Parameterization tips and tricks

Manuel Melo (m.n.melo@itqb.unl.pt)
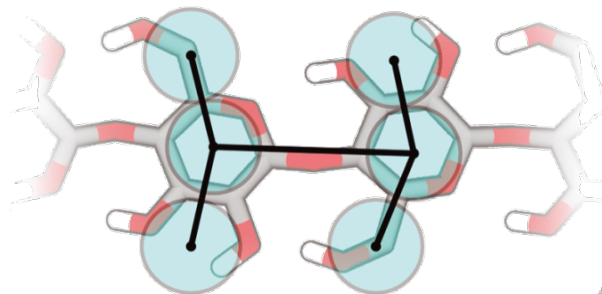
A new topology: what do we need?

Mapping (center of mass)
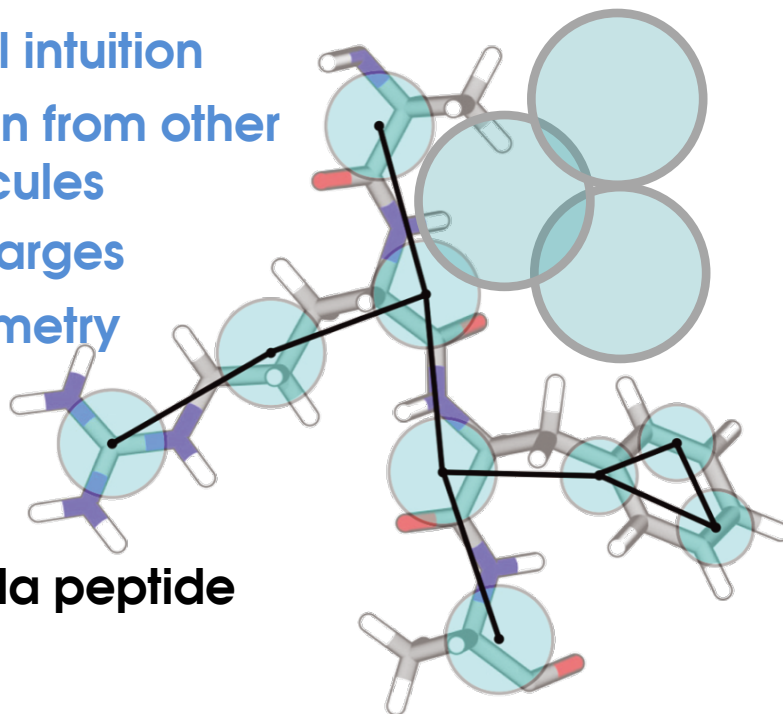
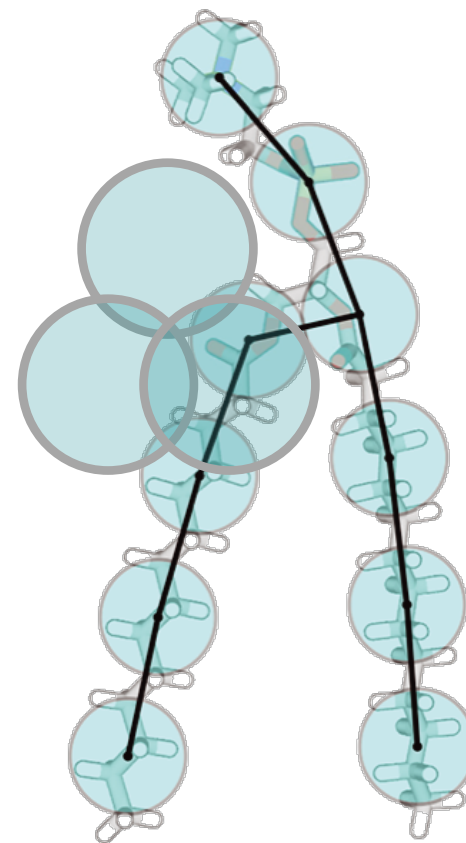*4 heavy atoms to 1 bead*
   (can be finer or coarser)

Not unique
   Use chemical intuition
   Get inspiration from other Martini molecules
   No partial charges
   Respect symmetry
   Be consistent

Cellulose

AlaArgPheAla peptide

Dlauroyl-PC

A new topology: what do we need?

Mapping (center of mass)

4 heavy atoms to 1 bead
(can be finer or coarser)
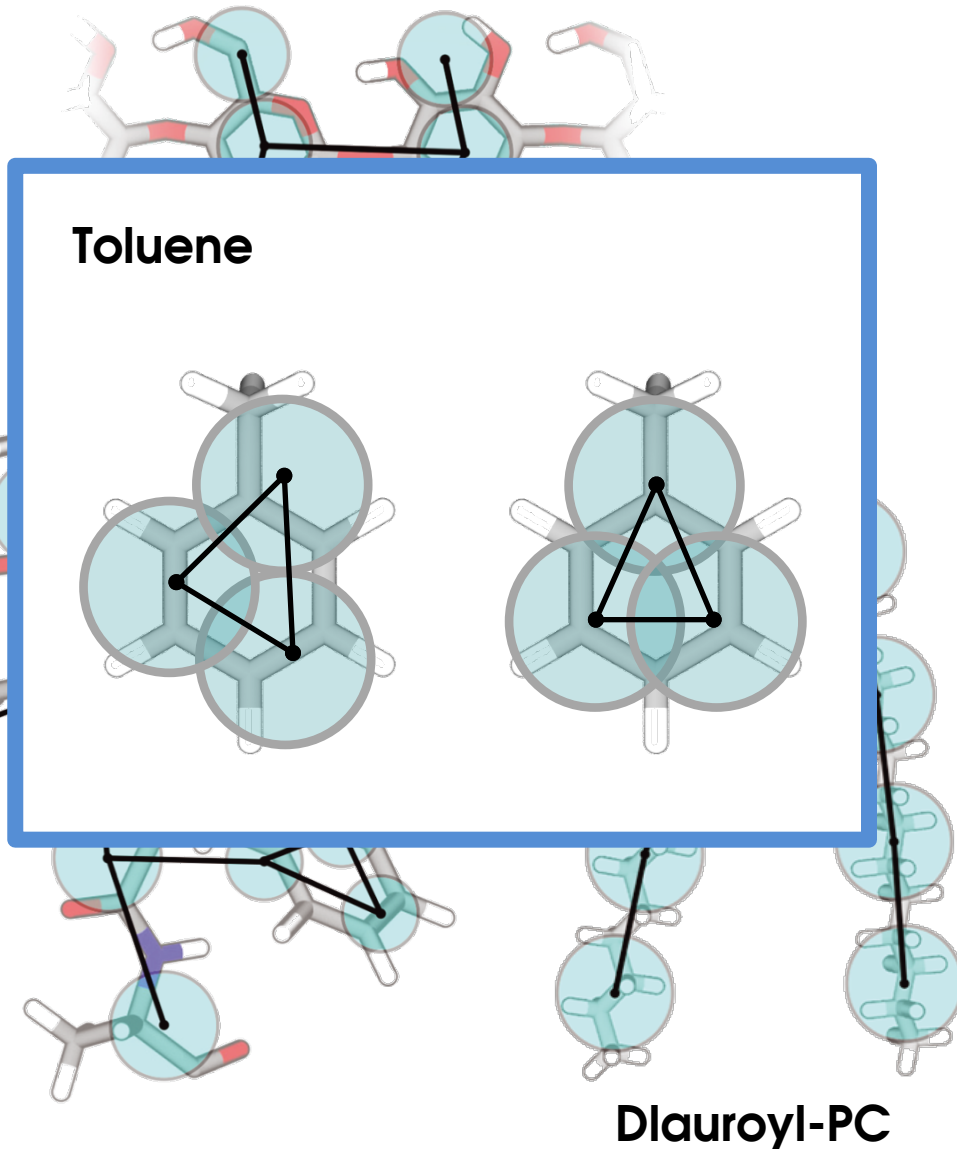
Not unique

Use chemical intuition

Get inspiration from other Martini molecules

No partial charges

Respect symmetry

Be consistent

Toluene

AlaArgPheAla peptide

Dlauroyl-PC
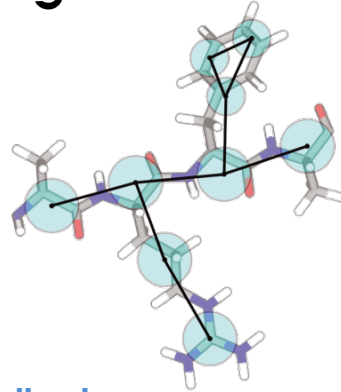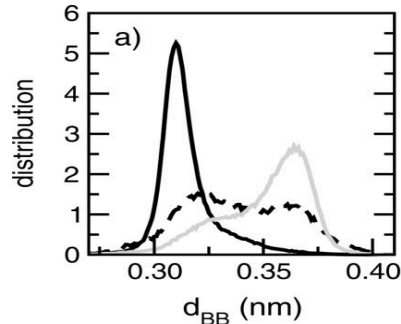
A new topology: what else do we need?

$$U(r) = \sum U_{bonded}(r) + \sum U_{nonbonded}(r)$$

A set of intramolecular potentials that recreate the correct distribution of relative configurations

A set of interparticle potentials that recreate the correct partitioning behavior (and density, surf. tension, etc.)

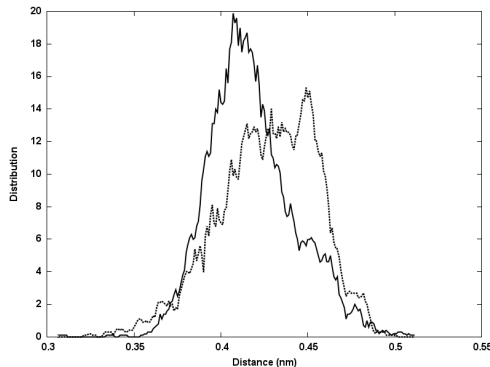## Where do the target distributions/partitions come from?

Experiments

CG-mapped PDB crystallographic/NMR coordinates

Partition data

Simulations with finer models

Bond stretching mapped from a GROMOS 54a6 simulation

ΔG, typically from alchemical modification simulations

Getting the right potentials

$$U_{bonded}(r) \qquad\qquad U_{nonbonded}(r)$$

Which bonded potentials? (besides Boltzmann-inverted ones)

Simple potentials ensure portability across simulation software

Are often optimized

May be insufficient (multimodal distributions, for instance)

Bias towards potentials implemented in GROMACS

Bonds:

$$U(x) = \frac{1}{2}k(x - x_0)^2$$

Simple

Numerically stable

Symmetric distribution (Gaussian)

Angles:

$$U(\theta) = \frac{1}{2}k(\cos(\theta) - \cos(\theta_0))^2$$

Numerically stable, unlike $U(\theta) = \frac{1}{2}k(\theta - \theta_0)^2$

Weak potential towards colinearity

Getting the right potentials

$$U_{bonded}(r) \qquad\qquad U_{nonbonded}(r)$$

Which bonded potentials? (besides Boltzmann-inverted ones)

Simple potentials ensure portability ac

Are often optimized

May be insufficient (multimodal dist

Bias towards potentials implemented i
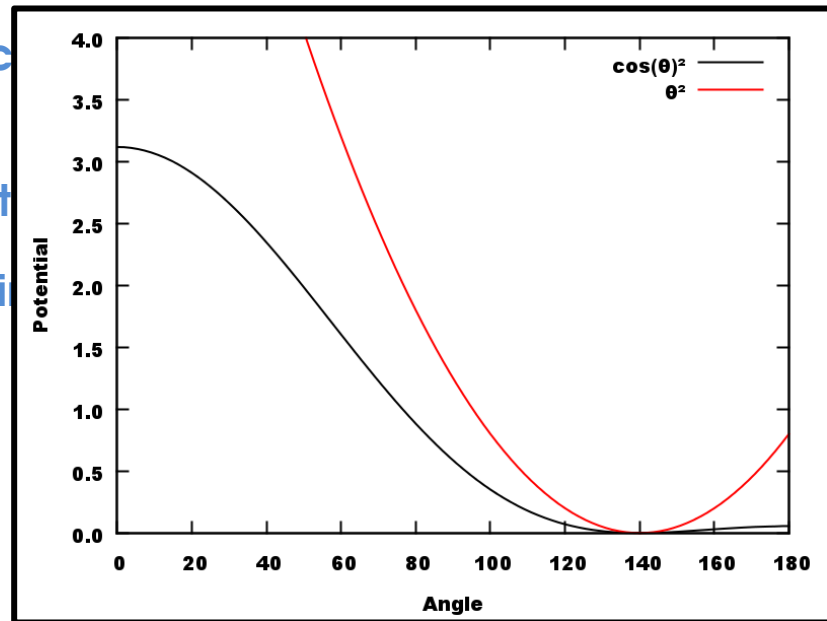
Bonds:

$$U(x) = \frac{1}{2}k(x - x_0)^2$$

Angles:

$$U(\theta) = \frac{1}{2}k\big(\cos(\theta) - \cos(\theta_0)\big)^2$$

Numerically stable, unlike $\quad U(\theta) = \frac{1}{2}k(\theta - \theta_0)^2$

Weak potential towards colinearity

Getting the right potentials

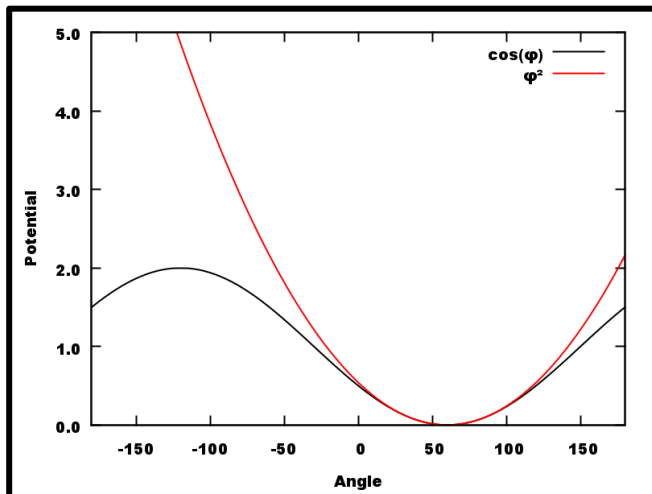$$U_{bonded}(r) \qquad\qquad U_{nonbonded}(r)$$

Which bonded potentials?

Dihedral angles:

$$U(\phi) = k\left(1 + \cos(n\phi - \phi_r)\right)$$

$$U(\phi) = \frac{1}{2}k\left(\phi - \phi_r\right)^2$$

Periodic

Can have multiple minima (*n*)

Suitable for keeping torsions that do not flip

Both types become unstable if any two of the constructing bonds become colinear!

Use together with restricted bending!

$$U(\theta) = \frac{1}{2}k\frac{\left(\cos(\theta) - \cos(\theta_0)\right)^2}{\sin^2(\theta)}$$

Getting the right potentials

$$U_{bonded}(r)$$

$$U_{nonbonded}(r)$$

Which bonded potentials?

Dihedral angles:

$$U(\phi) = k(1 + \cos(n\phi - \phi_r))$$

$$U(\phi) = \frac{1}{2}k(\phi - \phi_r)^2$$

o not flip

Both types become unstable if any two of the constructing bonds become colinear!

Use together with restricted bending!

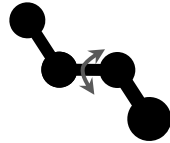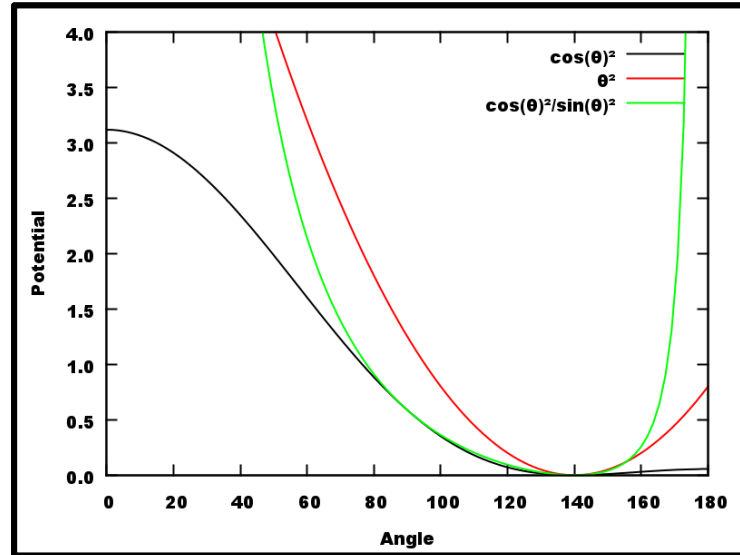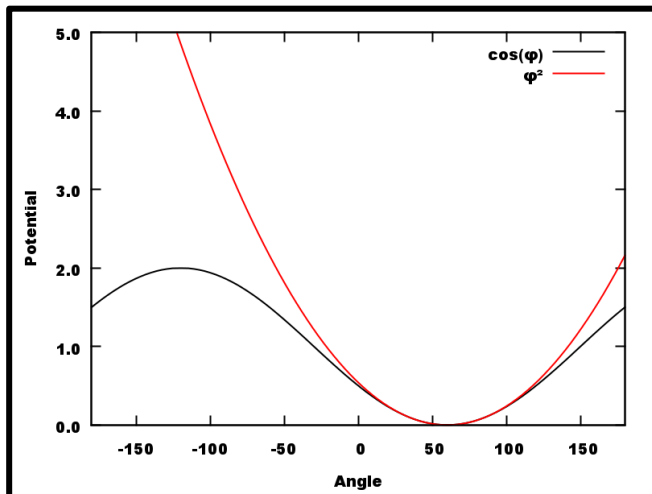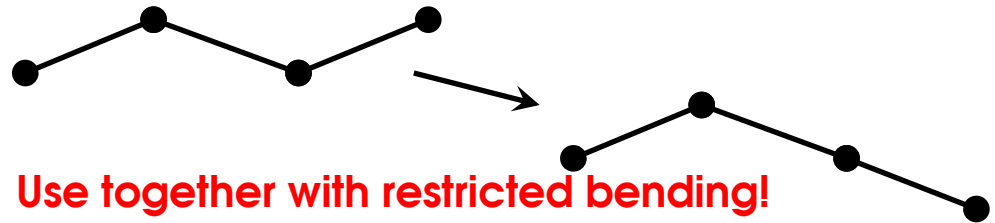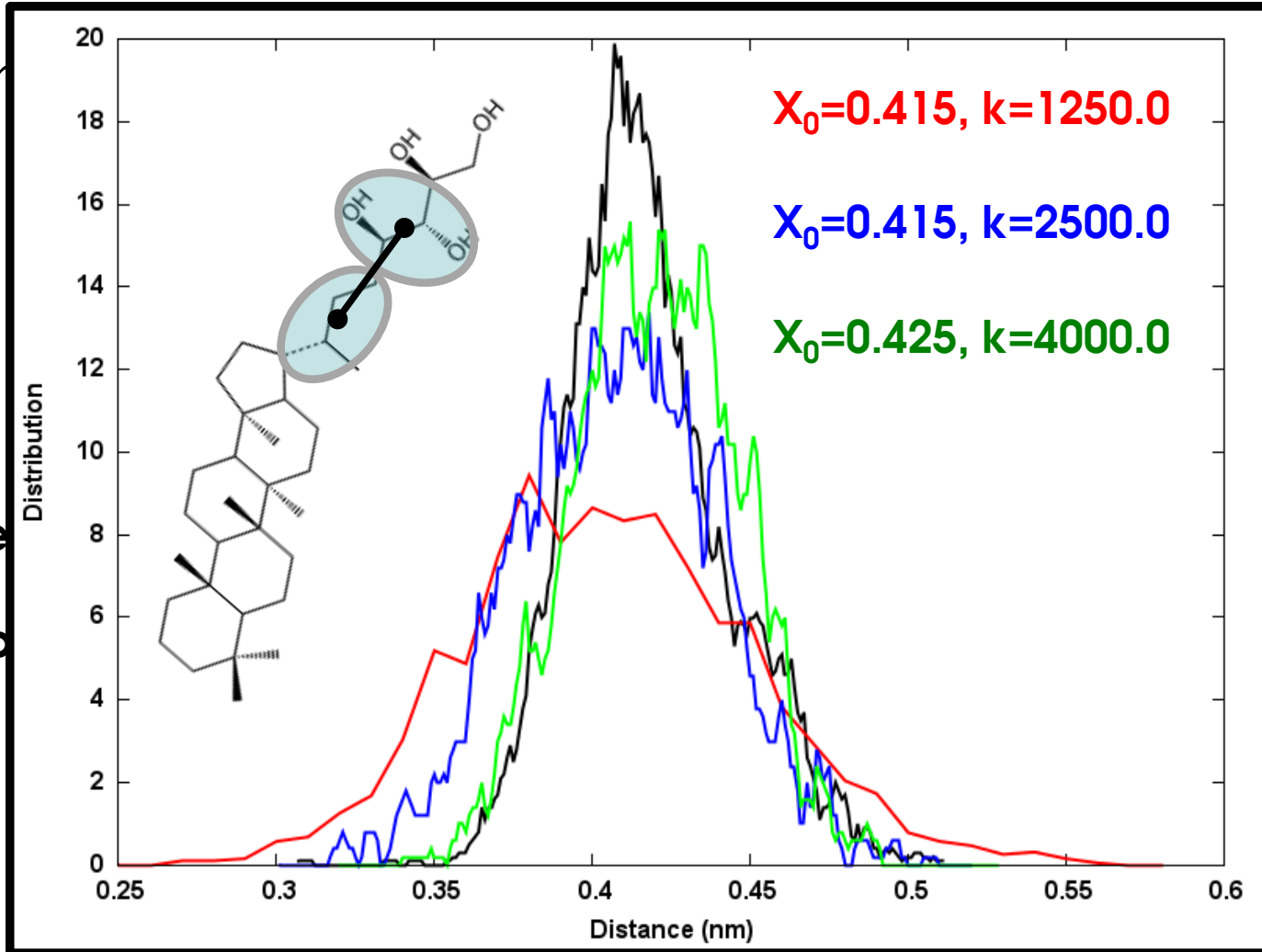$$U(\theta) = \frac{1}{2}k\frac{(\cos(\theta) - \cos(\theta_0))^2}{\sin^2(\theta)}$$

Getting the right potentials

$U_{bon\ldots}$ $\ldots(r)$

Repe…
until
happ…

X$_0$=0.415, k=1250.0

X$_0$=0.415, k=2500.0

X$_0$=0.425, k=4000.0

$$U(x) = \frac{1}{2}k(x - x_0)^2, \quad (k : \text{kJ/mol} \cdot \text{nm}^2, \ x : \text{nm})$$
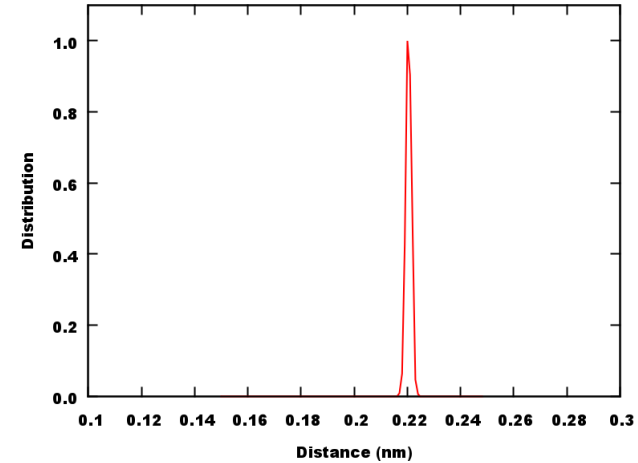
Constraints

Highly localized distributions

Narrow distributions require stiff potentials

Stiff potentials require short time steps

At the Martini scale the breadth of stiff distributions, and the high oscillation frequencies, become unimportant
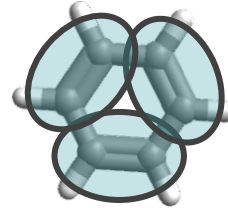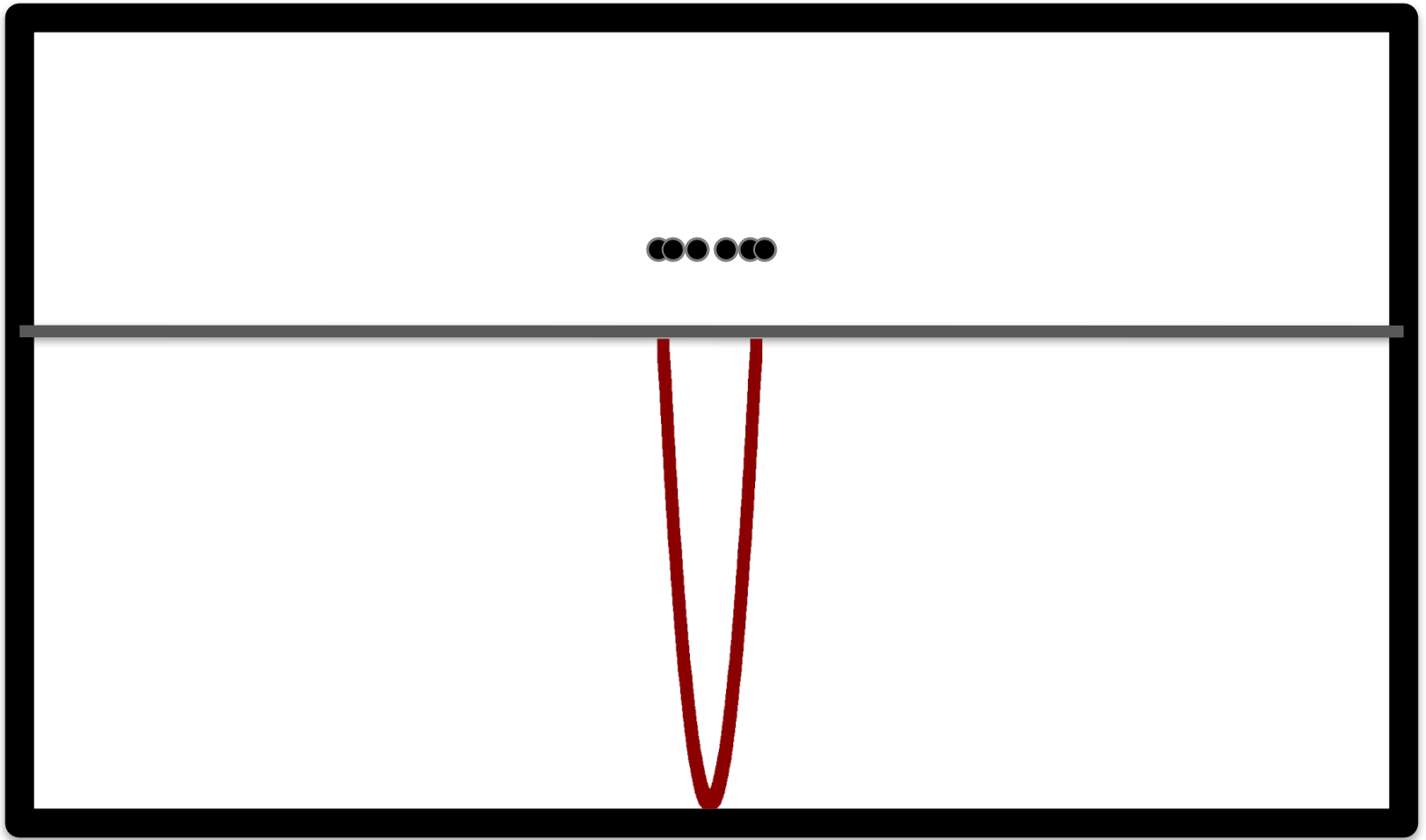
Constraints

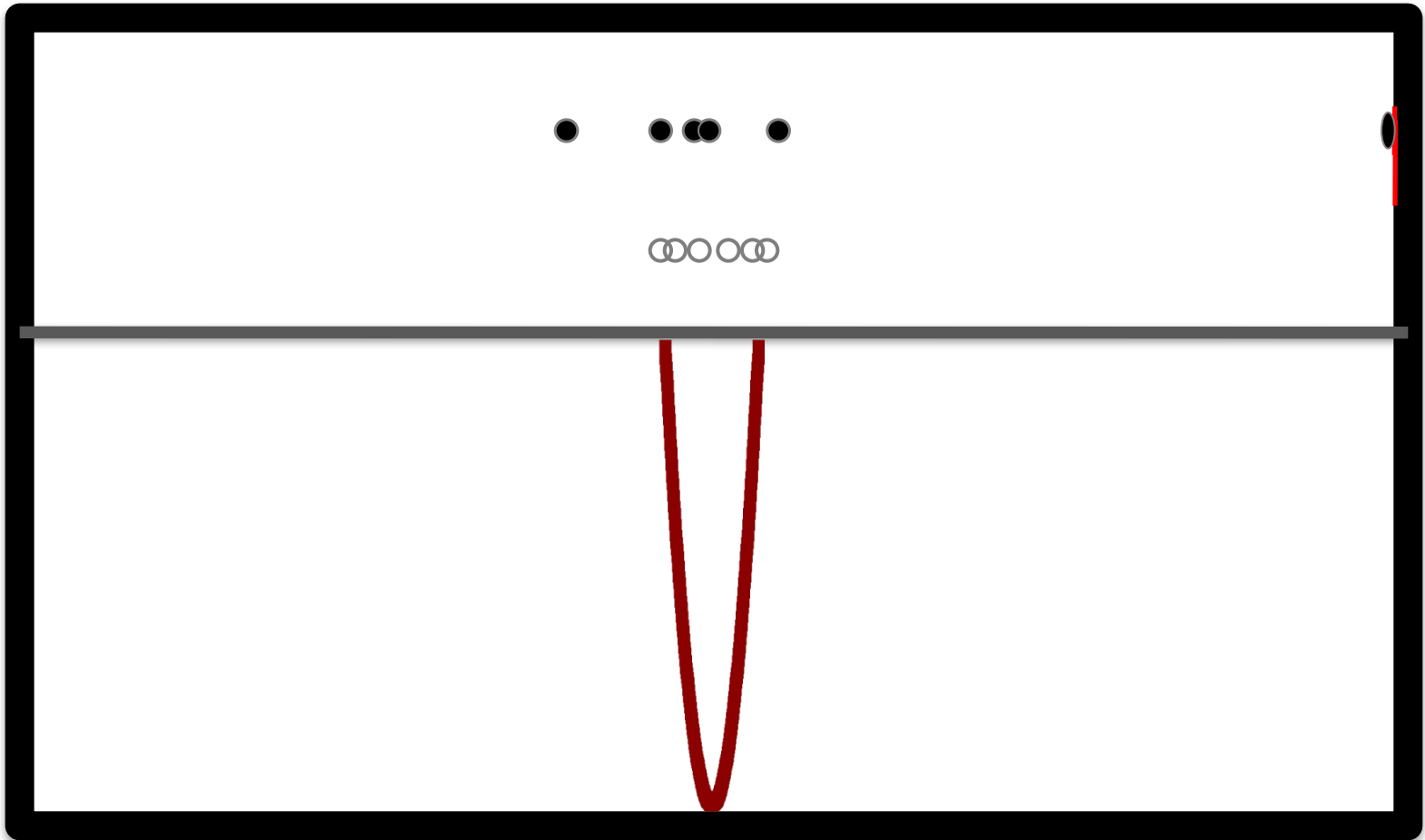## Highly localized distributions

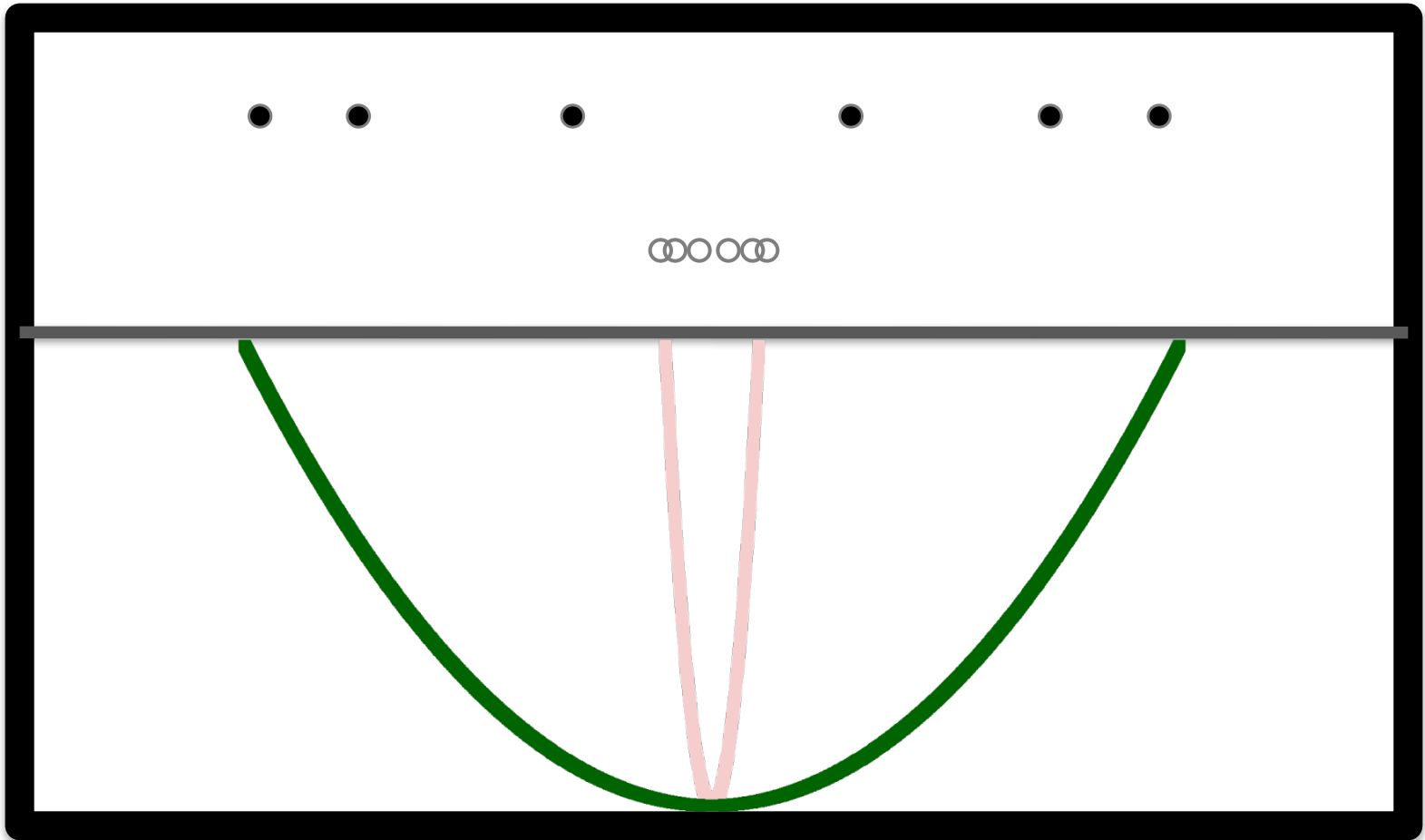Narrow distributions require stiff potentials

Stiff potentials require short time steps

At the Martini scale the breadth of stiff distributions, and the high oscillation frequencies, become unimportant
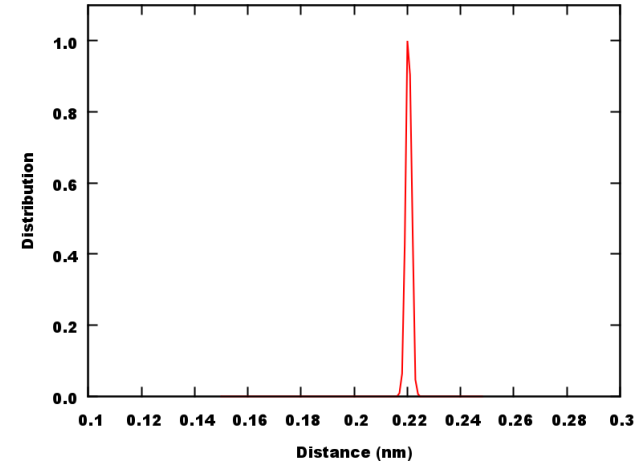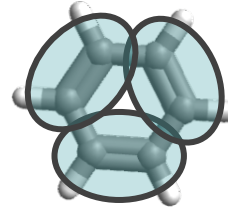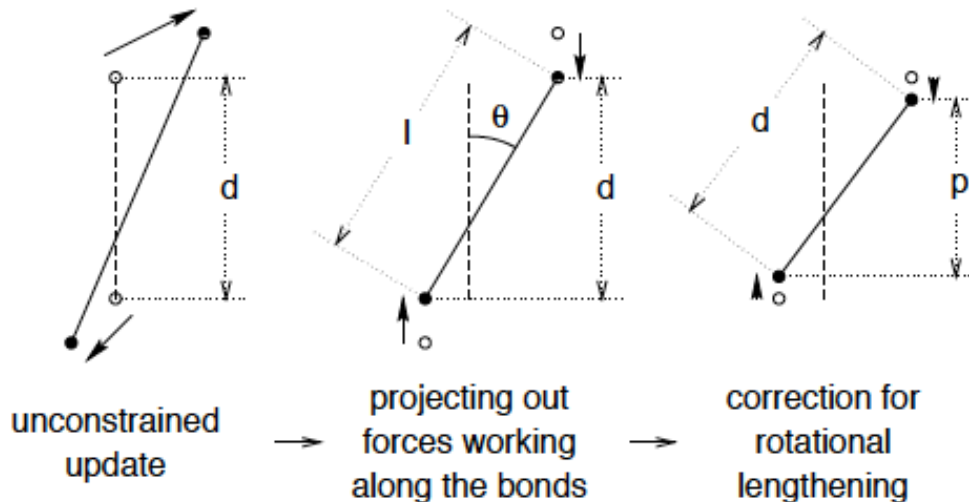

Use constraints

Interparticle distance becomes a system constant (1 DOF less)

unconstrained update → projecting out forces working along the bonds → correction for rotational lengthening

Getting the right potentials: bead types

$$U_{bonded}(r)$$

$$U_{nonbonded}(r)$$

TABLE 3: Thermodynamic Properties of the CG Particle Types[a]

| type	building block	examples	ΔG^{vap} exp	CG	ΔG^{hyd} exp	CG	ΔG^{part}_{HW} exp	CG	ΔG^{part}_{CW} exp	CG	ΔG^{part}_{EW} exp	CG	ΔG^{part}_{OW} exp	CG
Q_{da}	$H_3N^+-C_2-OH$	ethanolamine (protonated)				−25		< −30		−18		−13		−18
Q_d	$H_3N^+-C_3$	1-propylamine (protonated)				−25		< −30		−18		−13		−18
	NA^+OH	sodium (hydrated)				−25		< −30		−18		−13		−18
Q_a	PO_4^-	phosphate				−25		< −30		−18		−13		−18
	CL^-HO	chloride (hydrated)				−25		< −30		−18		−13		−18
Q_0	C_3N^+	choline				−25		< −30		−18		−13		−18
P_5	$H_2N-C_2=O$	acetamide	sol	sol	−40	−25	−27	−28	(−20)	−18	−15	−13	−8	−10
P_4	HOH (× 4)	water	−27	−18	−27	−18	−25	−23		−14	−10	−7	−8	−9
	$HO-C_2-OH$	ethanediol	−35	−18	−33	−18	−21	−23		−14		−7	−8	−9
P_3	$HO-C_2-O$	acetic acid	−31	−18	−29	−18	−19	−21	−9	−10	−2	−6	−1	−7
	$C-NH-C=O$	methylformamide	−35	−18		−18		−21		−10		−6	−5	−7
P_2	C_2-OH	ethanol	−22	−16	−21	−14	−13	−17	−5	−2	−3	1	−2	−2
P_1	C_3-OH	1-propanol	−23	−16	−21	−14	−9	−11	−2	−2	0	1	1	−1
		2-propanol	−22	−16	−20	−14	−10	−11	−2	−2	−1	1	0	−1
N_{da}	C_4-OH	1-butanol	−25	−16	−20	−9	−5	−7	2	0	4	2	4	3
N_d	H_2N-C_3	1-propylamine	−17	−13	−18	−9	(−6)	−7	(1)	0	(−3)	2	(3)	3
N_a	$C_3=O$	2-propanone	−17	−13	−16	−9	−6	−7	1	0	−1	2	−1	3
	$C-NO_2$	nitromethane	−23	−13	−17	−9	−6	−7		0		2	−2	3
	$C_3=N$	proprionitrile	−22	−13	−17	−9	−5	−7		0		2	1	3
	$C-O-C=O$	methylformate	−16	−13	−12	−9	(−6)	−7	(4)	0	(−1)	2	(0)	3
	$C_2HC=O$	propanal		−13	−15	−9	−4	−7		0	2	2	3	3
N_0	$C-O-C_2$	methoxyethane	−13	−10	(−8)	−2	(1)	−2		6	(3)	6	(3)	5
C_5	C_3-SH	1-propanethiol	−17	−10		1		5		10		10		6
	$C-S-C_2$	methyl ethyl sulfide	−17	−10	−6	1	(7)	5		10		10	(9)	6
C_4	$C_2=C_2$	2-butyne	−15	−10	−1	5		9		13		13	9	9
	$C=C-C=C$	1,3-butadiene		−10	2	5	11	9		13		13	11	9
	$C-X_4$	chloroform	−18	−10	−4	5	(7)	9	14	13		13	11	9
C_3	$C_2=C_2$	2-butene		−10		5		13		13		13	13	14
	C_3-X	1-chloropropane	−16	−10	−1	5	12	13		13		13	12	14
		2-bromopropane	−16	−10	−2	5		13		13		13	12	14
C_2	C_3	propane	gas	−10	8	10		16		15		14	14	16
C_1	C_4	butane	−11[b]	−10	9	14	18	18		18		14	16	17
		isopropane	gas	−10	10	14		18		18		14	16	17
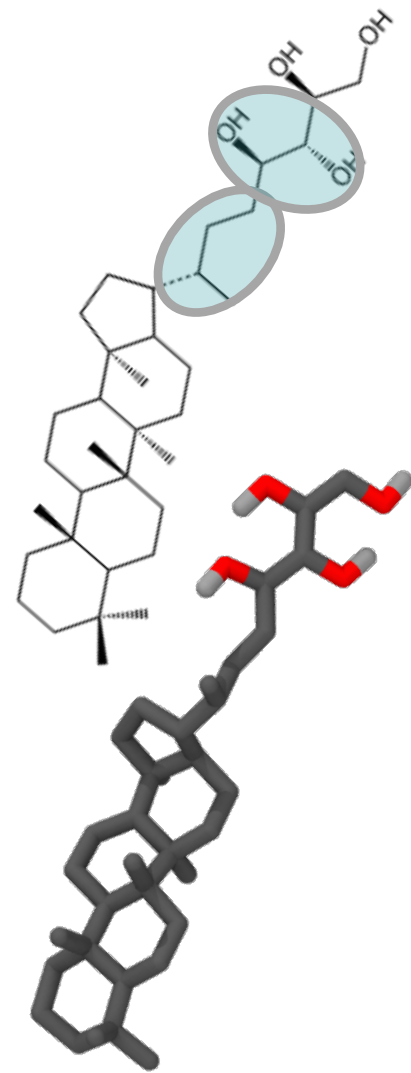
Getting the right potentials: bead types

$$U_{bonded}(r) \qquad\qquad U_{nonbonded}(r)$$

Verify the reproduction of partition free energies and adjust bead types accordingly
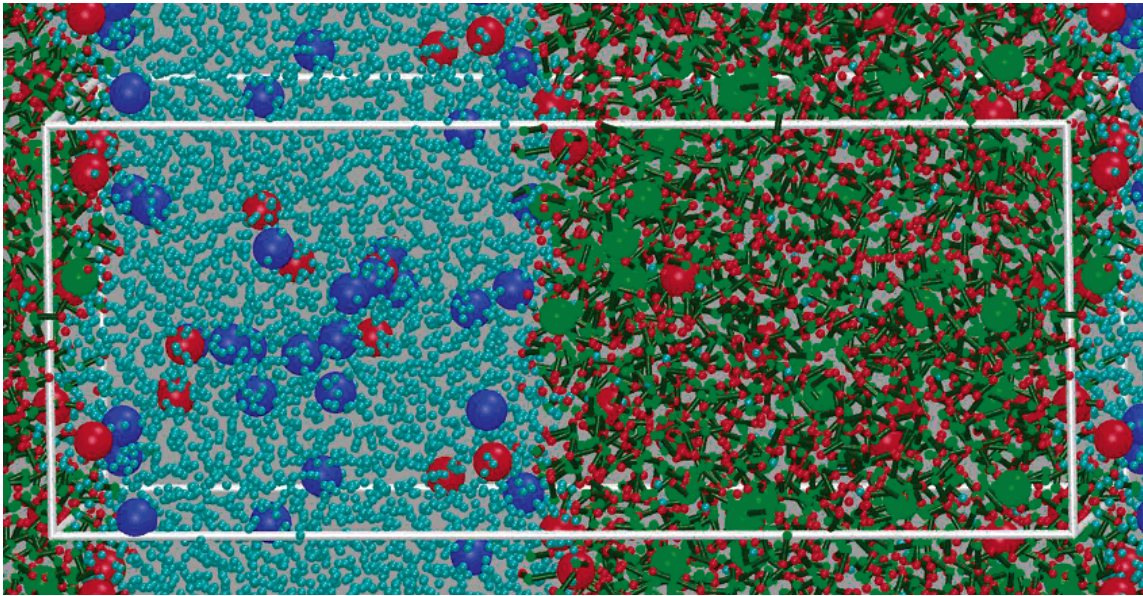
$$\Delta G_{part} = kT\ln\big([solute]_W / [solute]_O\big)$$

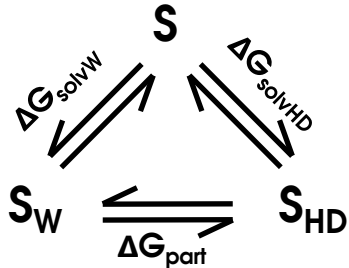Significant counts in both phases must be obtained

The interface may play a role

Too expensive to be used atomistically as a source of target free-energy data
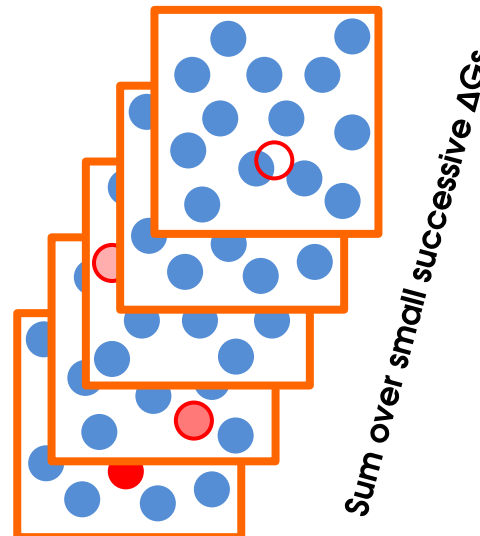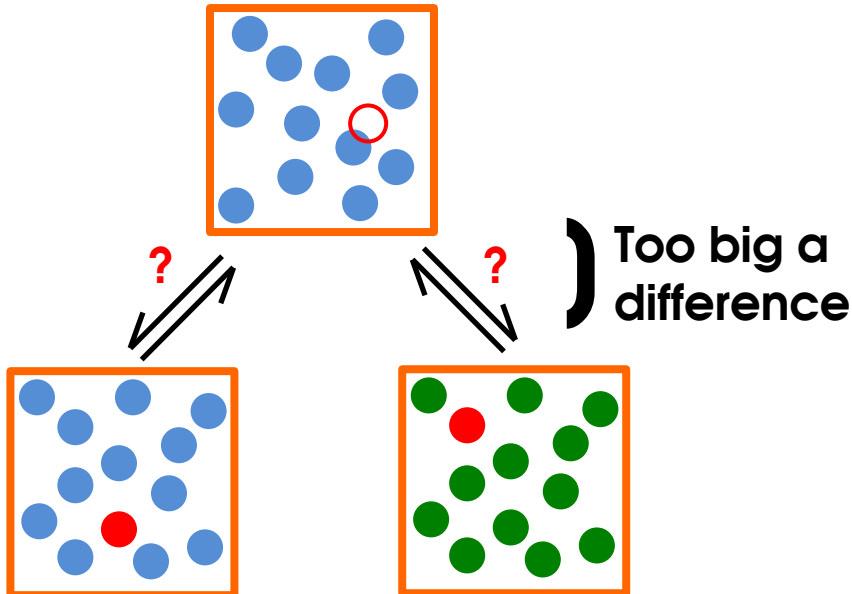
Getting the right potentials: bead types

$$U_{bonded}(r)$$

$$U_{nonbonded}(r)$$

A more efficient approach through the hydration free-energies

Alchemical decoupling

Too big a difference

Almost always more efficient than running a system with separated phases

Can be used with atomistic systems

Sum over small successive ΔGs

Getting the right potentials: bead types

$$U_{bonded}(r)$$ $$U_{nonbonded}(r)$$

Important: tailor the matched data to your applications!

Biomolecular applications: hydrophobic/hydrophilic

Not exclusively partitions
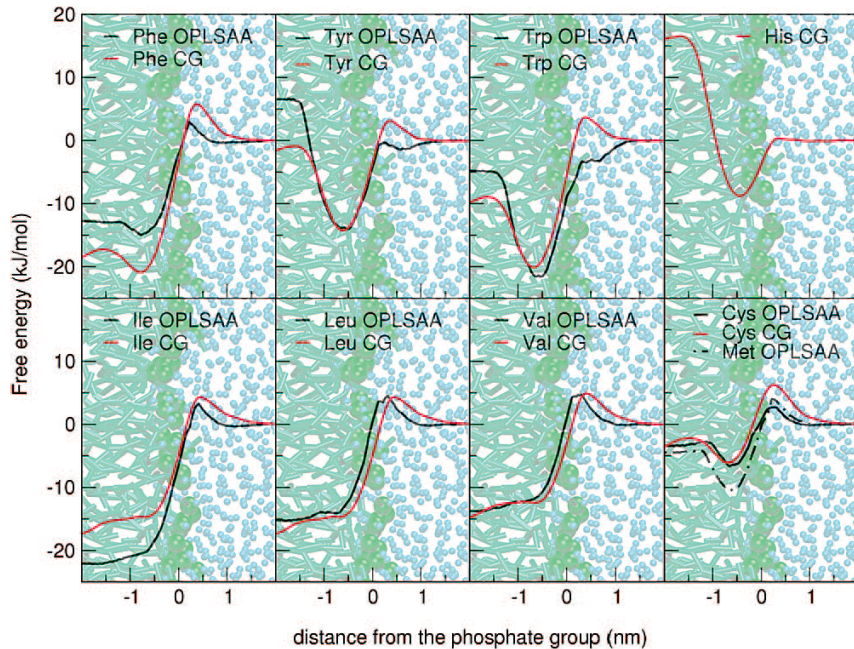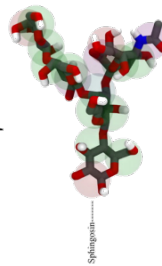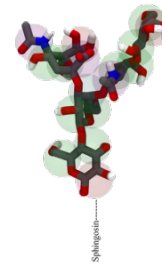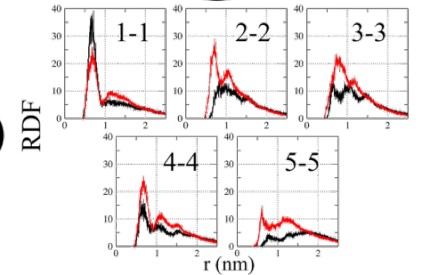
Getting the right potentials: bead types

$$U_{bonded}(r) \qquad\qquad U_{nonbonded}(r)$$

When a finer mapping is needed (to represent planar geometries, for instance)

Very high bead density

Effective very deep energy well (condensation)

Equilibrium distance of ~0.52 nm
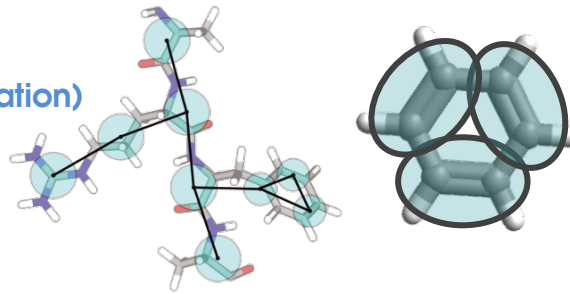
The S-beads

25% shallower potentials with shorter equilibrium distance (~0.43nm)

Allow the correct packing of rings

Prevent condensation of the system

Follow the same scale as regular Martini beads (SC1…SP5…SQda)

When not to use?

To make a higher resolution mapping (overmapping is BAD)

To fill in apparent gaps in your mapping (overmapping is BAD)

overmapping is BAD (maybe this will change in the near future)
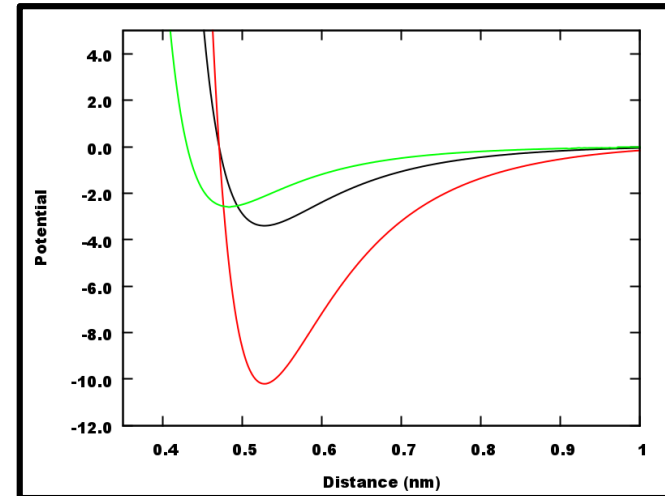
Interaction with regular beads still follows the regular potential: validate the free energy of interaction!

Important notes

Keep the Martini philosophy in mind but tailor it to your applications

If parameterizing solvents/melts, aim to reproduce bulk properties

For polymers
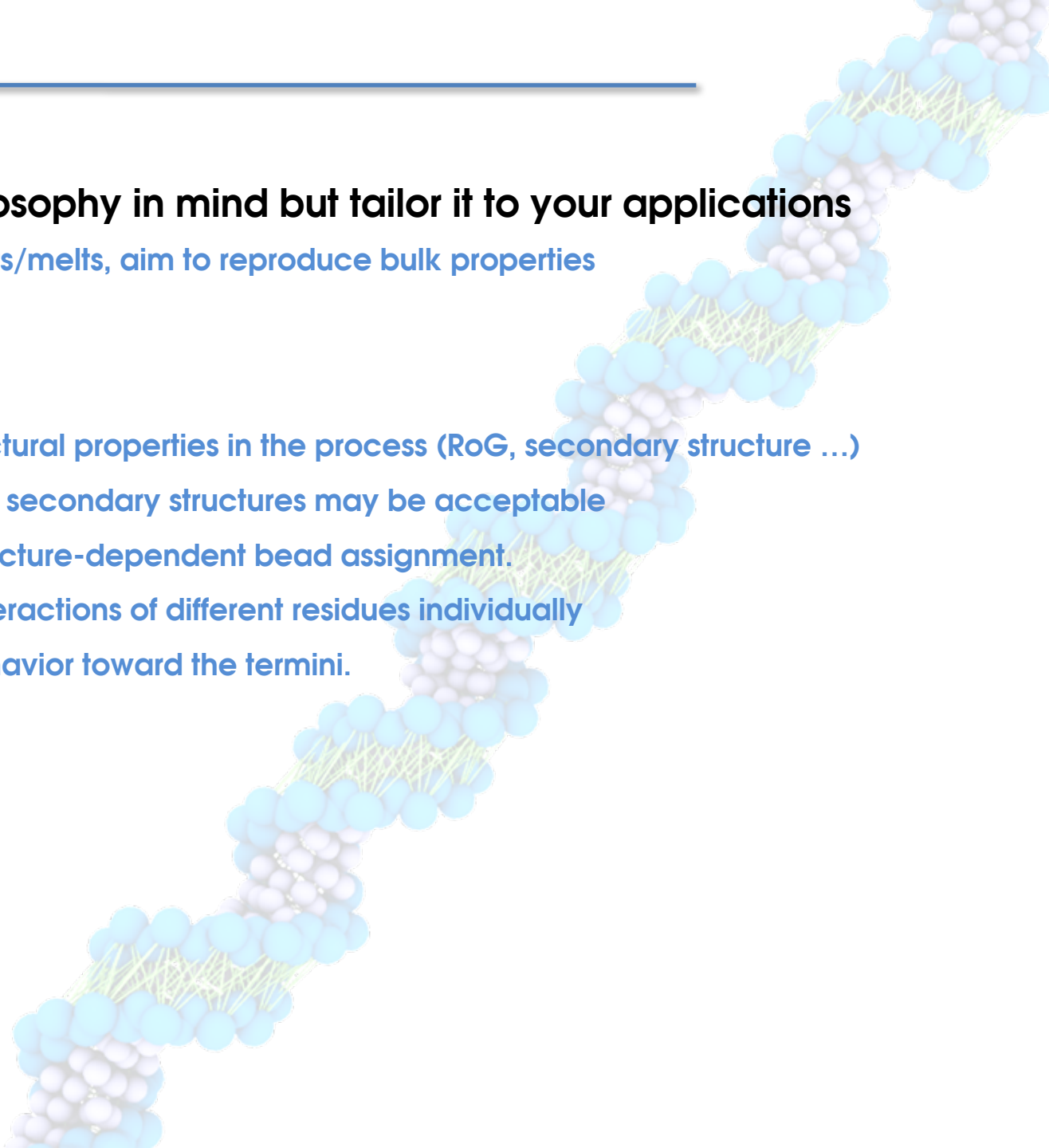
Include long-range structural properties in the process (RoG, secondary structure …)

Don't give up: restrained secondary structures may be acceptable

But may require structure-dependent bead assignment.

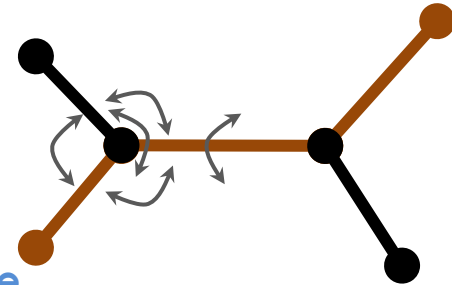Tune the nonbonded interactions of different residues individually

Beware of divergent behavior toward the termini.

More tips & tricks

Don't over-restrain the bonded interactions

Redundancy makes convergence difficult to achieve

Often the nonbonded interactions push a free angle/dihedral into place

Exclusions and fake bonds between nonbonded particles

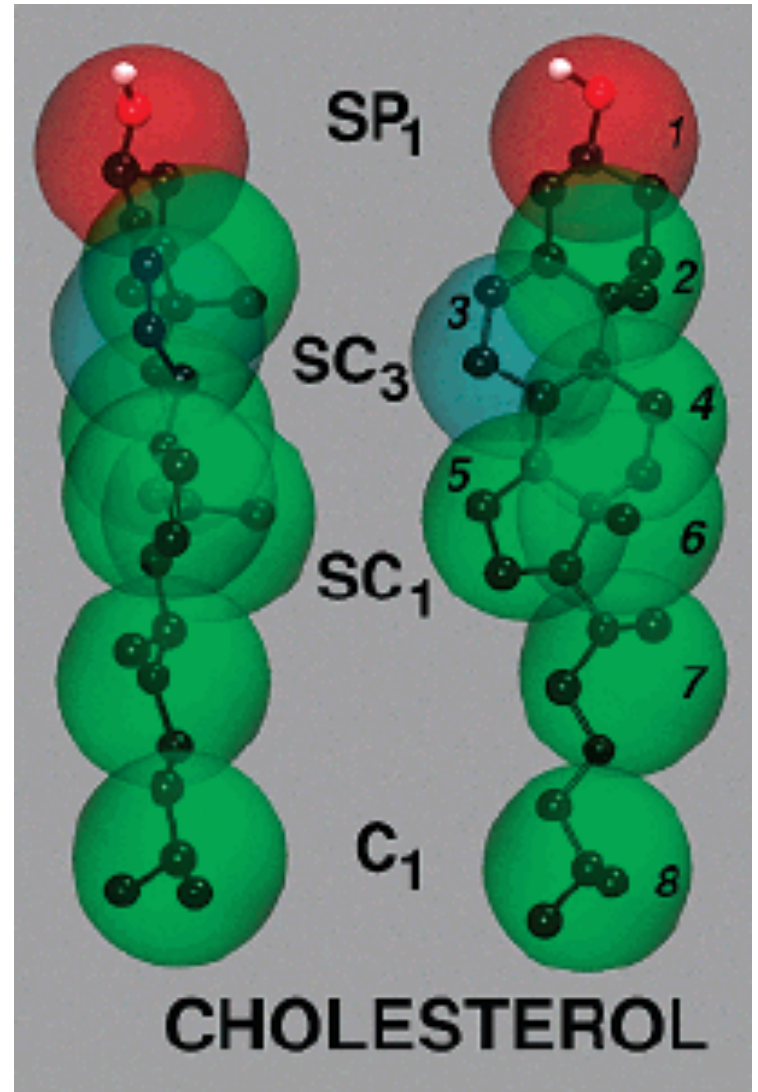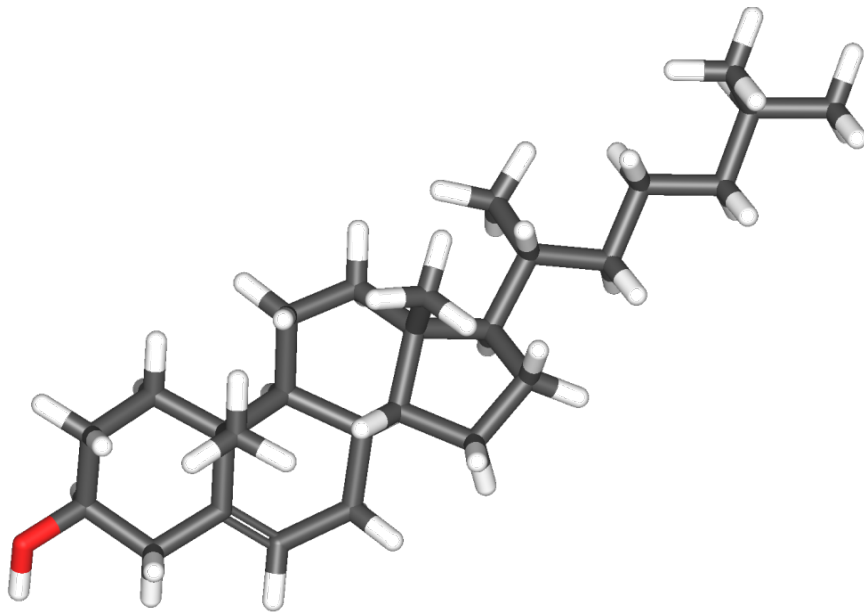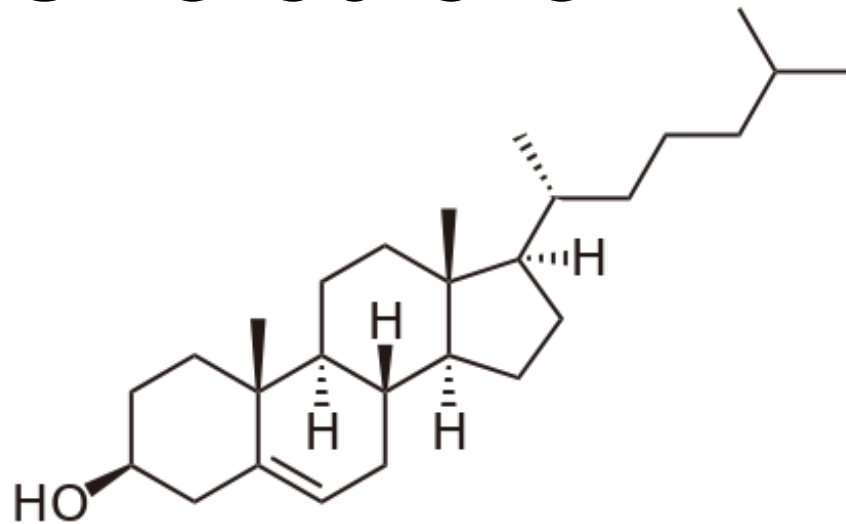Beware of beads at distances below the nonbonded repulsion limit — consider excluding them

Bonds can be made between nonconsecutive beads
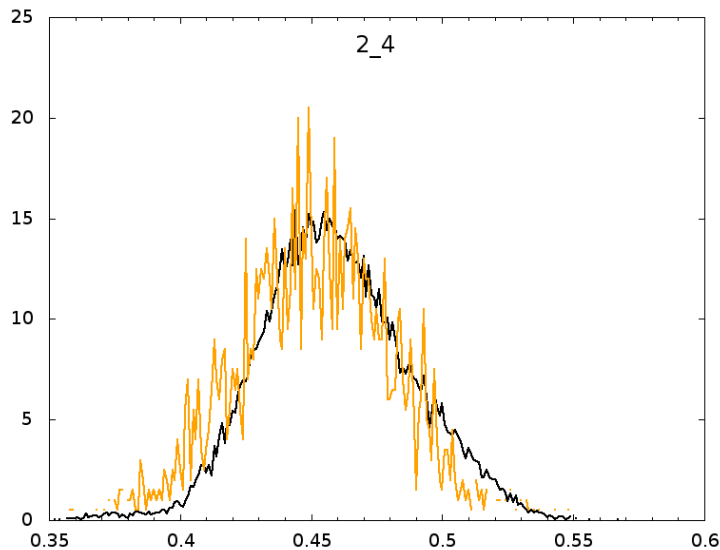
Advanced parameterization: Virtual sites

Cholesterol

SP₁

SC₃

SC₁

C₁

CHOLESTEROL

The Martini model
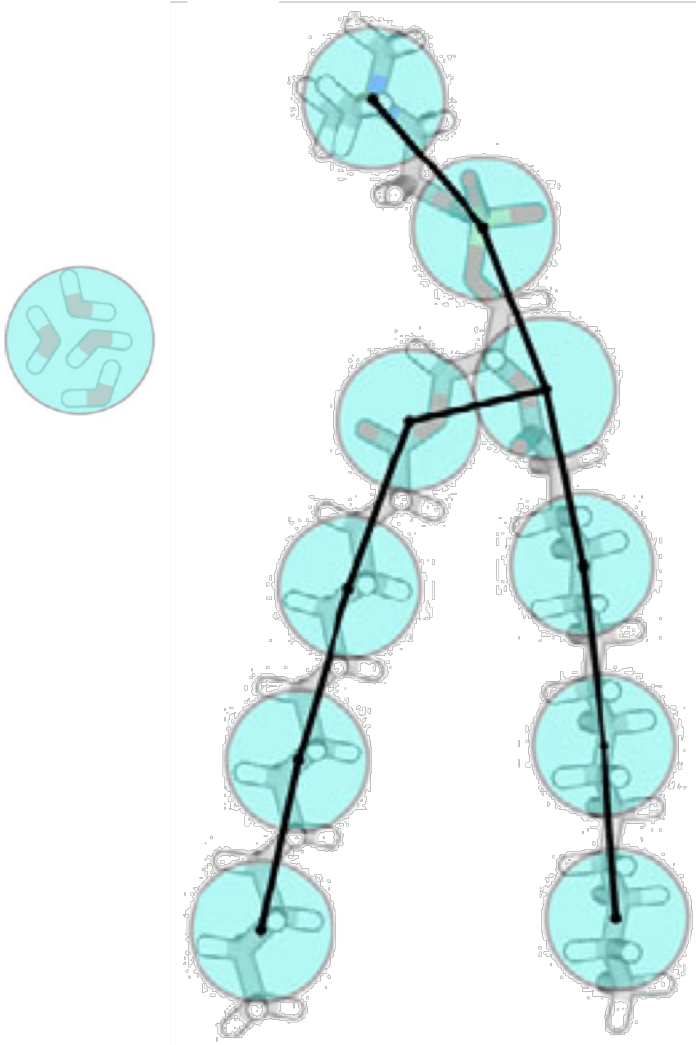
Fewer particles

Softer potentials

>1000x speedup

up to 40 fs timesteps
(20-40x larger than atomistic)

but...

```
bonds that rotated more than 30 degrees:
 atom 1 atom 2   angle   previous, current, constraint length
     3      5    33.9      0.4466   0.3470        0.3460
     5      7    56.6      0.4478   0.4017        0.4060
     4      5    34.9      0.3997   0.2857        0.2940

Step 8073, time 322.92 (ps)  LINCS WARNING
relative constraint deviation after LINCS:
rms 0.505886, max 1.177047 (between atoms 5 and 7)
bonds that rotated more than 30 degrees:
 atom 1 atom 2   angle   previous, current, constraint length
     3      4    83.5      0.2731   0.2632        0.2720
     3      5    83.8      0.3470   0.4081        0.3460
     5      7   117.0      0.4017   0.8839        0.4060
     4      5    47.2      0.2857   0.3689        0.2940
Wrote pdb files with previous and current coordinates

Step 8074, time 322.96 (ps)  LINCS WARNING
relative constraint deviation after LINCS:
rms 1662.365558, max 4036.032715 (between atoms 5 and 7)
bonds that rotated more than 30 degrees:
 atom 1 atom 2   angle   previous, current, constraint length
     3      4    91.6      0.2632  62.9947        0.2720
     3      5    86.5      0.4081 117.0789        0.3460
     5      7    89.6      0.8839 1639.0353       0.4060
     4      5    94.3      0.3689 101.7266        0.2940
     1      3   103.8      0.3853  11.2611        0.4930
     1      4    86.7      0.6435  41.9594        0.6040
Wrote pdb files with previous and current coordinates
Segmentation fault      (core dumped)
```
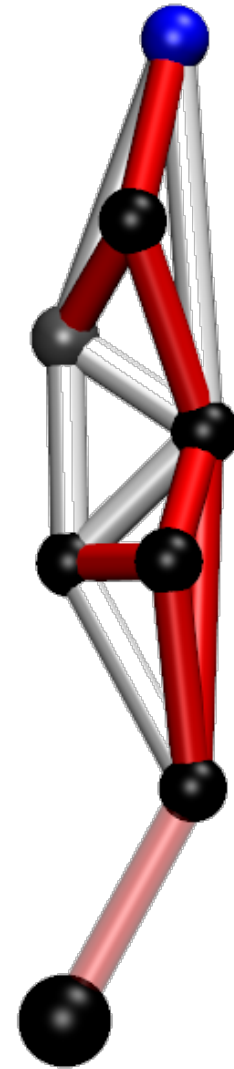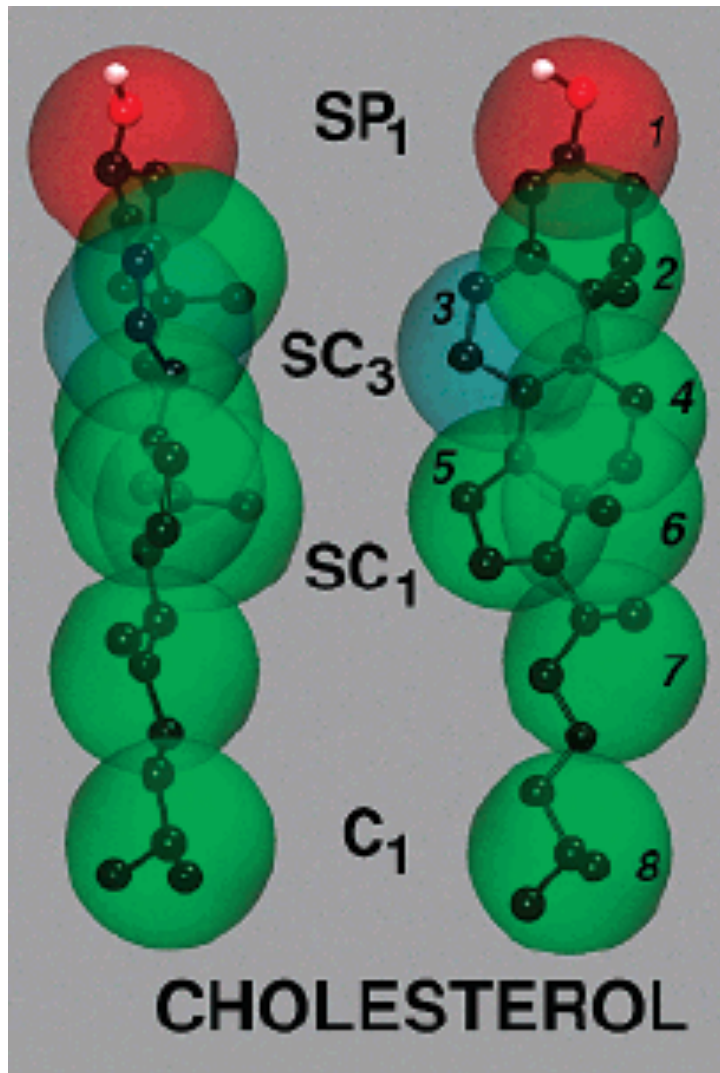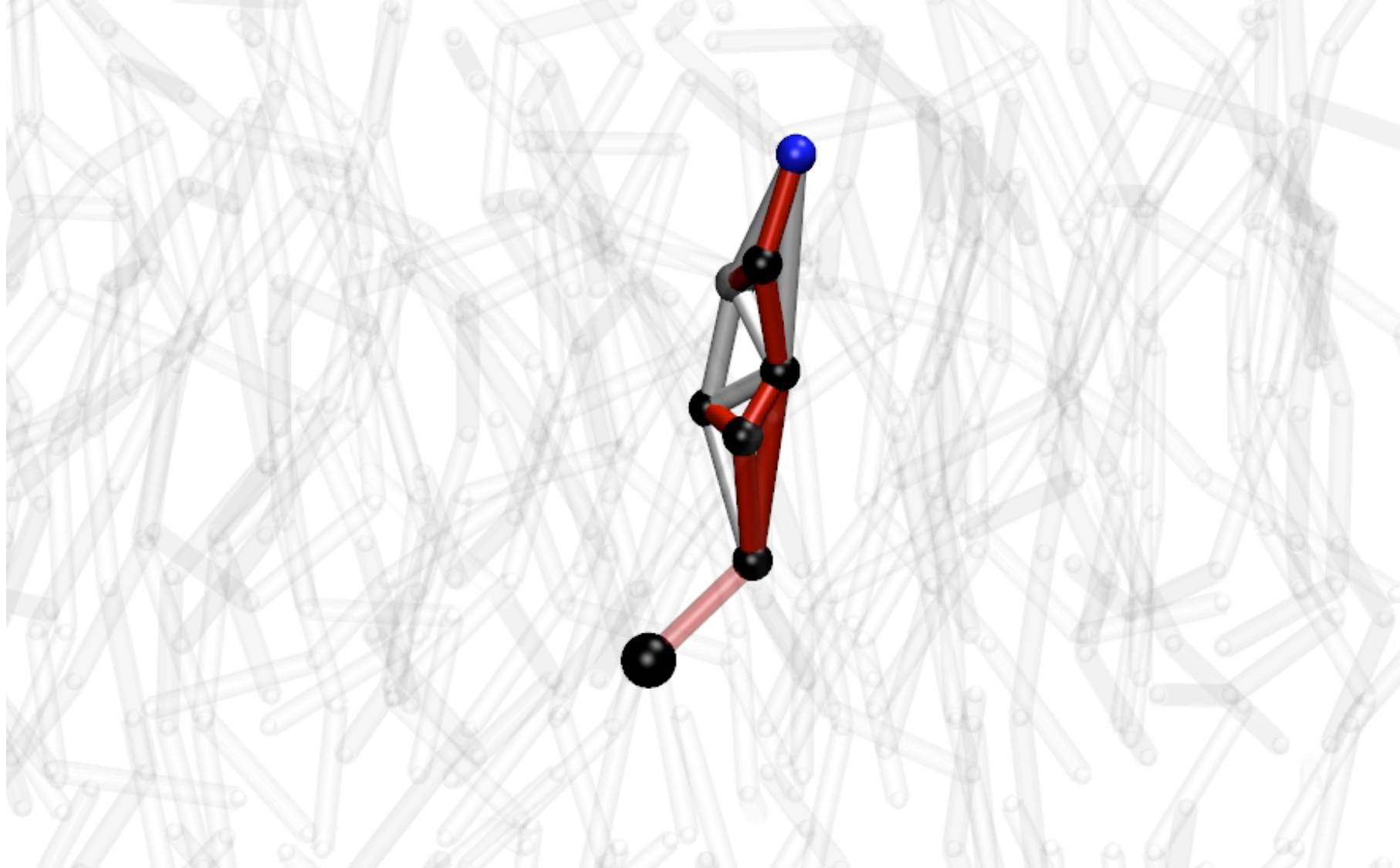
SP$_1$

1

2

3

SC$_3$

4

5

SC$_1$

6

7

C$_1$

8

CHOLESTEROL

```
#!/bin/bash
while gmx mdrun -v -cpi state.cpt -noappend -maxh 0.05
do
    rm -rf *part*
done
```
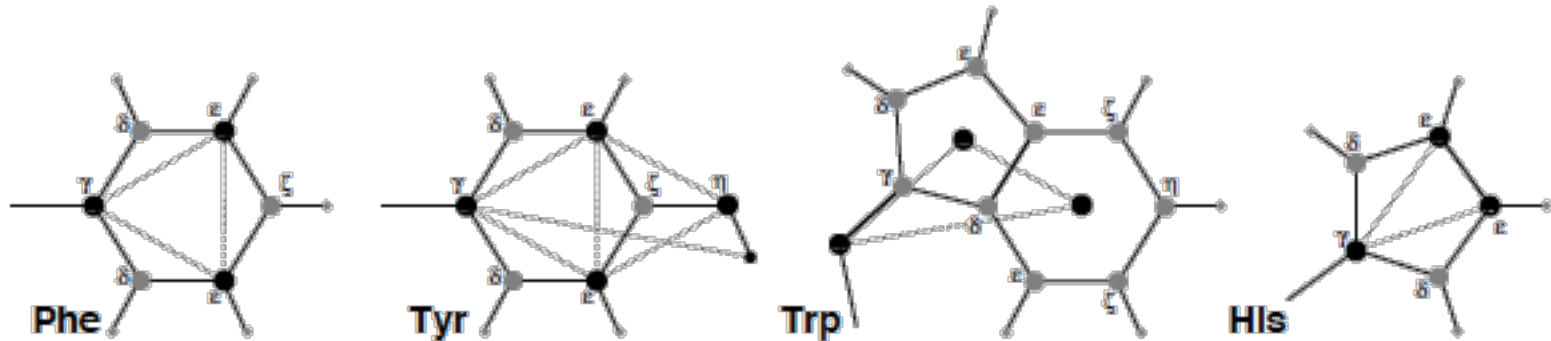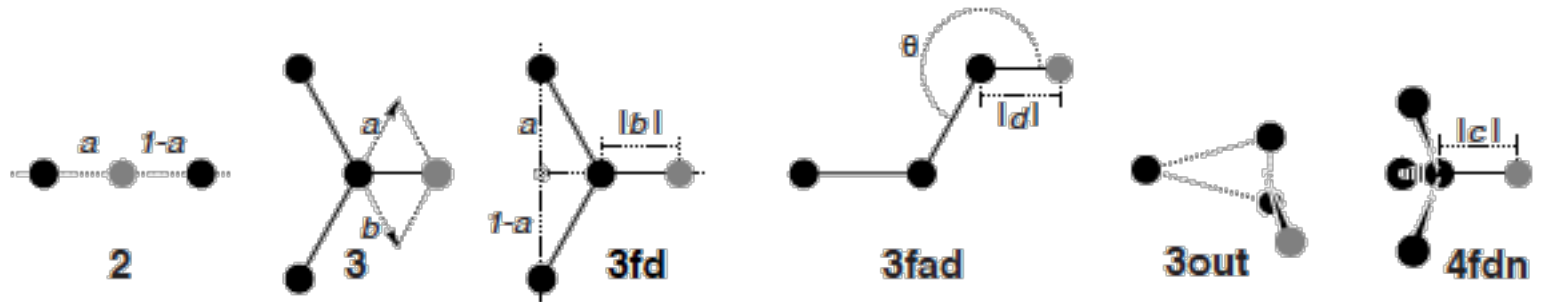
What can be done?

Decrease the time step

Constrain the whole thing

Increase the bead masses

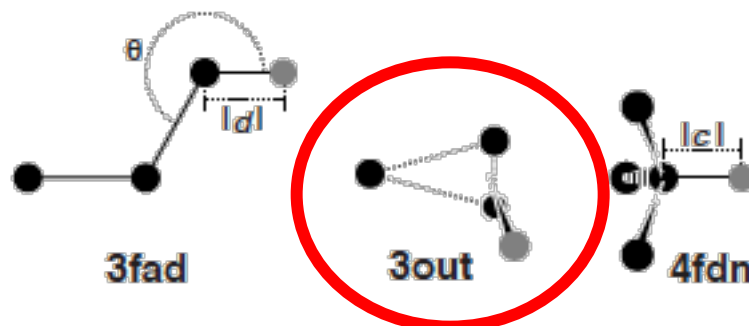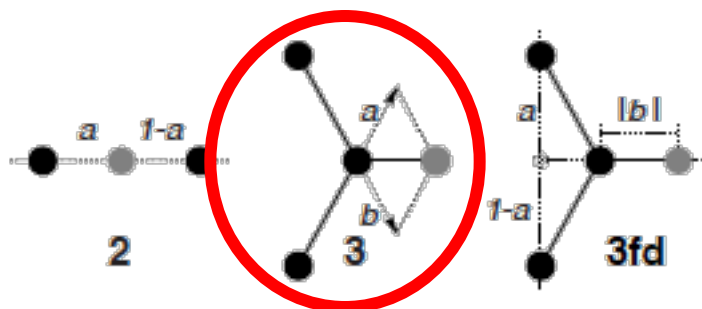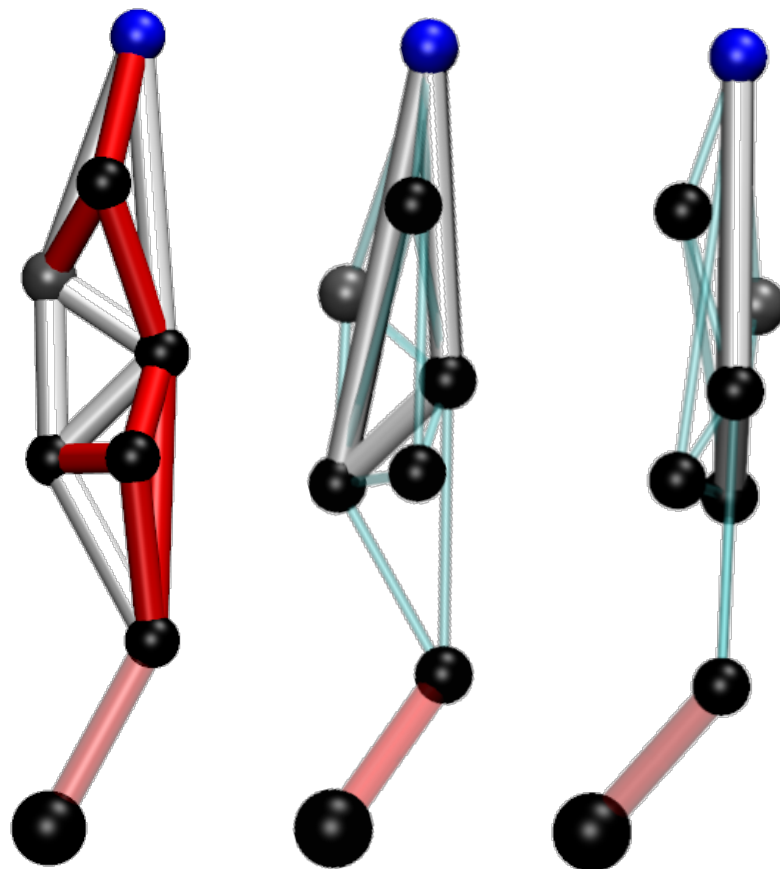Use virtual interaction sites

Virtual interaction sites?

What I did

Chose three atoms for my frame

Obtained the average positions
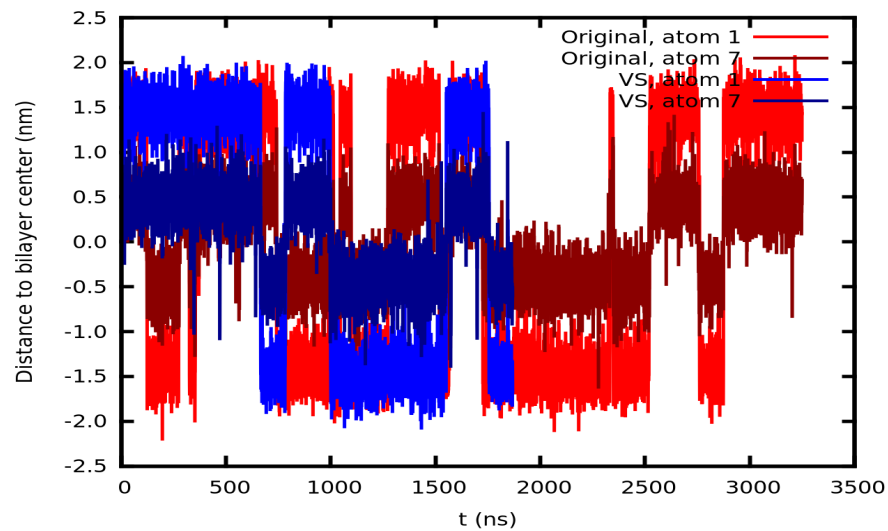of the remaining four atoms
relative to the frame

Defined those four atoms as
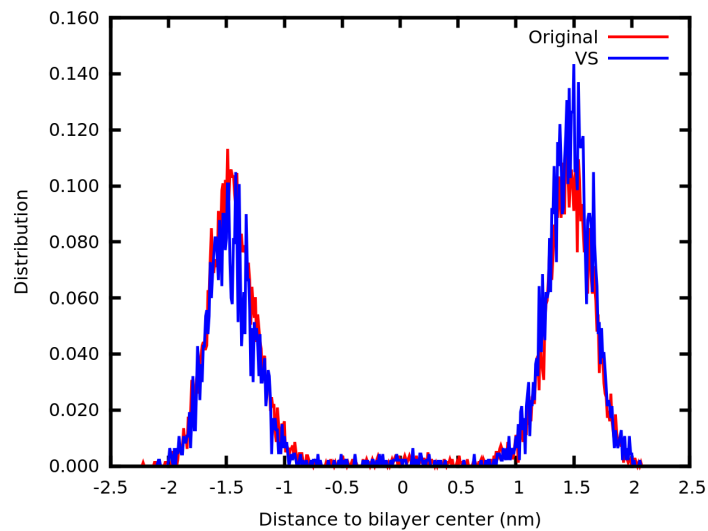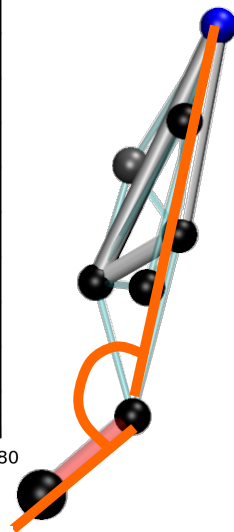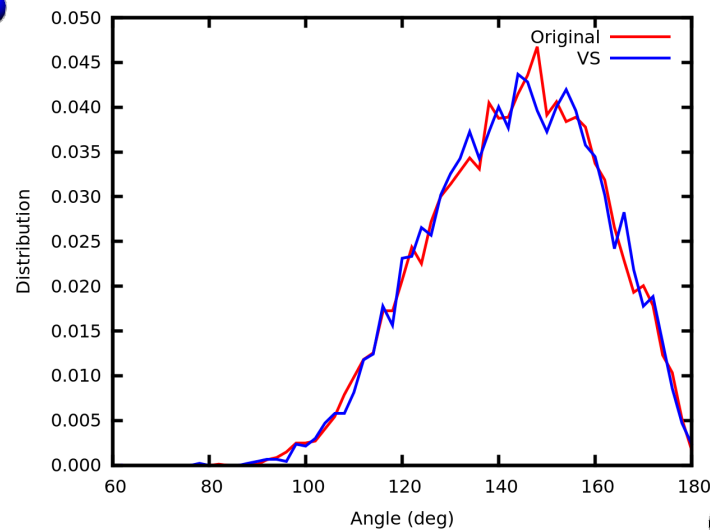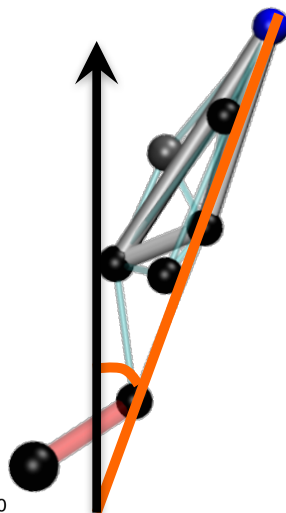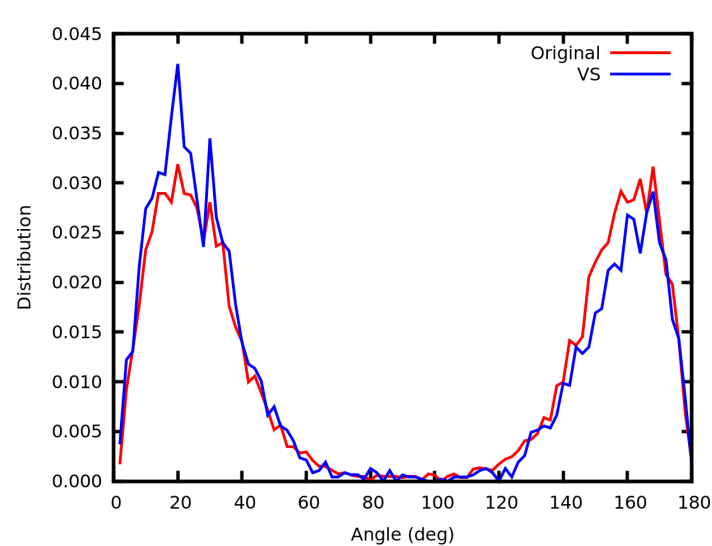different virtual sites

And it worked!
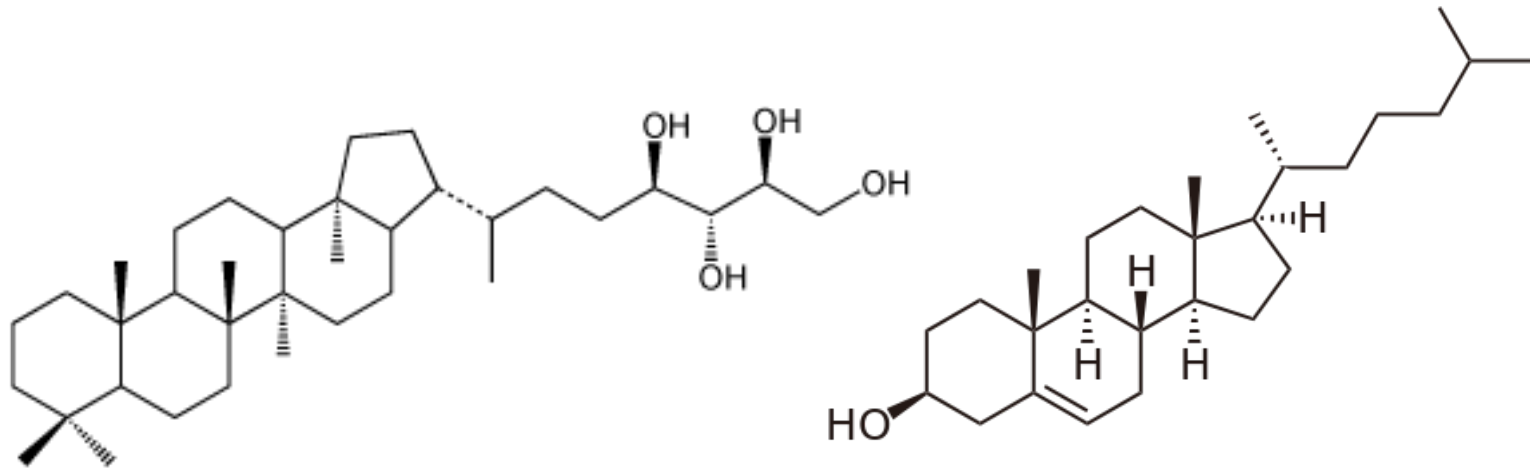
Virtual site version ran stable at 40fs (total 1.6μs)

Comparing to a simulation with the original topology ran at 20fs (total 3.5μs)

But large systems with many cholesterols simulated for long times still crashed...

Hopanoids

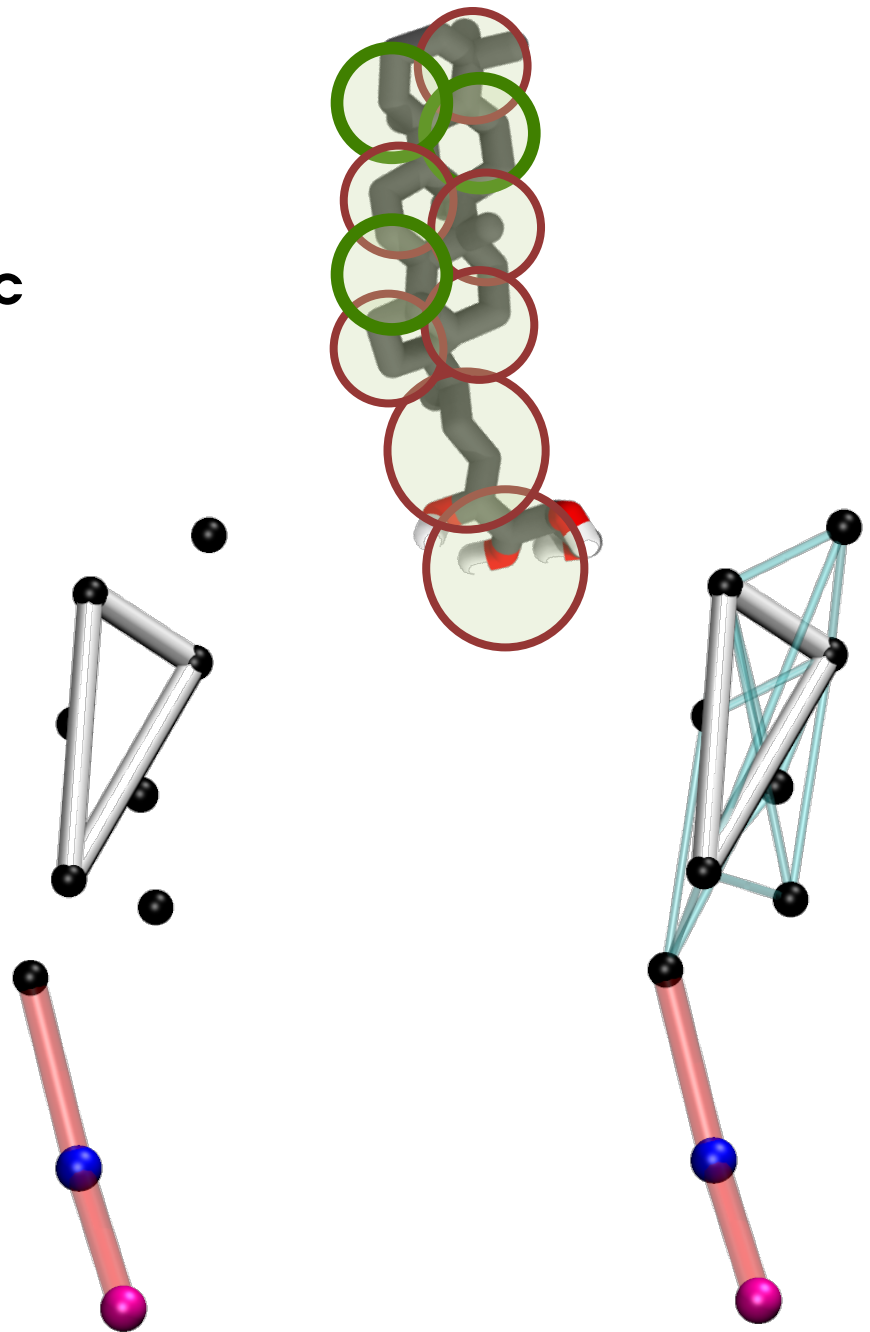bacteriohopanetetrol   cholesterol

Involved in different kinds of bacterial membrane adaptations

CG parameterization

Started from an existing atomistic topology

Chose a mapping scheme and a frame for virtual sites

Constructed virtual sites from the average positions of a mapped atomistic simulation

Stable?

```
step 6937800, will finish Mon Aug 19 15:15:59 2013imb F   2%
step 6937900, will finish Mon Aug 19 15:15:59 2013imb F   5%
step 6938000, will finish Mon Aug 19 15:15:59 2013imb F   4%
step 6938100, will finish Mon Aug 19 15:15:59 2013imb F   5%
step 6938200, will finish Mon Aug 19 15:15:59 2013imb F   4%
step 6938300, will finish Mon Aug 19 15:15:59 2013imb F   2%
step 6938400, will finish Mon Aug 19 15:15:59 2013imb F   1%
step 6938500, will finish Mon Aug 19 15:15:59 2013imb F   5%
step 6938600, will finish Mon Aug 19 15:15:59 2013imb F   2%
step 6938700, will finish Mon Aug 19 15:15:59 2013imb F   5%
step 6938800, will finish Mon Aug 19 15:15:59 2013imb F   2%
step 6938900, will finish Mon Aug 19 15:15:59 2013imb F   5%
step 6939000, will finish Mon Aug 19 15:15:59 2013imb F   3%
step 6939100, will finish Mon Aug 19 15:15:59 2013imb F   2%
step 6939200, will finish Mon Aug 19 15:15:59 2013
A list of missing interactions:
                Bond of      2 missing      -1
                Angle of     2 missing      -2


-------------------------------------------------
Program mdrun, VERSION 4.6.3
Source code file: /manel/gromacs-4.6.3/src/mdlib/domdec_top.c, line: 389

Software inconsistency error:
One or more interactions were multiple assigned in the domain decompostion
For more information and tips for troubleshooting, please check the GROMACS
website at http://www.gromacs.org/Documentation/Errors
-------------------------------------------------


"There's No Room For the Weak" (Joy Division)
```
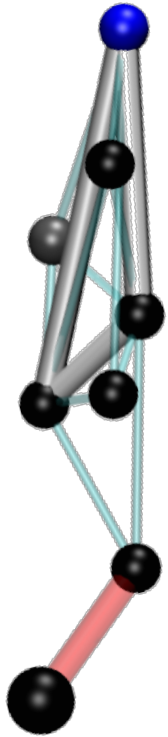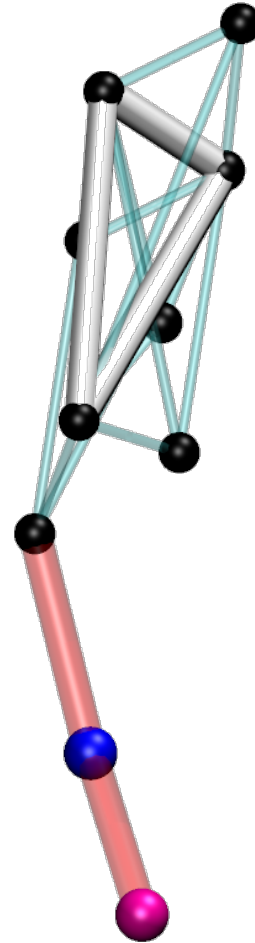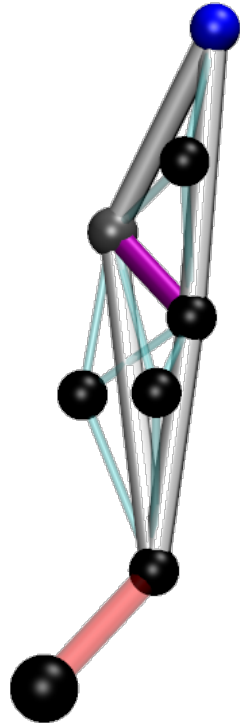
Is the bonded structure too rigid?

cholesterol
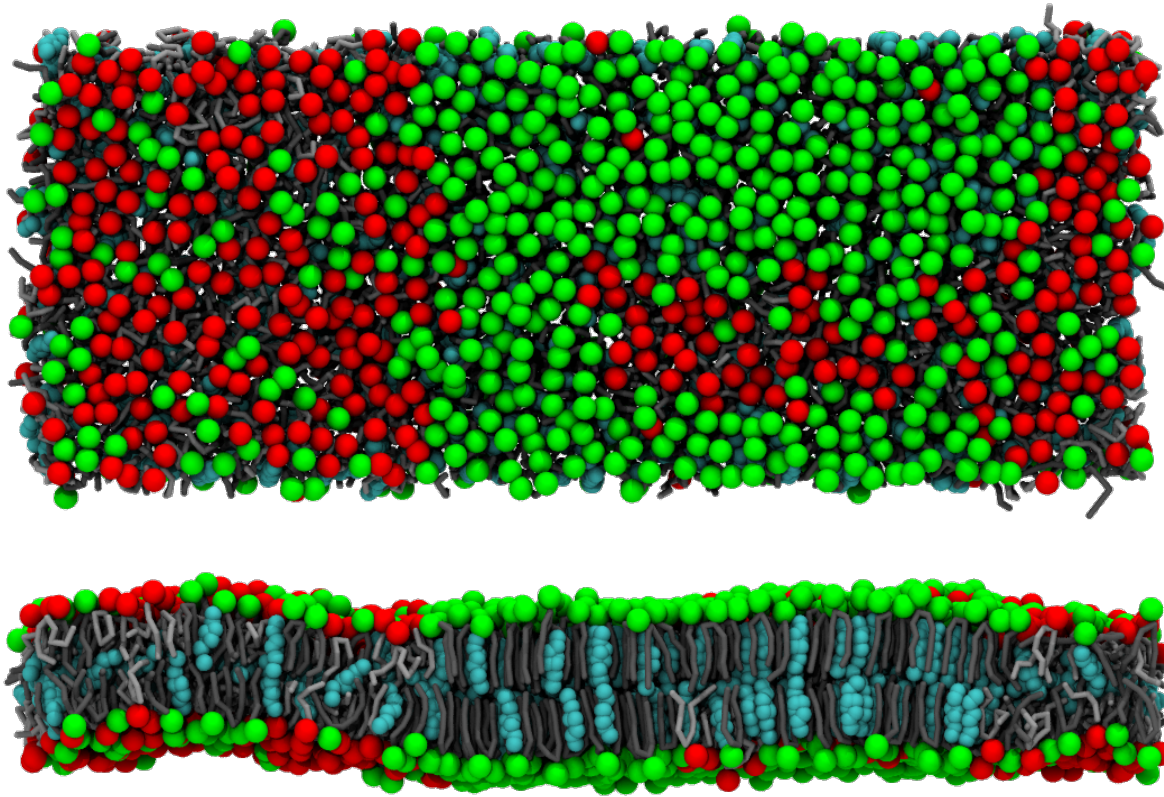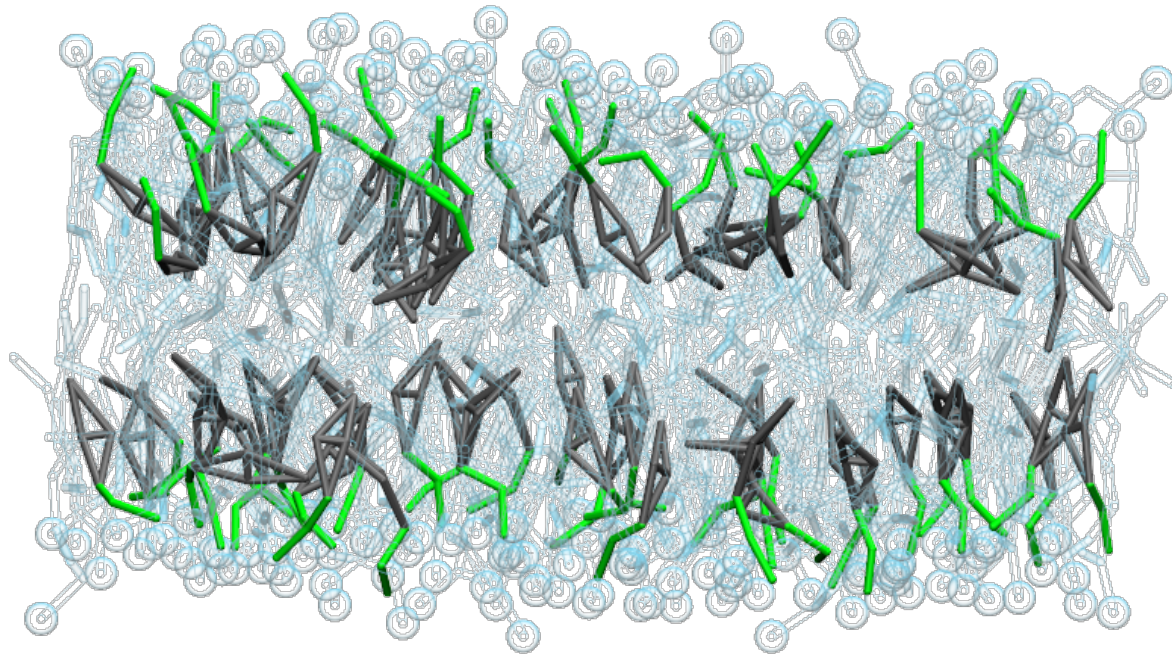
bacteriohopanetetrol

Success!

Cholesterol in bilayer (30μs)

Success!

Bacteriohopanetetrol in POPC bilayer (3μs)

3:1 POPC:STEROL