MD Simulations: Introduction and Ethylene Glycol as Example

- Michael Probst[1], Renat Nazmutdinov[2], Oksana Ismailova[1], Alexander Kaiser[1]

- Institute for Ion Physics and Applied Physics
- University of Innsbruck, Austria

- Inorganic Chemistry
- Kazan State Technological University
- Kazan,Tatarstan, Russian Federation

2012-08-28 Dubna

Why does one need MD – Simulations ?

- The interactions within a chemical system (Atoms, molecules, ions ...) result basically from its electronic structure which governs the interaction between the particles.

- Suppose, the total energy E(x) is known as a function of the coordinates x of all particles:

- How can one calculate thermodynmamic properties from E ?

- Only by numerical simulation.

Why numerical ?

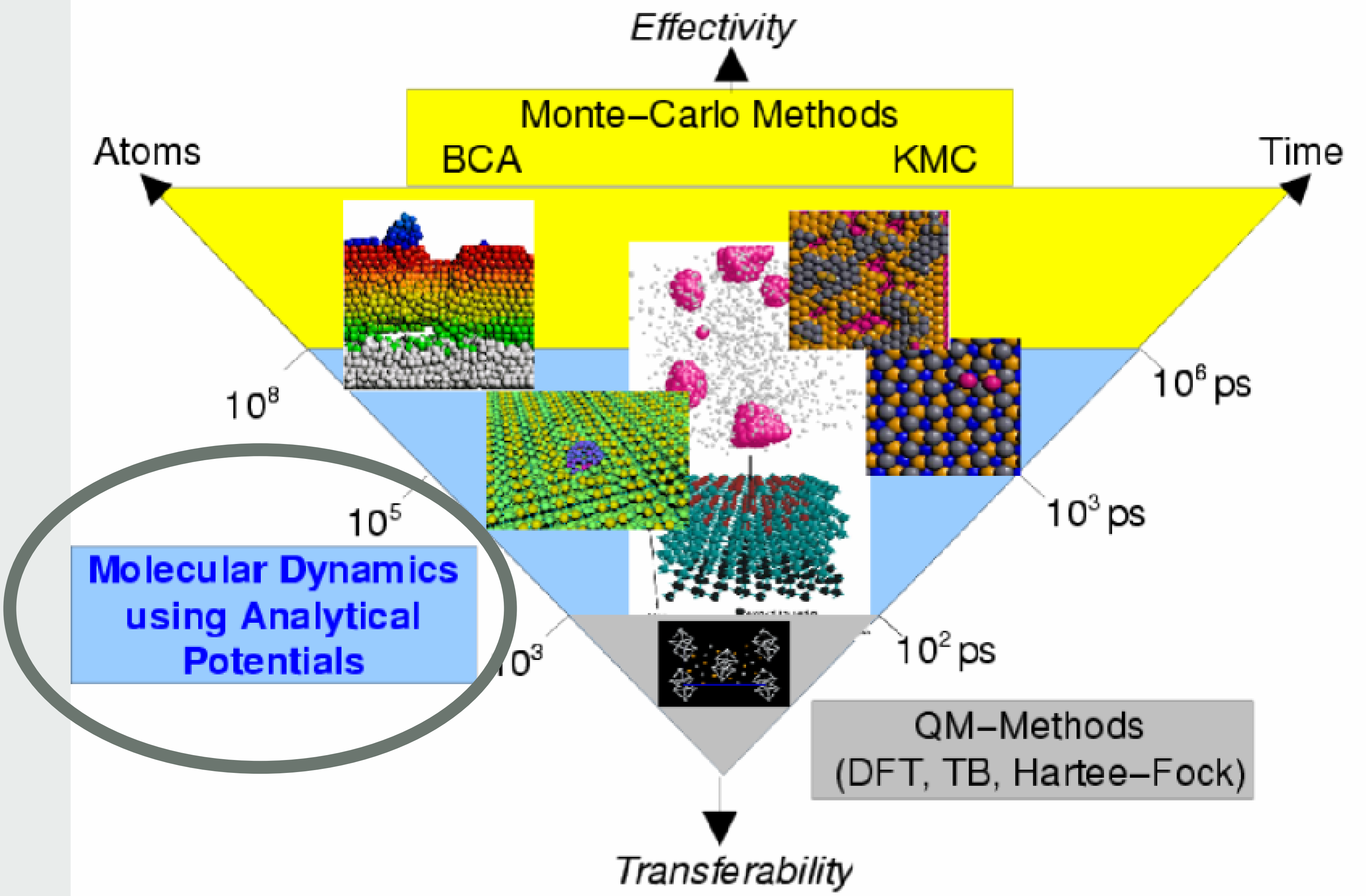
- Reason 1: Integration over phase space:
  - Integral over a high-dimensional weighted function
- Reason 2: Calculation of trajectories.
  - analytically solvable only for $E(x)=k\ x^2$:
    $X(t)=a\ \sin(wt)+b\ \cos(wt)$

Many variations:

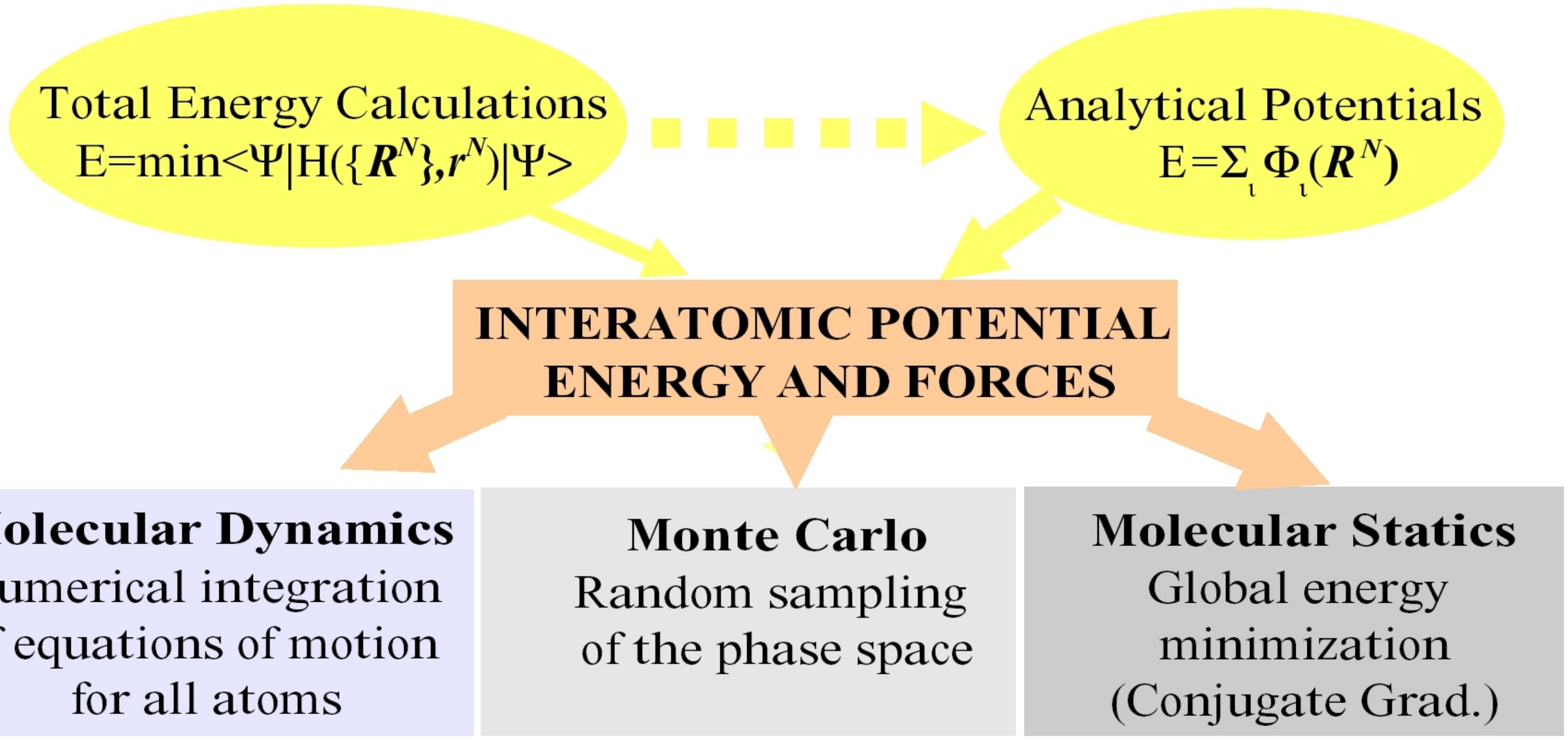
Accurate	Simple
Slow on the computer	Fast
Large Systems	Small Systems

Accurate	Simple
Exact QM	(Semi) Classical Mechanics Approximate QM
Electrons	Interaction potentials electron densities
Atoms	Continuum Point Charges

CC principle 4



Other Atomistic Computer Simulation Methods are very much related to MD



Total Energy Calculations
$E=\min\langle\Psi|H(\{\boldsymbol{R}^N\},\boldsymbol{r}^N)|\Psi\rangle$

Analytical Potentials
$E=\Sigma_i\,\Phi_i(\boldsymbol{R}^N)$

INTERATOMIC POTENTIAL ENERGY AND FORCES

Molecular Dynamics
Numerical integration of equations of motion for all atoms

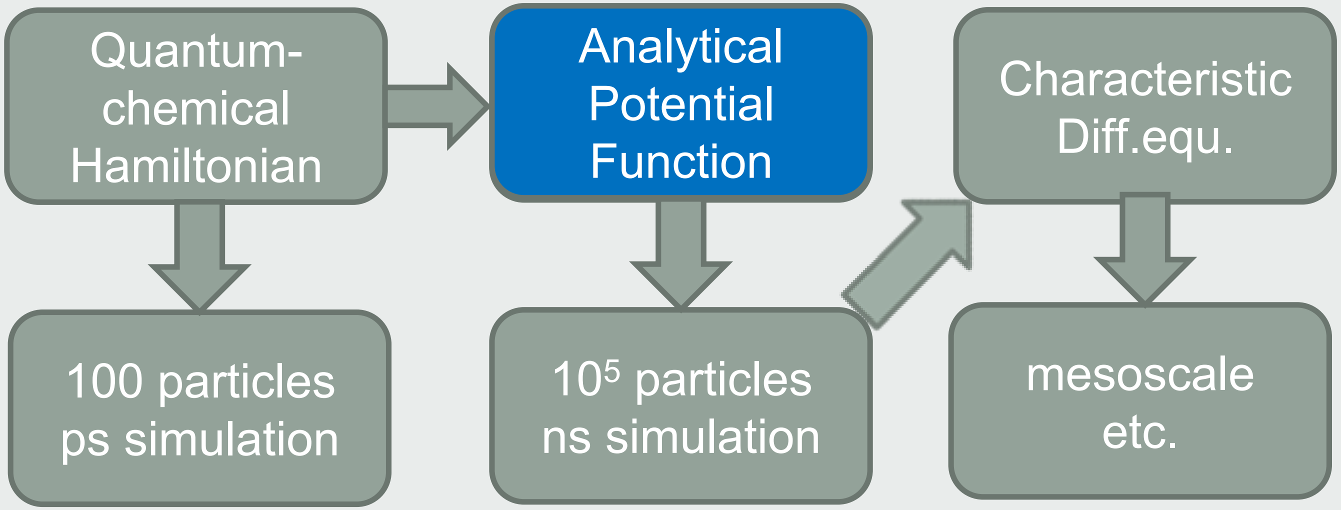
Monte Carlo
Random sampling of the phase space

Molecular Statics
Global energy minimization (Conjugate Grad.)

- Advantage:
  One can calculate large systems (thousands and millions of atoms).

- Disadvantage:
  The potential energy function must be provided. It‘s quality is crucial for the result.

- Bridging the time gap (like the spacial gap before):



Quantum-chemical Hamiltonian

Analytical Potential Function

Characteristic Diff.equ.

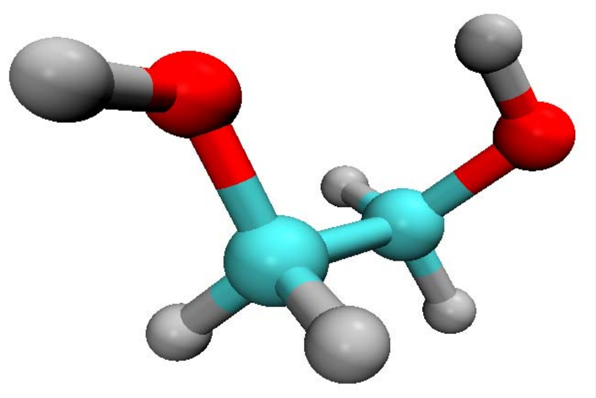
100 particles ps simulation

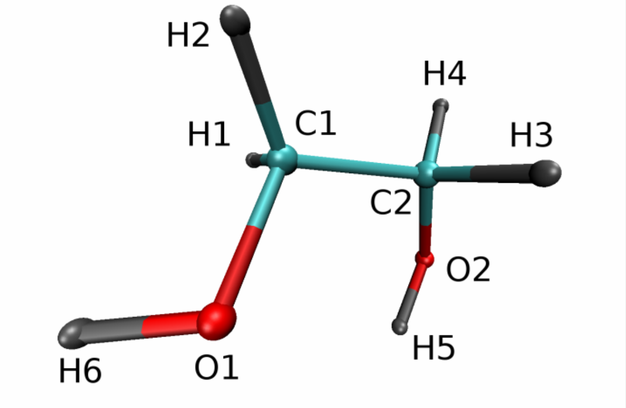
10^5 particles ns simulation

mesoscale etc.

Example:

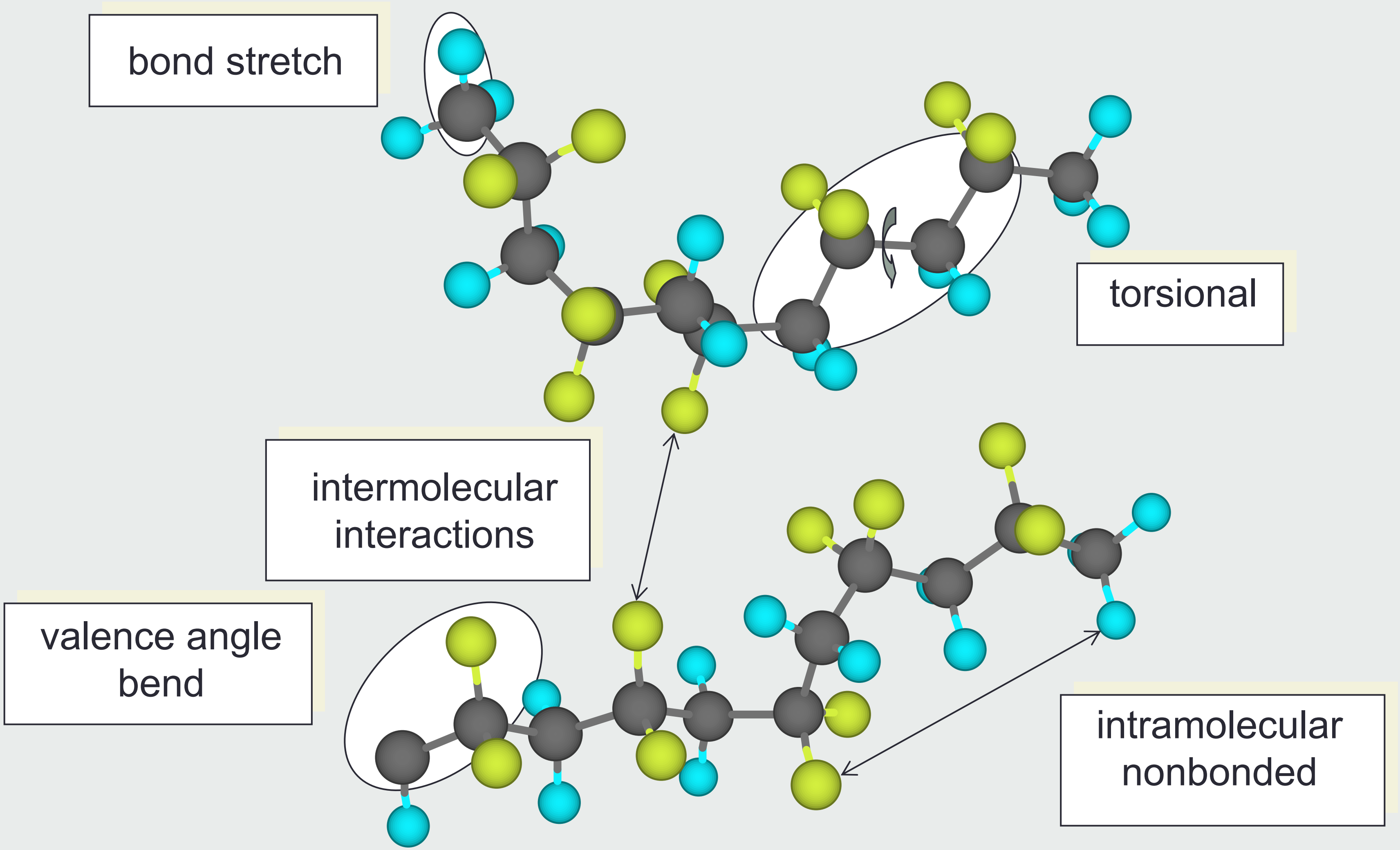
Properties of Ethylene glycol in the condensed phase





- Flexible and symmetric H-O-C-C-O-H chain structure
- Molecular physical properties of ethylene glycol depend on chain angles and aggregation

A) Interaction energy



bond stretch

torsional

intermolecular interactions

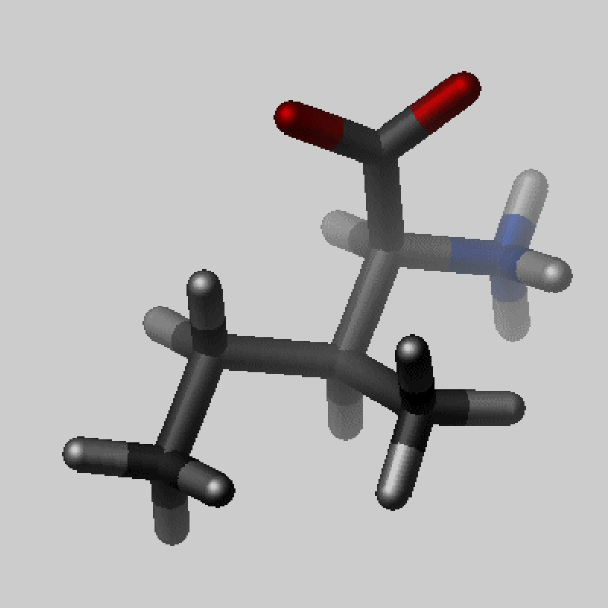
valence angle bend

intramolecular nonbonded

A bit more detail

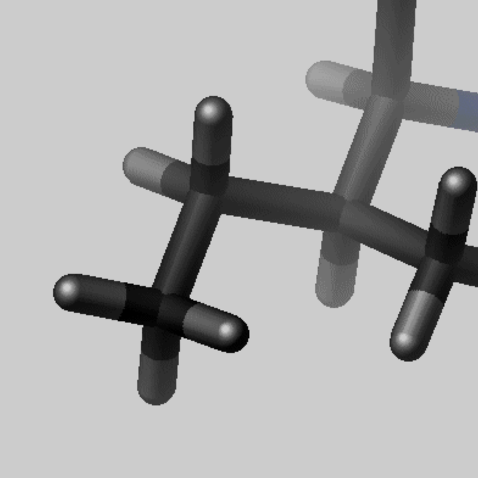
Intramolecular:

- Bonded
- Non-bonded

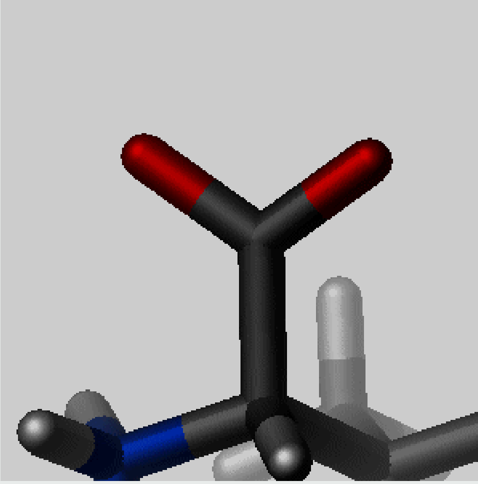


$$E_{bonded} = E_{bond-stretch} + E_{angle-bend} + E_{rotate-along-bond}$$

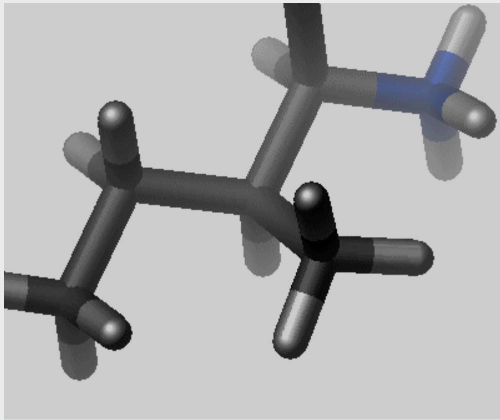
$$E_{bond-stretch} = \sum_{1,2\,pairs} K_b (b - b_0)^2$$



$$E_{bond-bend} = \sum_{angles} K_\theta (\theta - \theta_0)^2$$



$$E_{rotate-bond} = \sum_{1,4\,pairs} K_\phi (1 - \cos(n\phi))$$



Interactions between non-Bonded Atoms

- van der Waals
- Electrostatic

$$\boldsymbol{E_{non-bonded} = E_{van-der-Waals} + E_{electrostatic}}$$

Van der Waals:
A Lennard-Jones (LJ) form is a compromise. Powers can vary.

Electrostatic / Coulomb:

Partial charges on atoms.

$$E_{Lennard-Jones} = \sum_{\substack{nonbonded\\pairs}} \left(\frac{A_{ik}}{r_{ik}^{12}} - \frac{C_{ik}}{r_{ik}^{6}} \right)$$

$$E_{electrostatic} = \sum_{\substack{nonbonded\\pairs}} \frac{q_i q_k}{D r_{ik}}$$

For EG:

- “OPLS force field” parameters from literature:
- LJ-Parameters for pair interactions of different atoms from the Lorenz–Berthelot rules:

$$\sigma_{ij} = (\sigma_i + \sigma_j)/2 \quad \text{and } \varepsilon_{ij} = (\varepsilon_i \varepsilon_j)^{1/2}.$$

- Energies of bonds and bond angles like above (harmonic)
- Torsions:

$$\sum_{torsion} = \sum_i \frac{V_1^i}{2}[1 + \cos\phi_i] + \frac{V_2^i}{2}[1 - \cos 2\phi_i] + \frac{V_3^i}{2}[1 + \cos 3\phi_i]$$

(ϕ_i is the dihedral angle, V1, V2 and V3 are Fourier coefficients)

B) Simulation = Trajectory calculation

- We need the movements (=trajectories x(t)) of the atoms and have to solve (‘integrate’) their equations of motion.
- That means to calculate *x(t)* from

$$\frac{-dE(x(t))}{-dx(t)} = m\frac{d^2x(t)}{dt^2}$$

- Since *E(x)* is complicated, this can only be done numerically which is, however, quite easy.

The equations of motion

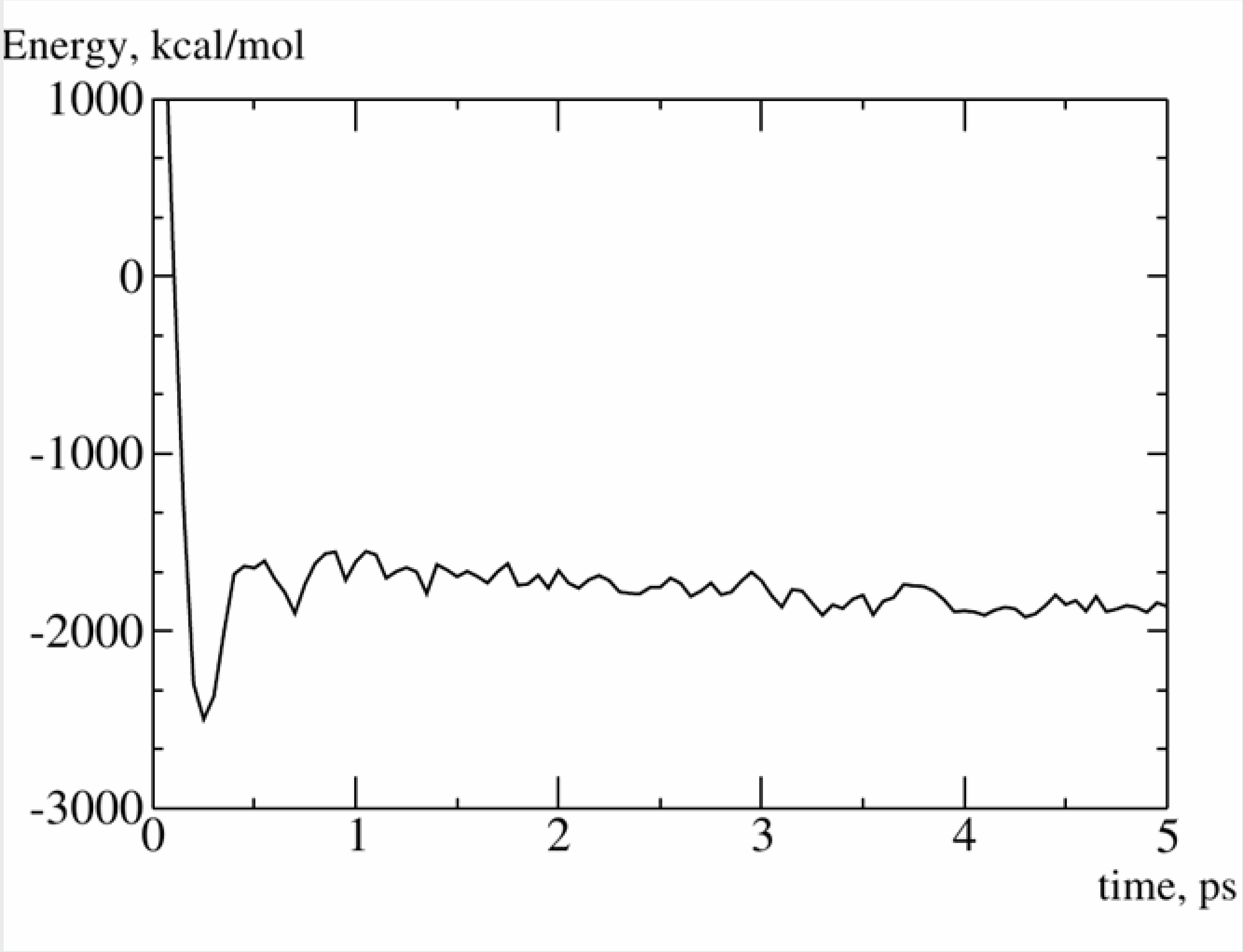
- Many possibilities:
- Verlet: $r(t+\delta t) = 2r(t) - r(t-\delta t) + \delta t2a(t)$
- Leapfrog: $r(t+\delta t) = r(t) + \delta t$
  $v(t+\delta t/2)v(t+\delta t/2) = v(t-\delta t/2) + \delta t\ a(t)$
- Gear …

- **Velocity verlet algorithm:**
  $\boldsymbol{r(t+\delta t) = r(t) + \delta t\ v(t) + \delta t2a(t)}$
  $\boldsymbol{v(t+\delta t) = v(t) + \delta t\ [a(t) + a(t+\delta t)]/2}$

Equilibration :

NVE ensemble at 298 K.

512 EG molecules.



Typical ***protocol:***

- 512 EG molecule in cubic periodic box.
- **1)** Equilibration in NVE ensemble at 298K for 5 ps.
- Time step of 0.5 fs.

- **2)** Equilibration in NVT ensemble at 298K for 5 ps with a weak-coupling Nose-Hoover thermostat(=modified EOMs).
- Simulation time 500 ps.

- **3)** Production run in NVT at 298K for 4 ns.
- Time step of 1 fs.
- Trajectory recording frequency of 2.

C) Analysis

Thermodynamic averages

Analysing the trajectories

‘normal’ averages – just statistics:

$$\langle A \rangle = \lim_{t\to\infty} \frac{1}{t} \int A(t)\,dt = \lim_{m\to\infty} \frac{1}{m} \sum_{i=1}^{m} A_i$$

Average over timesteps (and within a timestep)

Thermodynamic averages

Analysing the trajectories

Example:

Average value of dipole moment					
T (K)	**200**	**250**	**298**	**350**	**400**
Dipole moment (D)	**2.40**	**2.37**	**2.48**	**2.59**	**2.80**
Experiment			**2.25**		

Other example: Diffusion constant

Distribution functions: The radial distribution function

$$g(r) = \frac{1}{N\rho}\left\langle \sum_{t=1}^{N_{TS}} \sum_{j\neq i}^{N_{pair}} \delta(r - r_{ij}) \right\rangle$$

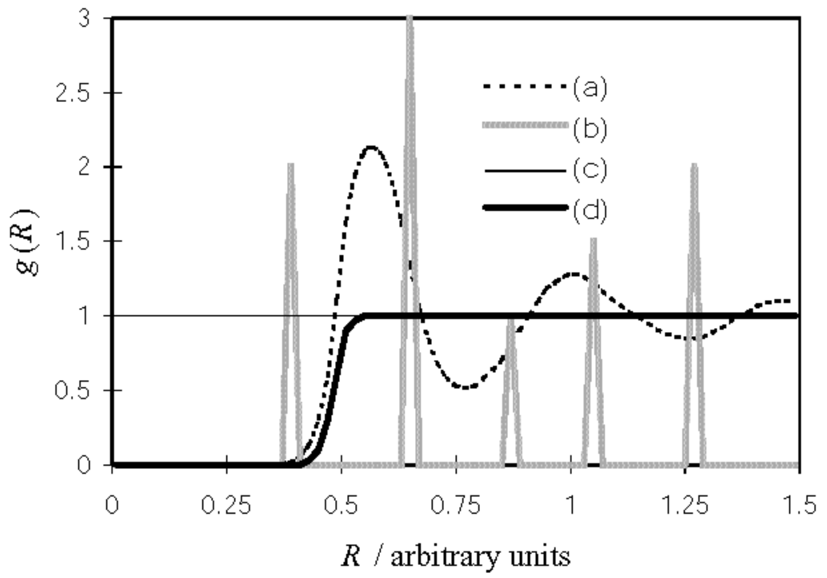
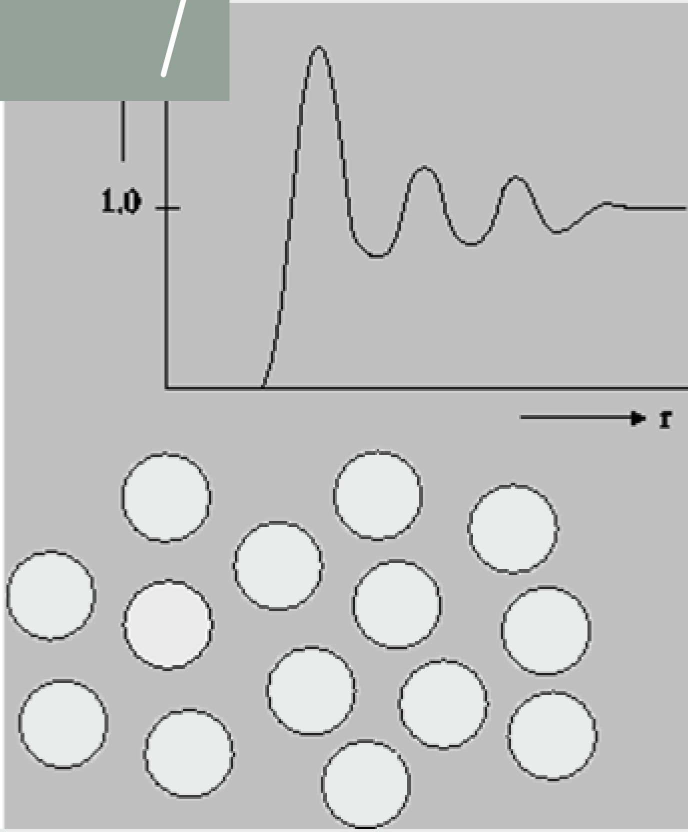
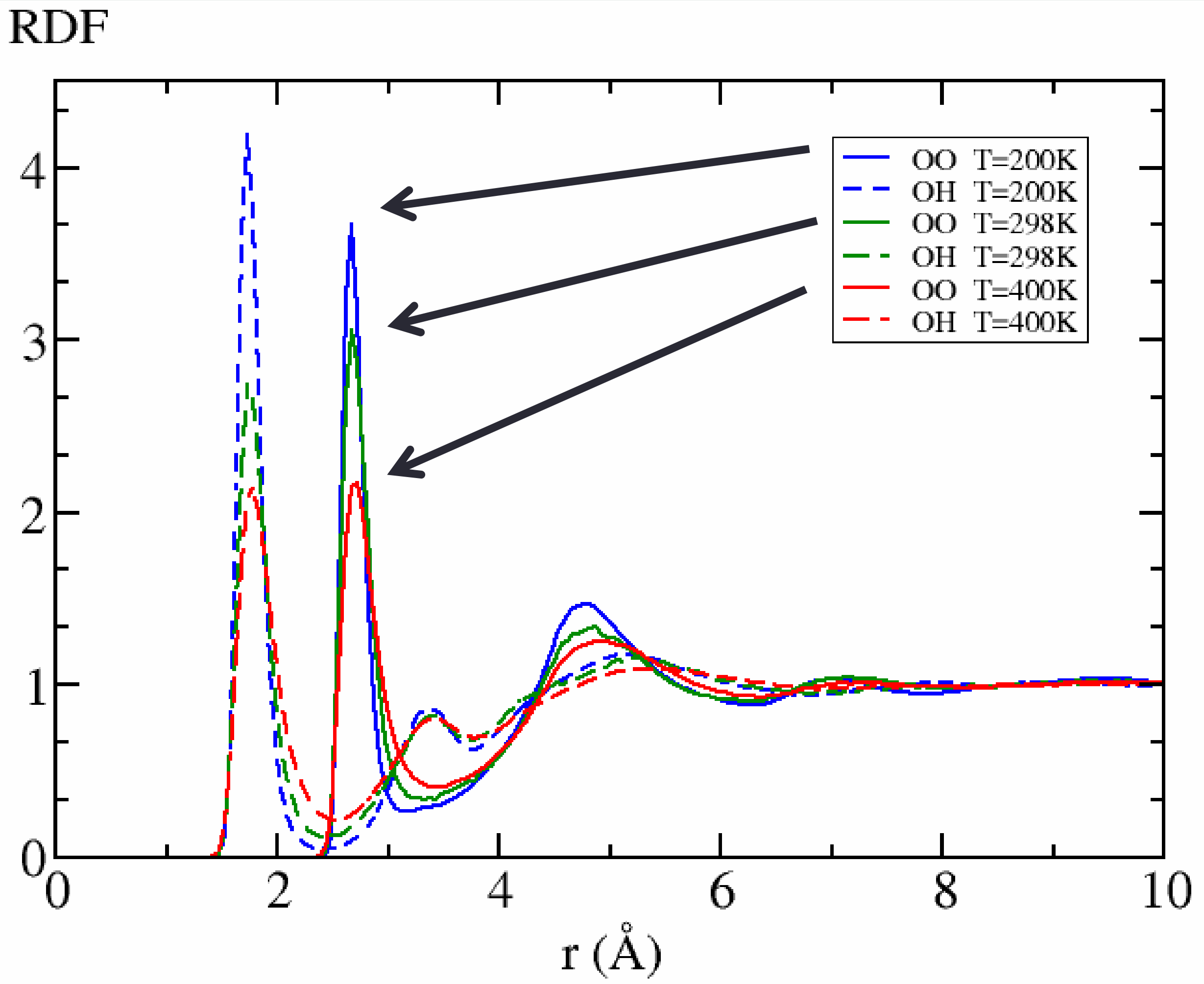


Figure 6. Typical radial distribution functions for (a) a liquid, (b) an ideal solid, (c) an ideal gas, and (d) a real gas.



Distribution functions: The radial distribution function



Work mode – computer related

1)

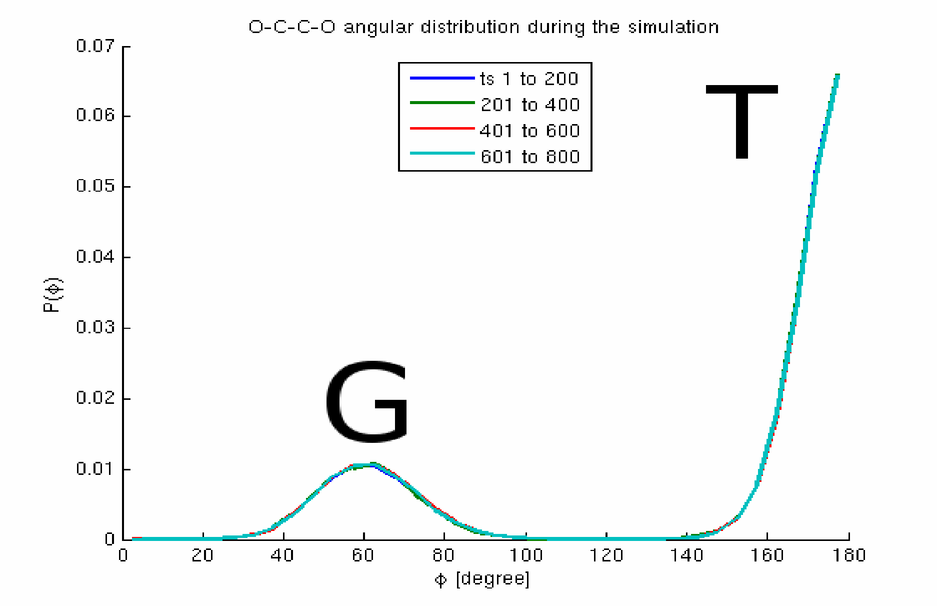
- Perform simulation, create trajectory file
- Calculate properties timestep by timestep

2)

- Perform simulation, create trajectory file
- Copy it all (as some kind of object, possibly in chunks) to memory
- Calculate properties by filtering operations.

Angular distributions are similar:

O-C-C-O

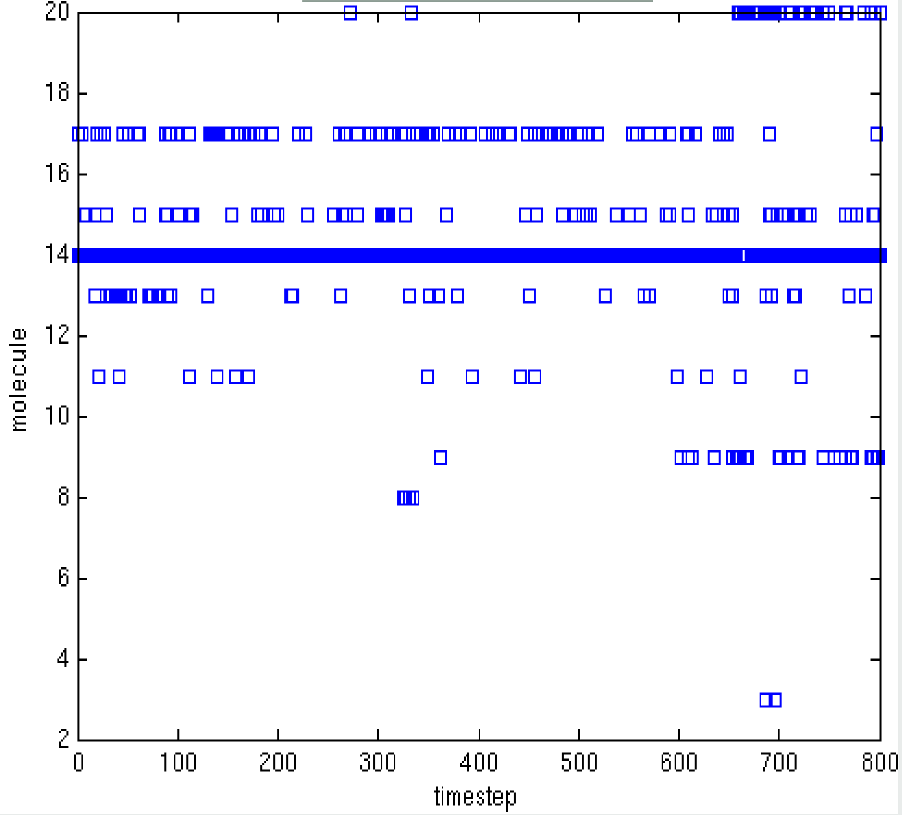


Distribution of the
O-C-C-O dihedral angle.

Dynamical properties:

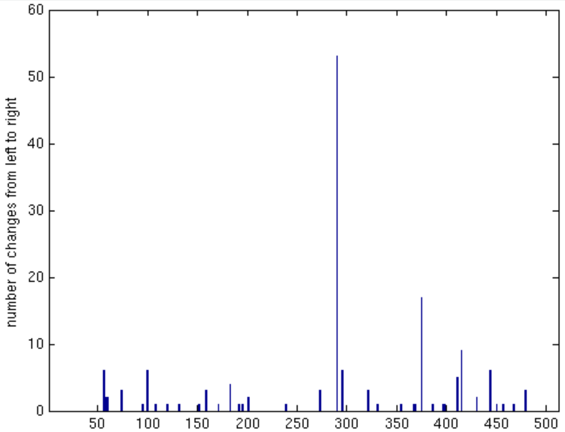
- Properties as a function of time.

H6-O3-C2-C1

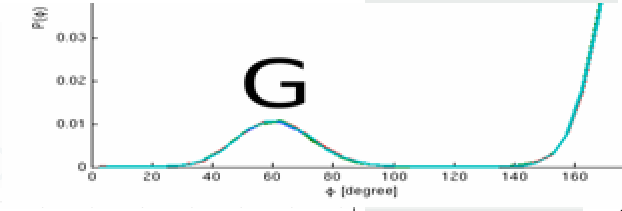


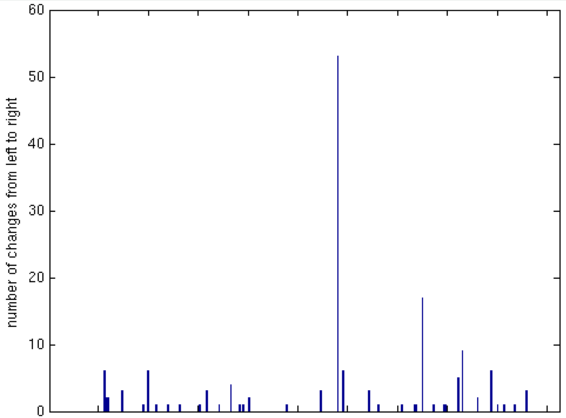
NPT ENSEMBLE AT 200K

O-C-C-O



a) 1.4%

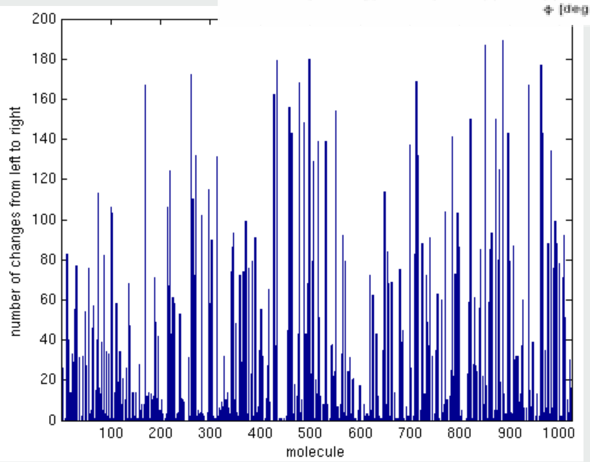


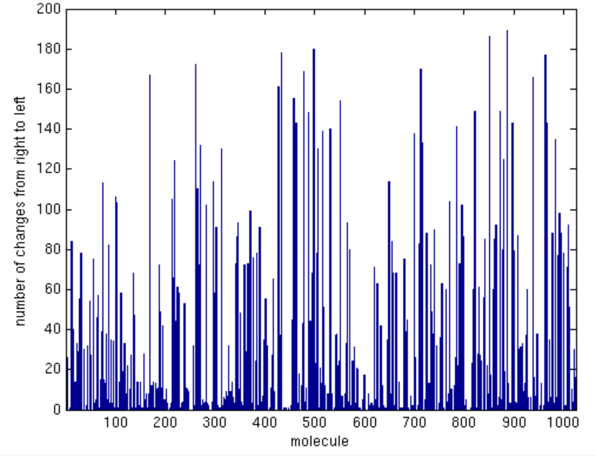


T

b) 0.8%

H-O-C-C



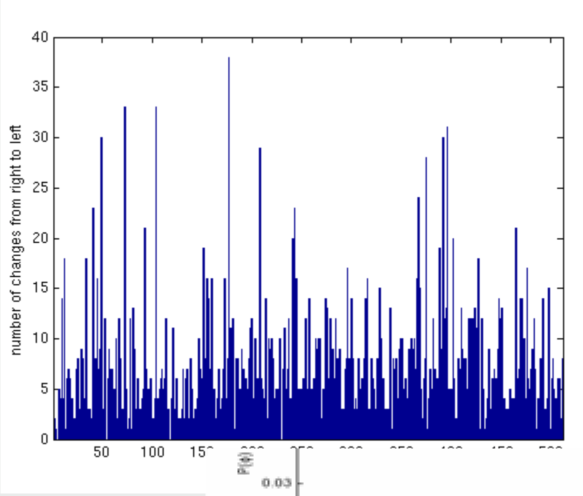


c) 4.3%

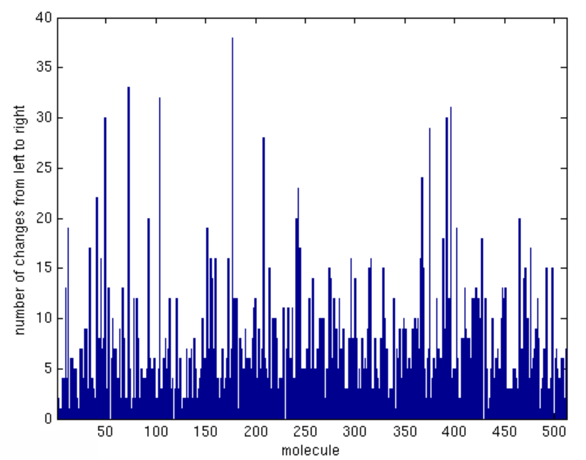
d) 4.5%

NPT ENSEMBLE AT 298K

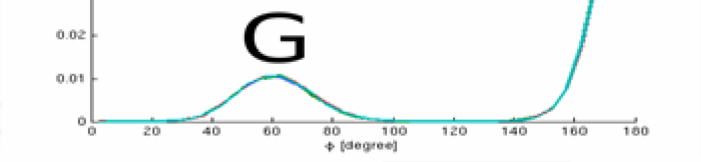
O-C-C-O



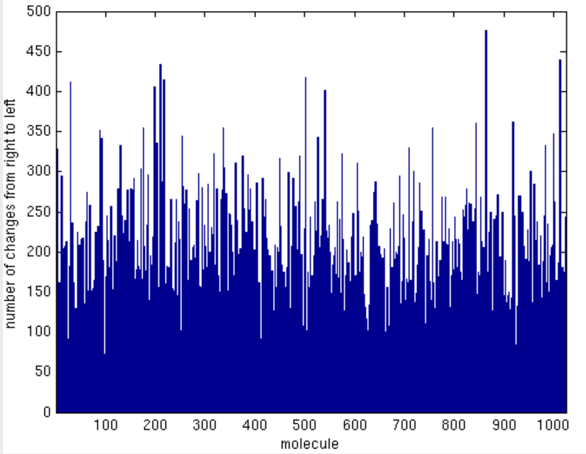
a) 20.3%



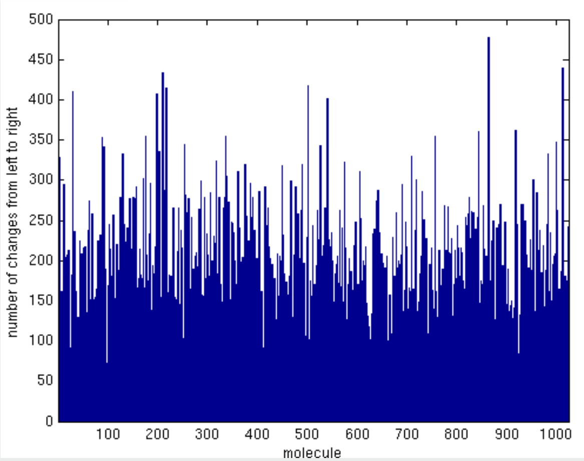
b) 16.8%



H-O-C-C



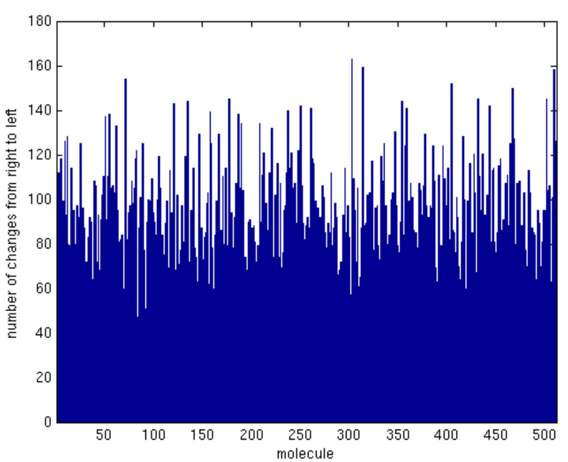
c) 11.7%



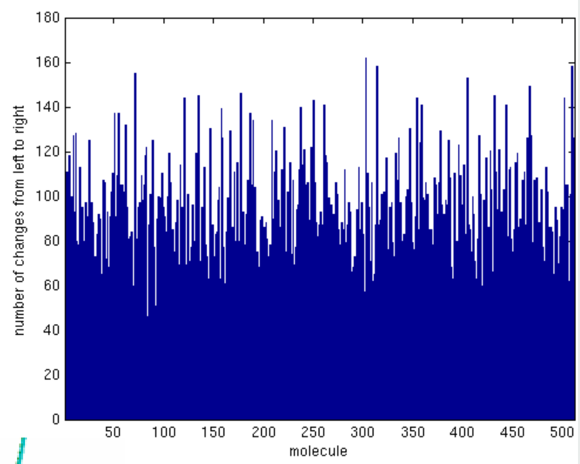
d) 10.1%

NPT ENSEMBLE AT 400K

O-C-C-O

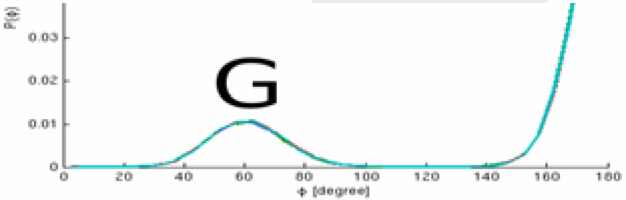


a) 25.0%

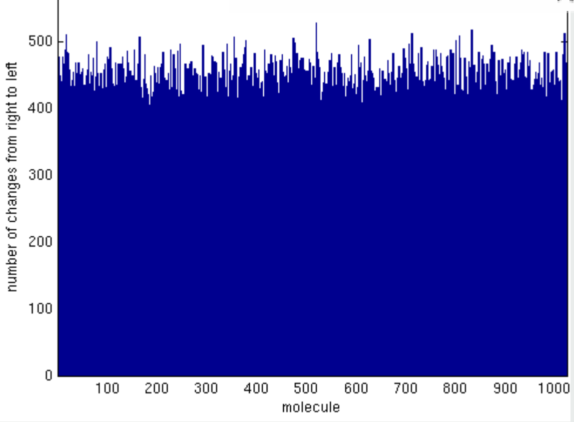


b) 21.7%

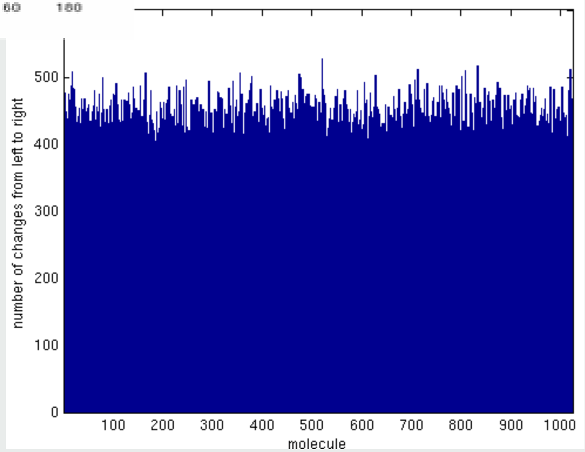
H-O-C-C



T



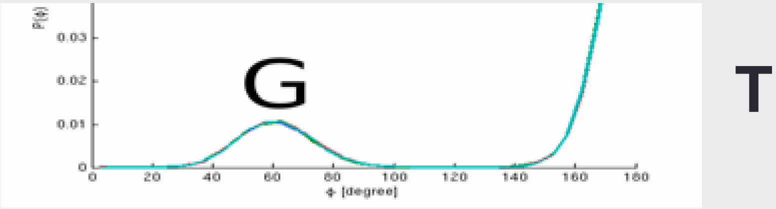
c) 15.9%



d) 17.7%

CC def 3

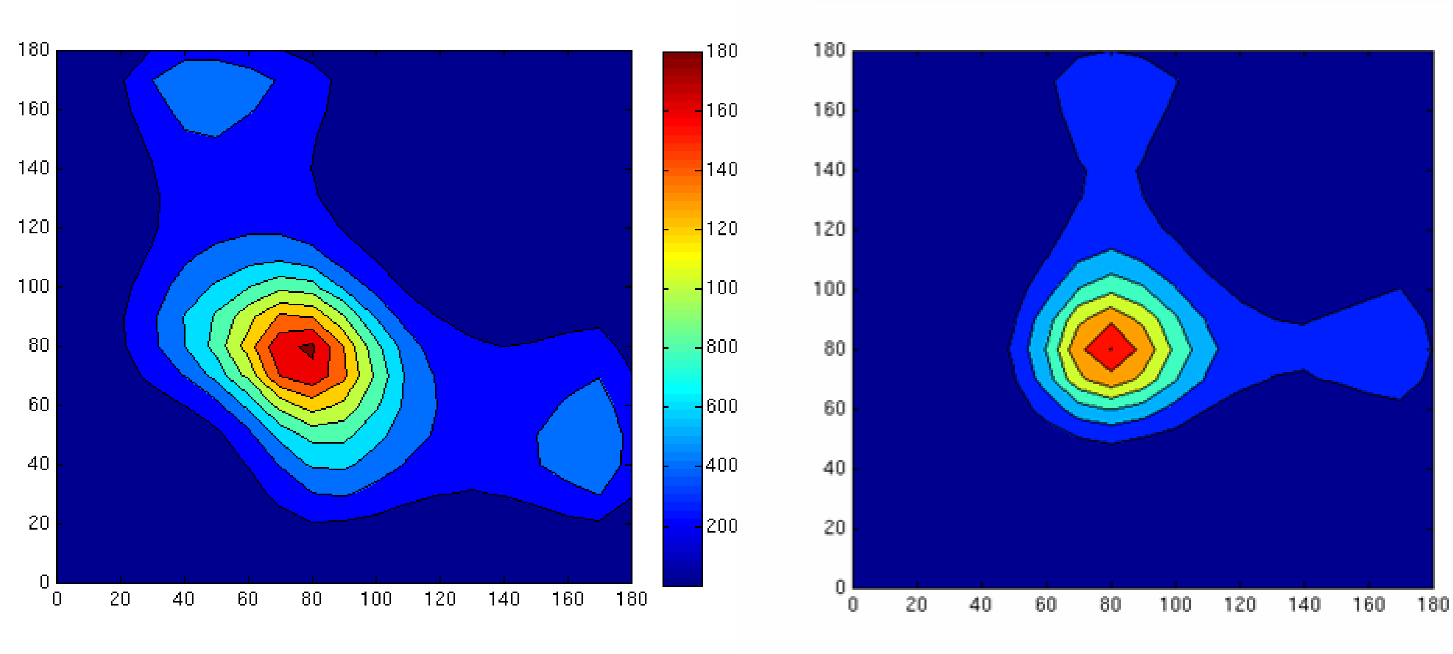
- General idea: Standard procedure:
  - A continuous distribution is discretized into (G and T)

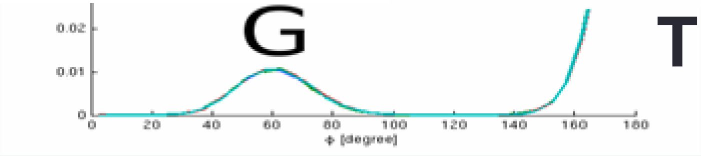


  - This frees up one dimension for analysing / visualizing it (‘Histogram’)
- Can be combined with another tool: nD histograms:

- DISTRIBUTION OF THE (static) CORRELATION BETWEEN THE 2 H-O-C-C DIHEDRAL ANGLES:

NPT ENSEMBLE AT 400K

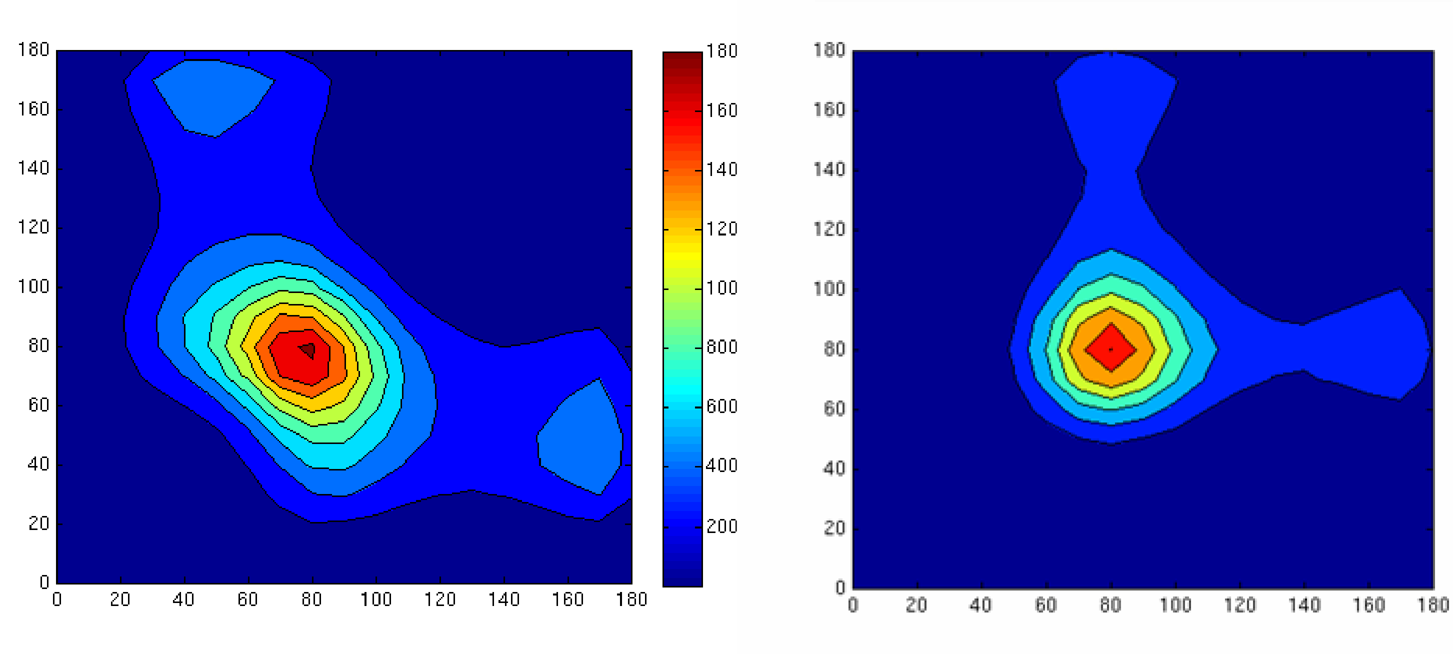


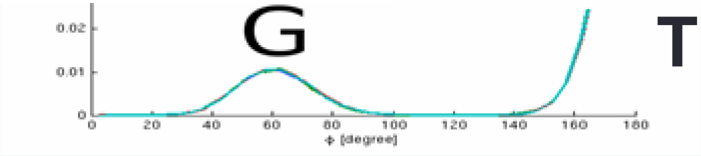


- **For O-C-C-O in G, the probability of finding one H-O-C-C in trans conformation is slightly larger.**
- **Both in trans are unlikely.**

- DISTRIBUTION OF THE (static) CORRELATION BETWEEN THE 2 H-O-C-C DIHEDRAL ANGLES:

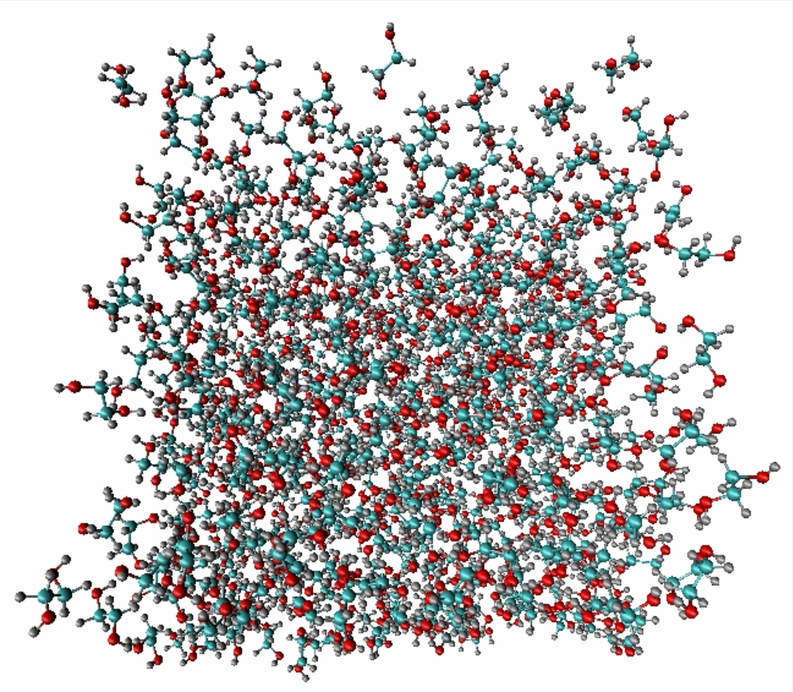
NPT ENSEMBLE AT 400K





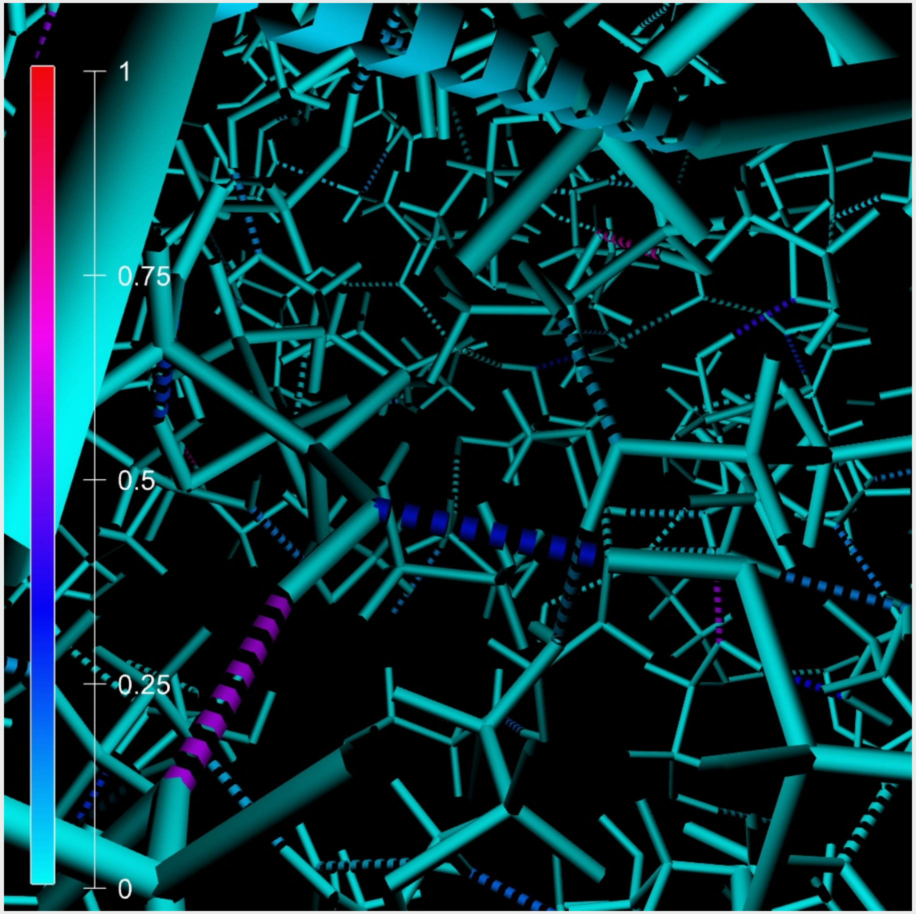
- **For O-C-C-O in G, the probability of finding one H-O-C-C in trans conformation is slightly larger.**
- **Both in trans are unlikely.**

- D) Visualisation of MD results
  - Serious science: Extracting information of high-dimensional data
  - One good program: VMD



In our EG – work:

- Extremely difficult problem:
  Undrstanding the H-bond network over space and time.
- Our own program:
  VISH (W. Benger et al.)



D) The interesting issues in electrochemistry

- Solvated molecules that can have different oxidation states
- $S_2O_8^{2-} + e \quad S_2O_8^{3-} \quad SO_4^{2-} + SO_4^{-}$
- Interesting are the reorganisation energies:

$$\lambda_s = E_{noneq}(S_2O_8^{2-}) - E_{eq}(S_2O_8^{2-})$$

$$\lambda_s = E_{noneq}(S_2O_8^{3-}) - E_{eq}(S_2O_8^{3-})$$

Oxidation / redaction reactions

- From the RDFs of S_2O_8 in EG / H_2O …

	water		EG	
	$S_2O_8^{2-}$	$S_2O_8^{3-}$	$S_2O_8^{2-}$	$S_2O_8^{3-}$
Coordination number	25	26	13	12
Distance of first minimum of RDF (nm)	0.58	0.56	0.673	0.638

- … one can calculate the **solvent reorganisation energies**:

x(EG)	$_s$ (red)	$_s$ (ox)	$_s$ (avg)
	kcal mol^{-1}		
0	22.4	54.5	33.2
1.0	38.3	32.2	34.9

- The increase of the average $_s$ for EG means that a dynamic solvent effect modulates the ET.

Thank you …