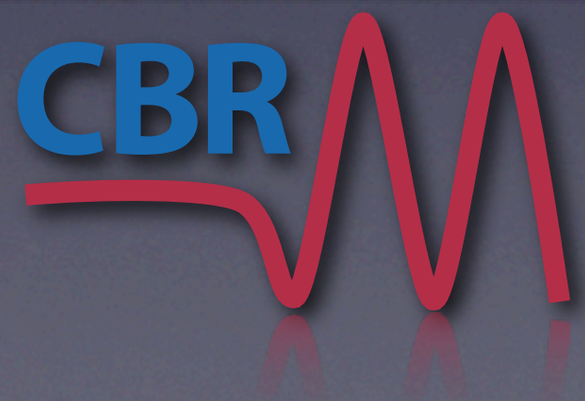
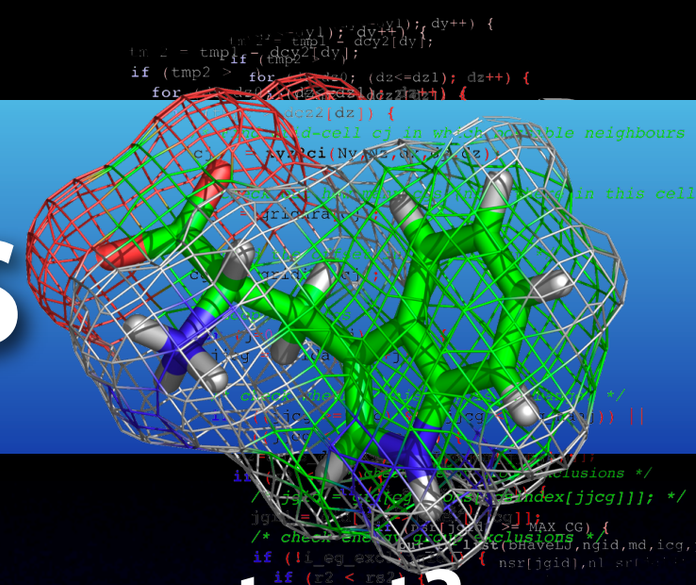
Simulations of Membranes & Membrane Proteins

Erik Lindahl

lindahl@cbr.su.se

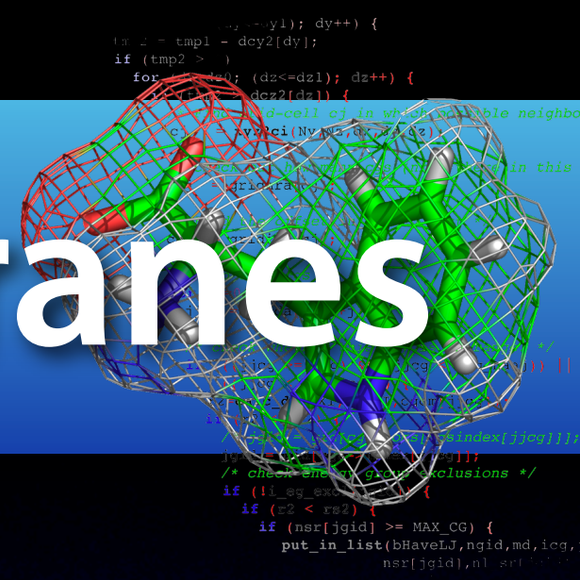


Outline: Membranes



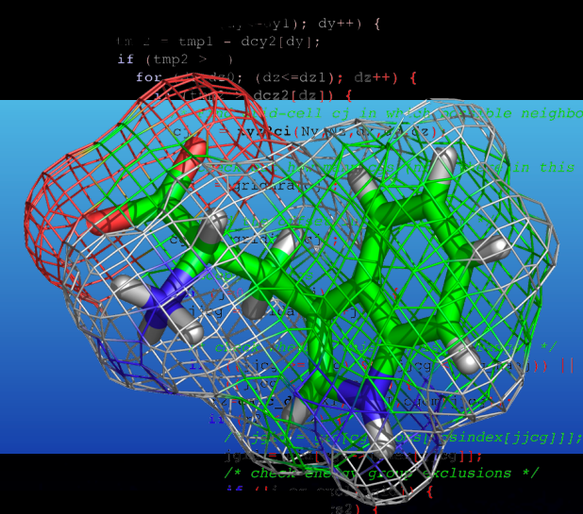
- Why are membranes interesting/important?
- Lipids - amphiphilic molecules
- Micelles, bilayers, vesicles, mixtures
- Bilayer phases (crystal, gel, liquid crystalline)
- Special difficulties for simulations
- Bilayer simulations & typical systems
- Membrane proteins, transmembrane helices
- Membrane protein simulations

Importance of Membranes

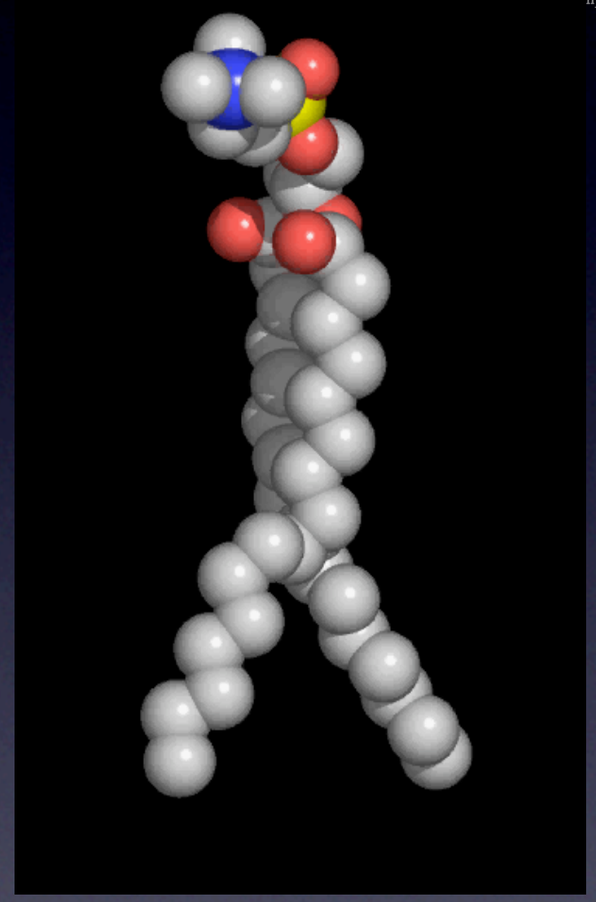


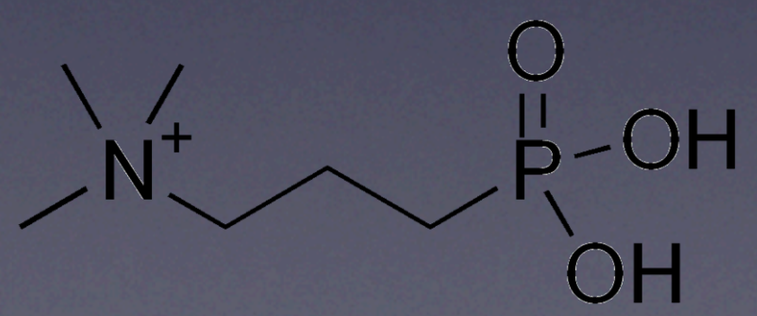
- Controls transport into/from cells
- 30% of eukaryotic proteins are associated with membranes (membrane proteins, receptors)
- 50% of current drugs target membrane proteins
- “Without membranes you die” (David van der Spoel, 2007)

Lipids

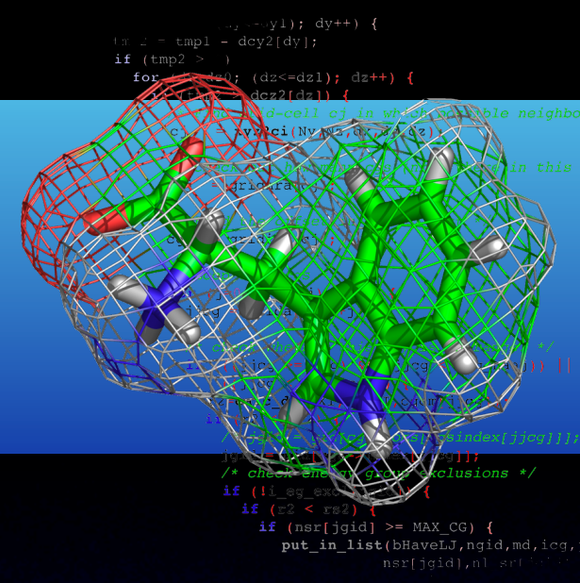


- Charged or strongly polar (zwitterionic) headgroups
- 1-2 Hydrophobic chain(s)
- Amphiphilic molecules
- Typical headgroups:
  Phosphocholine (PC)
  Phosphoetanolamine (PE)
- Glycerol link to acyl chain
- Typical chains:
  Palmitoyl (16 CH_2 groups)



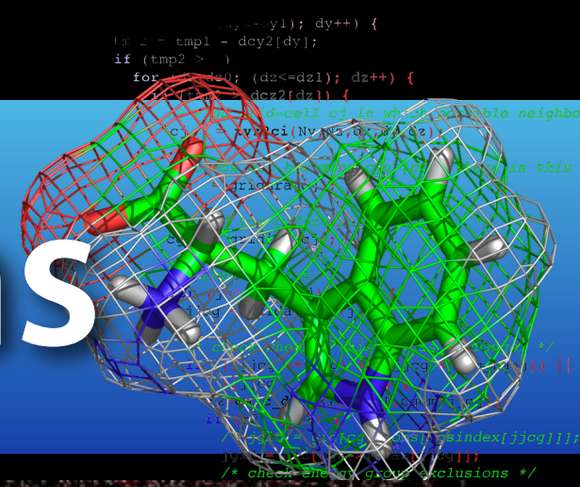


Lipid names...

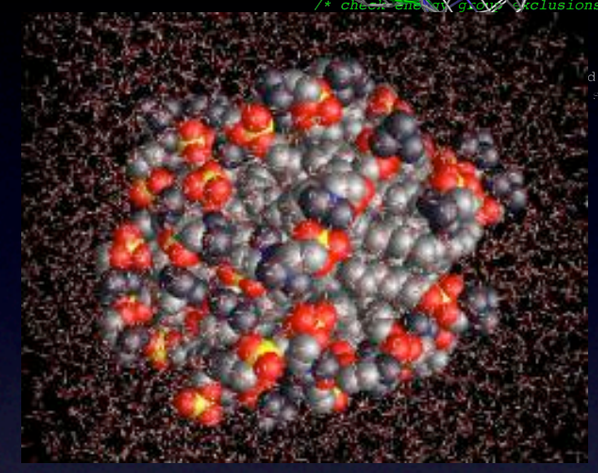


- **DPPC - Dipalmitoylphosphatidylcholine**
  1,2-Dipalmitoyl-sn-Glycero-3-phosphocholine
- **DOPG - Dioleoylphosphatidylglycerol**
  1,2-Dioleoyl-sn-Glycero-3-phosphoglycerol (negative charge!)
- **POPE - Palmitoyloleoylphosphatidyletanolamine**
  1-Palmitoyl-2-Oleoyl-sn-Glycero-3-phosphoethanolamine
- **DPC - Dodecylphosphocholine**
  (single chain only, forms micelles)
- **Cholesterol**

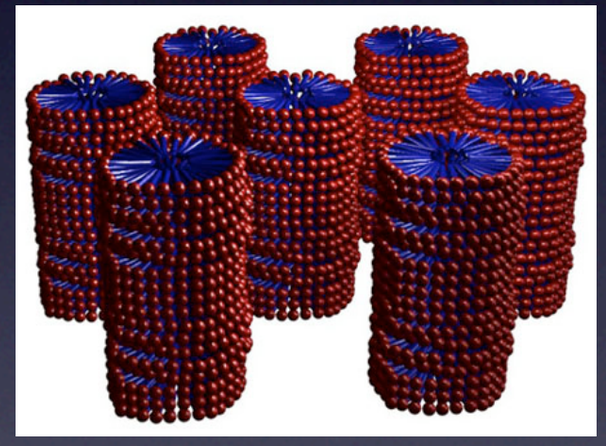
Lipid conformations



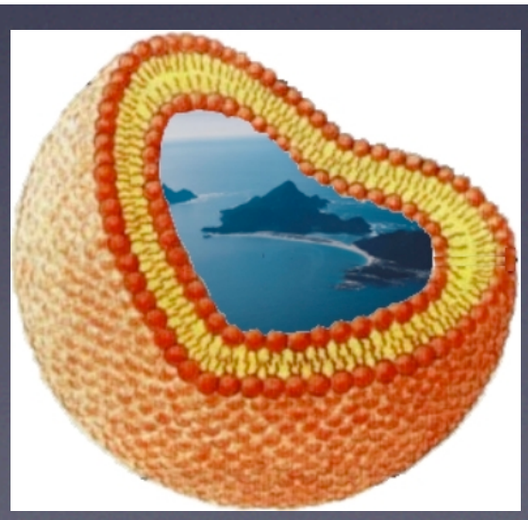
- Large headgroups, single chain: Lipids form micelles



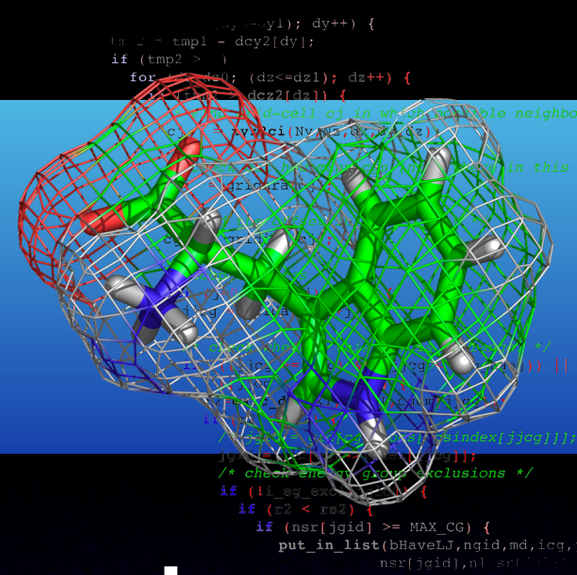
- Higher concentration: hexagonal phases



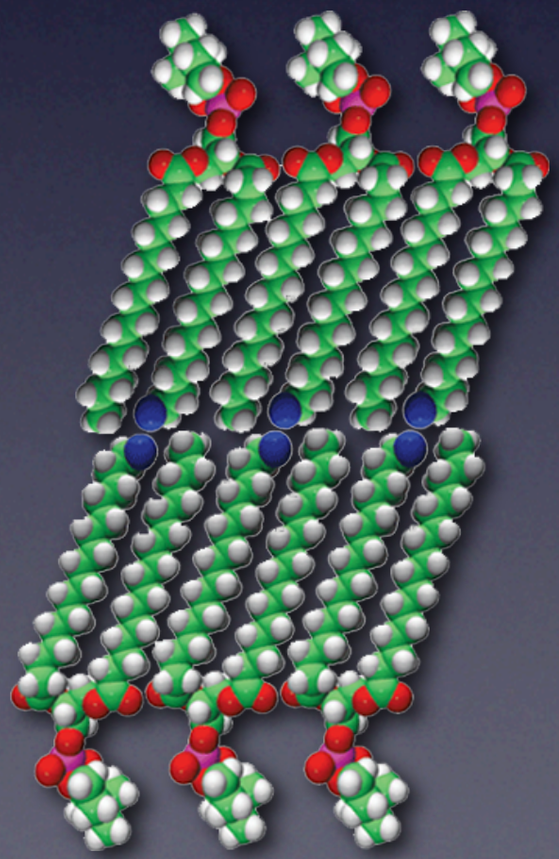
- There are also inverted micelles and hexagonal phases
- Similar size headgroup/chains: Bilayers or vesicles

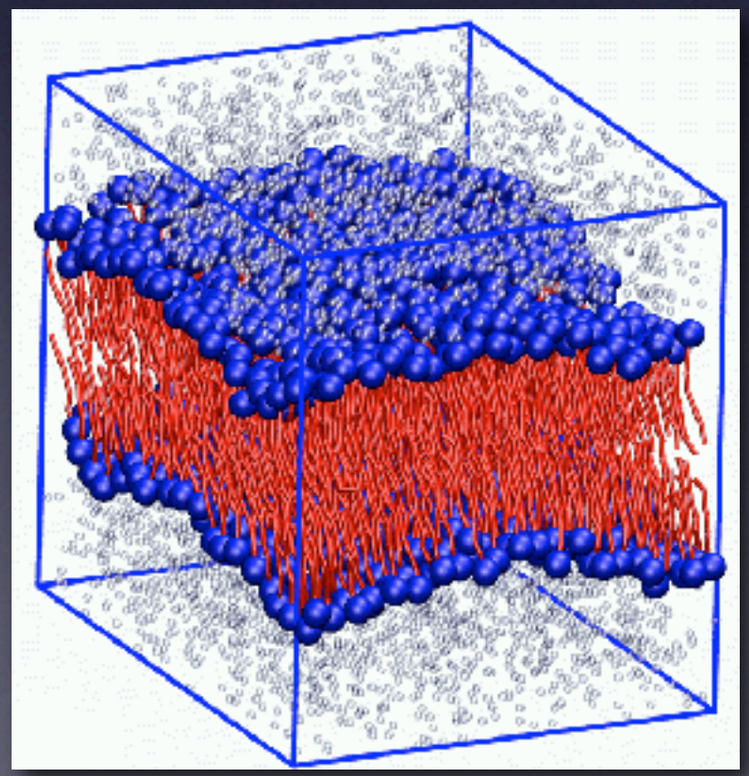


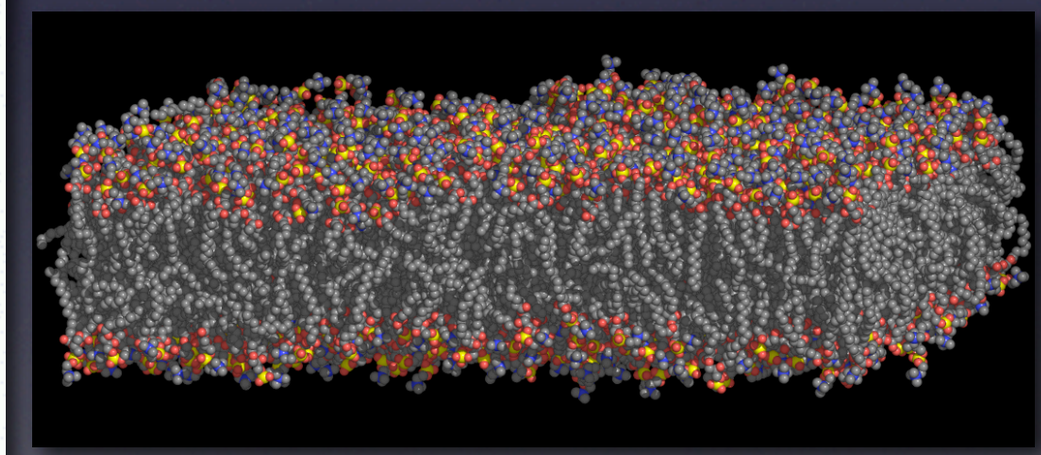
Bilayer phases



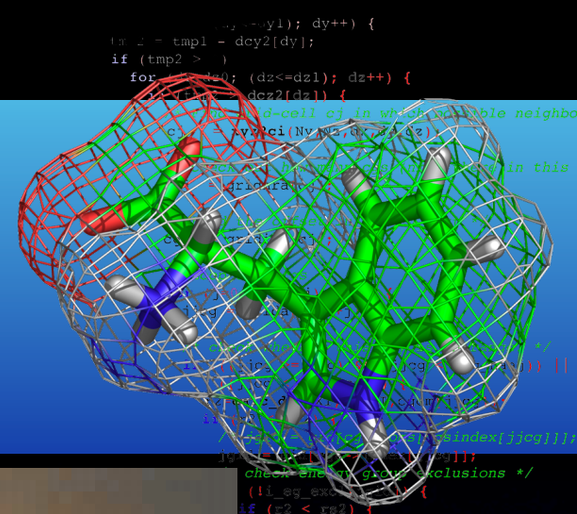
- Very low temperature: proper crystalline phase
- low temperature: gel phase
- room temperature: liquid crystalline phase (Lα)



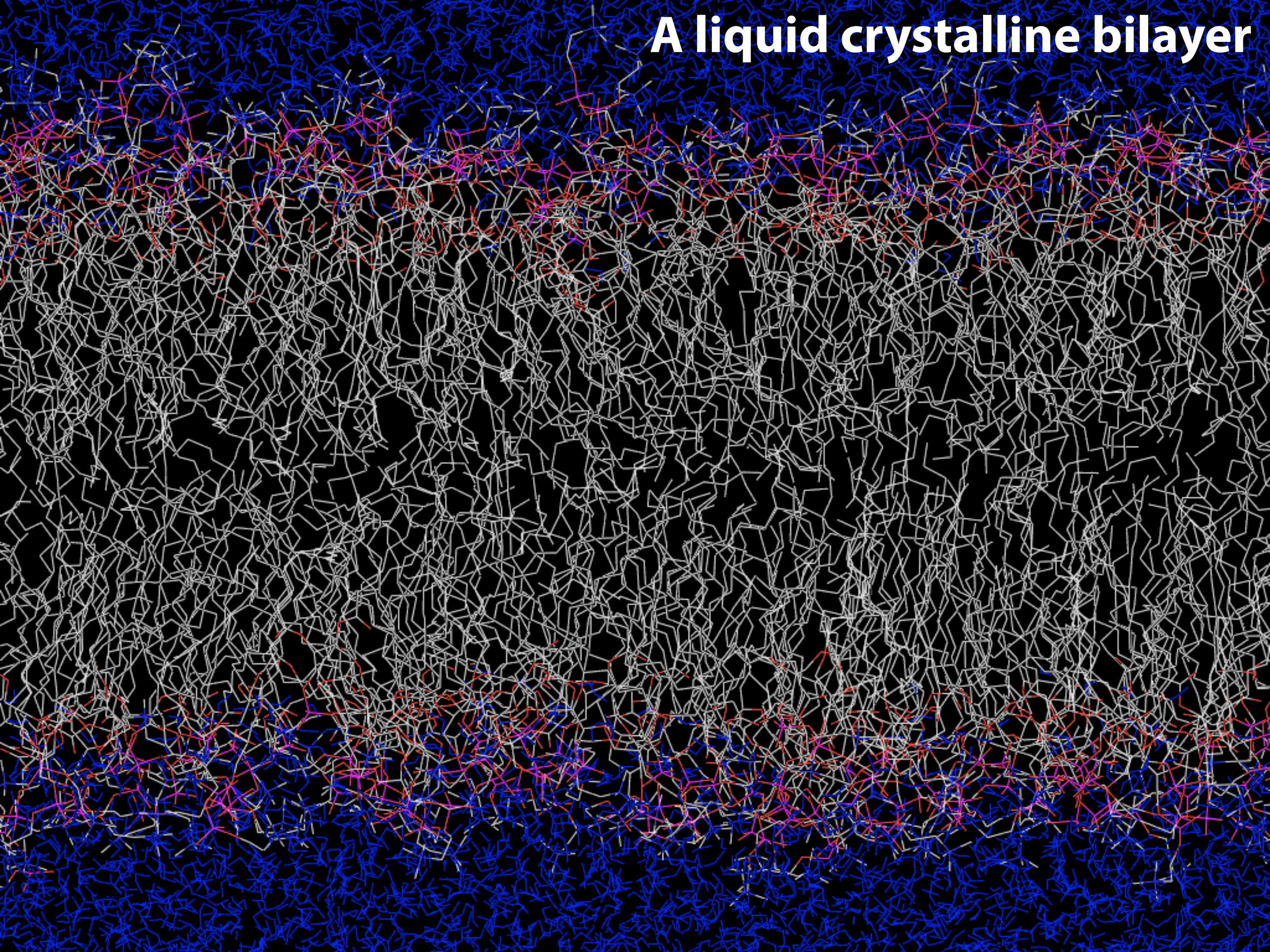




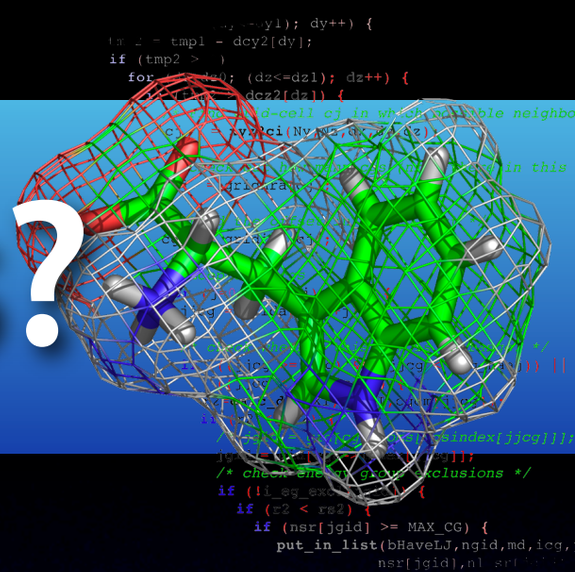
Cellular membranes



A liquid crystalline bilayer

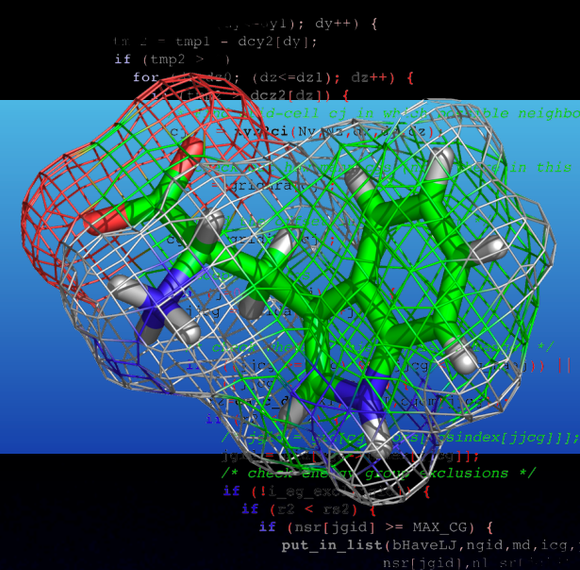


Lipid Bilayer analysis?



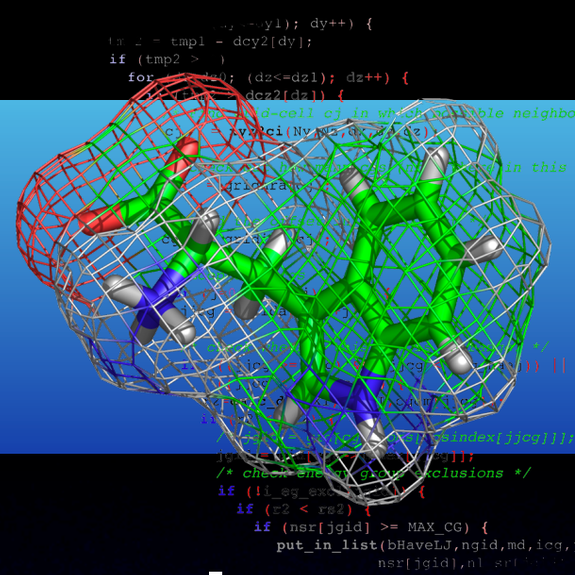
- Experimental techniques:
  Neutron scattering
  Liquid X-ray crystallography
  NMR, EPR, flourescence spectroscopy
- Average area/lipid
- Order parameters from spectroscopy
- Electron density over the bilayer
- Lipid diffusion
- Average properties -> great statistics!

Bilayer simulations



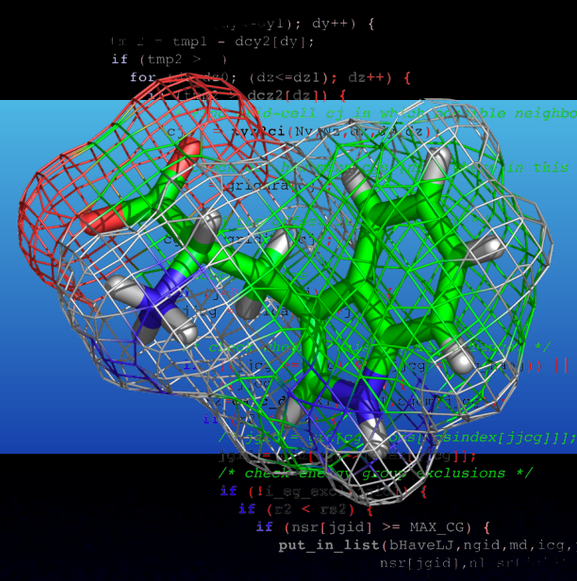
- Lipid parameters / force fields
- Topologies
- Getting initial conformations
- Simulation cell shapes
- Interaction parameters & setup
- Pressure coupling
- Center-of-mass motion

Lipid force fields



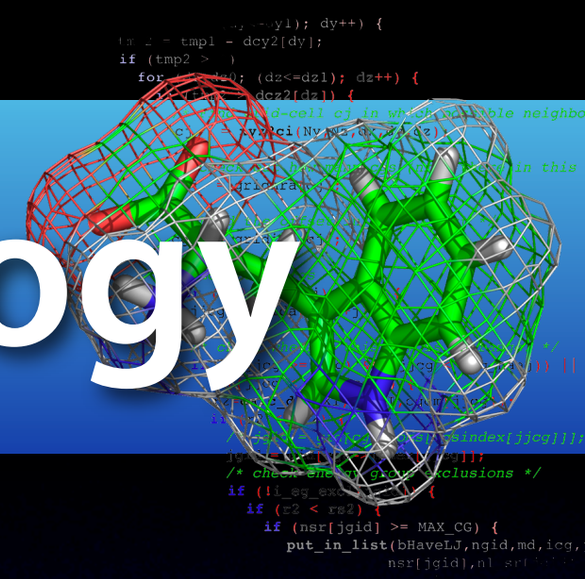
- Groups: phosphate, choline, glycerol, acyl
- Surprisingly hard to get accurate results
- Many force fields lead to lipid area & volumes that are 20% lower than experimental values!
- Do NOT use: Vanilla Gromacs FF, Gromos96, OPLS-AA/L, Amber, old Charmm FF
- Do USE: Berger lipid force field, Charmm27
- Berger is available on Gromacs site, Charmm27 in beta for Gromacs distributions

Lipid topologies



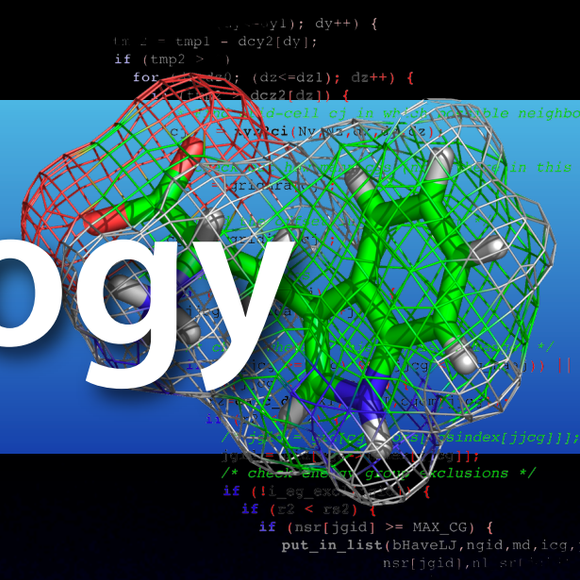
- No bilayer coordinates in PDB
- No standardized atom or chain names
- pdb2gmx does not work automatically
- Best option: Find existing topologies (Gromacs site, Tieleman, Lindahl, others)
- Alternative 1: Create RTP building blocks
- Alternative 2: Hack topologies manually

Example DPPC topology

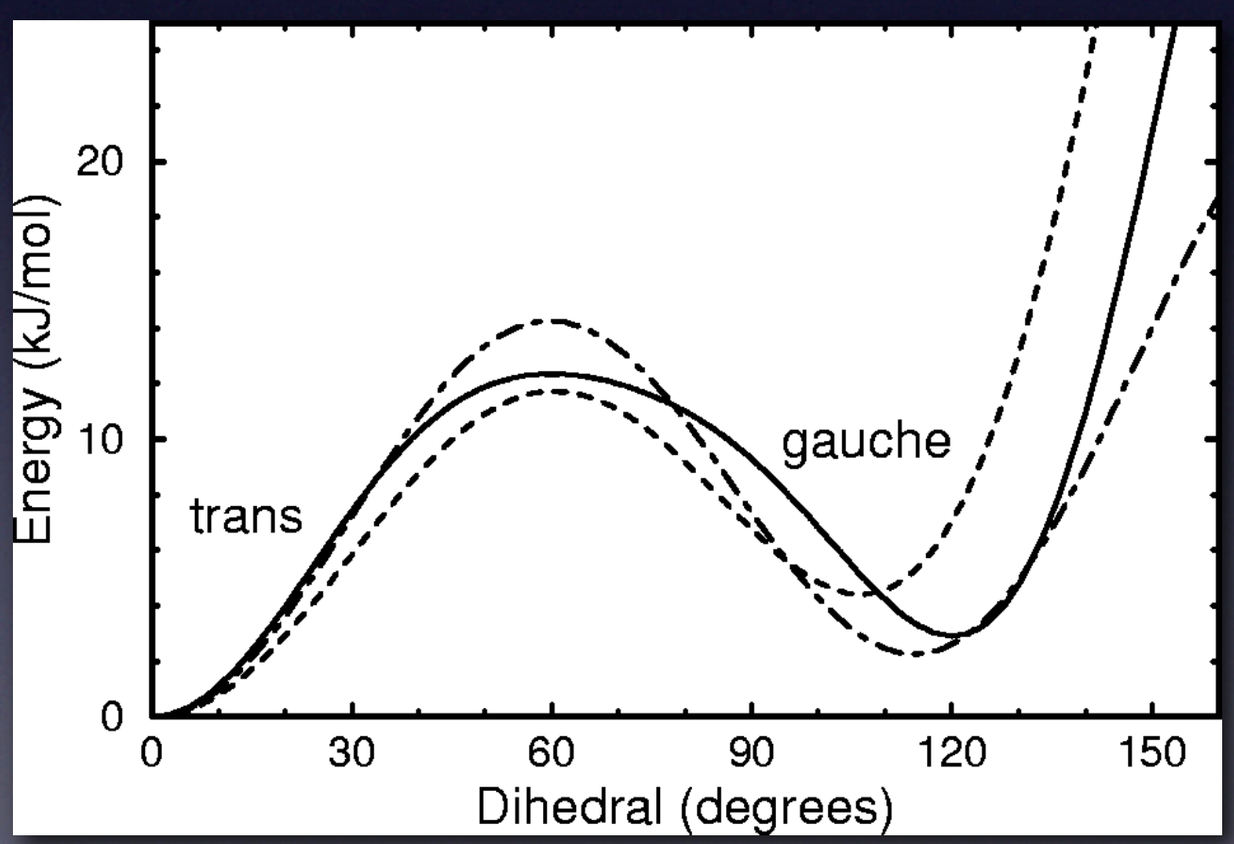


```
;Topology for united-atom DPPC
;Written by Erik Lindahl, version 980624
;After a gromos version by Olle Edholm

; include our own forcefield - dont use gromacs parameters!
;#include "ffDPPC.itp"

[moleculetype]
;Name            nrexcl
DPPC             3

[atoms]
;   nr       type      resnr   residu     atom     cgnr         charge      ; Chiu charges!
     1         C3N          1      DPPC       CN1          1        0.16
     2         C3N          1      DPPC       CN2          1         0.16
     3         C3N          1      DPPC       CN3          1         0.16
     4          NL          1      DPPC       NTM          1         0.26
...
    17         CH2          1      DPPC       C1B          6        -0.04
    18         CH2          1      DPPC       C1C          7         0.0
    19         CH2          1      DPPC       C1D          8         0.0
    20         CH2          1      DPPC       C1E          9         0.0
```

Example DPPC topology



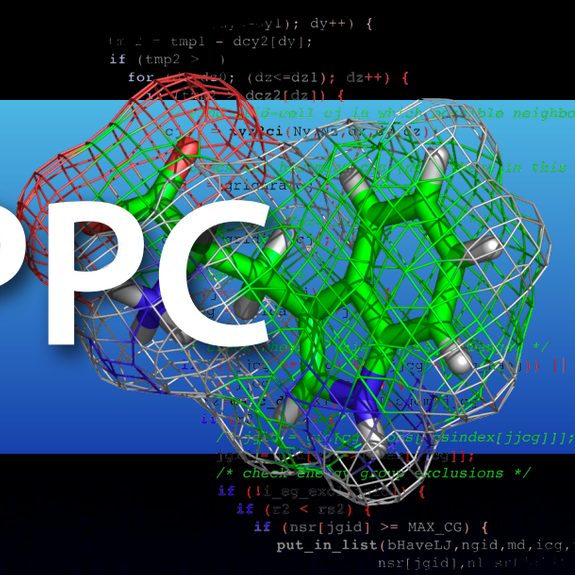
```
...
[ dihedrals ]
;  ai    aj    ak    al funct    (func=1 is normal dihedrals, 3=R-B)
    1     4     5     6    1
    4     5     6     7    1
...
   17    18    19    20    3 ; R-B torsions for acyl chains
    18    19    20    21    3
...
[ system ]
; Name
DPPC membrane with 28 waters per lipid

[ molecules ]
; Compound        #mols
DPPC              64
SOL               3000
```



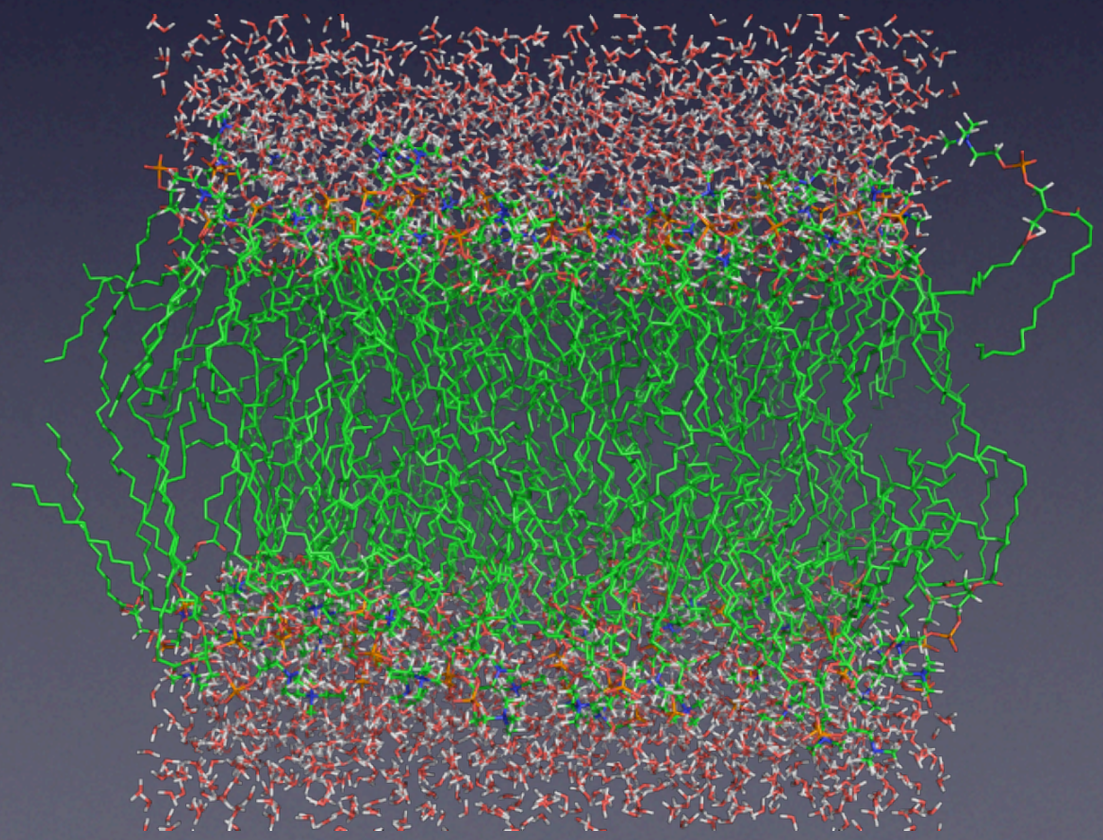
1,4 interactions *excluded*

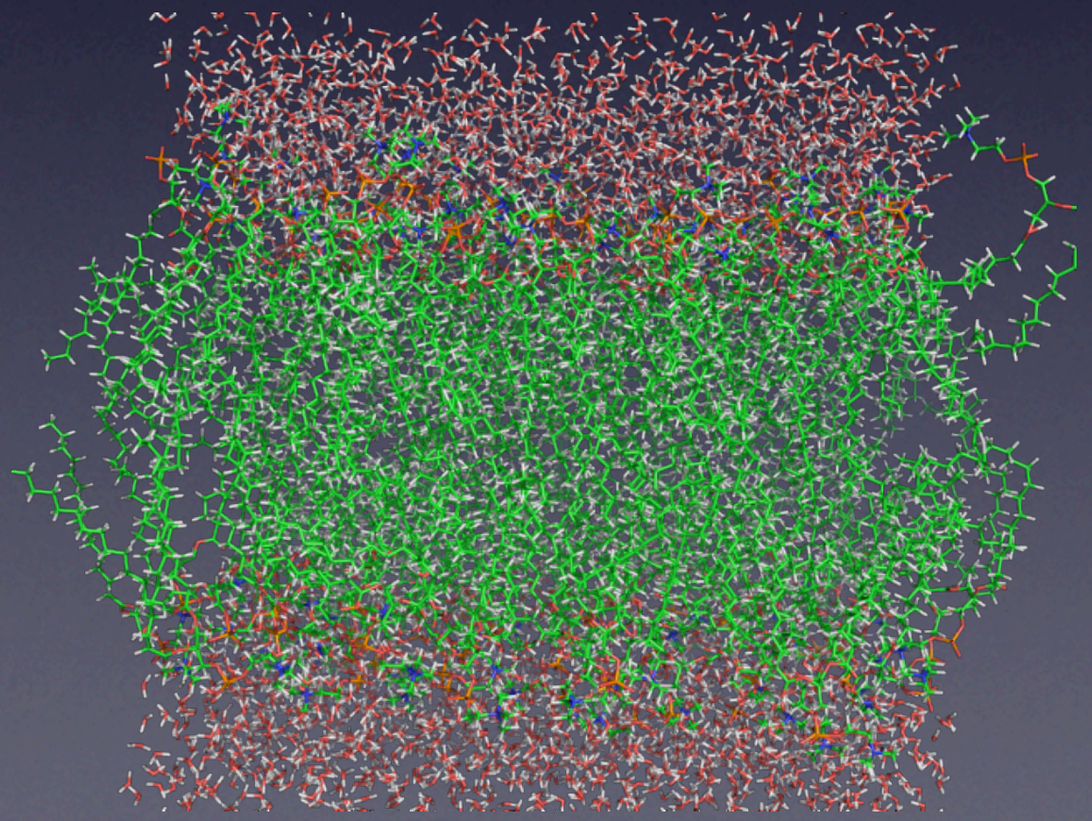
R-B/Kuwajima torsions

United vs. all atom DPPC

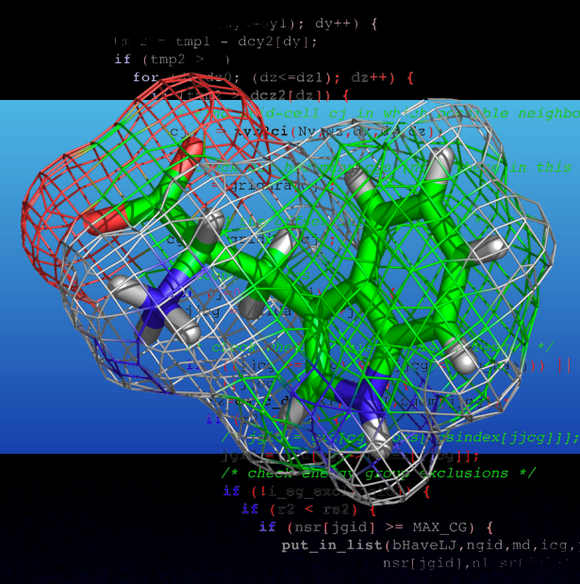


- 50 vs 130 atoms per lipid
- 3x atom density in hydrophobic region
- ~9x interaction density in hydrophobic region
- Roughly 4-5x simulation speed difference



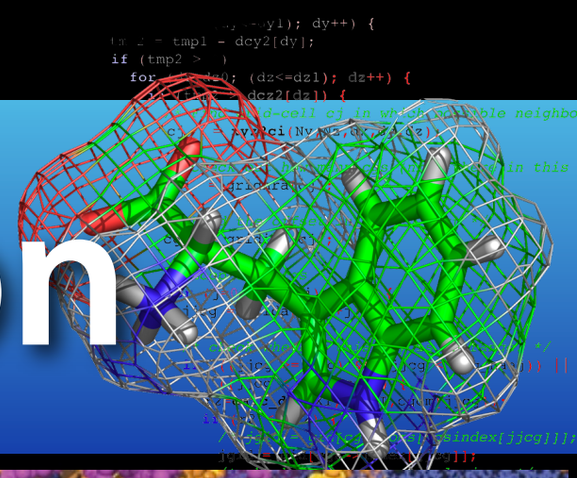


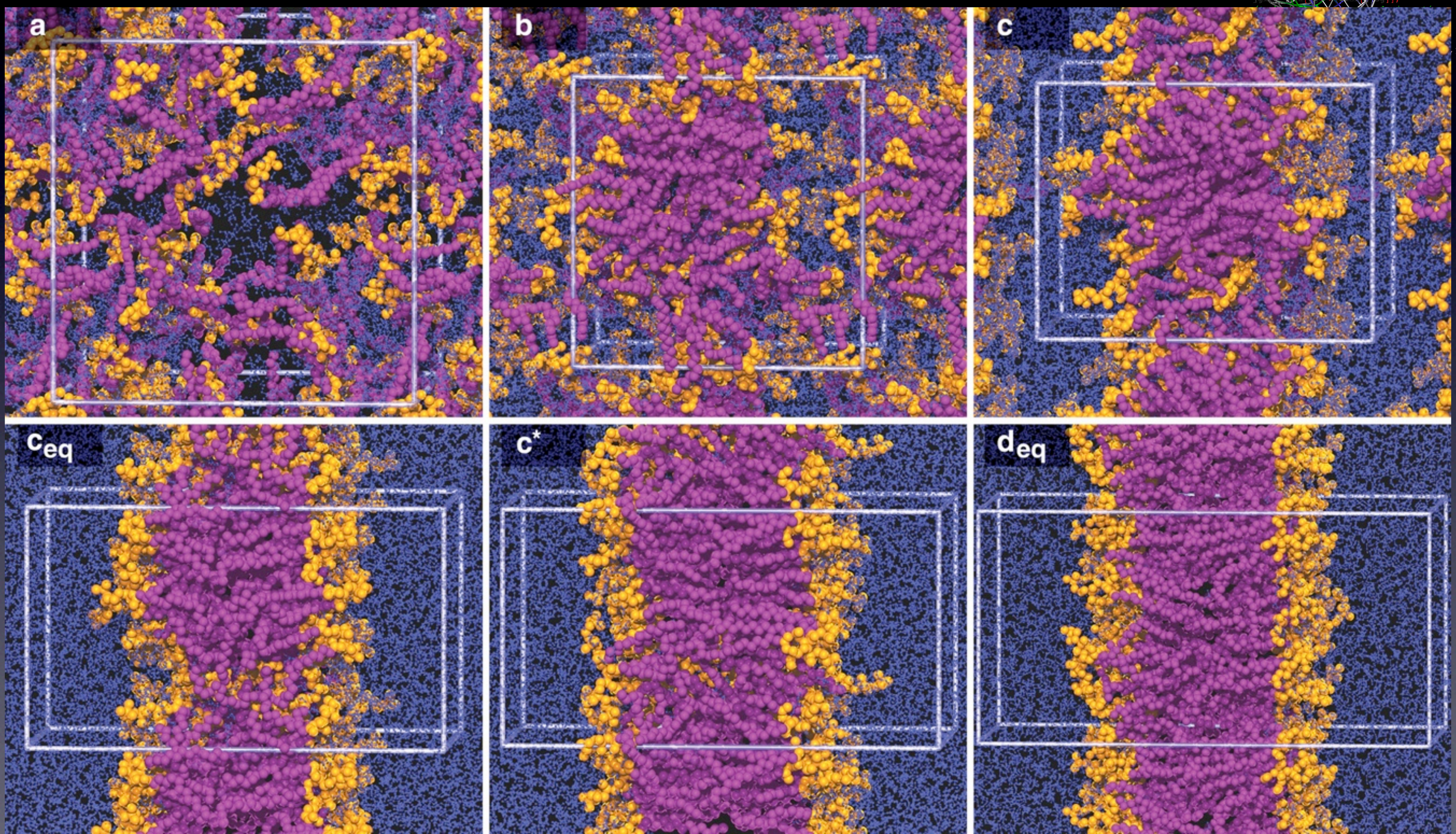
Bilayer structures



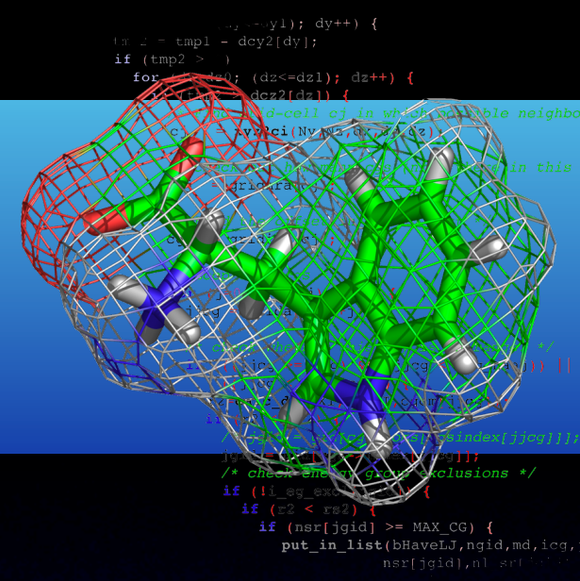
- Find pre-equilibrated conformations
- Ask authors of published papers
  (works great for Gromacs topologies too)
- Use genbox & genconf to change system size
- Repeat single lipid conformation on grid, equilibrate for a long time in vacuo
- Repeat single lipid(s), add random rotation and tilts, shorter equilibration
- Simulate bilayer formation ‘the natural way’

Membrane formation

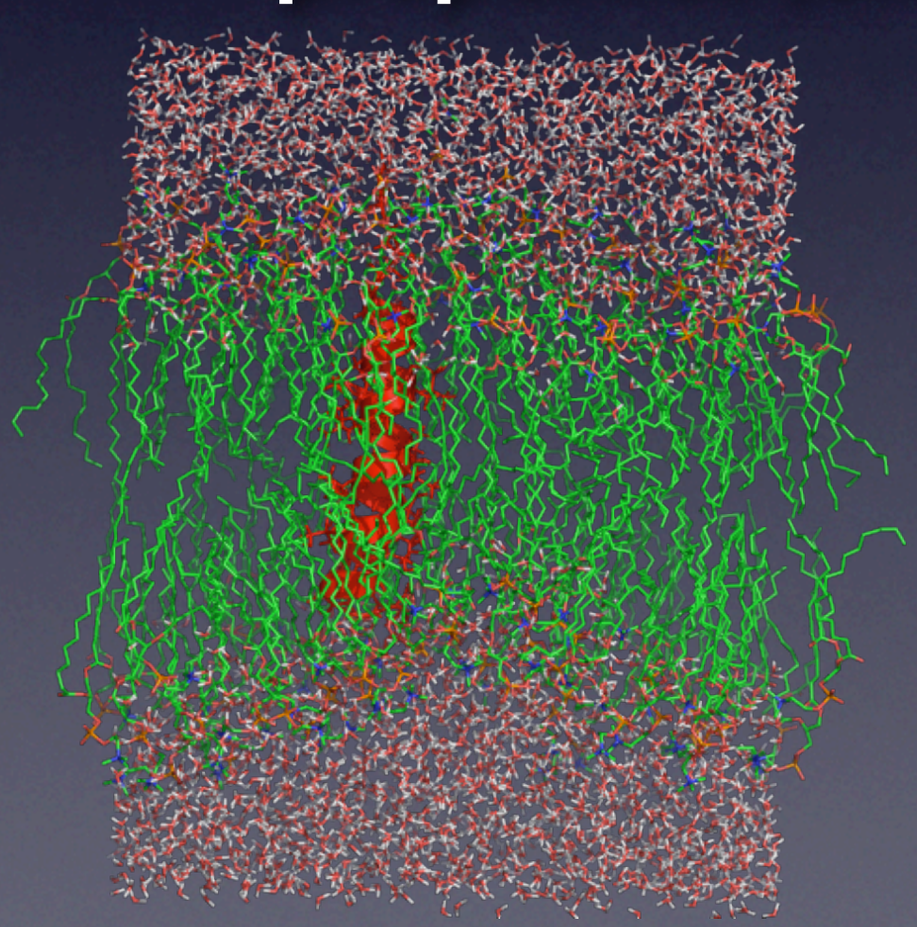


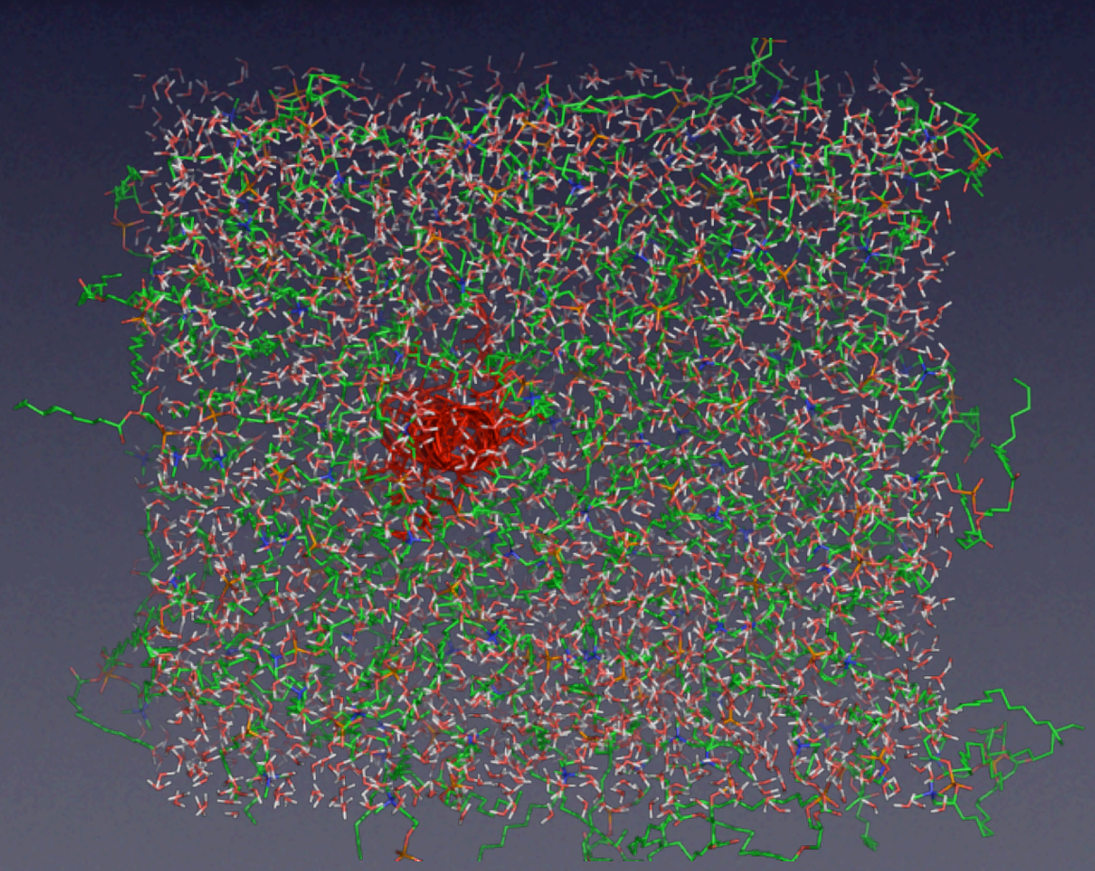


Cell shapes

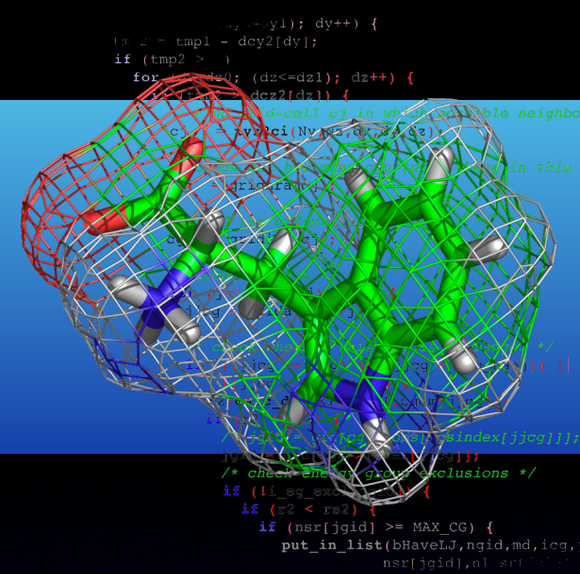


- Rectangular membranes
- Easiest option when the simulation goal is to calculate wave-vector dependent properties (undulations, etc.)

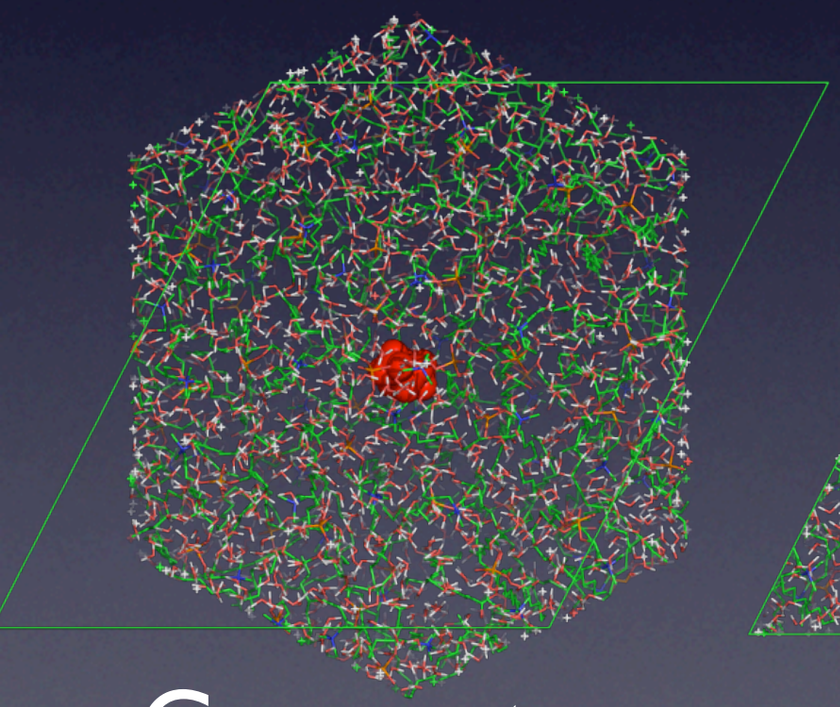




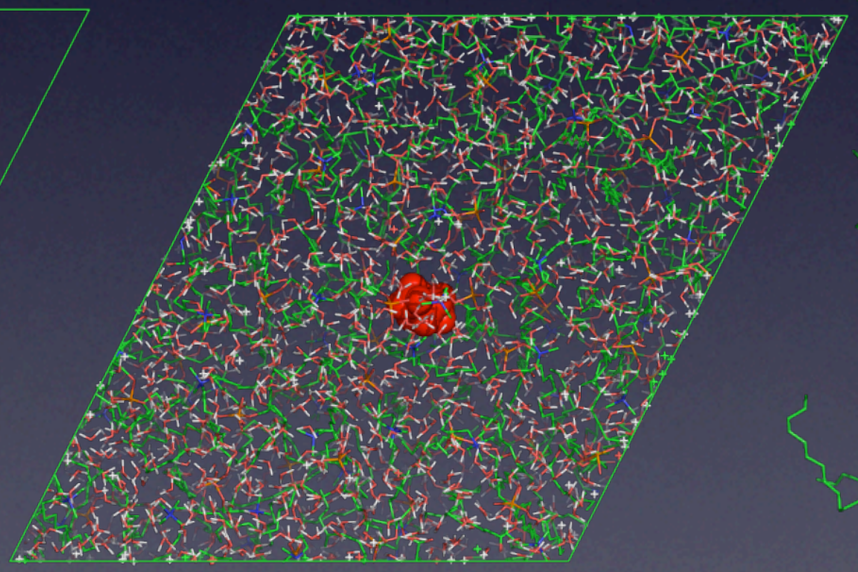
Cell shapes



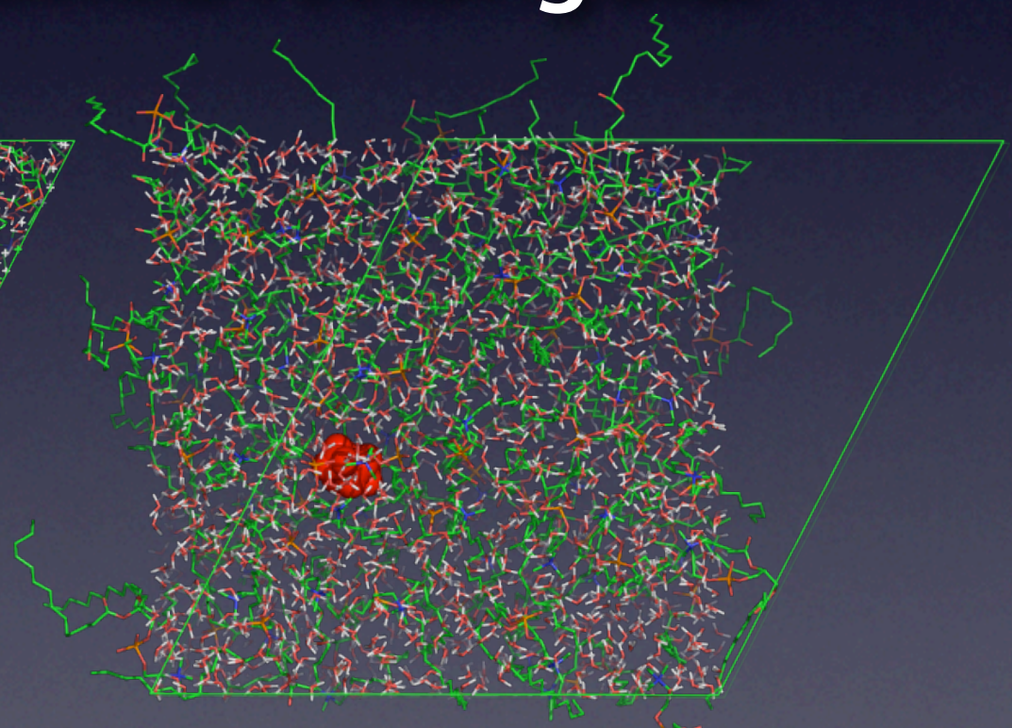
- Hexagonal membranes
- Maximizes periodic separation distance
- Can be viewed compact/triclinic/rectangular



Compact

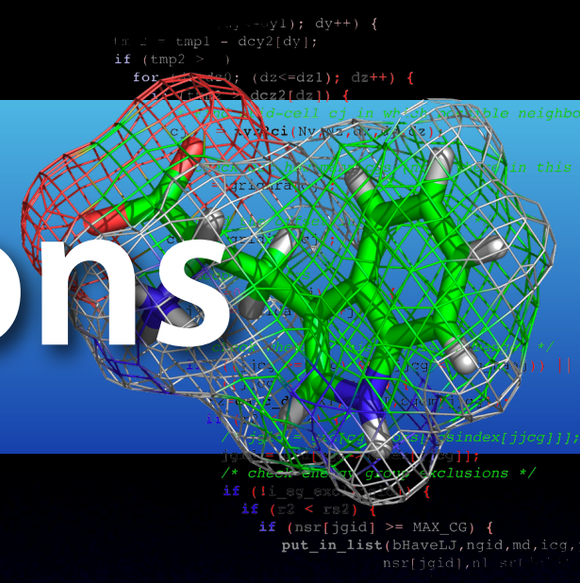


Triclinic
(actual periodicity)



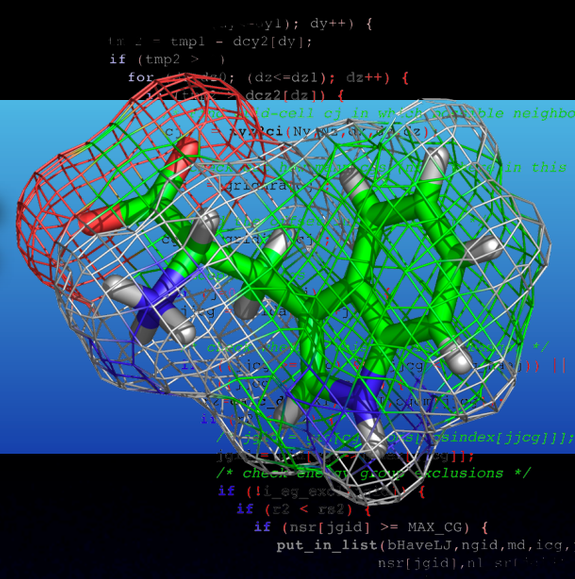
Rectangular & whole
(default output)

Coulomb interactions



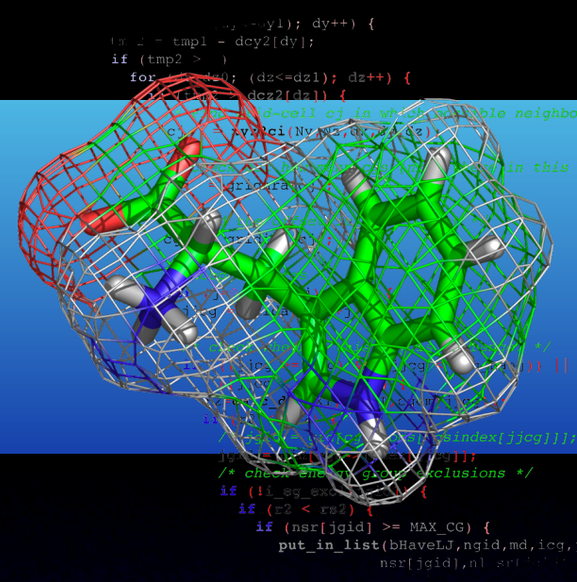
- Very large dipoles & parts with low dielectric screening (hydrophobic core)
- Cut-offs are bad
- Reaction-field would assume an isotropic and homogeneous system (not the case)
- Always use PME
- PME works fine with triclinic cells in Gromacs

Van der Waals stuff



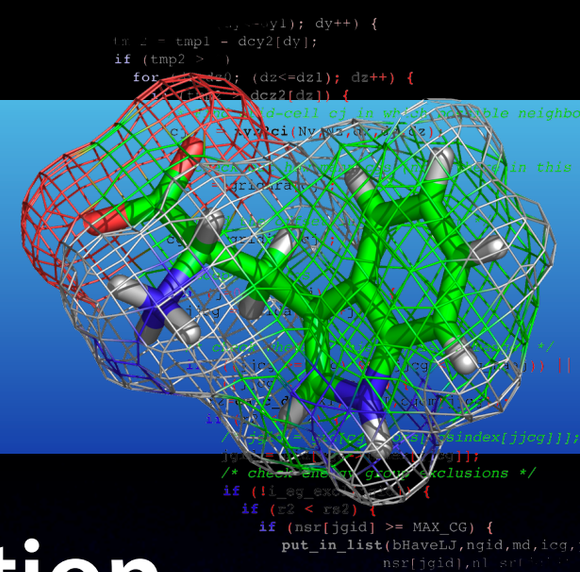
- One reason for the difficulties in simulating membranes is the sensitivity to nonbonded Van der Waals parameters
- Small changes will affect packing of chains, and thus the pressure and area/lipid
- Normal cut-off or switch around 1.0nm is OK, but turn on long-range dispersion correction to get correct pressure

Pressure coupling



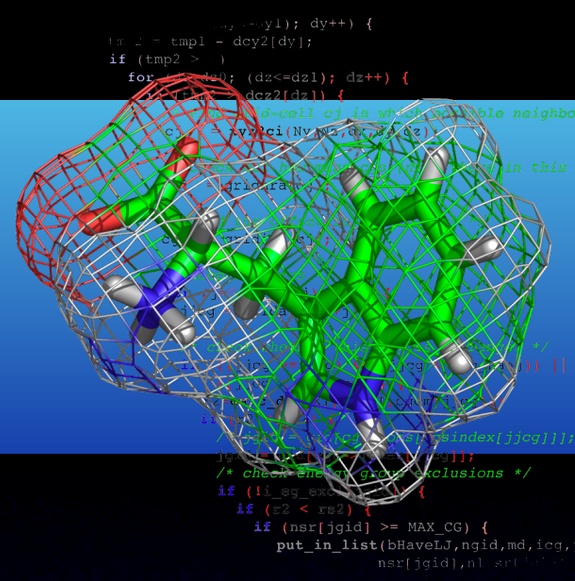
- Two-dimensional liquid crystals
- Bilayer should be able to deform in XY-plane
- Anisotropic pressure coupling should be enabled throughout production runs!
  - Neat trick: Semiisotropic coupling (XY+Z)
- Berendsen coupling does not provide a true NPT ensemble - Parinello-Rahman is better
- Slow relaxation (10ps) to avoid oscillations

Compressibility



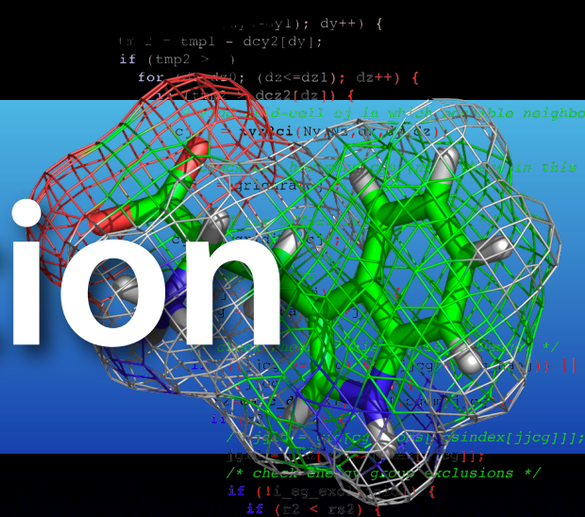
- Water value works fine as approximation
- Compressibility is really a 3x3 tensor
  Gromacs: 3 diagonal + 3 off-diagonal elements
- Off-diagonal elements zero:
  only *scale* cell along the box vectors
- Off-diagonal elements same as diagonal:
  Enable cell distortion/shear transformations
- You can also *force* the cell to distort by setting
  off-diagonal reference pressure to non-zero

Thermostats

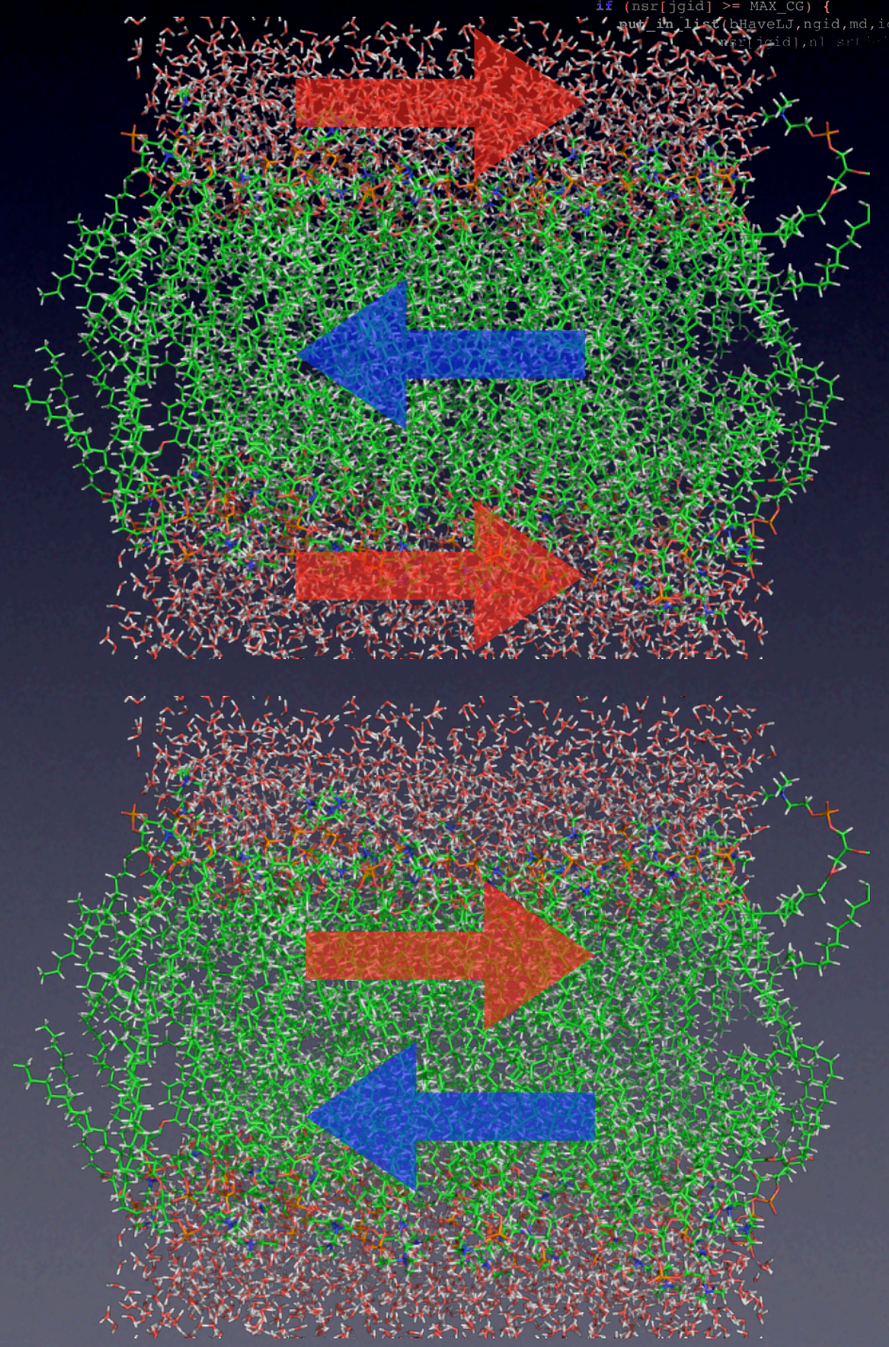


- Polar and non-polar parts of the system are not always strongly coupled
- Bond constraints remove kinetic energy: anisotropic temperature in acyl chains!
- Use separate thermostats for water & lipids, and also protein(s) if present

Center-of-mass motion

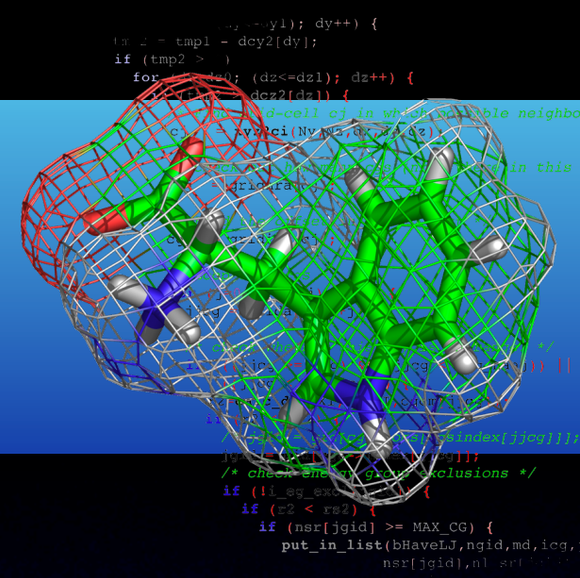


- Center-of-mass motion of entire system normally removed in simulations
- Weak z-coupling in bilayers:
  - Water moves right, lipids left
  - Upper layer right, lower left
- Remove center-of-mass motion separately for: water, upper layer, lower layer

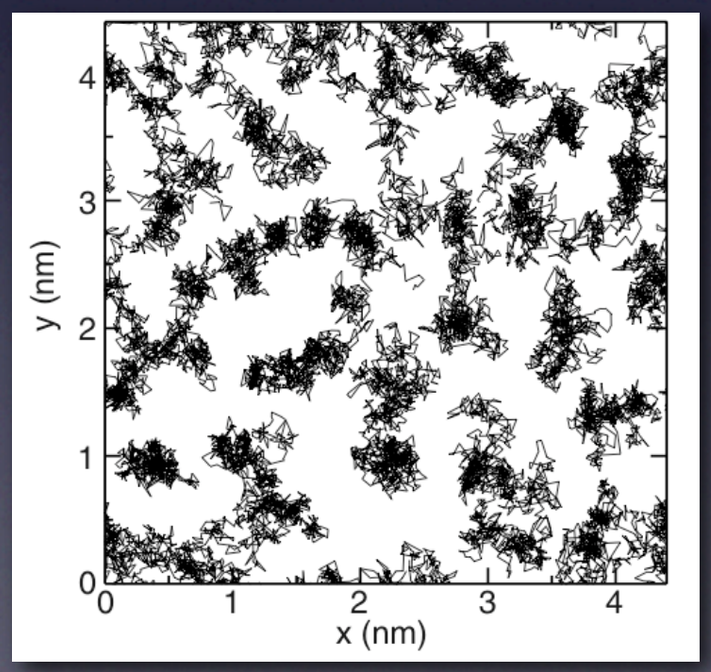


```
comm_grps = upper lower SOL
```

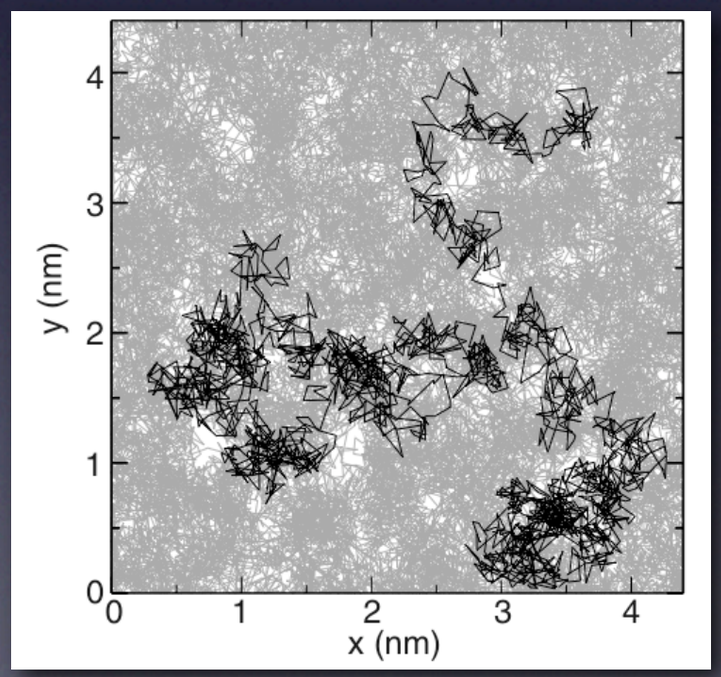
Lipid diffusion



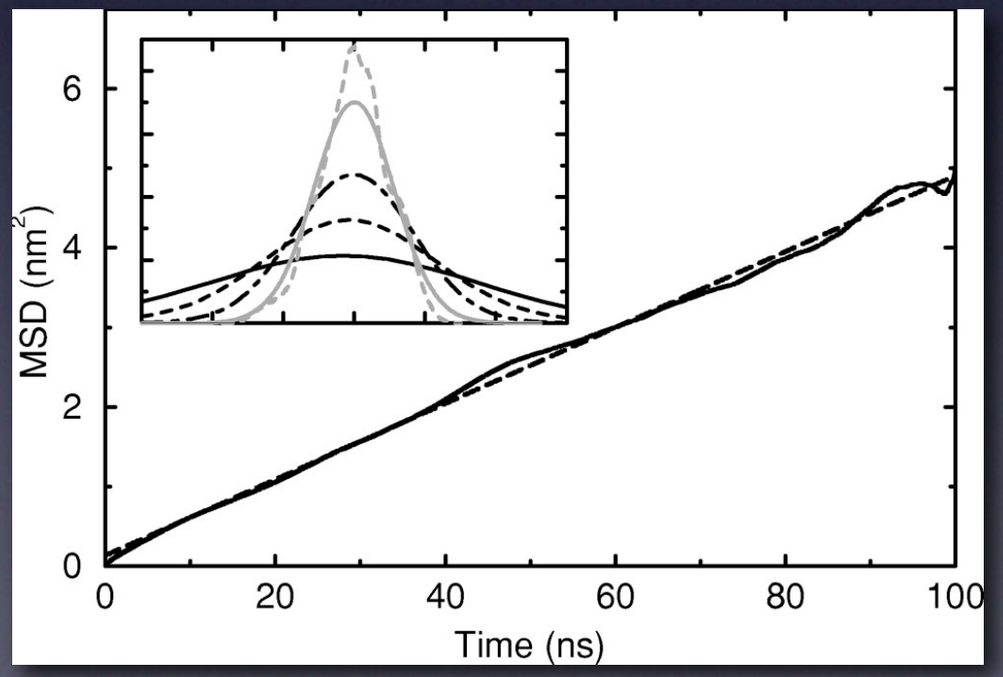
- Frequently overestimated by an order of magnitude due to layer c-o-m motion!
- Can be corrected after simulations too, though



10 ns

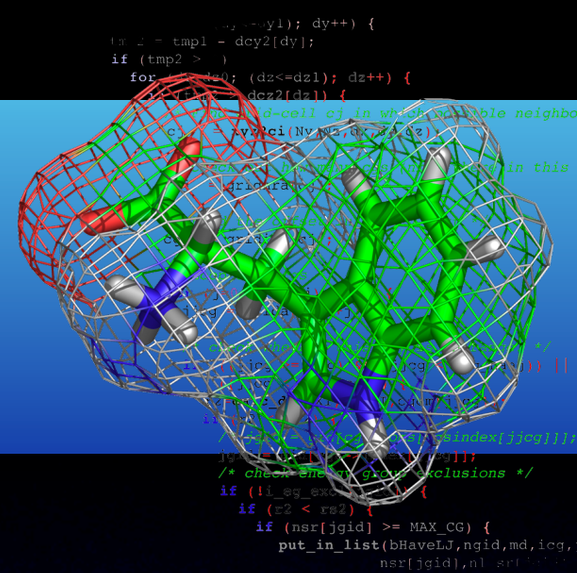


100 ns

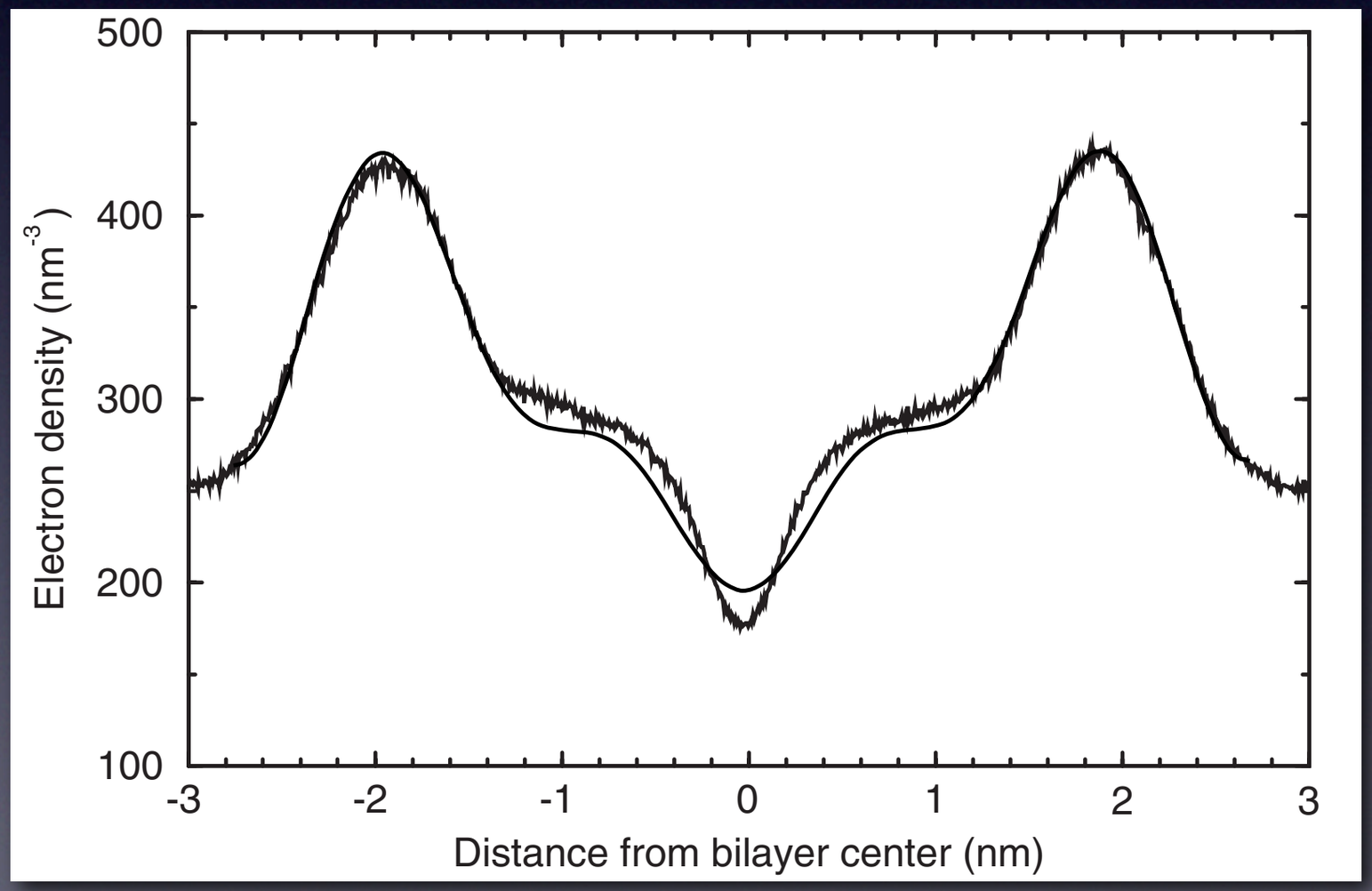


MSD (10x slower than H2O)

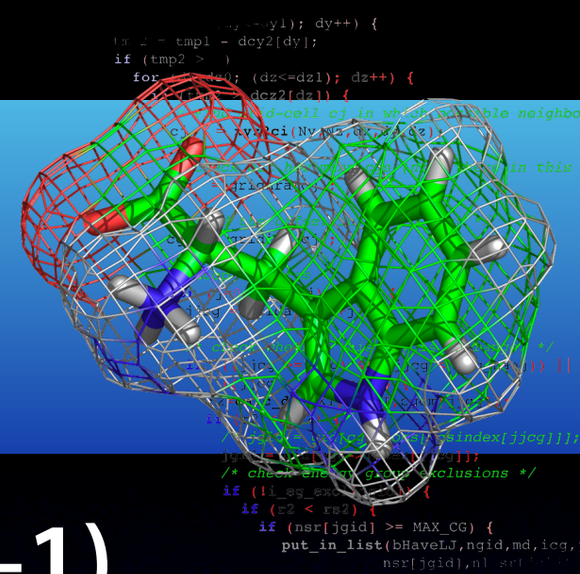
Electron density



- g_density
  You need to provide electrons.dat with number of electrons/atom, see -h flag.

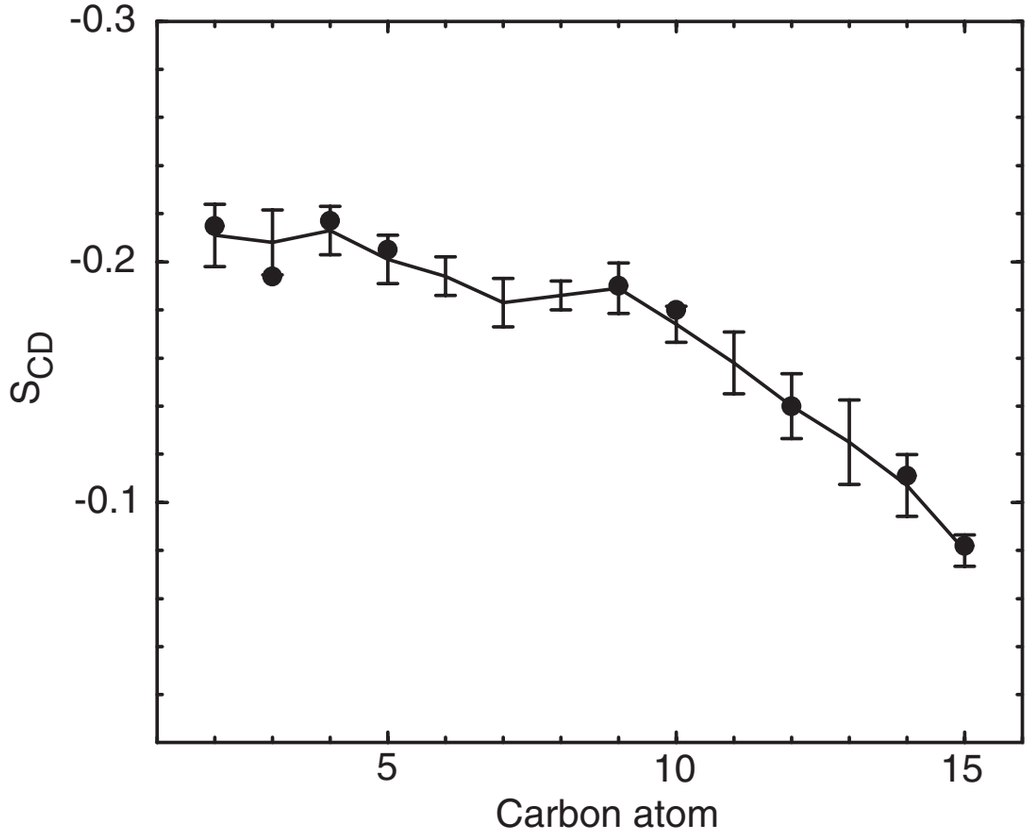


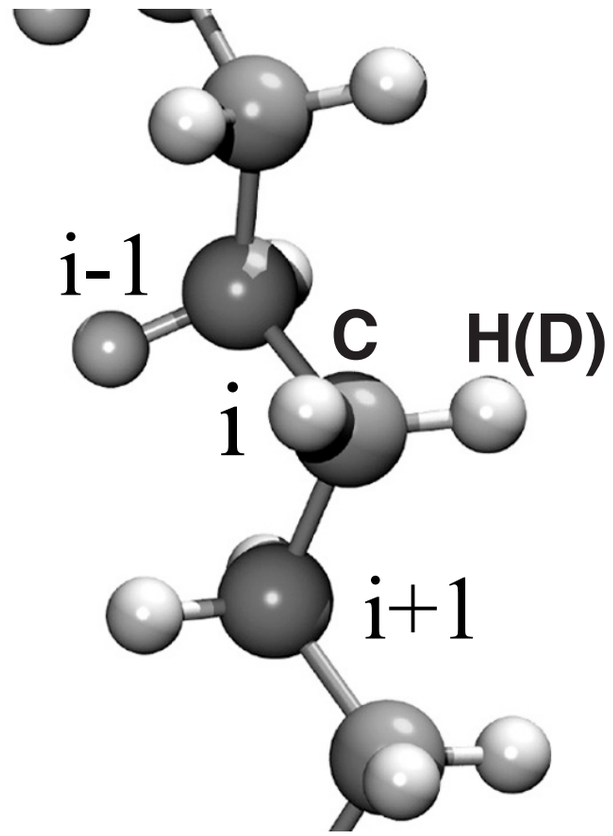
Order parameters



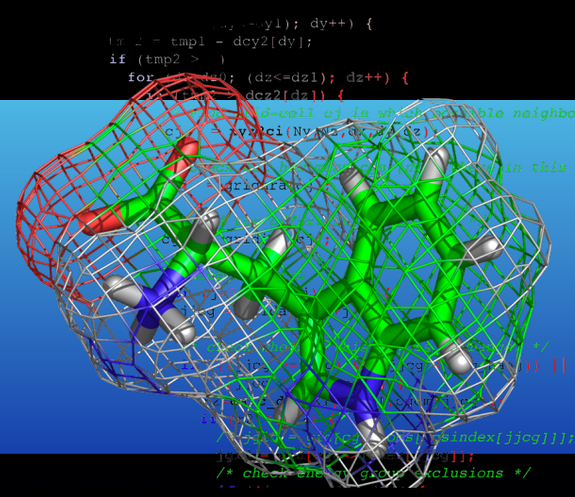
- g_order (requires index file with i-1, i+1)

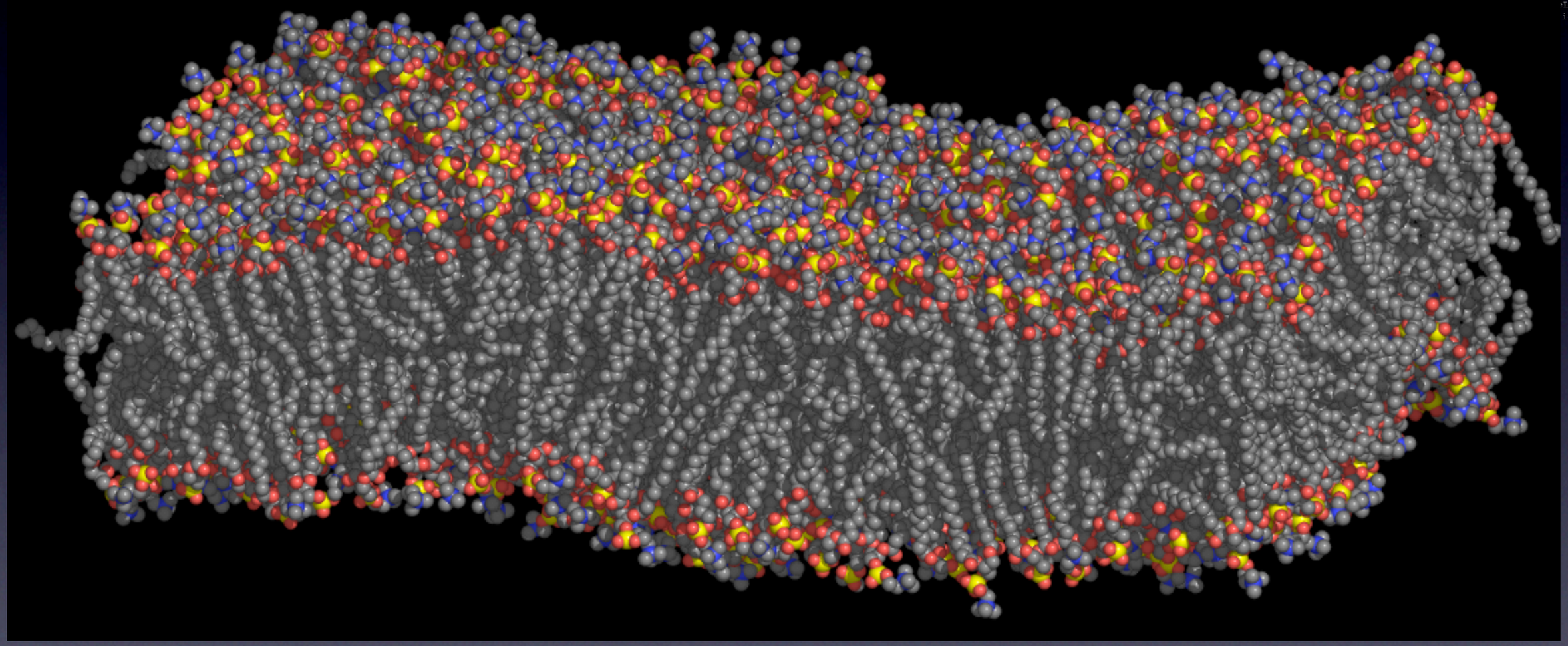
$$S_{CD} = \frac{3\left\langle \cos^2\theta \right\rangle - 1}{2}$$





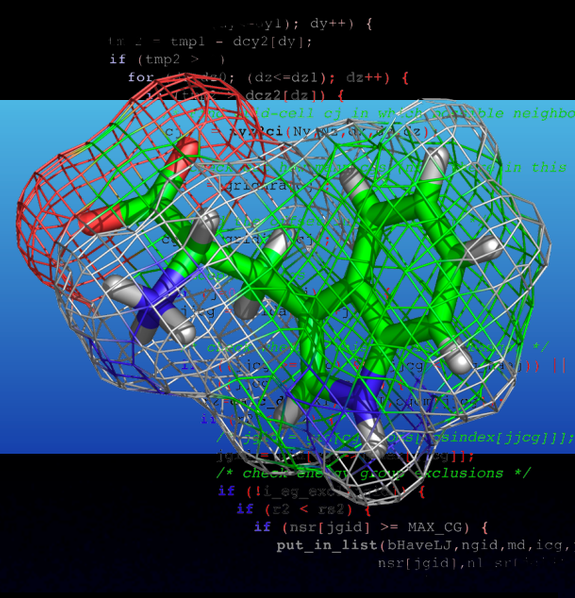
Undulations

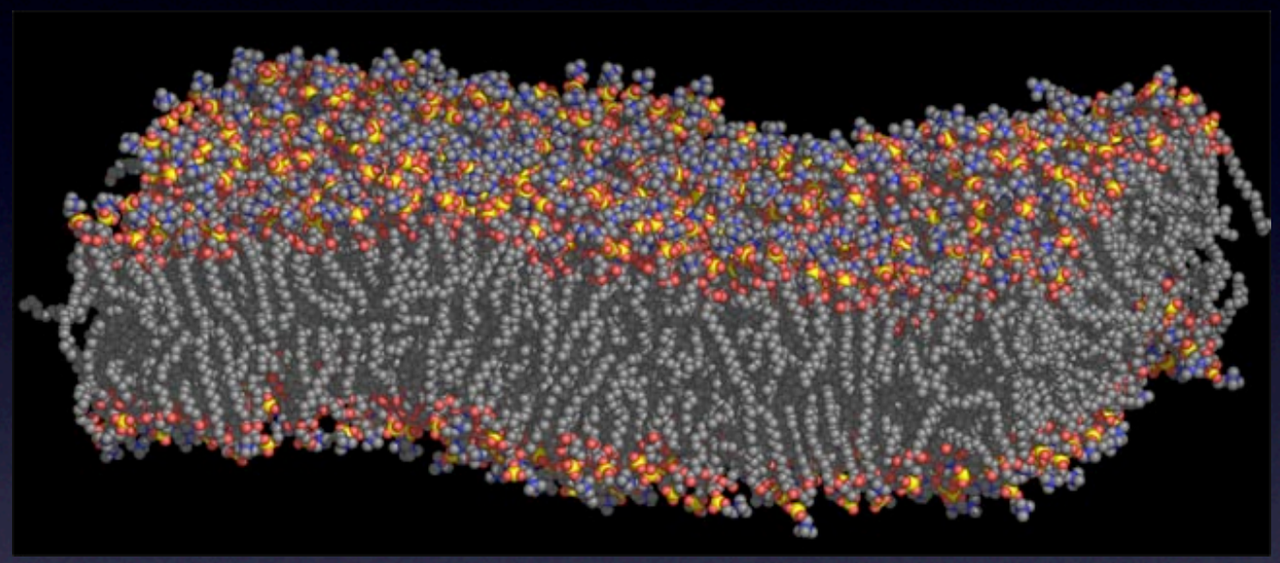


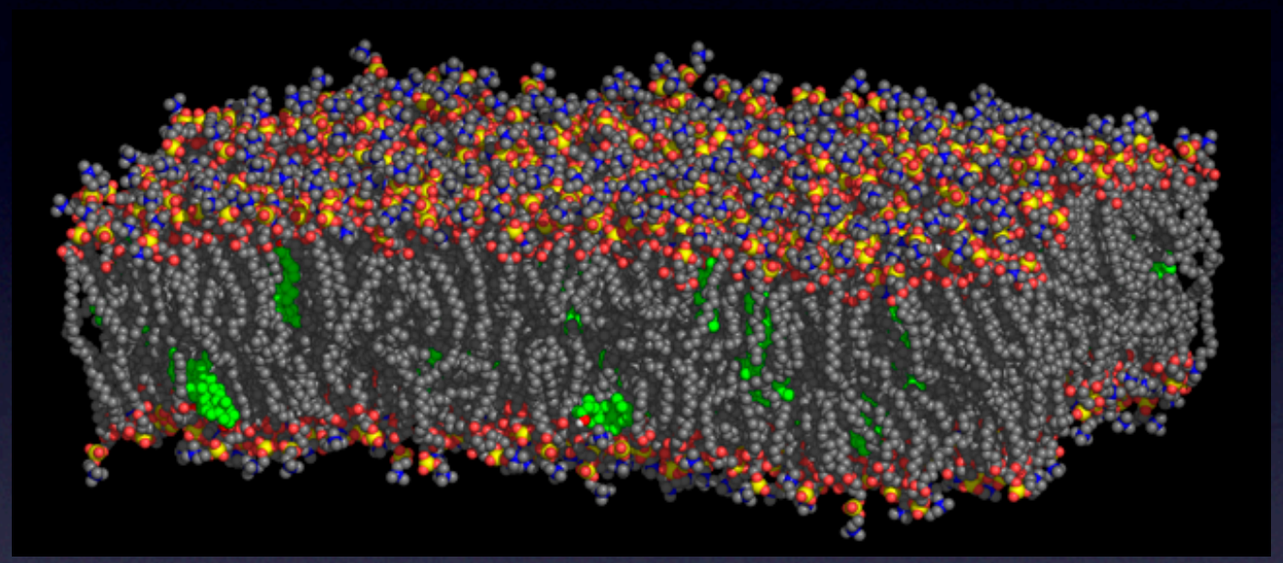


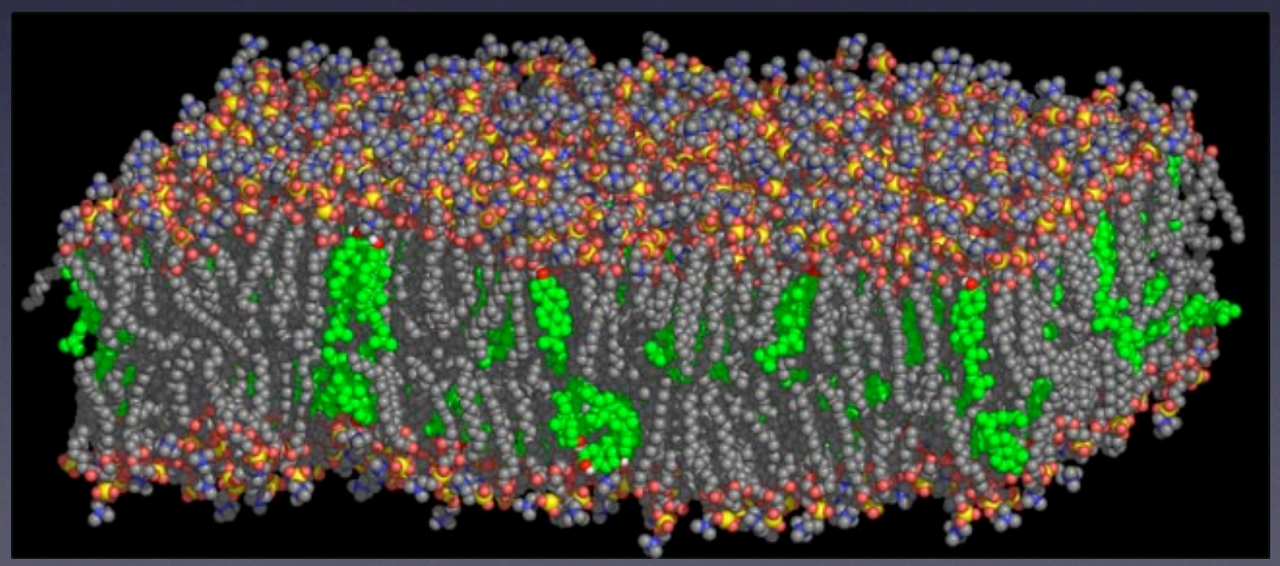
- Assign lipid z-coordinates to a grid
- Perform 2D Fourier transforms
- Plot amplitides vs. wave vector magnitude

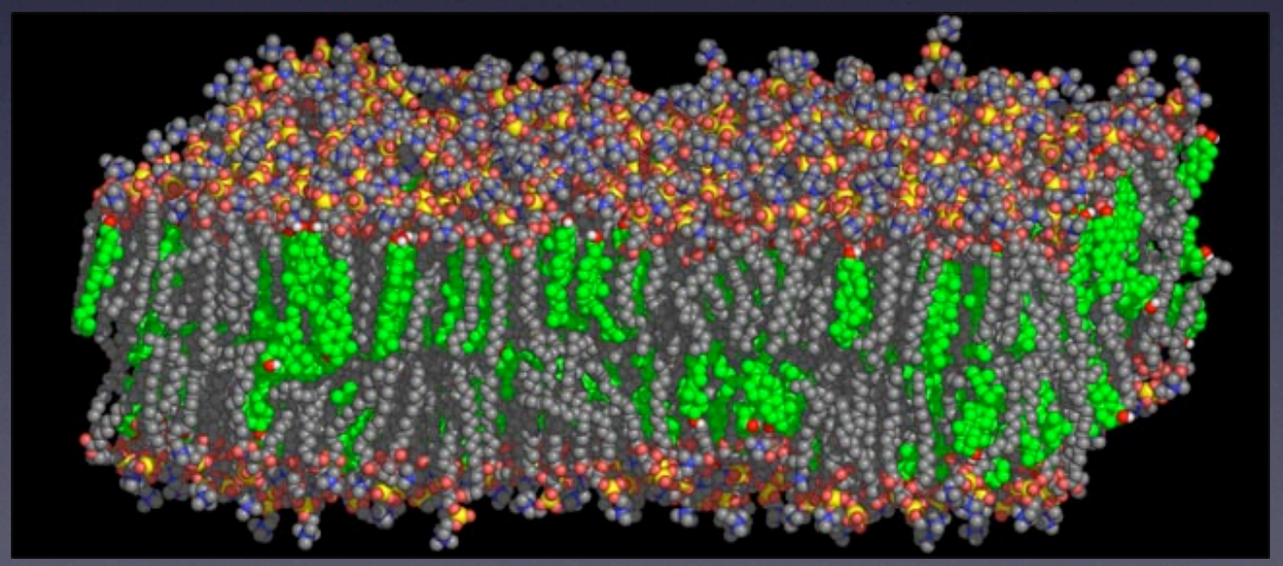
Example: cholesterol





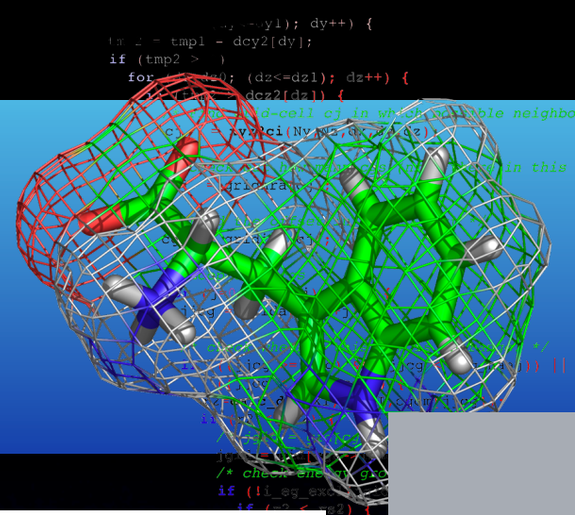


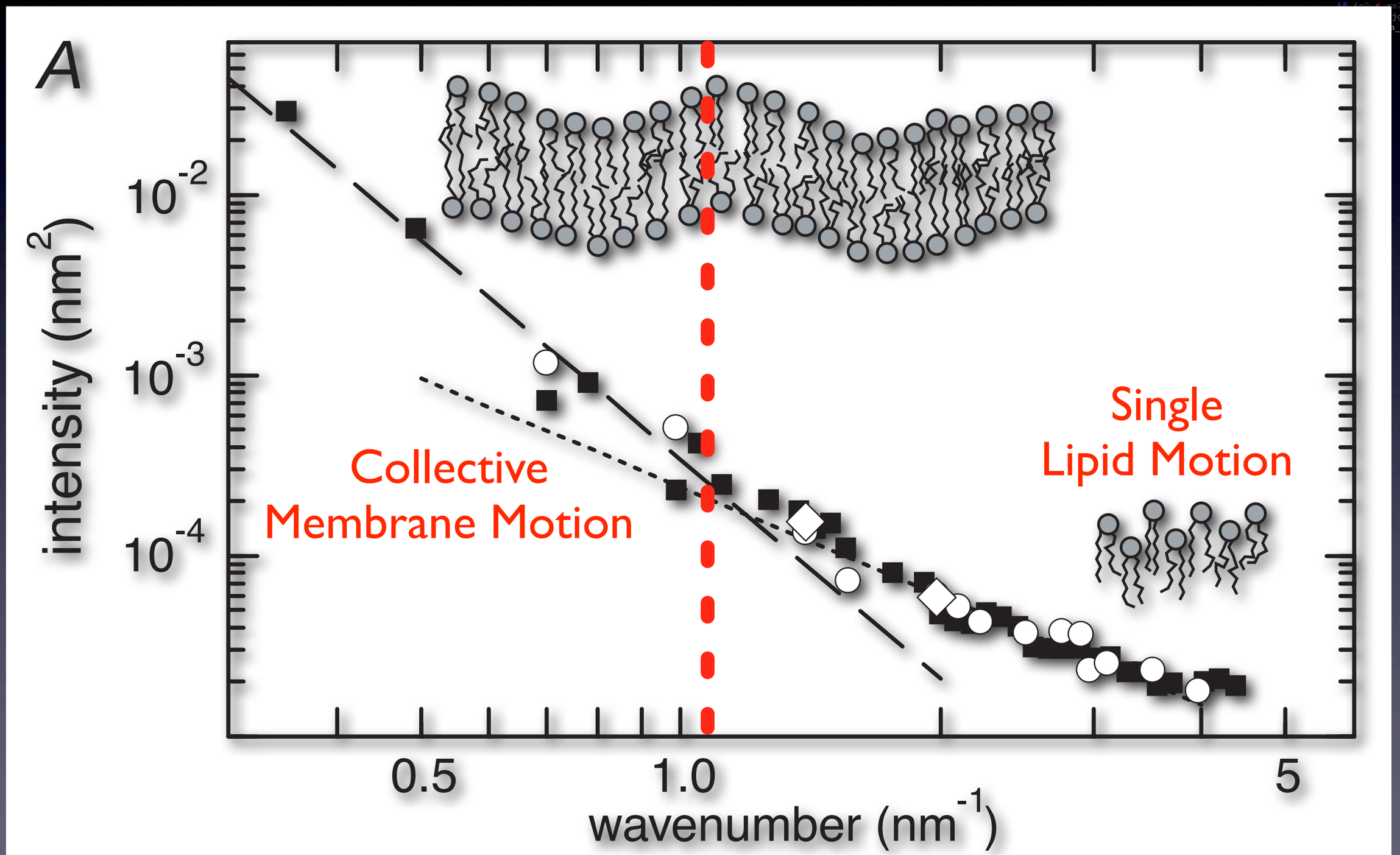




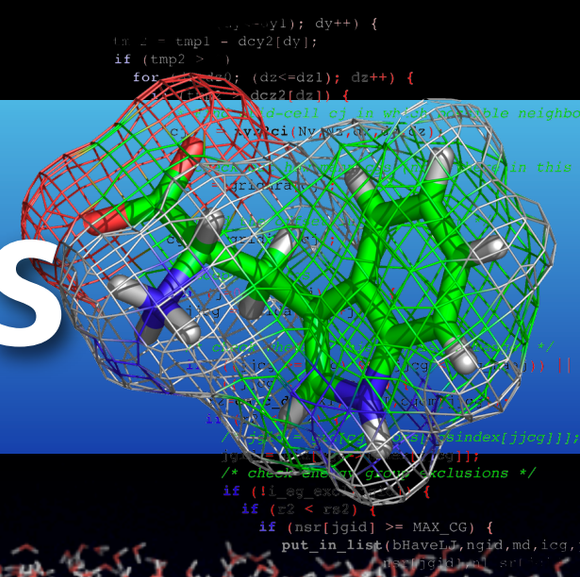
(Cells can control membrane stiffness with cholesterol)

Undulations





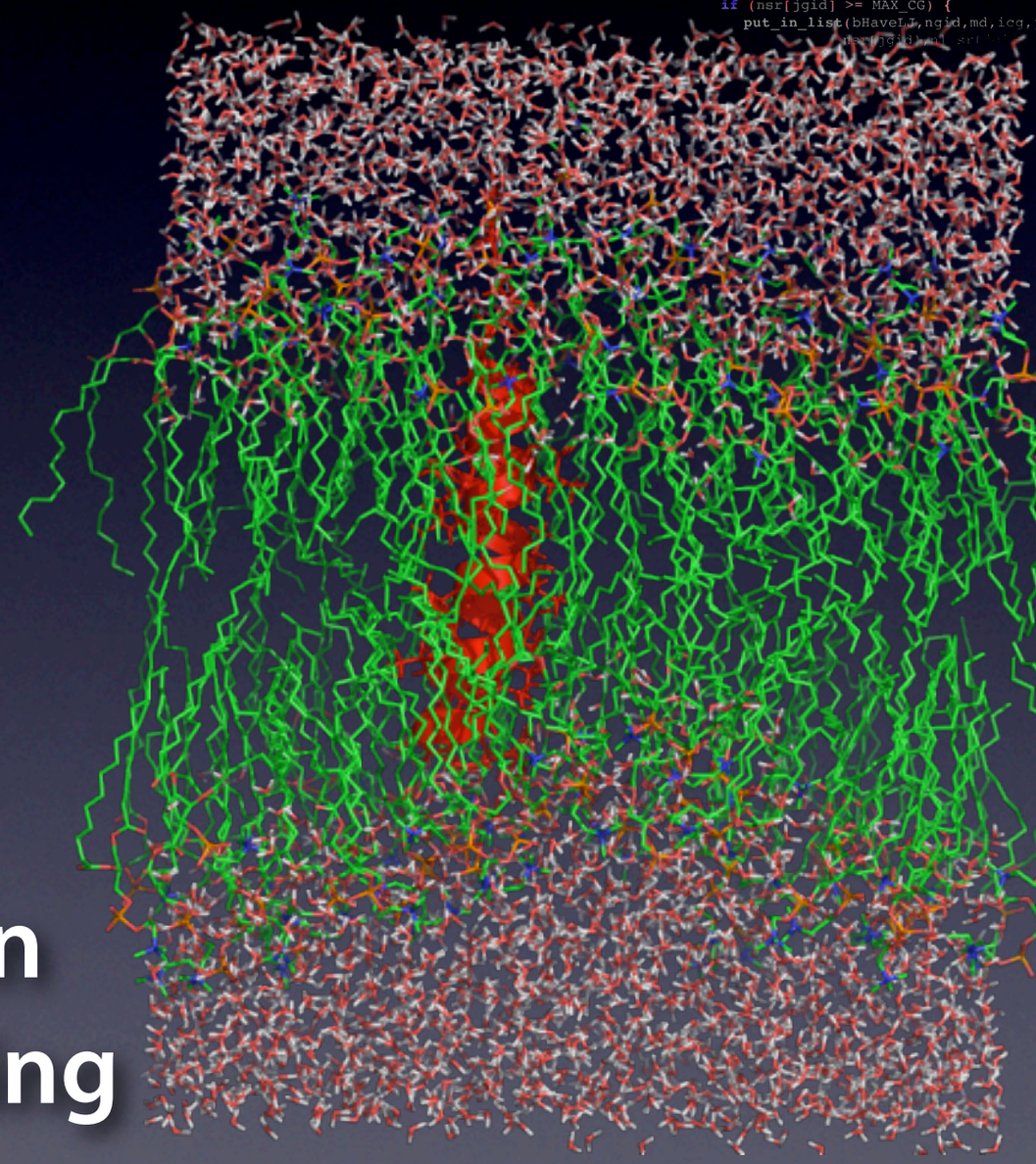
Membrane proteins



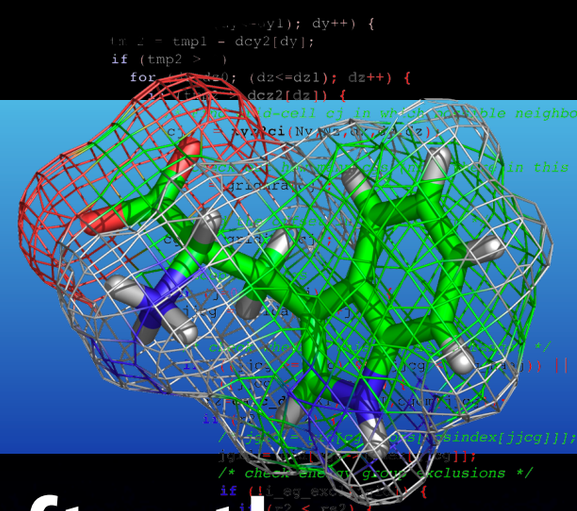
- Prepare structure without membrane
- Solvate protein in bilayer, e.g. with genbox
- Add solvent water
- Freeze protein for relaxation

```
freezegrps = protein
freezedim = y y y
```

- Avoid water entering between lipids and protein by restraining the water z coordinates



Water restraints

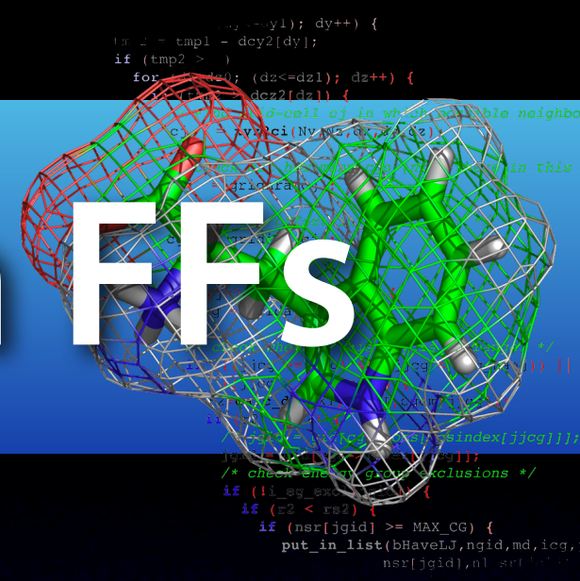


- Open the topology in a text file, and after the water #include line add

```
; Position restraint for each water oxygen
[ position_restraints ]
;  i funct       fcx        fcy        fcz
   1    1          0          0       1000
```

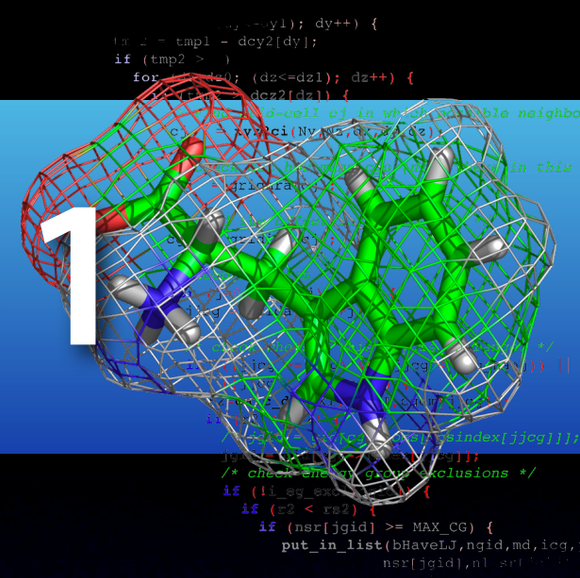
- Waters are free in XY-plane, but z-restrained
- Works for lipids too in really bad system
- Equlibrate lipids ~10ns, then water too
- If water enters the membrane, remove it
- Multiple equilibration cycles can be necessary

Mixed lipid & protein FFs



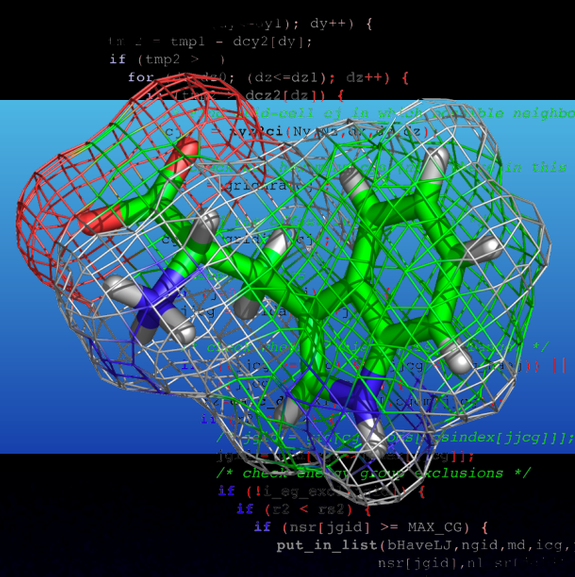
- Berger force field does not include proteins
  However: Based on OPLS, so it mixes nicely with OPLS-AA/L (all-atom) for proteins!
- Charmm27 OK too (but 4-5x more expensive)
- Pure Gromos96, OPLS-AA/L, Amber, etc: constrain cell dimensions for reasonable lipid density & area - can still be OK if lipids are mainly passive solvent for your protein

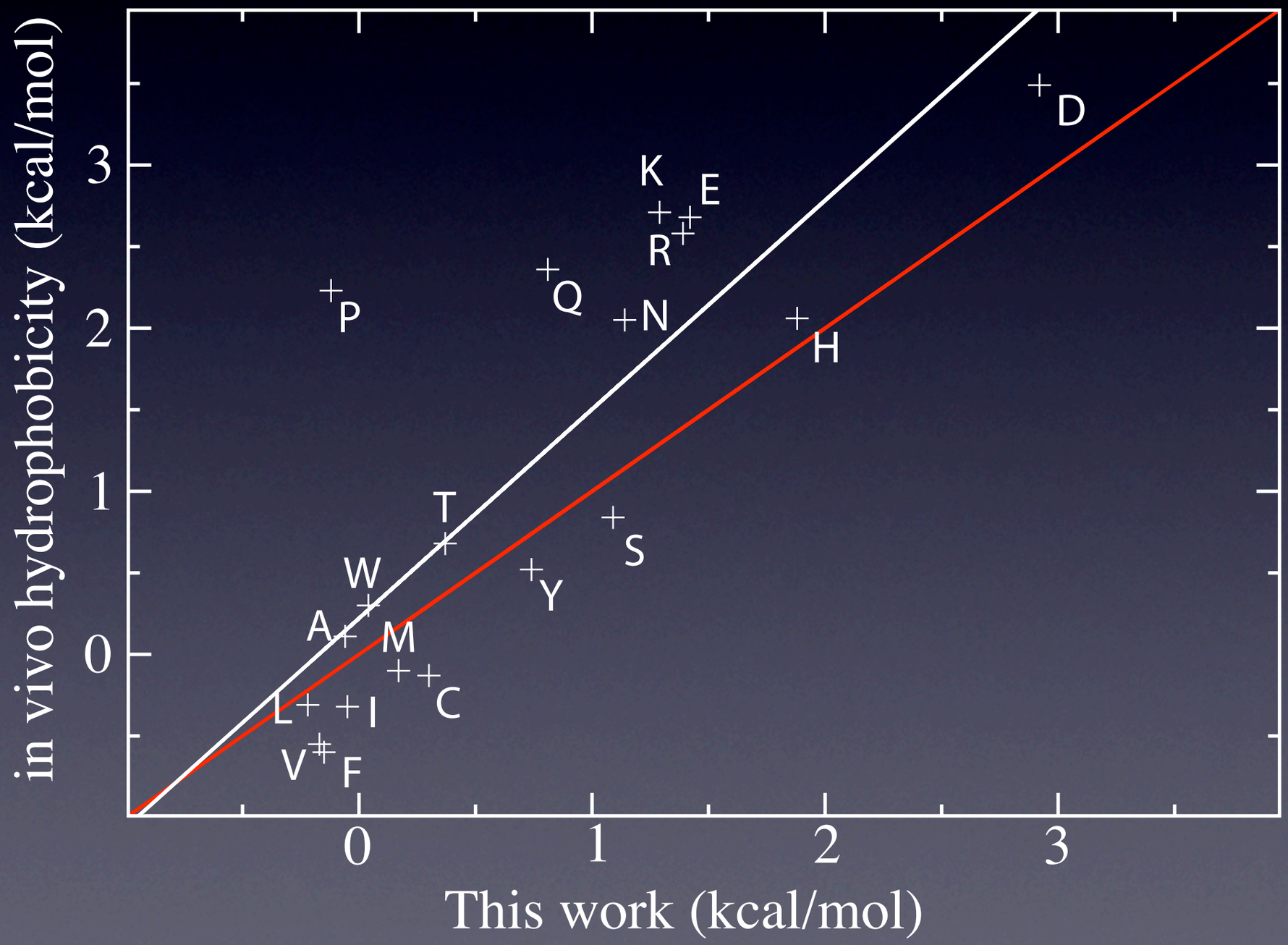
United vs. All-atom - 1



- The Berger force field performs equally well, if not better, compared to Charmm27
- Hydrogens on chains only have low charge No net charge - extremely weak dipoles!
- In principle the weak dipoles could aid the solvation of polar/charged groups
- However, in practice the Berger force field mixed with OPLS-AA reproduce these experimental values very well too!

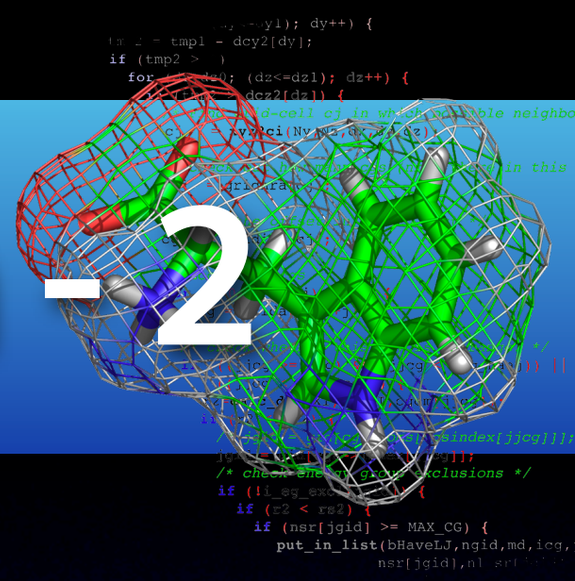
Bilayer solvation





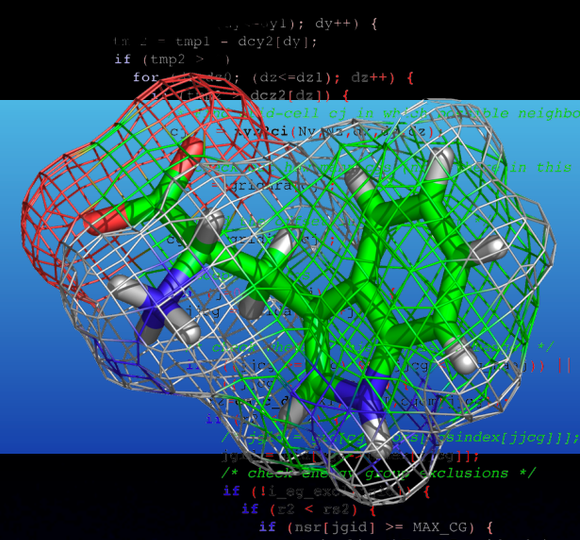
If anything, the solvation cost is slightly low compared to in vivo values!

United vs. All-atom -2



- Nothing wrong with all-atom force fields
- For proteins in water the difference is negligible
- But not for membranes - significant performace boost with united atom model
- “Proof is in the pudding”
- Hard to justify 4x increased cost without any indication the results would be better
- Sampling still worse issue than FF details!

Summary



- Easier than proteins: 256-1024 lipids provide way better statistics
- Harder than proteins: Quite sensitive to force field details, less work done
- Very slow processes (>100ns)
- Interesting non-equilibrium biology: ion transport, insertion, vesicle formation, etc.
- Membranes scale great (>1000 CPUs in CVS)