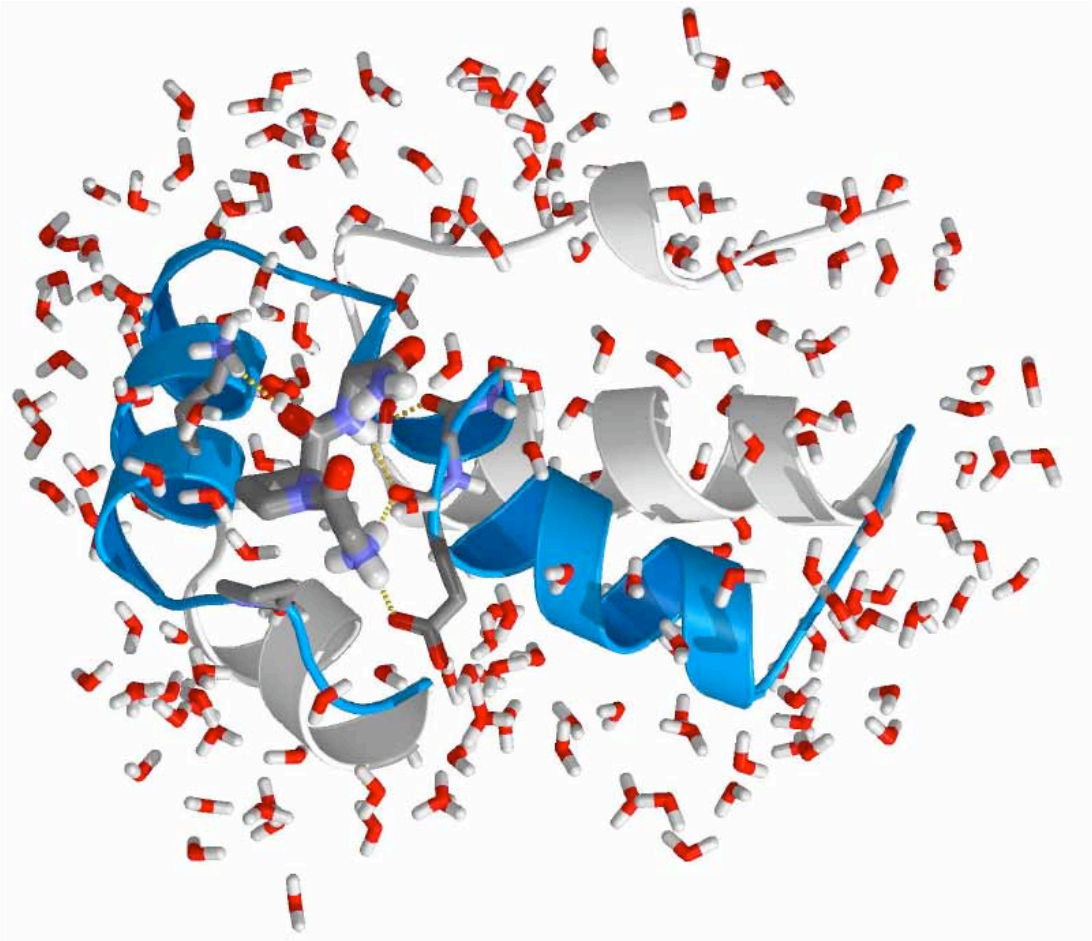
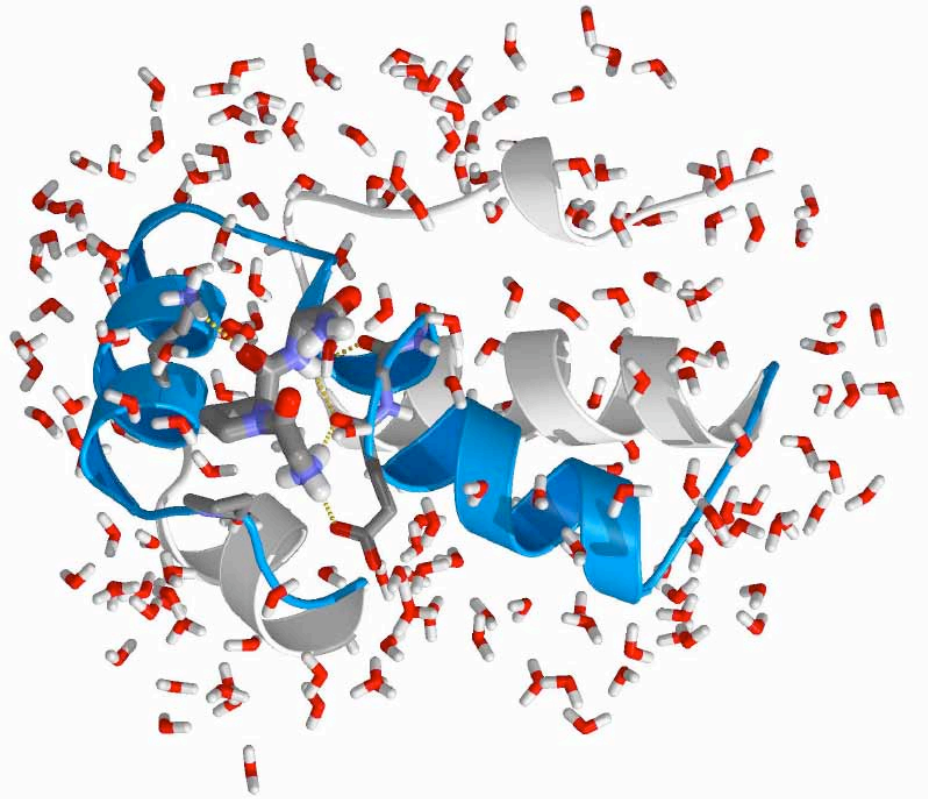
Molecular Dynamics Simulation

A Short Introduction



Michel Cuendet

Plan



- **Introduction**
- **The classical force field**
- **Setting up a simulation**
- **Connection to statistical mechanics**
- **Usage of MD simulation**

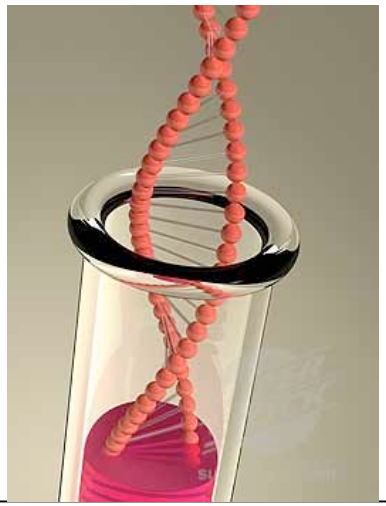
Why we do simulation

In some cases, experiment is :

1.	impossible	*Inside of stars* *Weather forecast*
2.	too dangerous	*Flight simulation* *Explosion simulation*
3.	expensive	*High pressure simulation* *Windchannel simulation*
4.	blind	*Some properties cannot be observed on very short time-scales and very small space-scales*

Simulation is a useful complement, because it can :

- ⇨ replace experiment
- ⇨ provoke experiment
- ⇨ explain experiment
- ⇨ aid in establishing intellectual property



Molecular modeling

What is Molecular Modeling?

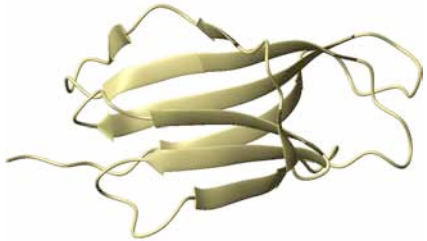
Molecular Modeling is concerned with the description of the atomic and molecular interactions that govern *microscopic* and *macroscopic* behaviors of physical systems

What is it good for?

The essence of molecular modeling resides in the connection between the *macroscopic* world and the *microscopic* world provided by the theory of statistical mechanics

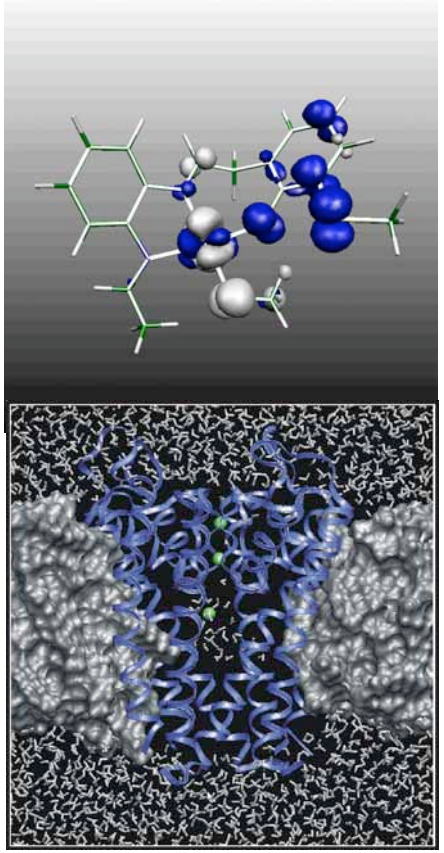
Macroscopic observable (Solvation energy, affinity between two proteins, H-H distance, conformation, ...)

Average of observable over selected microscopic states



Computational tools

- **Quantum Mechanics (QM)**
  Electronic structure (Schrödinger)
  - ACCURATE
  - EXPENSIVE ⇨ small system

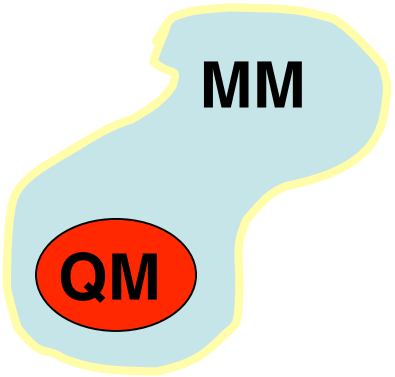


10-100 atoms
10-100 ps

- **Classical Molecular Mechanics (MM)**
  Empirical forces (Newton)
  - LESS ACCURATE
  - FAST

10^4-10^5 atoms
10-100 ns

- **Mixed Quantum/Classical (QM/MM)**



10^4-10^5 atoms
10-100 ps

Types of phenomena

Goal : simulate/predict ***processes*** such as

1. polypeptide folding
2. biomolecular association
3. partitioning between solvents
4. membrane/micelle formation

} *thermodynamic equilibria governed by weak (non-bonded) forces*

5. chemical reactions, enzyme catalysis
6. enzyme catalysis
7. photochemical ractions, electron transfer

} *chemical transformations governed by strong forces*

characteristics *(1-4)*:

- degrees of freedom: atomic (solute + solvent) → classical MD
- equations of motion: classical dynamics
- governing theory: statistical mechanics

characteristics *(5-7)*:

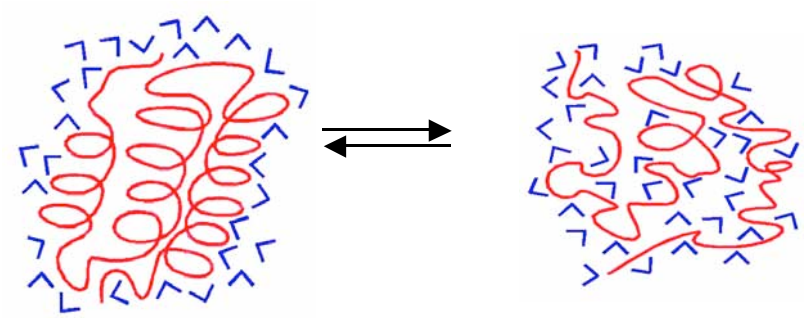
→ quantum MD

- degrees of freedom: electronic, nuclear
- equations of motion: quantum dynamics
- governing theory: quantum statistical mechanics

Processes: Thermodynamic Equilibria

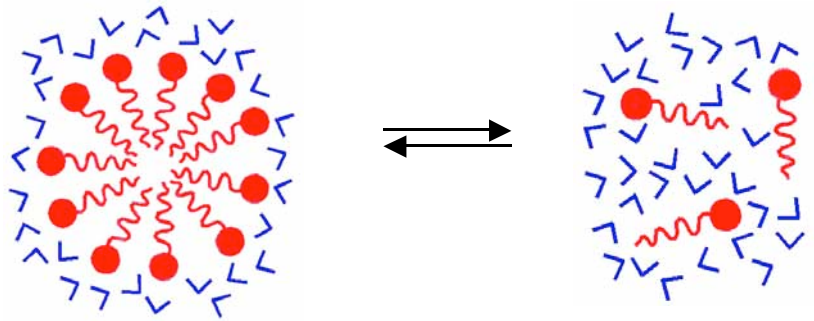
Folding

folded/native ⇌ denatured



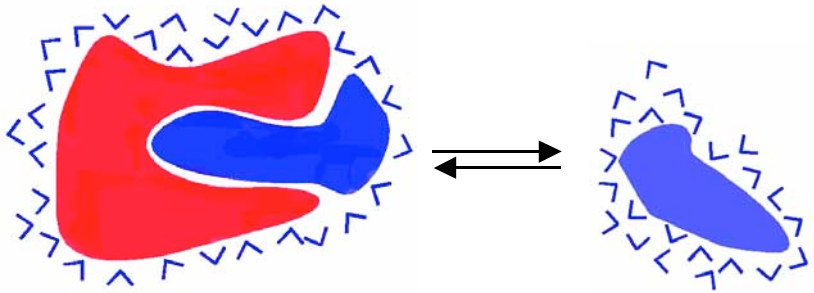
Micelle Formation

micelle ⇌ mixture



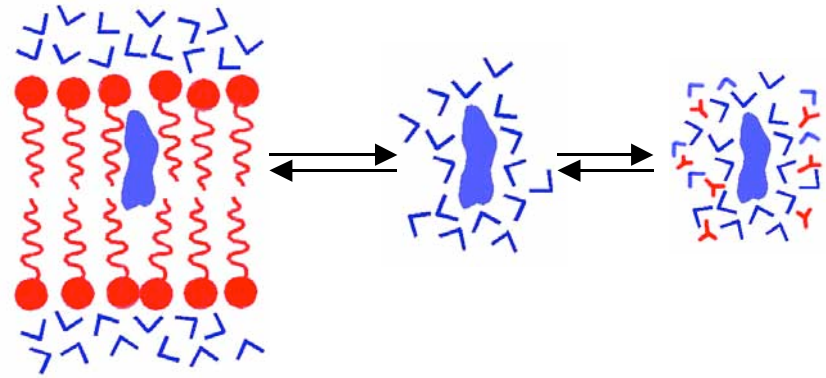
Complexation

bound ⇌ unbound

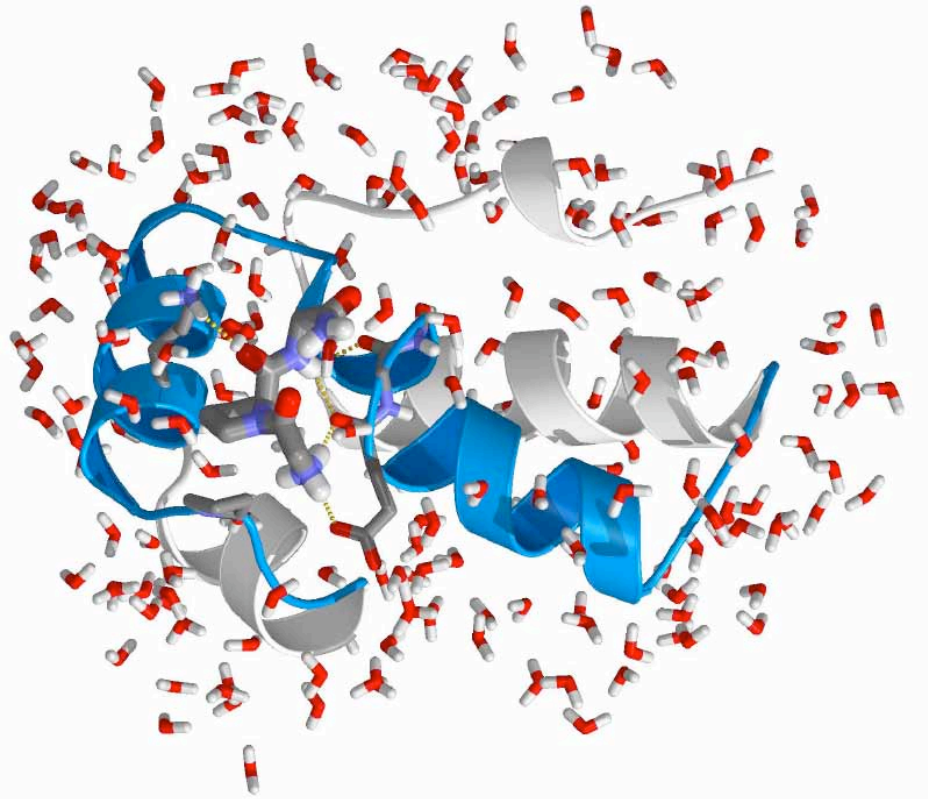


Partitioning

in membrane ⇌ in water ⇌ in mixtures



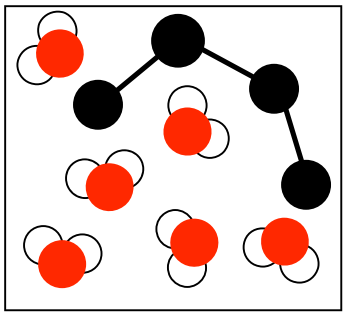
Plan



- Introduction
- **The classical force field**
- Setting up a simulation
- Connection to statistical mechanics
- Usage of MD simulation

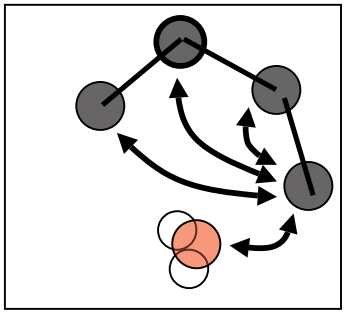
Definition of a model for molecular simulation

Every molecule consists of atoms that are very strongly bound to each other



Degrees of freedom: atoms as elementary particles + topology

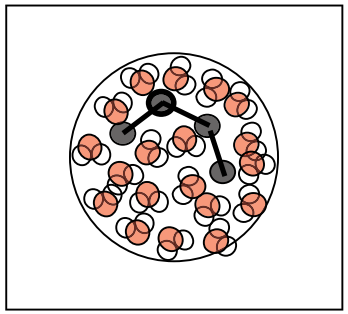
Forces or interactions between atoms



Force field = physico-chemical knowledge

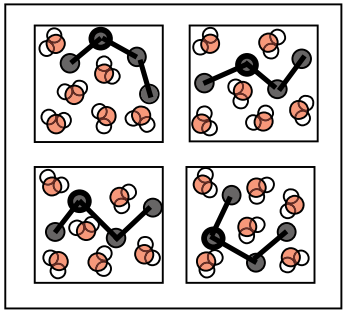
MOLECULAR MODEL

Boundary conditions



system
temperature
pressure
walls
external forces

Methods to generate configurations of atoms: Newton



Classical force fields

Goals of classical (semi-empirical) force fields

- Definition of empirical potential energy functions $V(\boldsymbol{r})$ to model the molecular interactions
- These functions need to be differentiable in order to compute the forces acting on each atom: $\boldsymbol{F} = -\boldsymbol{\nabla} V(\boldsymbol{r})$

Implementation of calssical potential energy functions

1. Theoretical functional forms are derived for the potential energy $V(\boldsymbol{r})$.

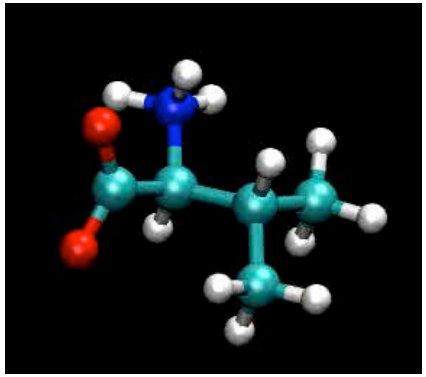
2. Definition of atom types that differ by their atomic number and chemical environment, e.g. the carbons in C=O or C-C are of different types.

3. Parameters are determined so as to reproduce the interactions between the various atom types by fitting procedures
   - experimental enthalpies (CHARMM)
   - experimental free energies (GROMOS, AMBER)

Parametrization available for proteins, lipids, sugars, ADN, ...

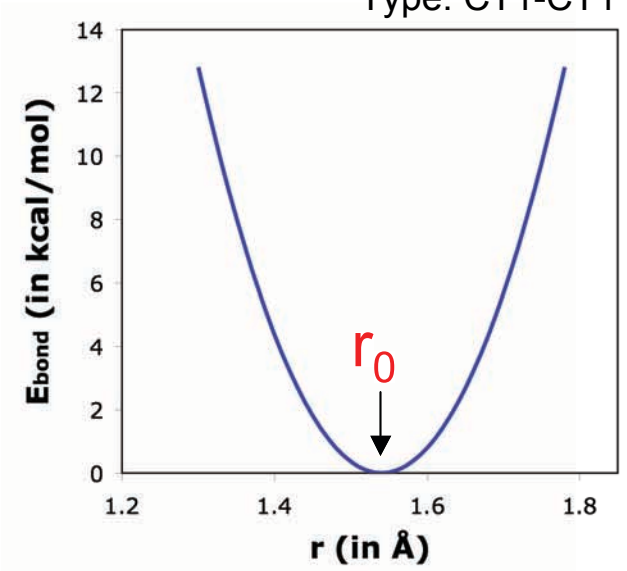
Covalent bonds and angles

Bonds

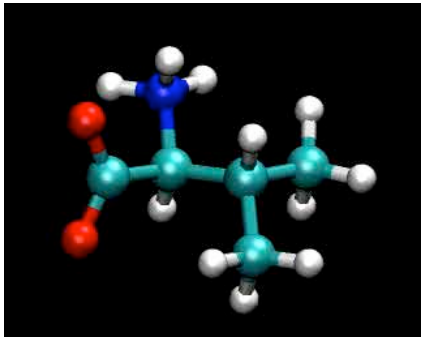


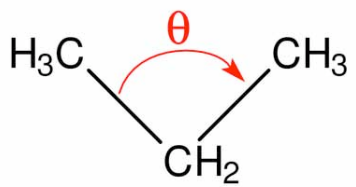
$$E_{bond} = K_b(r - r_0)^2$$

Type: CT1-CT1



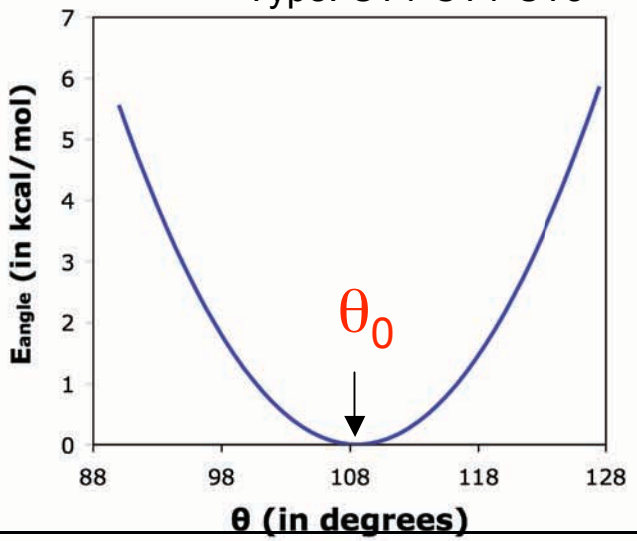
Angles





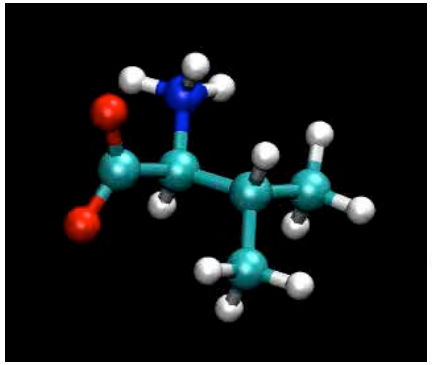
$$E_{angle} = K_\theta(\theta - \theta_0)^2$$

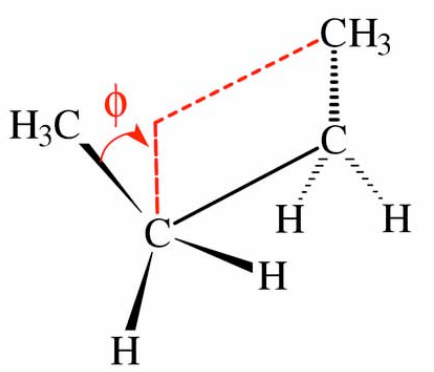
Type: CT1-CT1-CT3



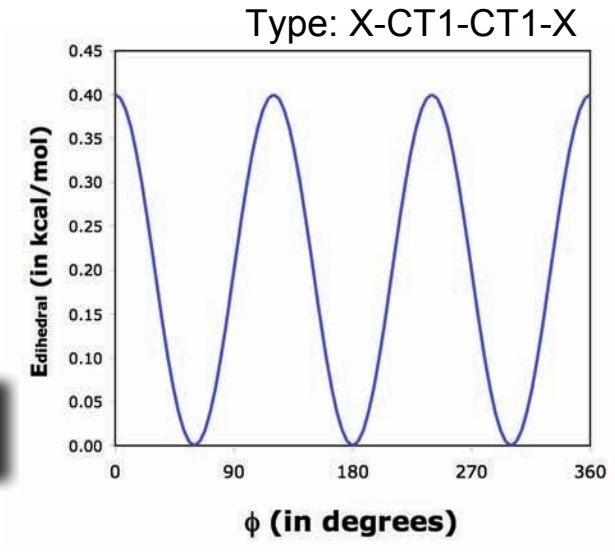
Dihedrals and Improper torsions

Dihedral angles

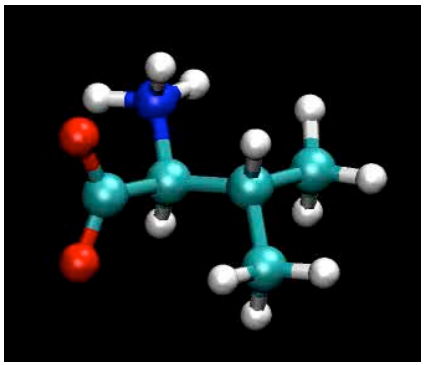


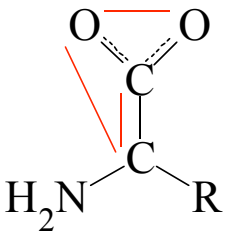


$$E_{dihedral} = K_{\phi}\left[1 + \cos\left(n\phi - \delta\right)\right]$$

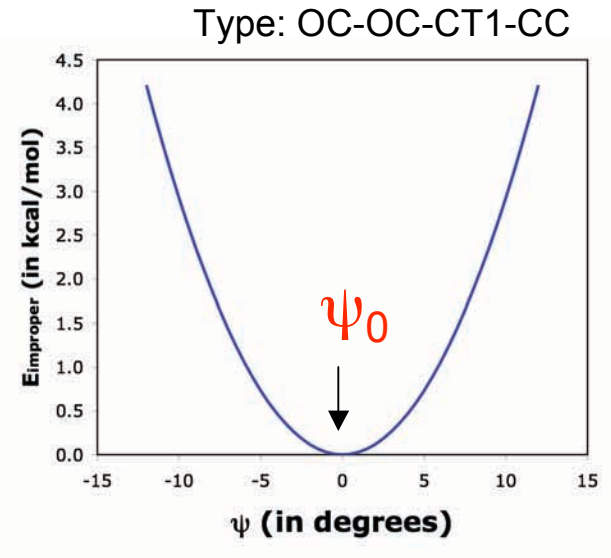


Improper angles





$$E_{improper} = K_{\phi}\left(\psi - \psi_0\right)^2$$



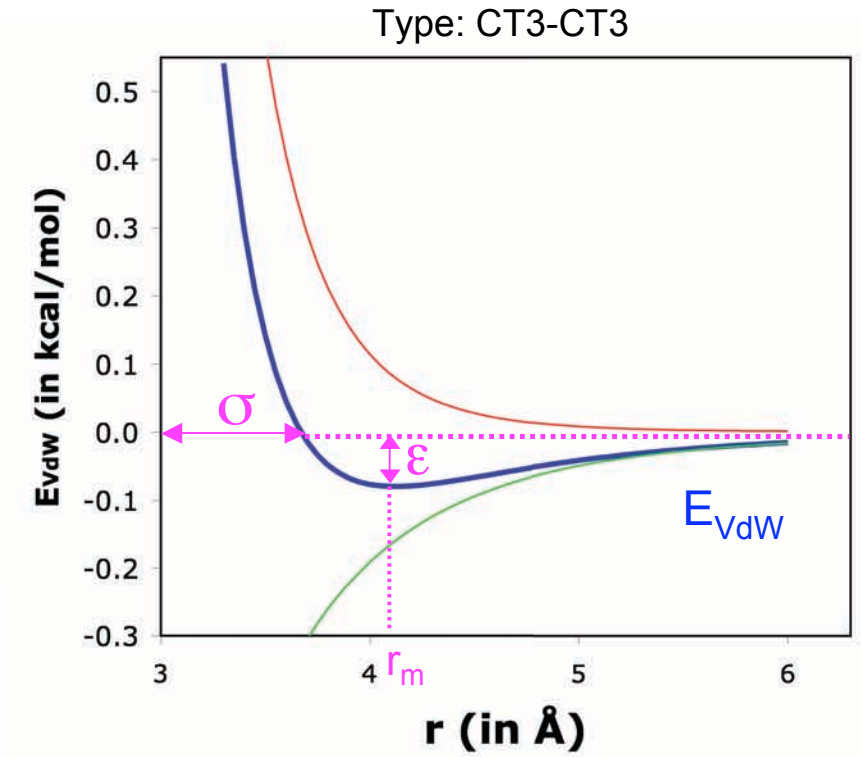
Van der Waals interactions

Lennard -Jones potential :

$$E_{VdW} = 4\varepsilon\left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6}\right] = \varepsilon\left[\left(\frac{r_m}{r}\right)^{12} - 2\left(\frac{r_m}{r}\right)^{6}\right]$$

σ : collision parameter
ε : well depth
r_m : distance at min $\quad r_m = 2^{1/6}\sigma$

Combination rule for two different atoms i, j : $\quad r_m = r_{m,i} + r_{m,j} \qquad \varepsilon = \sqrt{\varepsilon_i \varepsilon_j}$



Repulsive :
Pauli exclusion principle

$$\propto \frac{1}{r^{12}}$$

Attractive:
induced dipole / induced dipole

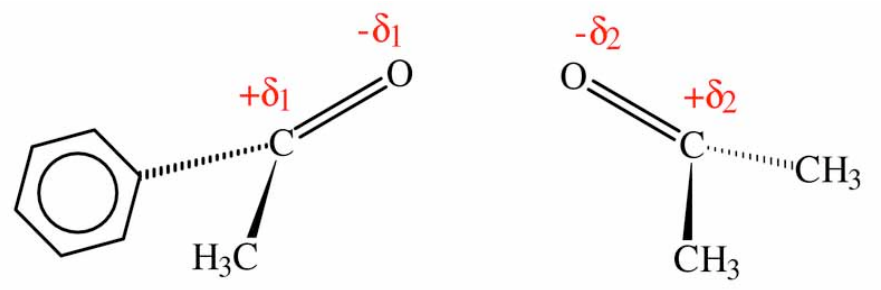
$$\propto -\frac{1}{r^{6}}$$

Electrostatic interactions

Coulomb law $$E_{elec} = \frac{q_i \, q_j}{4\pi\varepsilon_0 \varepsilon \, r_{ij}}$$

where ε is the dielectric constant :

1	for vacuum,
4-20	for protein core,
80	for water



The Coulomb energy decreases only as *1/r*

Despite dielectric shielding effects, it is a long range interaction

Special techniques to deal with this :

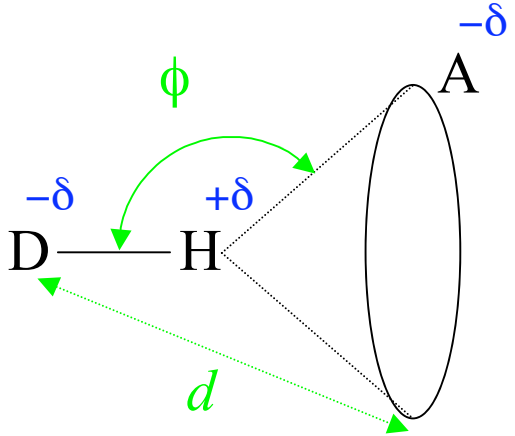
- PME : for stystems with periodic boundary conditions
- Reaction Field : suppose homogeneous dielectric outside cutoff

Derived Interactions

Some interactions are often referred to as particular interactions, but they result from the two interactions previously described, i.e. the electrostatic and the van der Waals interactions.

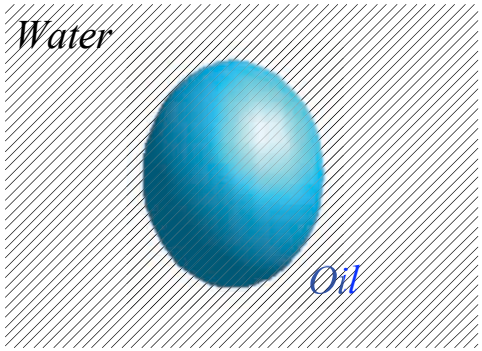
1) Hydrogen bonds (Hb)

- Interaction of the type D-H ··· A
- The origin of this interaction is a dipole-dipole attraction
- Typical ranges for distance and angle:
  $2.4 < d < 4.5$Å and $180° < \phi < 90°$



2) Hydrophobic effect

- Collective effect resulting from the energetically unfavorable surface of contact between the water and an apolar medium (loss of water-water Hb)
- The apolar medium reorganizes to minimize the water exposed surface (compaction, association...)



The total potential energy function

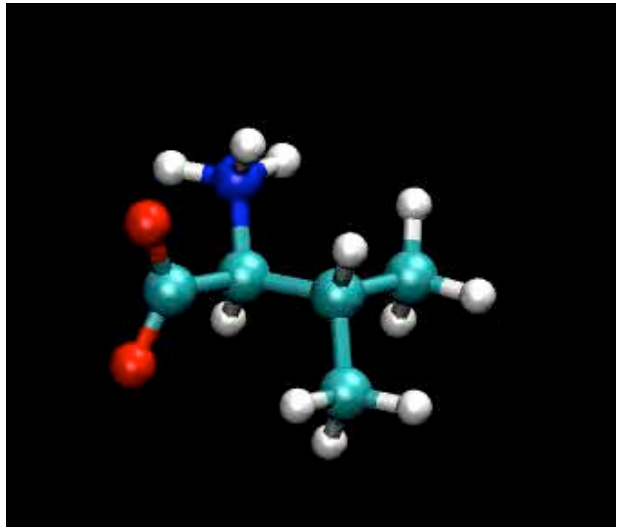
$$E = \sum_{bonds} K_b (r - r_0)^2 + \sum_{angles} K_\theta (\theta - \theta_0)^2 + \sum_{dihedrals} K_\phi \left[1 + \cos(n\phi - \delta)\right] + \sum_{impropers} K_\psi (\psi - \psi_0)^2$$

$$+ \sum_{i>j} \varepsilon \left[\left(\frac{r_m}{r} \right)^{12} - 2 \left(\frac{r_m}{r} \right)^{6} \right] + \sum_{i>j} \frac{q_i q_j}{4\pi\varepsilon_0 \varepsilon\, r}$$

For a system with 1500 atoms

~ 10^6 pairs of interacting atoms

Introduction of cutoff

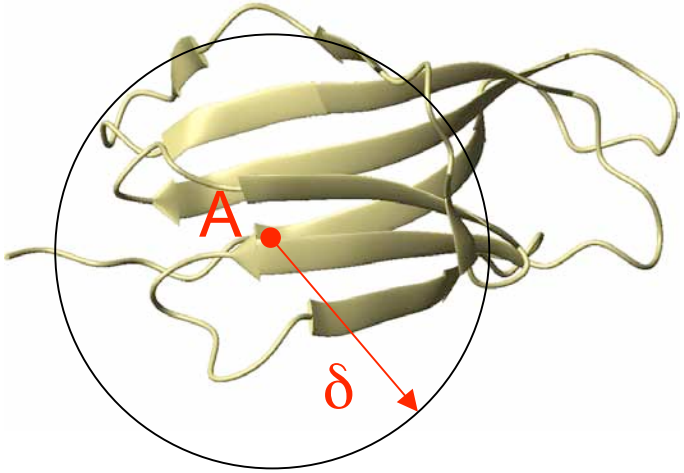


Cutoff for non-bonded interactions

For an atom A, only non-bonded interactions with atoms within δ Å are calculated

➡ Non-bonded neighbour lists

Generally, δ = 8 to 14 Å

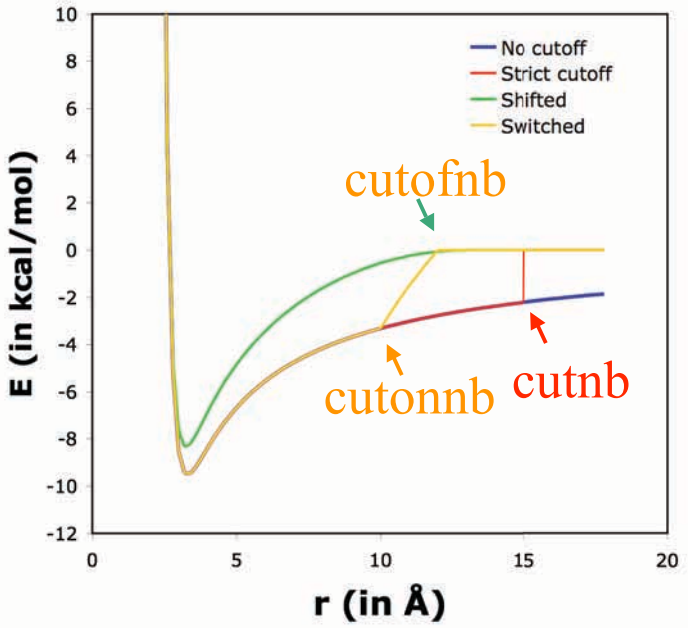


Three cutoff schemes: strict, shift, switch

Shift and switch:

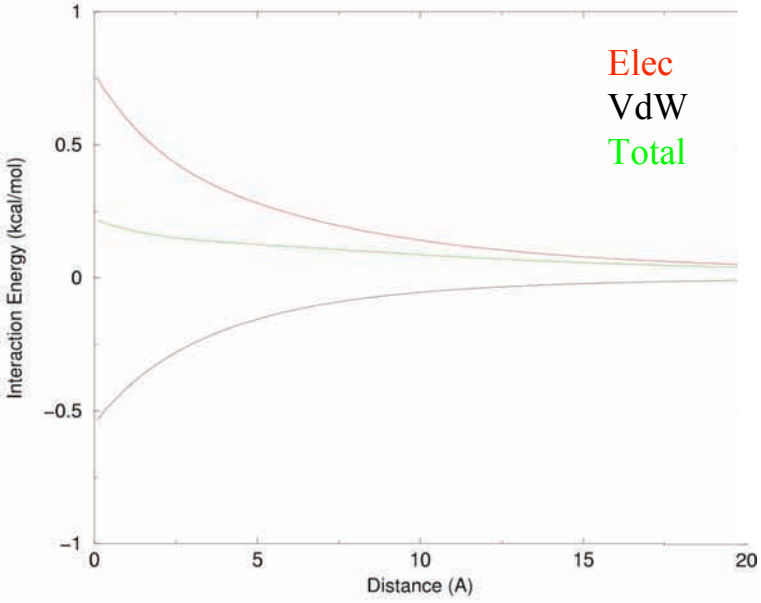
$$E'(r) = E(r) \times S(r)$$

S(r) differentiable

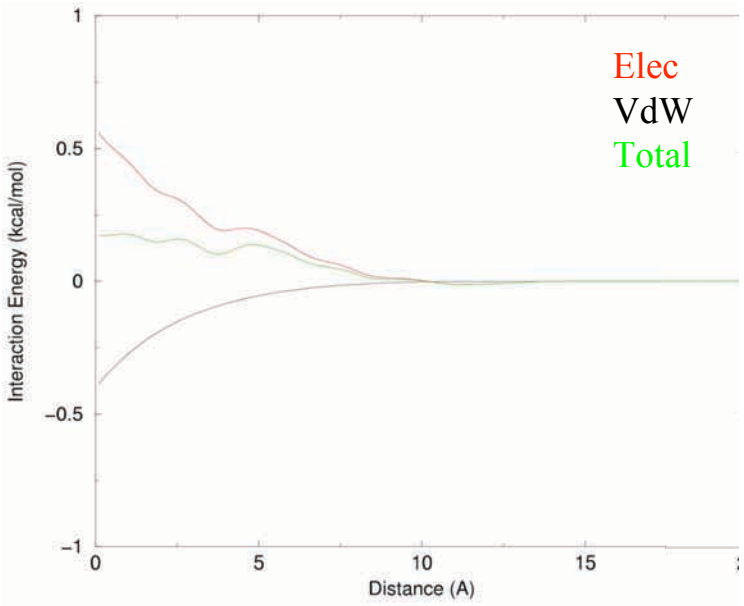


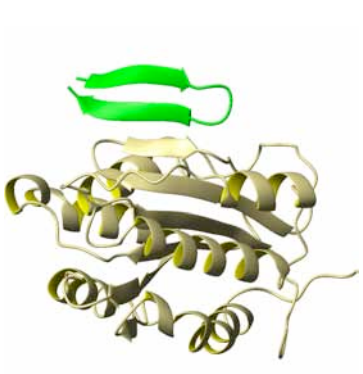
Effect of cutoff

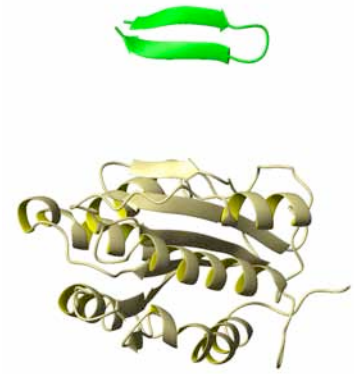
No cutoff

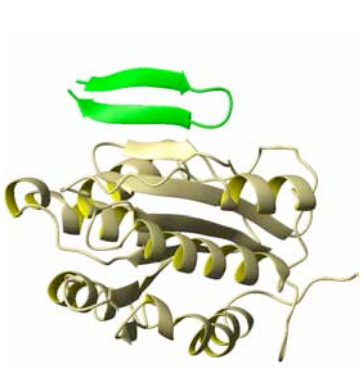


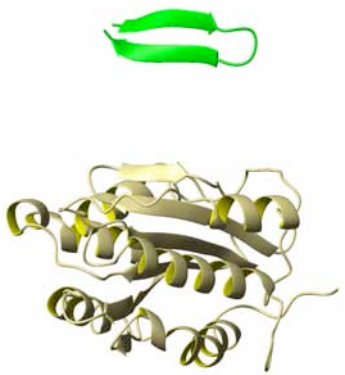
8 Å cutoff











Force field parametrization

$$E = \sum_{bonds} K_b\left(r - r_0\right)^2 + \sum_{angles} K_\theta\left(\theta - \theta_0\right)^2 + \sum_{dihedrals} K_\phi\left[1 + \cos\left(n\phi - \delta\right)\right] + \sum_{impropers} K_\psi\left(\psi - \psi_0\right)^2 + \sum_{i>j} 4\varepsilon\left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6}\right] + \sum_{i>j} \frac{q_i q_j}{4\pi\varepsilon_0 \varepsilon\, r}$$

Type of data	Type of system	Phase	Type of properties	Force field parameters
structural data (exp.)	small molecules	crystalline solid phase	molecular geometry: bond lengths, angles	r_0, θ_0, ψ_0
spectroscopic data (exp.)	small molecules	gas phase	intra-molecular vibrations: force constants	K_b, K_θ, K_ψ
quantum-chemical calculations : energy profiles (theor.)	small molecules	gas phase	torsional-angle rotational profiles	K_ϕ, δ, n
quantum-chemical calculations : electron densities (theor.)	small molecules	gas phase	atom charges	charges q_i (initial)
thermodynamic data (exp.)	molecules in solution, mixtures	condensed phase	heat of vaporisation, density, free energy of solvation	v. d. Waals : σ_i, ε_i charges q_i (final)
dielectric data (exp.)	small molecules	condensed phase	dielectric permittivity, relaxation	charges q_i
transport data (exp.)	small molecules	condensed phase	transport coefficients: diffusion, viscosity	v. d. Waals : σ_i, ε_i charges q_i

Solvation

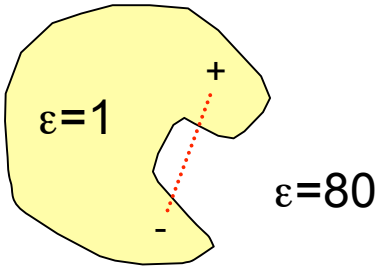
Fundamental influence on the structure, dynamics and thermodynamics of biological molecules

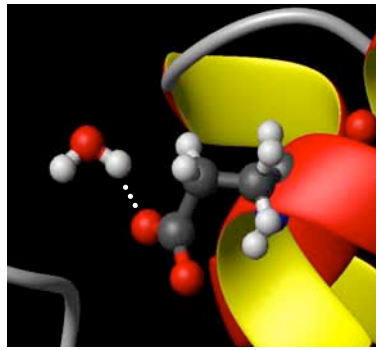
Effect through:

- Solvation of charge
- Screening of charge - charge interactions

$$E_{elec} = \frac{q_i \, q_j}{4\pi\varepsilon_0 \varepsilon(r) \, r}$$

- Hydrogen bonds between water molecules and polar functions of the solute





Taken into account via:

- ❑ Explicit solvation. Water molecules are included.
  - Stochastic boudary conditions
  - Periodic boundary conditions
- ❑ Implicit solvation. Water effect is modeled.
  - Screening constant
  - Implicit solvation models (Poisson Boltzmann, Generalized Born)

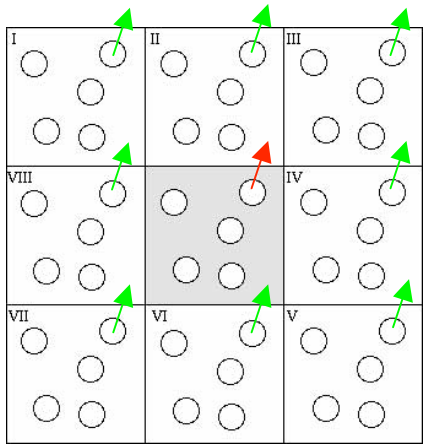
Explicit solvation schemes

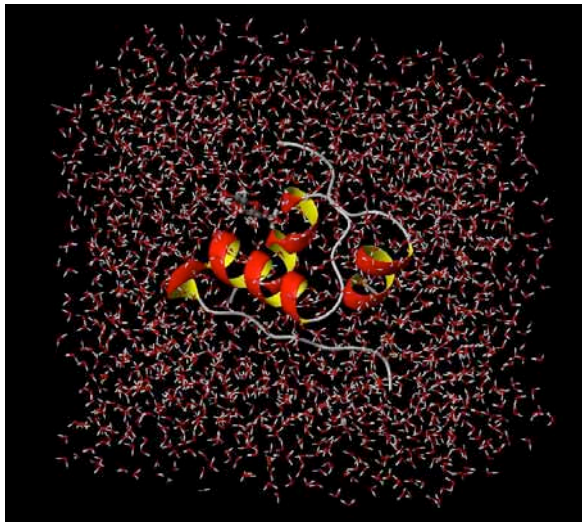
Stochastic boundary conditions

The region of interest is solvated in a water sphere at 1atm. The water molecules are submitted to an additional force field that restrain them in the sphere while maintaining a strong semblance to bulk water.

Periodic boundary conditions

The fully solvated central cell is simulated, in the environment produced by the repetition of this cell in all directions.





Implicit solvation

Screening constant

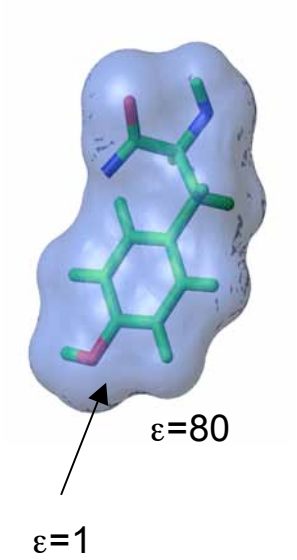
$$\varepsilon = Nr$$

N=4,8. The dielectric constant is a function of atom distance.
Mimic screening effect of solvent. Simple, unphysical but efficient.

In CHARMM:

```
NBOND RDIE EPS 4.0
```

Implicit solvent models



- Poisson Boltzmann (PB) equation.

$$\nabla \bullet \{\varepsilon(r)\nabla\phi(r)\} - \kappa' \sinh[\phi(r)] = -4\pi\rho(r)$$

f(r) : electrostatic potential,
r(r) : charge density

Equation solved numerically. Very time consuming.
In CHARMM: PBEQ module.

- Generalized born (GB) equation.

$$G_{elec} = \sum_{i>j} \frac{q_i q_j}{4\pi\varepsilon_0 r} - \frac{1}{2}\left(1 - \frac{1}{\varepsilon}\right) \sum_i \sum_j \frac{q_i q_j}{\sqrt{r^2 + a_i a_j \exp\left(-r^2/4a_i a_j\right)}}$$

a_i : Born radius

solvation energy

Others: EEF1, SASA, etc...

Explicit hydrogen bonds with water molecules are lost!

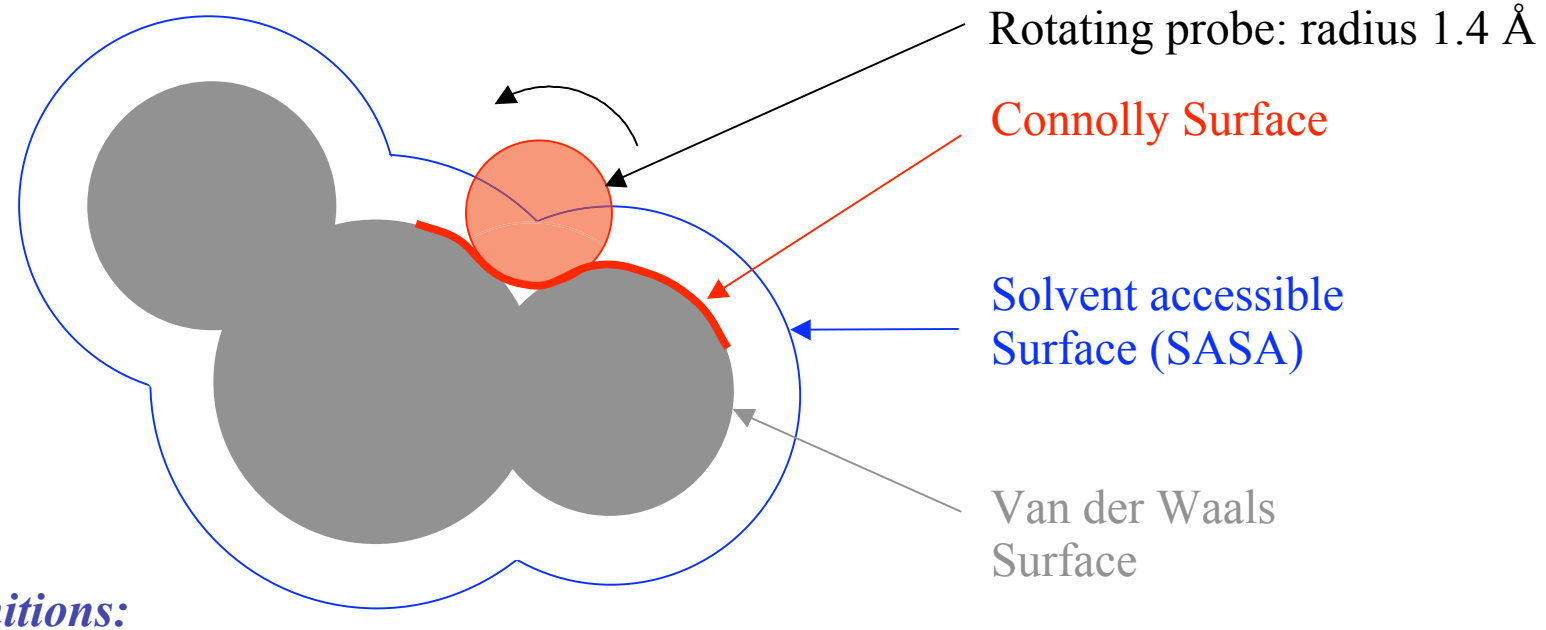
Introduction to molecular surfaces

Hydrophobic effect :

Simply modelled by a non-polar solvation (free) energy term, proportional to the solvent accessible surface area (SASA) :

$$\Delta G_{np,\,solv} = \gamma SASA + b$$

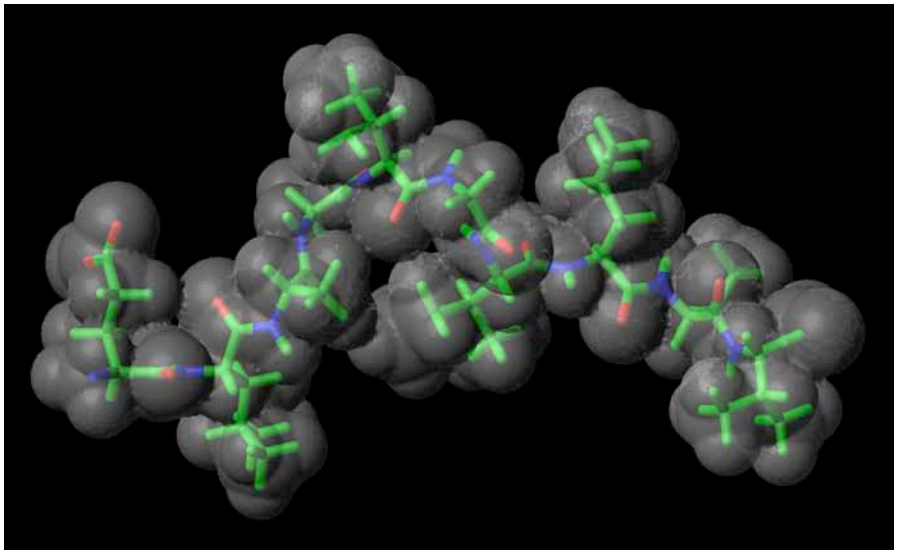
γ = 0.00542 kcal mol^{-1} Å^{-2}
b = 0.92 kcal mol^{-1}



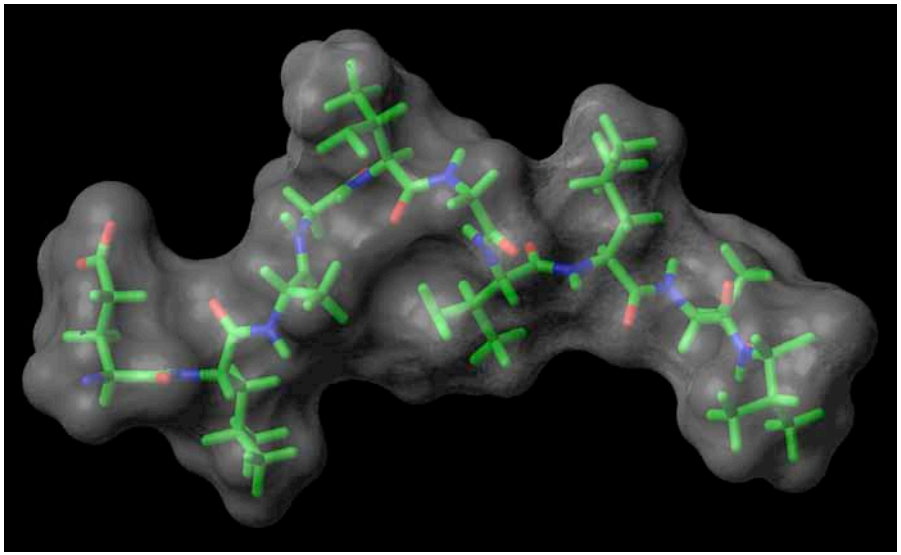
Definitions:

- Van der Waals: ensemble of van der Waals sphere centered at each atom
- Connolly: ensemble of contact points between probe and vdW spheres
- Solvent: ensemble of probe sphere centers

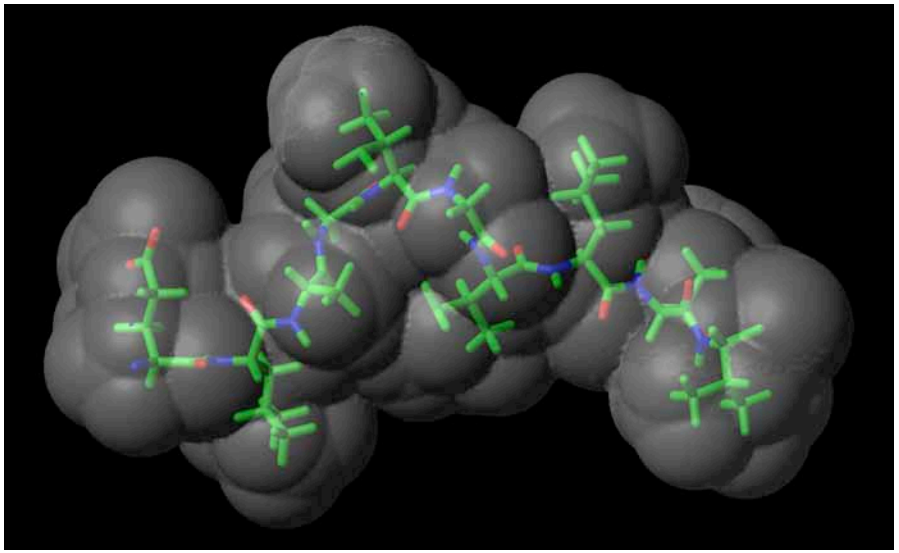
Examples of molecular surfaces



Van der Waals



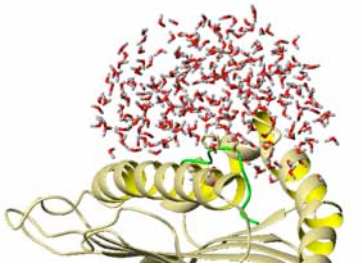
Connolly
(Contact)



Solvent accessible

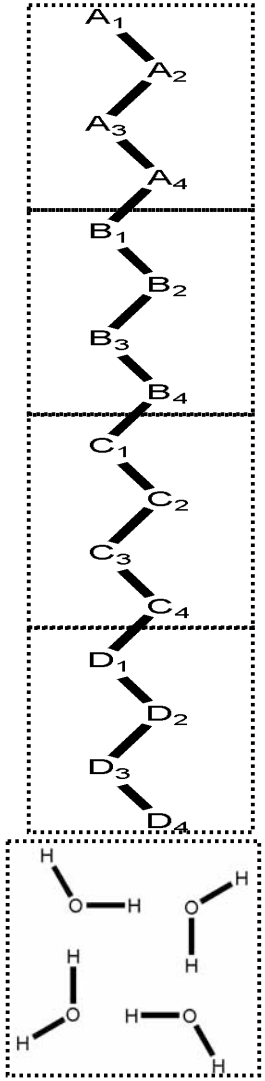
Limitations of classical MD

Problems	Solutions
1) **Fixed set of atom types**	
2) **No electronic polarization:** - fixed partial charges allow for *conformational* polarization but not *electronic* polarization	· Fluctuating charges treated as dynamical parameters · Charges on springs representing e^- clouds · QM-MM · Full QM simulations
3) **Quadratic form of potentials:** - problematic far from equilibrium values - no bond creation or deletion	· QM-MM · Full QM simulations



Coarse grain models

All-atom model
16 (CH_2 or CH_3) atoms



Map
to all-atom configurations

Centre of mass
$A_1 - A_4$

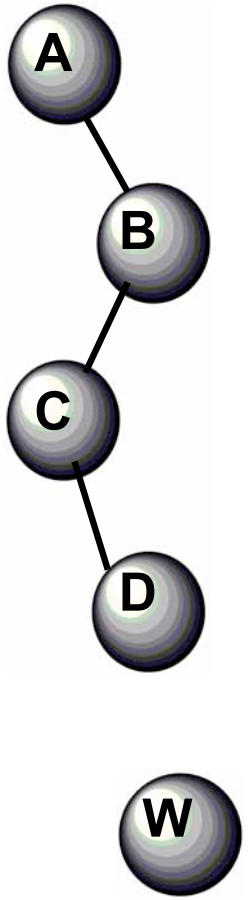
Centre of mass
$B_1 - B_4$

Centre of mass
$C_1 - C_4$

Centre of mass
$D_1 - D_4$

Centre of mass
$W_1 - W_4$

Coarse-grained model
4 atoms

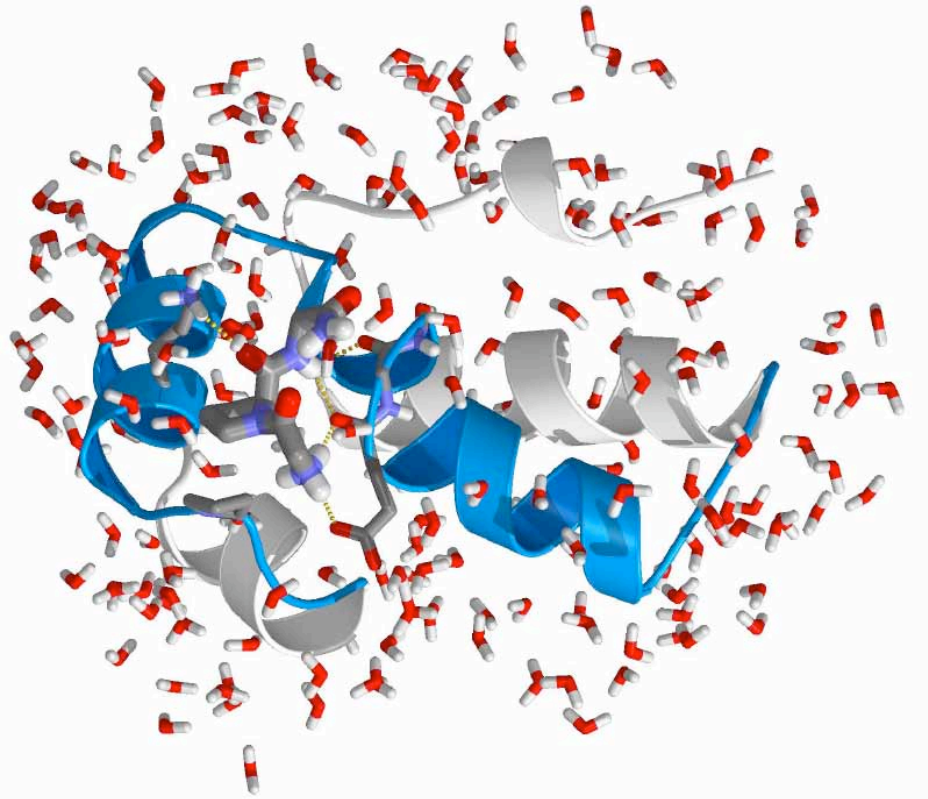


Molecular dynamics software

Package name	supported force fields
• CHARMM www.charmm.org	CHARMM (E / I; AA / UA), Amber
• Amber amber.scripps.edu	Amber (E / I ; AA)
• GROMOS www.igc.ethz.ch/GROMOS	Gromos (E / vacuum ; UA)
• Gromacs www.gromacs.org	Amber, Gromos, OPLS - (all E)
• NAMD www.ks.uiuc.edu/Research/namd	CHARMM, Amber, Gromos, ...

E = explicit solvent	AA = all atom
I = implicit solvent	UA = united atom (apolar H omitted)

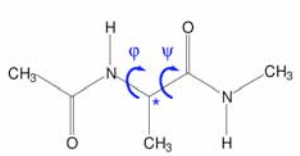
Plan



- Introduction
- The classical force field
- **Setting up a simulation**
- Connection to statistical mechanics
- Usage of MD simulation

Minimal input for MM

1) Topological properties:

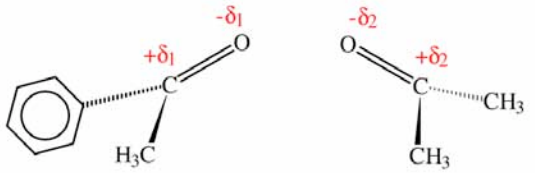


description of the covalent connectivity of the molecules to be modeled

2) Structural properties:

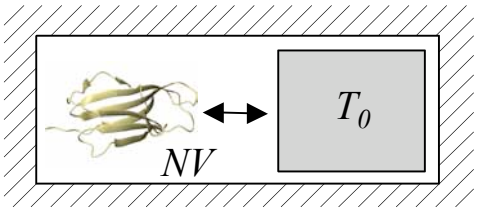
the starting conformation of the molecule, provided by an X-ray structure, NMR data or a theoretical model

3) Energetical properties:



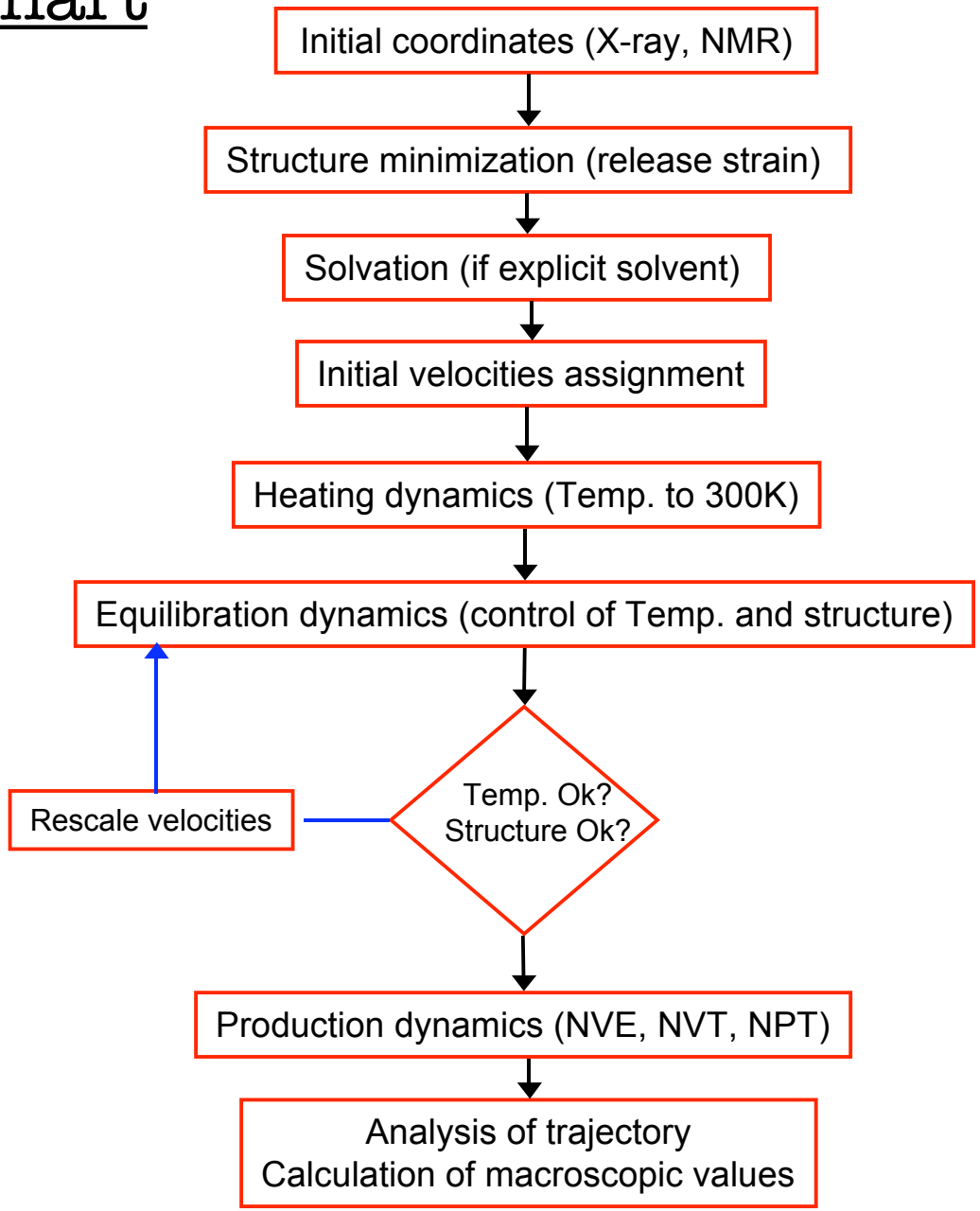
a force field describing the force acting on each atom of the molecules

4) Thermodynamical properties:



a sampling algorithm that generates the thermodynamical ensemble that matchese experimental conditions for the system, e.g. N,V,T , N,P,T, ...

MD flowchart



Coordinate files

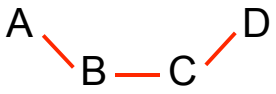
Cartesian coordinates

(x,y,z)

```
ATOM      1  N    VAL     1      -0.008  -0.022  -0.030  1.00  0.00      PEP
ATOM      2  HT1  VAL     1      -0.326   0.545   0.778  1.00  0.00      PEP
ATOM      3  HT2  VAL     1      -0.450  -0.956  -0.084  1.00  0.00      PEP
ATOM      4  HT3  VAL     1      -0.172   0.566  -0.876  1.00  0.00      PEP
ATOM      5  CA   VAL     1       1.477  -0.077   0.073  1.00  0.00      PEP
ATOM      6  HA   VAL     1       1.777  -0.598   0.971  1.00  0.00      PEP
ATOM      7  CB   VAL     1       2.038  -0.740  -1.193  1.00  0.00      PEP
```

Internal coordinates (IC)

Given 4 consecutive atoms A-B-C-D, the IC are:



R_{AB}, θ_{ABC}, ϕ_{ABCD}, θ_{BCD}, R_{CD}

```
   1    1 N      1 C      1 *CA    1 CB     1.4896  105.11  117.88  111.68   1.5353
   2    1 N      1 C      1 *CA    1 HA     1.4896  105.11 -118.25  108.30   1.0807
   3    1 N      1 CA     1 CB     1 CG1    1.4896  108.86  176.05  110.92   1.5421
   4    1 CG1    1 CA     1 *CB    1 CG2    1.5421  110.92  121.67  110.44   1.5454
   5    1 CG1    1 CA     1 *CB    1 HB     1.5421  110.92 -118.15  109.36   1.1177
   6    1 CA     1 CB     1 CG1    1 HG11   1.5353  110.92   56.96  111.11   1.1099
   7    1 HG11   1 CB     1 *CG1   1 HG12   1.1099  111.11 -119.80  110.60   1.1134
   8    1 HG11   1 CB     1 *CG1   1 HG13   1.1099  111.11  120.81  110.60   1.1103
```

It is possible to calculate missing cartesian coordinates from the existing ones and the IC

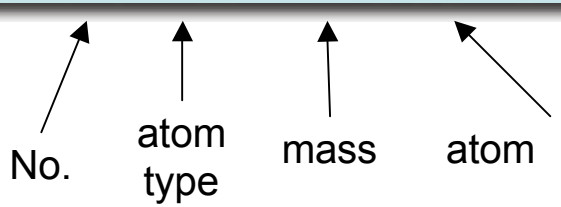
Definition of atom types

The residue topology file (RTF) contains the atom types and the standard topology of residues.

Example: the CHARMM force field, version 22
file : top_all22_prot.inp

Atom types section :

```
MASS     1 H      1.00800 H ! polar H
MASS     2 HC     1.00800 H ! N-ter H
MASS     3 HA     1.00800 H ! nonpolar H
MASS     4 HT     1.00800 H ! TIPS3P WATER HYDROGEN
MASS     5 HP     1.00800 H ! aromatic H
MASS     6 HB     1.00800 H ! backbone H
[...]
MASS    20 C     12.01100 C ! carbonyl C, peptide backbone
MASS    21 CA    12.01100 C ! aromatic C
MASS    22 CT1   12.01100 C ! aliphatic sp3 C for CH
MASS    23 CT2   12.01100 C ! aliphatic sp3 C for CH2
MASS    24 CT3   12.01100 C ! aliphatic sp3 C for CH3
MASS    25 CPH1  12.01100 C ! his CG and CD2 carbons
```



90 atom types

Decomposition into residues

In CHARMM, molecules are decomposed into residues.

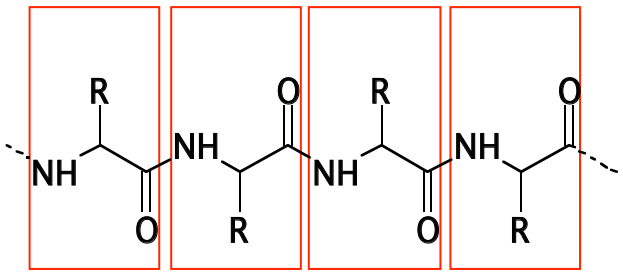
A molecule may be composed of one to several hundreds or thousands of residues.

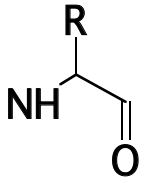
Residues correspond to amino acids for proteins or to nucleotides for DNA.

Topologies for individual residues are pre-defined in CHARMM

➡ Easy construction of the protein topology from the sequence.

Only informations about 20 amino acids needed to construct the topology of all proteins.





Residue definition in CHARMM

Residue topology definition

file : top_all22_prot.inp

```
RESI ALA          0.00
GROUP
ATOM N    NH1    -0.47  !      |
ATOM HN   H       0.31  !   HN-N
ATOM CA   CT1     0.07  !      |      HB1
ATOM HA   HB      0.09  !      |     /
GROUP                   !   HA-CA--CB-HB2
ATOM CB   CT3    -0.27  !      |     \
ATOM HB1  HA      0.09  !      |      HB3
ATOM HB2  HA      0.09  !    O=C
ATOM HB3  HA      0.09  !      |
GROUP                   !
ATOM C    C       0.51
ATOM O    O      -0.51
BOND CB CA  N   HN  N   CA
BOND C  CA  C   +N  CA HA  CB HB1  CB HB2  CB HB3
DOUBLE O  C
IMPR N -C CA HN  C CA +N O
DONOR HN N
ACCEPTOR O C
IC -C   CA   *N   HN    1.3551 126.4900  180.0000 115.4200  0.9996
IC -C   N    CA   C     1.3551 126.4900  180.0000 114.4400  1.5390
IC N    CA   C    +N    1.4592 114.4400  180.0000 116.8400  1.3558
IC +N   CA   *C   O     1.3558 116.8400  180.0000 122.5200  1.2297
IC CA   C    +N   +CA   1.5390 116.8400  180.0000 126.7700  1.4613
IC N    C    *CA  CB    1.4592 114.4400  123.2300 111.0900  1.5461
IC N    C    *CA  HA    1.4592 114.4400 -120.4500 106.3900  1.0840
IC C    CA   CB   HB1   1.5390 111.0900  177.2500 109.6000  1.1109
IC HB1  CA   *CB  HB2   1.1109 109.6000  119.1300 111.0500  1.1119
IC HB1  CA   *CB  HB3   1.1109 109.6000 -119.5800 111.6100  1.1114
```

- Total charge
- Atom name
- Atom type
- Atom charge
- Bond
- Improper angle
- Previous residue
- Next residue
- IC

Angles and dihedrals can be generated automatically from this.

Patching residues

Treat protein special features.

To make N- and C-termini:

```
PRES NTER         1.00 ! standard N-terminus
GROUP                  ! use in generate statement
ATOM N    NH3    -0.30 !
ATOM HT1  HC      0.33 !         HT1
ATOM HT2  HC      0.33 !     (+)/
ATOM HT3  HC      0.33 ! --CA--N--HT2
ATOM CA   CT1     0.21 !   |    \
ATOM HA   HB      0.10 !   HA    HT3
DELETE ATOM HN
BOND HT1 N HT2 N HT3 N
DONOR HT1 N
DONOR HT2 N
DONOR HT3 N
IC HT1  N    CA   C     0.0000  0.0000  180.0000  0.0000  0.0000
IC HT2  CA   *N   HT1   0.0000  0.0000  120.0000  0.0000  0.0000
IC HT3  CA   *N   HT2   0.0000  0.0000  120.0000  0.0000  0.0000
```

To make disulfide bridges:

```
PRES DISU        -0.36 ! patch for disulfides. Patch must be 1-CYS and 2-CYS.
                       ! use in a patch statement
                       ! follow with AUTOgenerate ANGLes DIHEdrals command
GROUP
ATOM 1CB  CT2    -0.10 !
ATOM 1SG  SM     -0.08 !           2SG--2CB--
GROUP                  !          /
ATOM 2SG  SM     -0.08 ! -1CB--1SG
ATOM 2CB  CT2    -0.10 !
DELETE ATOM 1HG1
DELETE ATOM 2HG1
BOND 1SG 2SG
IC 1CA  1CB  1SG  2SG   0.0000  0.0000  180.0000  0.0000  0.0000
IC 1CB  1SG  2SG  2CB   0.0000  0.0000   90.0000  0.0000  0.0000
IC 1SG  2SG  2CB  2CA   0.0000  0.0000  180.0000  0.0000  0.0000
```

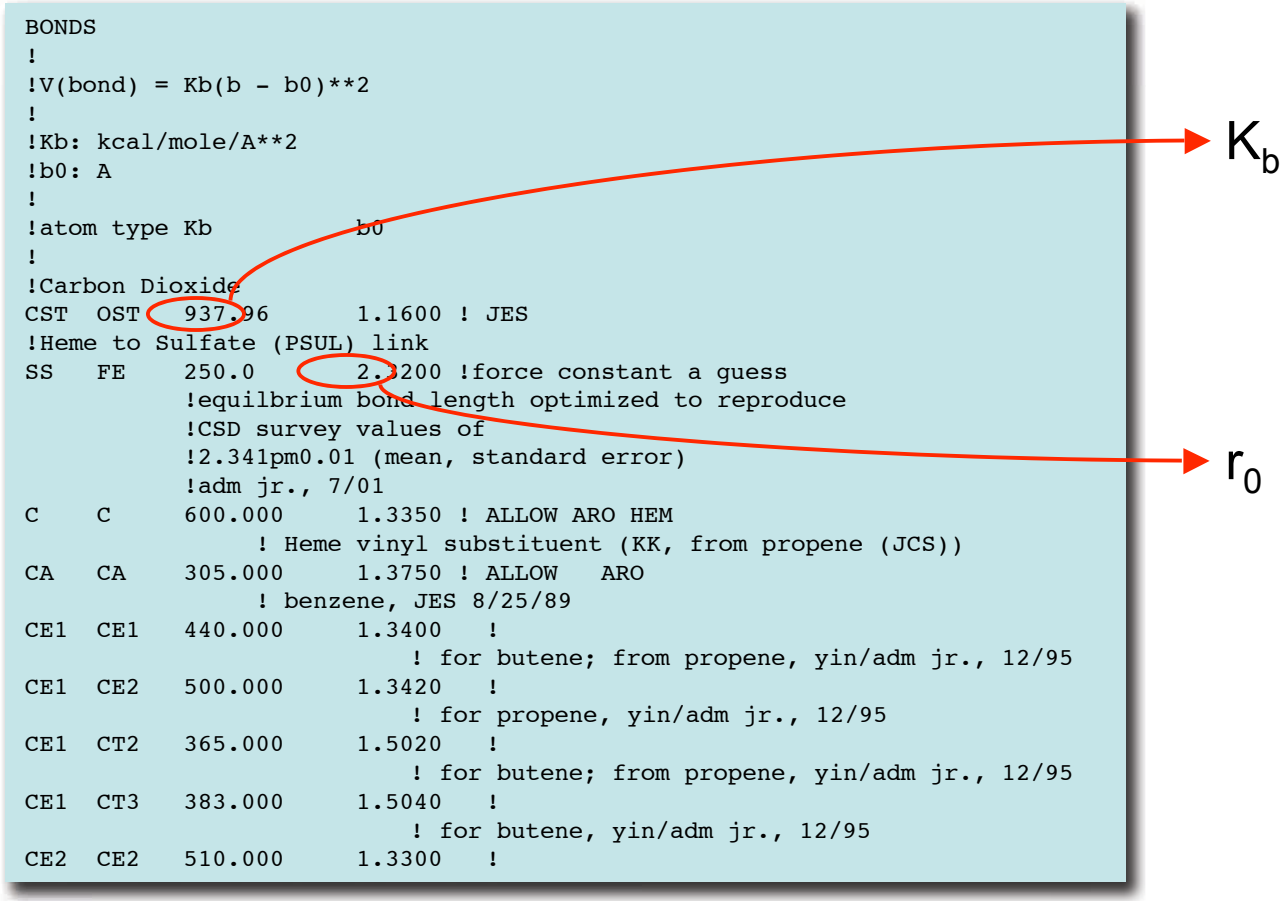
Etc...

The parameter file

Contains force field parameters. file : par_all22_prot.inp

Bonds:

```
BONDS
!
!V(bond) = Kb(b - b0)**2
!
!Kb: kcal/mole/A**2
!b0: A
!
!atom type Kb          b0
!
!Carbon Dioxide
CST  OST    937.96      1.1600 ! JES
!Heme to Sulfate (PSUL) link
SS   FE     250.0       2.3200 !force constant a guess
            !equilbrium bond length optimized to reproduce
            !CSD survey values of
            !2.341pm0.01 (mean, standard error)
            !adm jr., 7/01
C    C      600.000     1.3350 ! ALLOW ARO HEM
                ! Heme vinyl substituent (KK, from propene (JCS))
CA   CA     305.000     1.3750 ! ALLOW   ARO
                ! benzene, JES 8/25/89
CE1  CE1    440.000     1.3400   !
                            ! for butene; from propene, yin/adm jr., 12/95
CE1  CE2    500.000     1.3420   !
                            ! for propene, yin/adm jr., 12/95
CE1  CT2    365.000     1.5020   !
                            ! for butene; from propene, yin/adm jr., 12/95
CE1  CT3    383.000     1.5040   !
                            ! for butene, yin/adm jr., 12/95
CE2  CE2    510.000     1.3300   !
```



Idem for angles, dihedrals, impropers, ...

Setting up a system in CHARMM

```
$  charmm < input_file.inp > output_file.out
```

```
*

!------ Standard Topology and Parameters
OPEN UNIT 1 CARD READ NAME top_all22_prot.inp
READ RTF CARD UNIT 1
CLOSE UNIT 1

OPEN UNIT 1 CARD READ NAME par_all22_prot.inp
READ PARA CARD UNIT 1
CLOSE UNIT 1

! Generate actual topology
OPEN UNIT 1 READ CARD NAME 1aho-xray.pdb
READ SEQUENCE PDB UNIT 1
GENE 1S SETUP FIRST NTER LAST CTER
REWIND UNIT 1
READ COOR PDB UNIT 1
CLOSE UNIT 1

! Make disulfide bridges
PATCH DISU 1S 12 1S 63
PATCH DISU 1S 16 1S 36
PATCH DISU 1S 22 1S 46
PATCH DISU 1S 26 1S 48
AUTOgenerate ANGLes DIHEdrals
```

```
! Fill IC table and build missing coordinates
IC FILL
IC PARA ALL
IC BUILD

! Build better coordinates for hydrogens
HBUILD SELE TYPE H* END

! Write coordinates in CHARMM format
OPEN UNIT 1 WRITE CARD NAME 1aho.crd
WRITE COOR CARD UNIT 1
CLOSE UNIT 1

! Write coordinates in PDB format
OPEN UNIT 1 WRITE CARD NAME 1aho.pdb
WRITE COOR PDB UNIT 1
CLOSE UNIT 1

! Write Protein Structure File for subsequent use
OPEN UNIT 1 WRITE CARD NAME 1aho.psf
WRITE PSF CARD UNIT 1
CLOSE UNIT 1
```

```
PSFSUM> Summary of the structure file counters :
        Number of segments      =        1   Number of residues   =       64
        Number of atoms         =      962   Number of groups     =      298
        Number of bonds         =      980   Number of angles     =     1753
        Number of dihedrals     =     2591   Number of impropers  =      169
        Number of cross-terms   =        0
        Number of HB acceptors  =       98   Number of HB donors  =      118
        Number of NB exclusions =        0   Total charge =    1.00000
```

The protein structure file (PSF)

The PSF is generated by CHARMM from the sequence of the proteins, the ligands, the water molecules, etc...,
using the information present in the residue topology file (RTF)

It contains all the information needed for future simulations :

- Residues and segments. How the system is divided into residues and segments.
- Atom information. Names, types, charges, masses.
- Bond, angle, dihedral and improper dihedral lists
- Electrostatic groupings. How some numbers of atoms are grouped for the purpose of calculating long range electrostatic

A segment is a group of molecules, for example:

- one single protein
- a collection of water molecules
- a collection of ions
- a ligand

Performing a calculation in CHARMM

Second step : perform calculation, e.g. energy evaluation

input

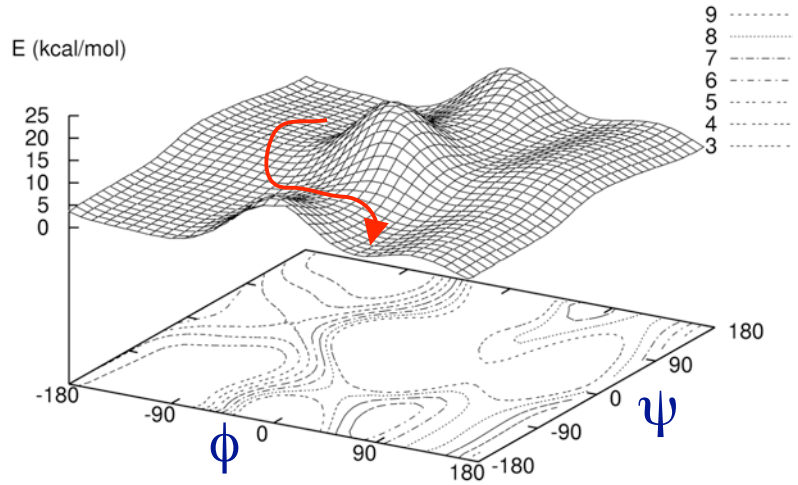
```
*

!------ Standard Topology and Parameters
OPEN UNIT 1 CARD READ NAME top_all22_prot.inp
READ RTF CARD UNIT 1
CLOSE UNIT 1

OPEN UNIT 1 CARD READ NAME par_all22_prot.inp
READ PARA CARD UNIT 1
CLOSE UNIT 1

!------ Actual topology
OPEN UNIT 1 READ CARD NAME 1aho.psf
READ PSF CARD UNIT 1
CLOSE UNIT 1

!------ Coordinates
OPEN UNIT 1 READ CARD NAME 1aho.pdb
READ COOR PDB UNIT 1
CLOSE UNIT 1

!------ Energy calculation
ENERGY

!------ End of input file
STOP
```

output

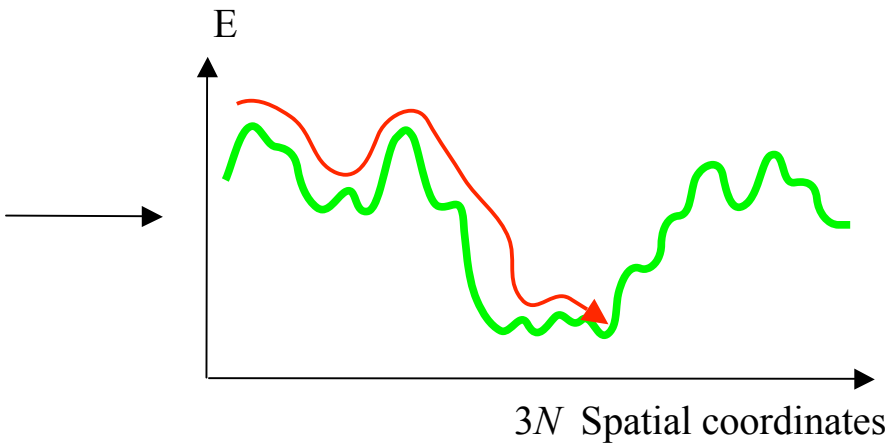
```
 CHARMM>    ENERGY

 NONBOND OPTION FLAGS:
     ELEC     VDW      ATOMs    CDIElec  SHIFt    VATOm    VSWItch
     BYGRoup  NOEXtnd  NOEWald
 CUTNB  = 14.000 CTEXNB =999.000 CTONNB = 10.000 CTOFNB = 12.000
 WMIN   =  1.500 WRNMXD =  0.500 E14FAC =  1.000 EPS    =  1.000
 NBXMOD =      5
 There are        0 atom  pairs and        0 atom  exclusions.
 There are        0 group pairs and        0 group exclusions.
 <MAKINB> with mode   5 found   2733 exclusions and   2534 interactions(1-4)
 <MAKGRP> found    886 group exclusions.
 Generating nonbond list with Exclusion mode = 5
 == PRIMARY == SPACE FOR   276432 ATOM PAIRS AND        0 GROUP PAIRS

 General atom nonbond list generation found:
   224678 ATOM PAIRS WERE FOUND FOR ATOM LIST
     9439 GROUP PAIRS REQUIRED ATOM SEARCHES

ENER ENR:  Eval#     ENERgy      Delta-E         GRMS
ENER INTERN:          BONDs       ANGLes       UREY-b    DIHEdrals    IMPRopers
ENER EXTERN:        VDWaals         ELEC       HBONds          ASP         USER
 ----------       ---------    ---------    ---------    ---------    ---------
ENER>        0  -1222.13834      0.00000      4.26768
ENER INTERN>       29.79474     88.24015      5.92868    239.13668      2.18868
ENER EXTERN>     -302.00504  -1285.42224      0.00000      0.00000      0.00000
 ----------       ---------    ---------    ---------    ---------    ---------
```

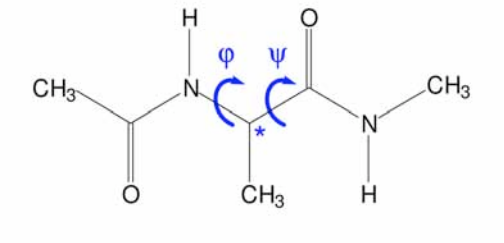
Energy landscape

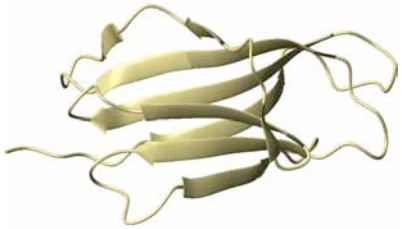
Landscape for ϕ/ψ plane of dialanine

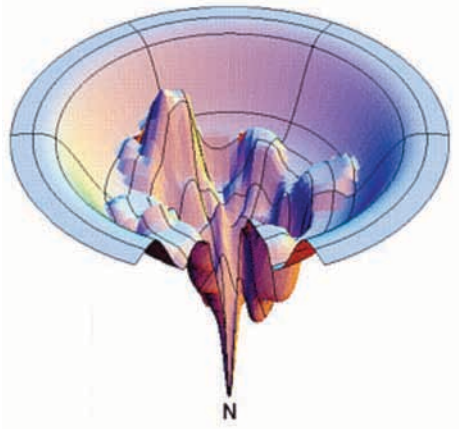


Complex landscape for a protein



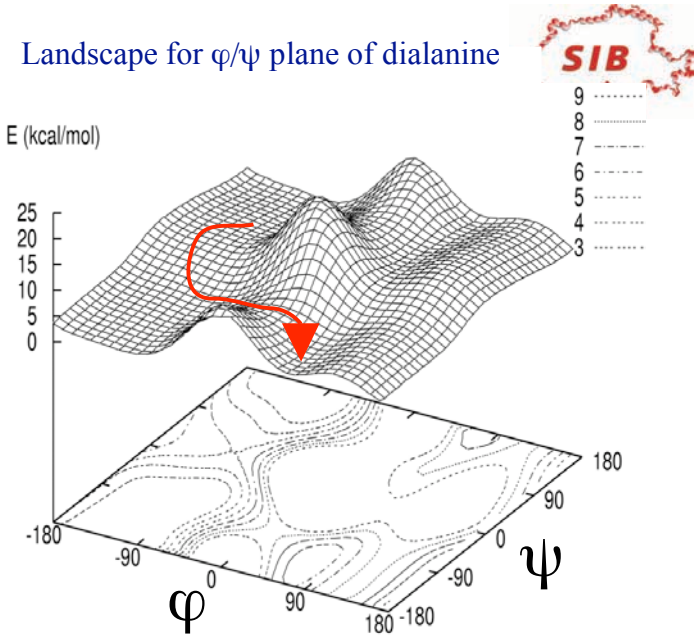






Minimization

Landscape for φ/ψ plane of dialanine



Finding minimum energy conformations given a potential energy function:

$$\frac{\partial E}{\partial x_i} = 0 \quad \text{and} \quad \frac{\partial^2 E}{\partial x_i^2} > 0$$

Used to:

- relieve strain in experimental conformations
- find (energetically) stable states

Huge number of degrees of freedom in macromolecular systems.

↳ Huge amount of local minima
Impossible to find true global minimum

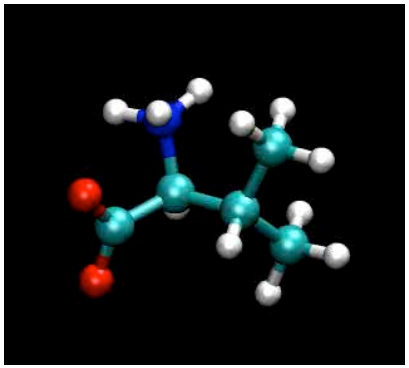
Different minimization methods available:

- Steepest descent (SD) ➜ relieve strain, find close local minimum
- Conjugated gradient (CONJ)
- Adopted Basis Newton Raphson (ABNR)

} Find lower energy mimima

CHARMM input for minimization

```
*

!------ Standard Topology and Parameters
OPEN UNIT 1 CARD READ NAME top_all22_prot.inp
READ RTF CARD UNIT 1
CLOSE UNIT 1

OPEN UNIT 1 CARD READ NAME par_all22_prot.inp
READ PARA CARD UNIT 1
CLOSE UNIT 1

!------ Actual topology
OPEN UNIT 1 READ CARD NAME val.psf
READ PSF CARD UNIT 1
CLOSE UNIT 1

!------ Coordinates
OPEN UNIT 1 READ CARD NAME val.pdb
READ COOR PDB UNIT 1
CLOSE UNIT 1

!------ ABNR minimization
MINI ABNR NSTEP 200

!------ End of input file
STOP
```



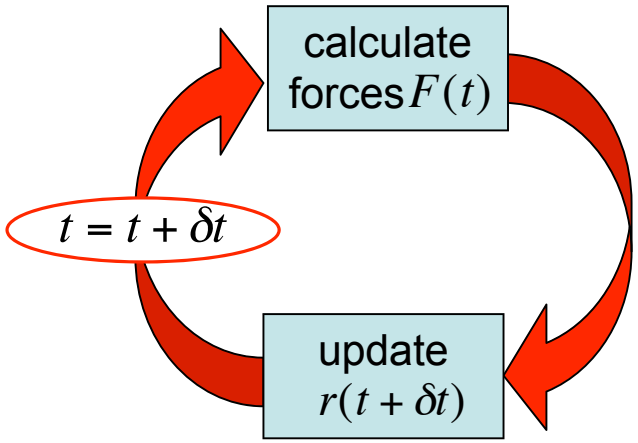
```
 ABNER> An energy minimization has been requested.

 EIGRNG =    0.0005000   MINDIM =           5
 NPRINT =           50   NSTEP  =         200
 PSTRCT =    0.0000000   SDSTP  =    0.0200000
 STPLIM =    1.0000000   STRICT =    0.1000000
 TOLFUN =    0.0000000   TOLGRD =    0.0000000
 TOLITR =          100   TOLSTP =    0.0000000
 FMEM   =    0.0000000
MINI MIN: Cycle      ENERgy      Delta-E         GRMS    Step-size
MINI INTERN:          BONDs       ANGLes       UREY-b    DIHEdrals    IMPRopers
MINI EXTERN:        VDWaals         ELEC       HBONds          ASP         USER
 ----------      ---------   ---------   ---------   ---------   ---------
MINI>        0    -18.69945     0.00000     3.66846     0.00000
MINI INTERN>         0.49721     2.67684     0.28823     6.34896     0.15433
MINI EXTERN>         7.68221   -36.34723     0.00000     0.00000     0.00000
 ----------      ---------   ---------   ---------   ---------   ---------
MINI>       50    -26.40712     7.70767     0.60709     0.00545
MINI INTERN>         0.65246     3.04259     0.34085     0.97825     0.05940
MINI EXTERN>         4.90369   -36.38436     0.00000     0.00000     0.00000
 ----------      ---------   ---------   ---------   ---------   ---------
MINI>      100    -27.90707     1.49995     0.38850     0.00279
MINI INTERN>         0.84536     3.91951     0.53225     1.87179     0.09659
MINI EXTERN>         6.11013   -41.28270     0.00000     0.00000     0.00000
 ----------      ---------   ---------   ---------   ---------   ---------
 UPDECI: Nonbond update at step       103
 Generating nonbond list with Exclusion mode = 5
 == PRIMARY == SPACE FOR      172 ATOM PAIRS AND        0 GROUP PAIRS

 General atom nonbond list generation found:
      120 ATOM PAIRS WERE FOUND FOR ATOM LIST
        0 GROUP PAIRS REQUIRED ATOM SEARCHES
MINI>      150    -28.09742     0.19035     0.04299     0.00027
MINI INTERN>         0.93508     3.79489     0.58387     2.19377     0.06973
MINI EXTERN>         6.18507   -41.85983     0.00000     0.00000     0.00000
 ----------      ---------   ---------   ---------   ---------   ---------
MINI>      200    -28.09805     0.00062     0.00826     0.00005
MINI INTERN>         0.92928     3.80437     0.58415     2.18679     0.06957
MINI EXTERN>         6.19000   -41.86221     0.00000     0.00000     0.00000
 ----------      ---------   ---------   ---------   ---------   ---------

 ABNER> Minimization exiting with number of steps limit (  200) exceeded.

ABNR MIN: Cycle      ENERgy      Delta-E         GRMS    Step-size
ABNR INTERN:          BONDs       ANGLes       UREY-b    DIHEdrals    IMPRopers
ABNR EXTERN:        VDWaals         ELEC       HBONds          ASP         USER
 ----------      ---------   ---------   ---------   ---------   ---------
ABNR>      200    -28.09805     0.00062     0.00826     0.00005
ABNR INTERN>         0.92928     3.80437     0.58415     2.18679     0.06957
ABNR EXTERN>         6.19000   -41.86221     0.00000     0.00000     0.00000
 ----------      ---------   ---------   ---------   ---------   ---------
```

Simple Molecular dynamics

Newton's law of motion:

$$F_i = m_i a_i = -\frac{dE_i}{dr_i}$$

MD algorithm



In discete time : integration algorithm.
Example: Verlet algrorithm

$$r(t+\delta t) = r(t) + v(t) \times \delta t + \frac{1}{2} a(t) \times \delta t^2$$

$$r(t-\delta t) = r(t) - v(t) \times \delta t + \frac{1}{2} a(t) \times \delta t^2$$

$$\Rightarrow \quad r(t+\delta t) = 2r(t) - r(t-\delta t) + \frac{F(t)}{m} \times \delta t^2$$

$\delta t \sim 1$ fs $= 10^{-15}$ s

Propagation of time: position at time t+dt is a determined by position at time t and t-dt, and by the acceleration at time t (i.e., the forces at time t)

The equations of motion are deterministic, e.g., the positions and the velocities at time zero determine the positions and velocities at all other times, t.

CHARMM input file for MD simulation

```
*

!------ Standard Topology and Parameters
OPEN UNIT 1 CARD READ NAME top_all22_prot.inp
READ RTF CARD UNIT 1
CLOSE UNIT 1

OPEN UNIT 1 CARD READ NAME par_all22_prot.inp
READ PARA CARD UNIT 1
CLOSE UNIT 1

!------ Actual topology
OPEN UNIT 1 READ CARD NAME val.psf
READ PSF CARD UNIT 1
CLOSE UNIT 1

!------ Coordinates
OPEN UNIT 1 READ CARD NAME val.pdb
READ COOR PDB UNIT 1
CLOSE UNIT 1

!------ MD simulation
OPEN WRITE UNIT 31 CARD NAME traj/dyna.rst  ! Restart file
OPEN WRITE UNIT 32 FILE NAME traj/dyna.dcd  ! Coordinates file
OPEN WRITE UNIT 33 CARD NAME traj/dyna.ene  ! Energy file

DYNA  VERLET START NSTEP 1000 TIMESTEP 0.001 -
      IUNWRI 31 IUNCRD 32 KUNIT 33 -
      IPRFRQ 100 NPRINT 100 NSAVC 100 NSAVV 100 ISVFRQ 2000 -
      FIRSTT 300.0 FINALT 300.0 ICHEW 1 -
      IEQFRQ 10 IASORS 0 ISCVEL 0 IASVEL 1 ISEED 8364127 -
      INBFRQ 10


!------ End of input file
STOP
```

Generated files:
- Restart file (rst). To restart after crash or to continue MD sim.
- Collection of instantaneous coordinates along trajectory (dcd)

Control propagation of time. Timestep in ps (i.e. 1fs here)

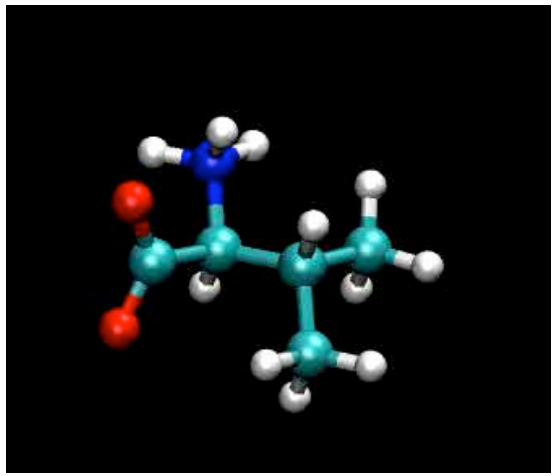
Control output

Control temperature

Control non bonded list update frequency

Output of MD simulation

```
DYNA DYN: Step          Time      TOTEner       TOTKe       ENERgy  TEMPerature
DYNA PROP:              GRMS      HFCTote       HFCKe       EHFCor       VIRKe
DYNA INTERN:           BONDs       ANGLes       UREY-b    DIHEdrals    IMPRopers
DYNA EXTERN:         VDWaals         ELEC       HBONds          ASP         USER
DYNA PRESS:             VIRE         VIRI       PRESSE       PRESSI       VOLUme
 ----------         ---------    ---------    ---------    ---------    ---------
DYNA>         0      0.00000     -6.35935     19.31479    -25.67414    381.16244
DYNA PROP>           1.79744     -6.35862     19.31700      0.00074     -3.31909
DYNA INTERN>         0.49721      2.67684      0.28823      3.69866      0.15433
DYNA EXTERN>         4.45183    -37.44124      0.00000      0.00000      0.00000
DYNA PRESS>          0.00000      2.21273      0.00000      0.00000      0.00000
 ----------         ---------    ---------    ---------    ---------    ---------

[...]

DYNA DYN: Step          Time      TOTEner       TOTKe       ENERgy  TEMPerature
DYNA PROP:              GRMS      HFCTote       HFCKe       EHFCor       VIRKe
DYNA INTERN:           BONDs       ANGLes       UREY-b    DIHEdrals    IMPRopers
DYNA EXTERN:         VDWaals         ELEC       HBONds          ASP         USER
DYNA PRESS:             VIRE         VIRI       PRESSE       PRESSI       VOLUme
 ----------         ---------    ---------    ---------    ---------    ---------
DYNA>        10      0.01000      7.81721     16.21830     -8.40109    320.05568
DYNA PROP>          14.53782      7.92338     16.53681      0.10617    -73.15251
DYNA INTERN>         3.28801     14.20105      3.70007      3.31010      0.21480
DYNA EXTERN>         9.54825    -42.66337      0.00000      0.00000      0.00000
DYNA PRESS>          0.00000     48.76834      0.00000      0.00000      0.00000
 ----------         ---------    ---------    ---------    ---------    ---------
 UPDECI: Nonbond update at step        10

[...]
```



Newtonian dynamics in practice

In theory, Newtonian dynamics conserves the total energy (isolated system) :

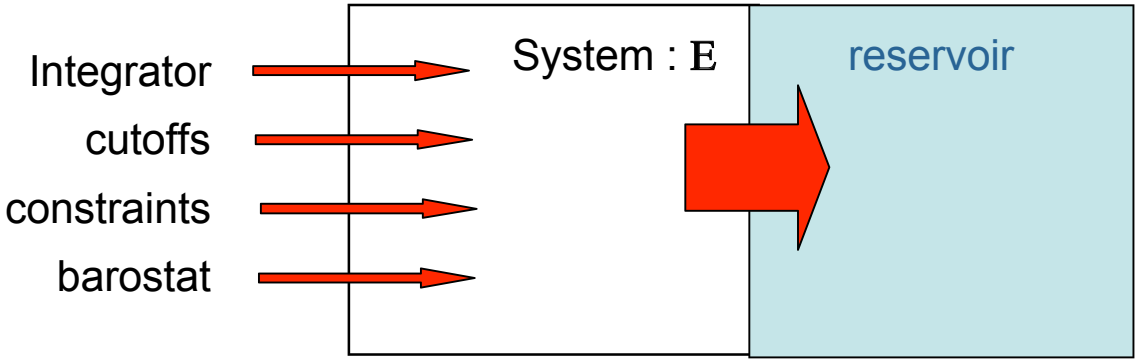
$$\dot{r} = \frac{p}{m}$$

$$\dot{p} = F(r)$$

$$H(r,p) = \sum_i \frac{p_i^2}{2m_i} + V(r) = cste$$

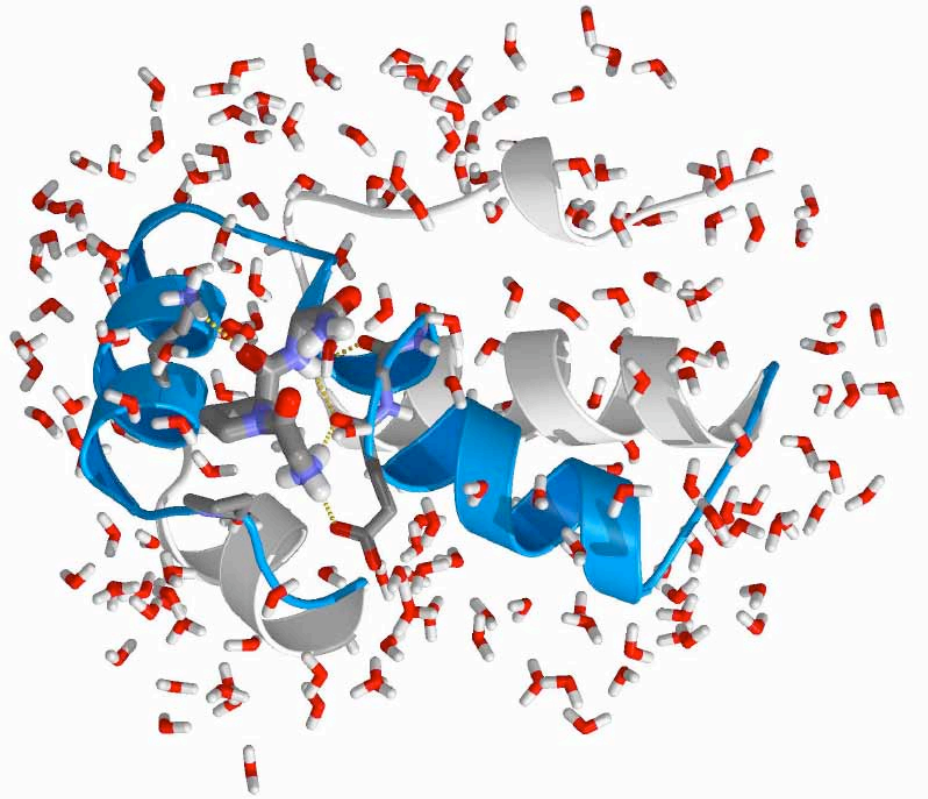
In practice, constant energy dynamics is not often used for two reasons :

1) Inaccuracies of the MD algorithm tend to heat up the system
We can couple the system to a heat reservoir to absorb the excess heat



2) The constant energy dynamics (NVE) does rearely represents the experimental conditions for the system simulated.

Plan



- Introduction
- The classical force field
- Setting up a simulation
- **Connection to statistical mechanics**
- Usage of MD simulation

Thermodynamic ensembles

A macroscopic state is described by :

- number of particles :	N	- chemical potential :	μ
- volume :	V	- pressure :	P
- energy :	E	- temperature :	T

Definition: a *thermodynamical ensemble* is a collection of microscopic states that all realize an identical macroscopic state

A microscopic state of the system is given by a point $(\boldsymbol{r}, \boldsymbol{p})$ of the phase space of the system, where $\boldsymbol{r} = (\boldsymbol{r}_1, ..., \boldsymbol{r}_N)$ and $\boldsymbol{p} = (\boldsymbol{p}_1, ..., \boldsymbol{p}_N)$ are the positions and the momenta of the N atoms of the system.

Examples of thermodynamical ensembles:

- Microcanonical:	fixed N, V, E	
- Canonical:	fixed N, V, T	often used in MD
- Constant P-T:	fixed N, P, T	often used in MD
- Grand Canonical:	fixed μ, P, T	

Bolzmann (canonical) distribution

Boltzmann showed that the *canonical* probability of the microstate *i* is given by

$$P_i = \frac{1}{Z} e^{-\beta E_i}$$

$\beta = 1/K_B T$
K_B = Boltzmann constant

Z is the *partition function*,

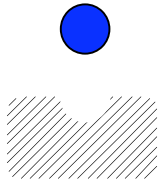
$$Z = \sum_j e^{-\beta E j}$$

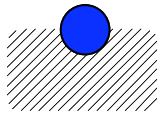
such that : $\sum_i P_i = 1$

The *partition function* is a very complex function to compute, because it represents a measure of the whole space accessible to the system.

Illustration:

If a system can have two unique states, state (1) and state (2), then the ratio of systems in state (1) and (2) is





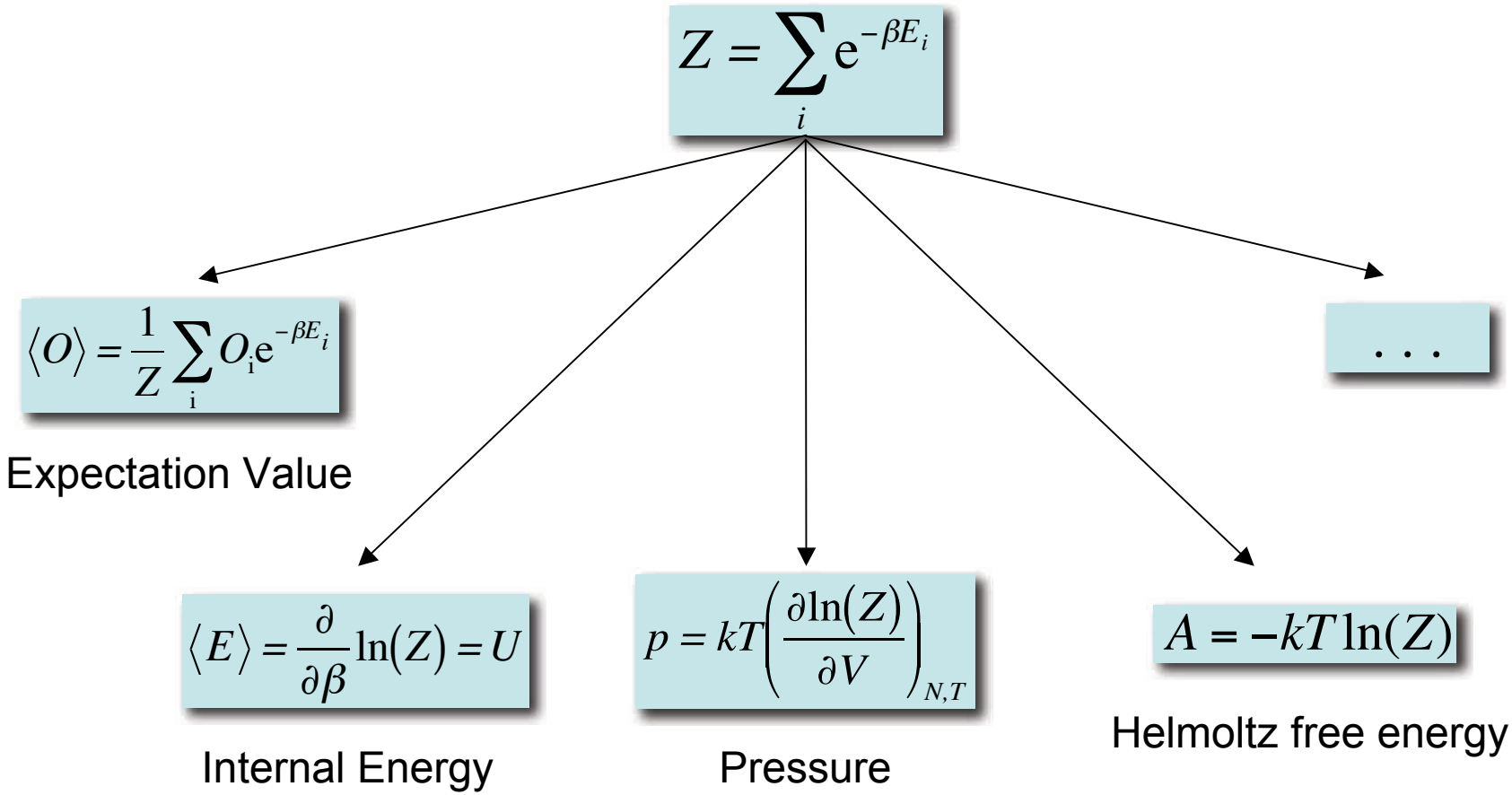
$$\frac{P_1}{P_2} = \frac{e^{-E_1}}{e^{-E_2}} = e^{-E_1 - E_2} = e^{-E}$$

at 300K, a ΔE of 1.3 kcal/mol results in a P1/P2 of 1 $\log_{10}$.

Cave: if state (1) and (2) are composed of several microscopic states, ΔE ≠ ΔG

The partition function

The determination of the macroscopic behavior of a system from a thermodynamical point of view is tantamount to computing the ***partition function***, *Z*, from which all the properties can be derived.



The ensemble average

Microscopic

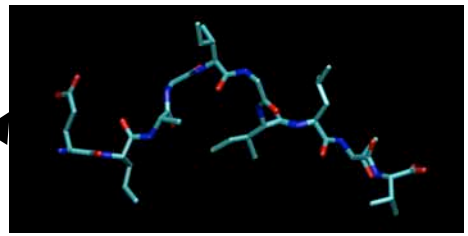
Macroscopic

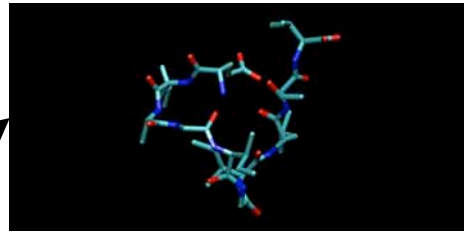
Expectation value

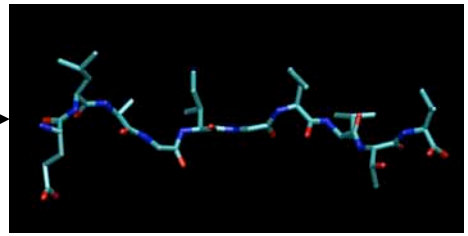
$$\langle O \rangle = \frac{1}{Z} \sum O_i e^{-\beta E_i}$$

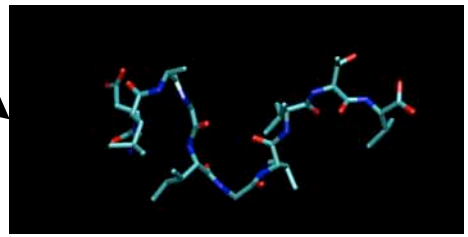
Where $Z = \sum e^{-\beta E_i}$

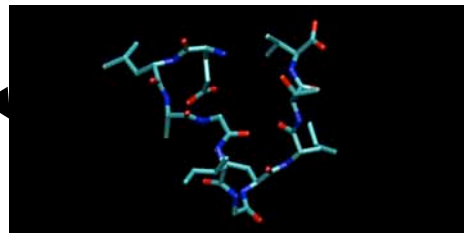
is the partition function











$E_1, P_1 \sim e^{-\beta E_1}$

$E_2, P_2 \sim e^{-\beta E_2}$

$E_3, P_3 \sim e^{-\beta E_3}$

$E_4, P_4 \sim e^{-\beta E_4}$

$E_5, P_5 \sim e^{-\beta E_5}$

Ergodic Hypothesis

The ***ergodic hypothesis*** is that the ensemble averages used to compute expectation values can be replaced by time averages over the simulation.

$$\langle O \rangle_{ensemble} \overset{\text{Ergodicity}}{=} \langle O \rangle_{time}$$

$$\frac{1}{Z}\int O(r,p)e^{-\beta E(r,p)}drdp = \frac{1}{\tau}\int_0^{\tau} O(t)dt$$

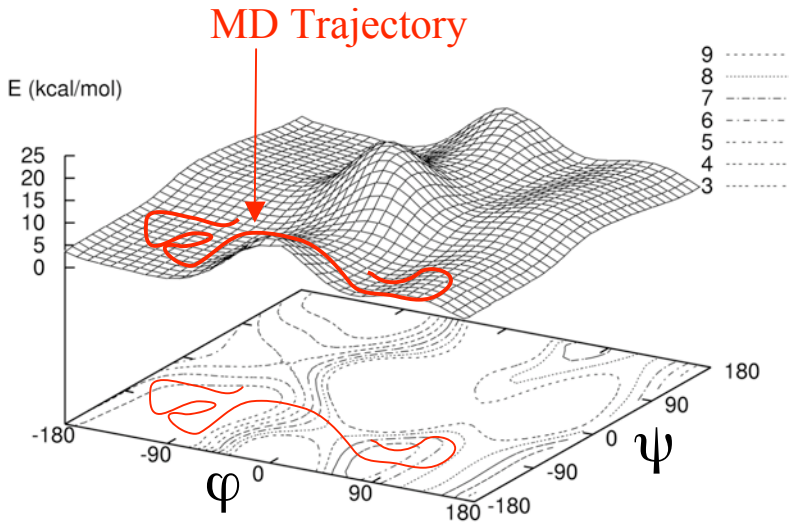
The microstates sampled by molecular dynamics are usually a small subset of the entire thermodynamical ensemble.

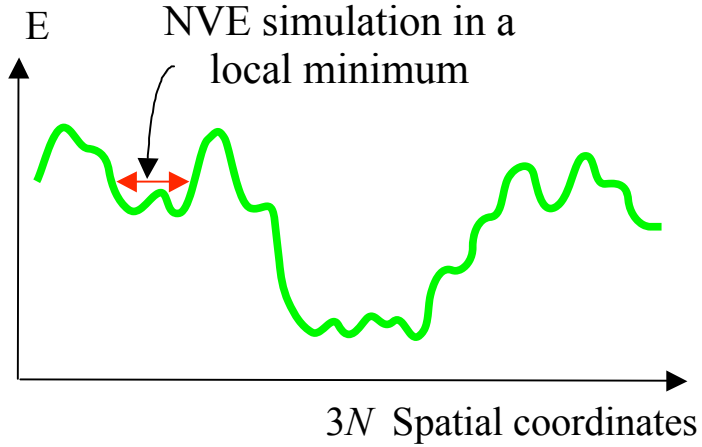
The validity of this hypothesis depends on the quality of the sampling produced by the molecular modelling technique. The sampling should reach all important minima and explore them with the correct probability,

- *NVE* simulations ➪ Microcanonic ensemble ➪ $P = cst.$
- *NVT* simulations ➪ Canonical ensemble ➪ $P(E) = e^{-\beta E}$
- *NPT* simulations ➪ Isothermic-isobaric ensemble ➪ $P(E) = e^{-\beta(E+PV)}$

Note that the Bolzmann weight $e^{-\beta E}$ is not present in the time average because it is assumed that conformations are sampled from the right probability.

Ergodic hypothesis : intuitive view





$$\langle O \rangle_{ensemble} = \frac{1}{Z} \int O(r,p) e^{-\beta E(r,p)} dr dp \quad \overset{?}{=} \quad \frac{1}{\tau} \int_0^{\tau} O(t) dt = \langle O \rangle_{time}$$

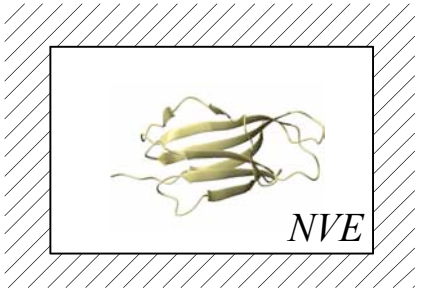
Two main requirements for MD simulation :

1) Accurate energy function *E(r,p)*

2) Appropriate algorithm, which
- generates the right ensemble
- samples efficiently

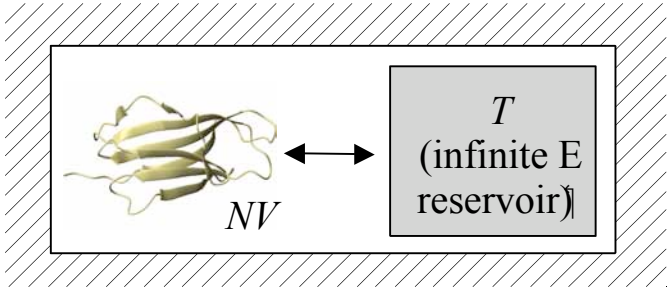
Sampling of the various ensembles

1) **Microcanonical ensemble** (constant N,V,E) sampling is obtained by simple integration of the Newtonian dynamics:
- Verlet, Leap-Frog, Velocity Verlet, Gear

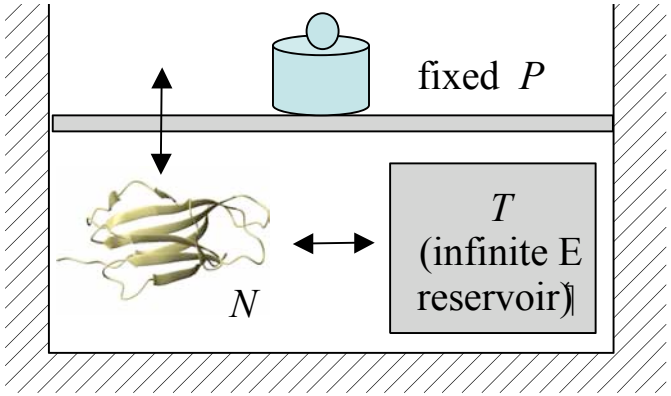


2) **Canonical ensemble** (constant N,V,T) sampling is obtained using thermostats :

1) ***Berendsen***: scaling of velocities to obtain an exponential relaxation of the temperature to T
2) ***Nose-Hoover***: additional degree of freedom coupled to the physical system acts as heat bath.



3) **Isothermic-isobaric ensemble** (constant N,P,T) In addition to the thermostat, the volume of the system is allowed to fluctuate, and is regulated by barostat algorithms.



The Nosé-Hoover thermostat

Phase space extended by two extra variables : (r, p, η, p_η)

physical variables

$$\dot{r}_i = \frac{p_i}{m_i}$$

$$\dot{p}_i = -\frac{\partial \Phi}{\partial r_i}(r,t) - \frac{p_\eta}{Q} p_i$$

$$\dot{\eta} = \frac{p_\eta}{Q}$$

$$\dot{p}_\eta = \sum_i \frac{p_i^2}{m_i} - N_{df} k_B T \ .$$

Newton

friction term

temperature regulation

- One can demonstrate that the canonical distribution is reproduced for the physical variables

$$Z(N,V,T) = \int dr dp \ e^{-\beta H(r,p)}$$

- Conserved quantity :

$$H'(\Gamma,t) = \sum_{i=1}^{N} \frac{p_i^2}{2m_i} + \Phi(r) + \frac{p_\eta^2}{2Q} + k_B T \bar{\eta}$$

- Non-Hamiltonian dynamics...

Other sampling methods I

Langevin Dynamics (LD)

In Langevin Dynamics, two additional forces are added to the standard force field:

- a ***friction*** force: $-\gamma_i \boldsymbol{p}_i$
  whose direction is opposed to the velocity of atom *i*

- a ***stochastic*** (random) force: $\zeta(t)$ such that $\langle\zeta(t)\rangle = 0$.

This leads to the following equation for the motion of atom *i*:

$$\dot{r}_i = \frac{p_i}{m_i} \qquad \dot{p}_i = F_i(r) + \gamma p_i + \zeta(t)$$

This equation can for example simulate the friction and stochastic effect of the solvent in implicit solvent simulations. The temperature is adjusted via γ and ζ, using the *dissipation-fluctuation* theorem.
The stochastic term can improve barrier crossing and hence sampling.

LD does ***not*** reproduce dynamical properties

Other sampling methods II

Monte Carlo Simulations and the Metropolis criterion

In this sampling method, instead of computing the forces on each atom to solve its time evolution, random movements are assigned to the system and the potential energy of the resulting conformer is evaluated.

To insure Boltzmann sampling, additional criteria need to be applied on the new conformer. Let C be the initial conformer and C' the randomly modified:

- if $V(C') < V(C)$, the new conformer is kept and C' becomes C for next step

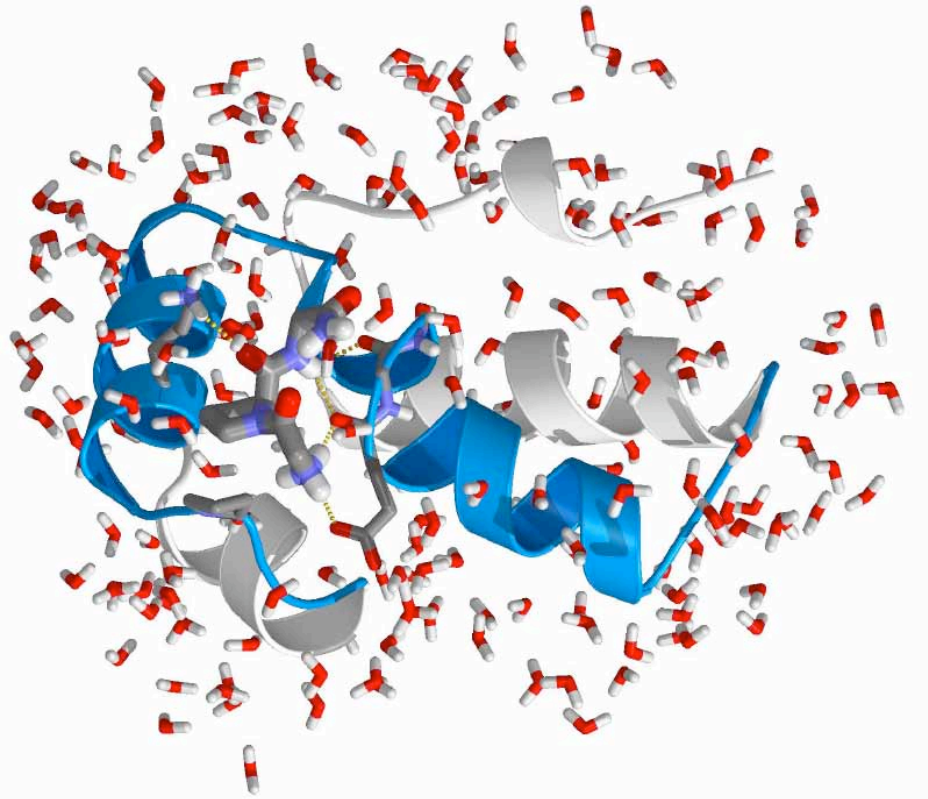
- if $V(C') > V(C)$, a random number, ρ, in the [0,1] interval is generated and

  if $e^{-\beta(V'-V)} > \rho$, the new conformer is kept and C' becomes C for next step

Using this algorithm, one insures Boltzmann statistics,

$$\frac{P(C')}{P(C)} = e^{-\beta(V'-V)}$$

Plan



- Introduction
- The classical force field
- Setting up a simulation
- Connection to statistical mechanics
- **Usage of MD simulation**

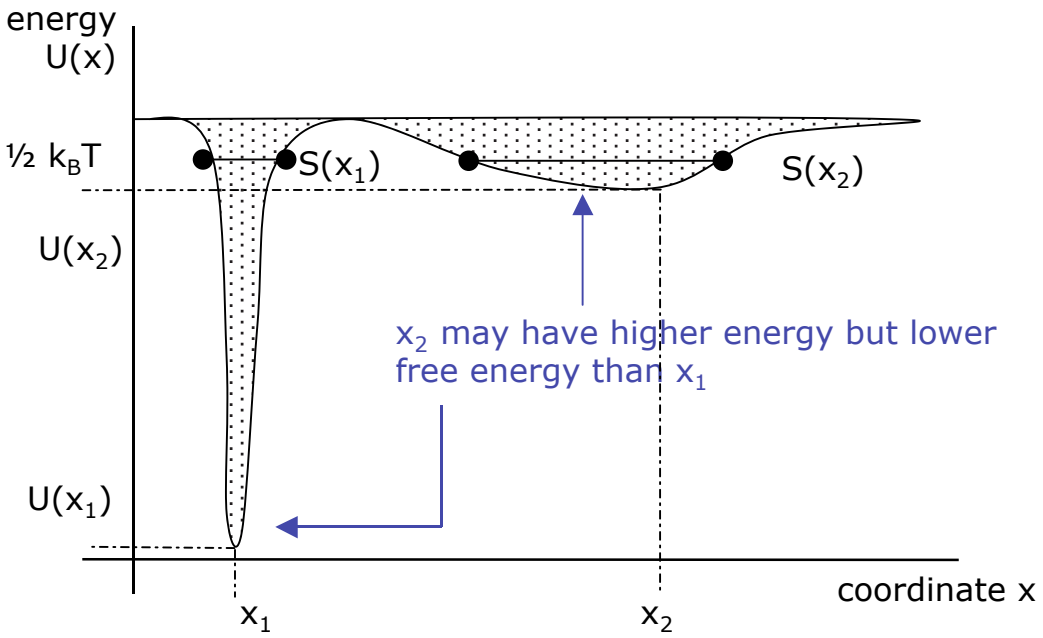
The importance of entropy

Mechanics: A state is characterised by ***one*** **minimum energy structure** (global minimum)

Statistical mechanics: A state is characterised by ***an ensemble*** **of structures**

Very small energy differences between states (~k_BT = 2.5 kJ/mol) resulting from summation over very many contributions

Entropic effects : Not only energy minima are of importance but whole range of x-values with energies ~k_BT



The free energy (F) governs the system

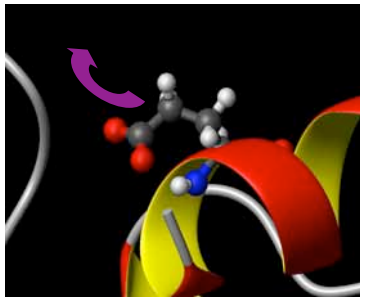
$F = U - TS$

Energy (U) – entropy (S) compensation at finite temperature T

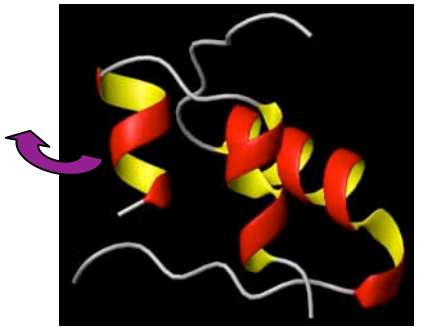
Dynamical behavior of proteins

Biological molecules exhibit a wide range of time scales over which specific processes occur; for example:

- Local Motions (0.01 to 5 Å, 10^{-15} to 10^{-1} s)
  - Atomic fluctuations
  - Sidechain Motions
  - Loop Motions



- Rigid Body Motions (1 to 10 Å, 10^{-9} to 1 s)
  - Helix Motions
  - Domain Motions
  - Subunit motions



- Large-Scale Motions (> 5 Å, 10^{-7} to 10^{4} s)
  - Helix coil transitions
  - Dissociation/Association
  - Folding and Unfolding

http://www.ch.embnet.org/MD_tutorial/

Types of problems

Molecular dynamics simulations permit the study of complex, dynamic processes that occur in biological systems. These include, for example:

- Protein stability
- Conformational changes
- Protein folding
- Molecular recognition: proteins, DNA, membranes, complexes
- Ion transport in biological systems

and provide the mean to carry out the following studies,

- Drug Design
- Structure determination: X-ray and NMR

http://www.ch.embnet.org/MD_tutorial/

Historical perspective

Theoretical milestones:

Newton (1643-1727):	Classical equations of motion: $F(t)=m\,a(t)$
Boltzmann(1844-1906):	Foundations of statistical mechanics
Schrödinger (1887-1961):	Quantum mechanical eq. of motion: $-ih\,\partial t\,\Psi(t)=H(t)\,\Psi(t)$

Molecular mechanics milestones:

Liquids

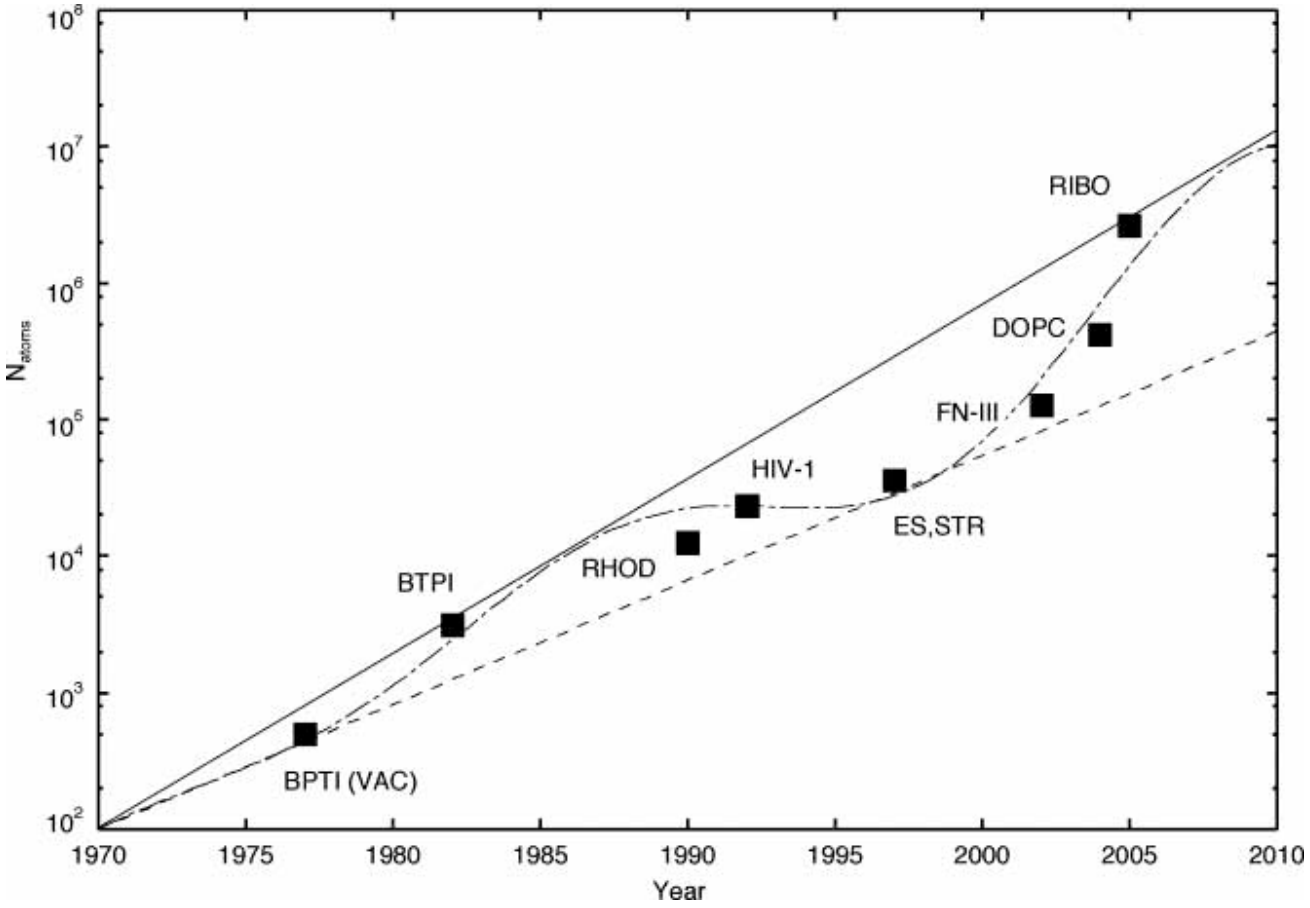
Metropolis (1953):	First Monte Carlo (MC) simulation of a liquid (hard spheres)
Wood (1957):	First MC simulation with Lennard-Jones potential
Alder (1957):	First Molecular Dynamics (MD) simulation of a liquid (hard spheres)
Rahman (1964):	First MD simulation with Lennard-Jones potential

Proteins

Karplus (1977) & McCammon (1977)	First MD simulation of proteins
Karplus (1983):	The CHARMM general purpose FF & MD program
Kollman(1984):	The AMBER general purpose FF & MD program
Car-Parrinello(1985):	First full QM simulations
Kollmann(1986):	First QM-MM simulations

System sizes

Simulations in explicit solvent



BPTI (VAC), bovine pancreatic trypsin inhibitor without solvent

BPTI, bovine pancreatic trypsin inhibitor with solvent

RHOD, photosynthetic reaction center of Rhodopseudomonas viridis

HIV-1, HIV-1 protease

ES, estrogen–DNA

STR, streptavidin

DOPC, DOPC lipid bilayer
RIBO, ribosome

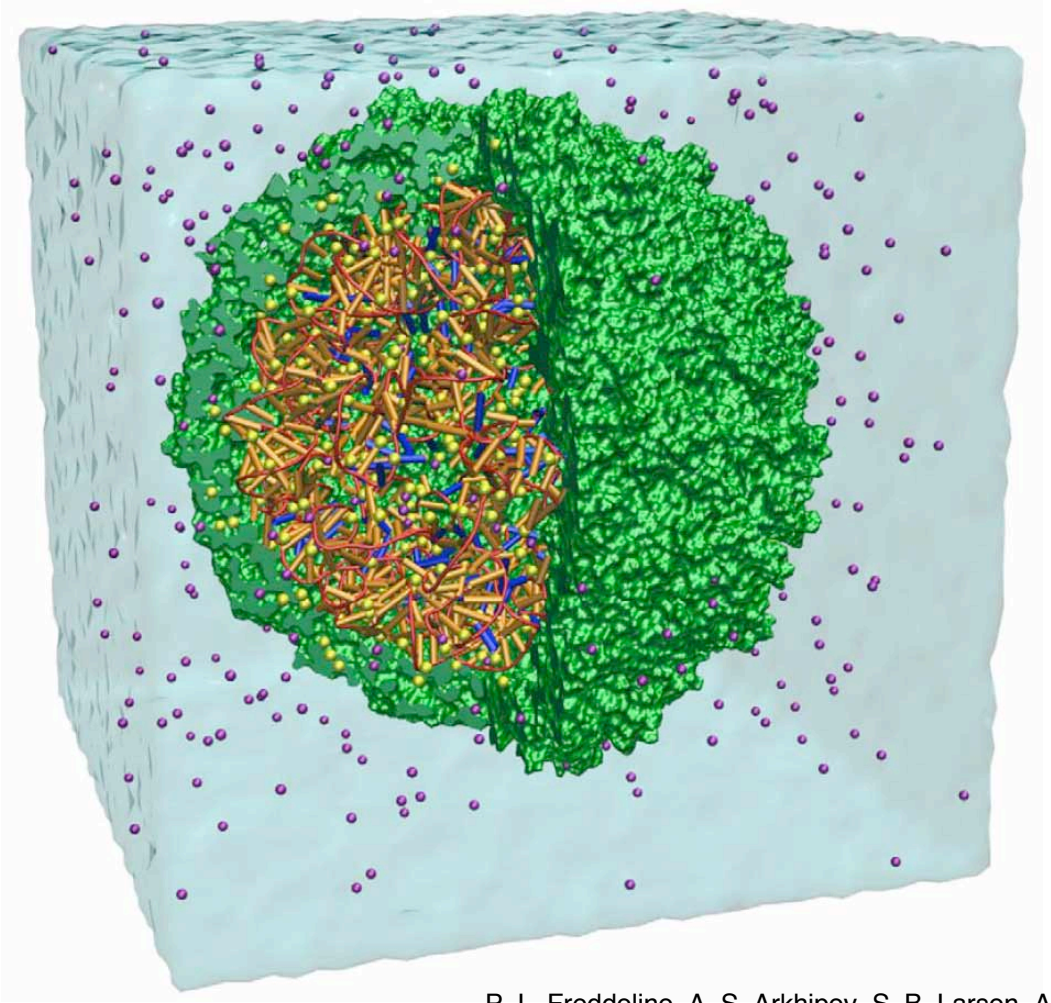
Solid curve, Moore's law doubling every 28.2 months.

Dashed curve, Moore's law doubling every 39.6 months.

K.Y. Sanbonmatsu, C.-S. Tung, Journal of Structural Biology 157(3) : 470, 2007

Largest system 2006

satellite tobacco mosaic virus



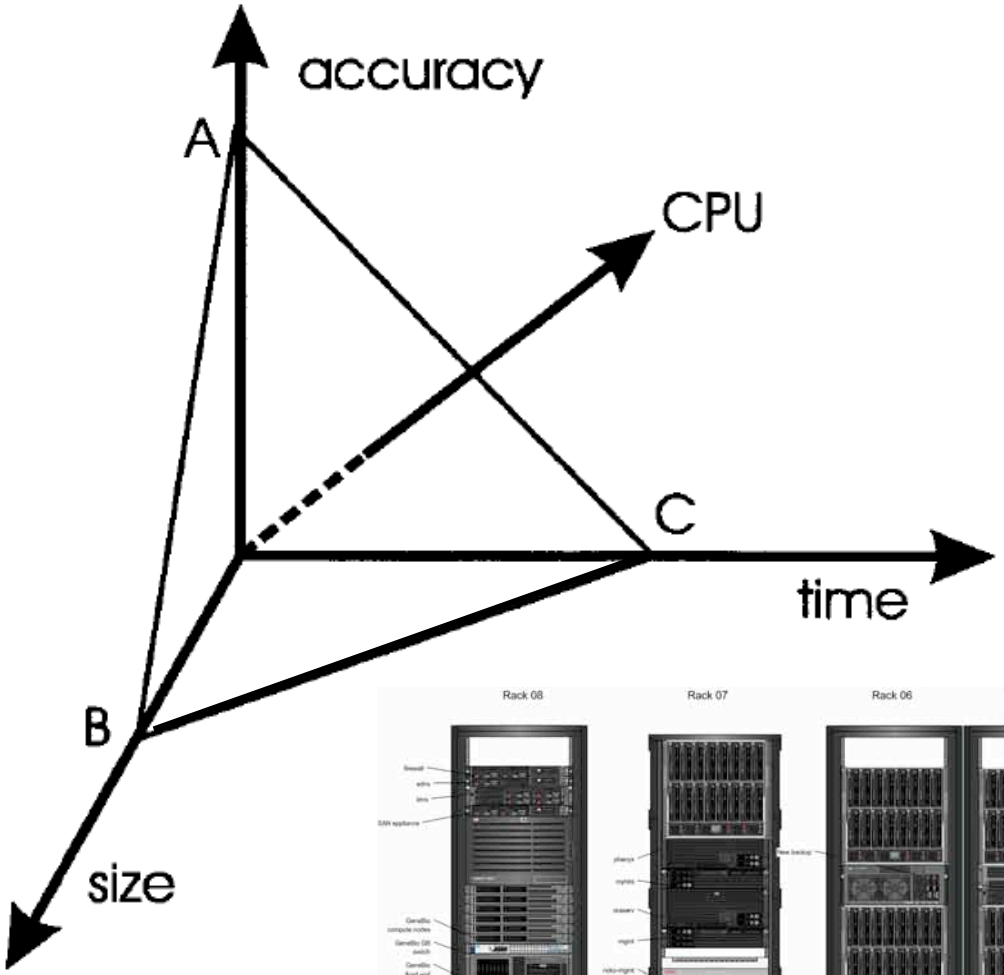
1 million atoms

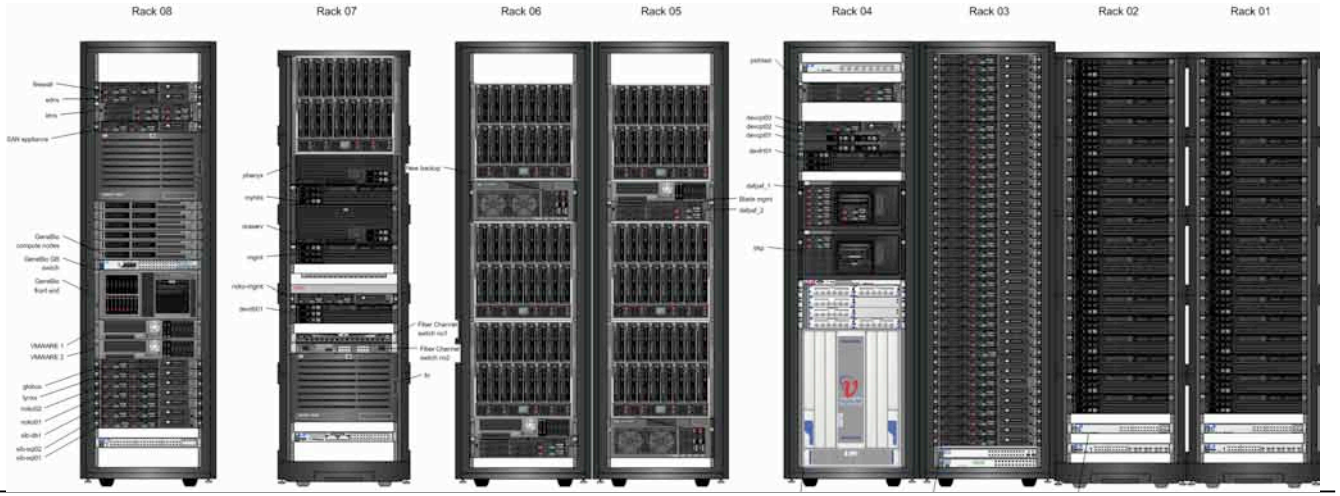
Simulation time : 50 ns

system size : 220 A

P. L. Freddolino, A. S. Arkhipov, S. B. Larson, A. McPherson, and K. Schulten, Molecular dynamics simulations of the complete satellite tobacco mosaic virus, Structure 14 (2006), 437.

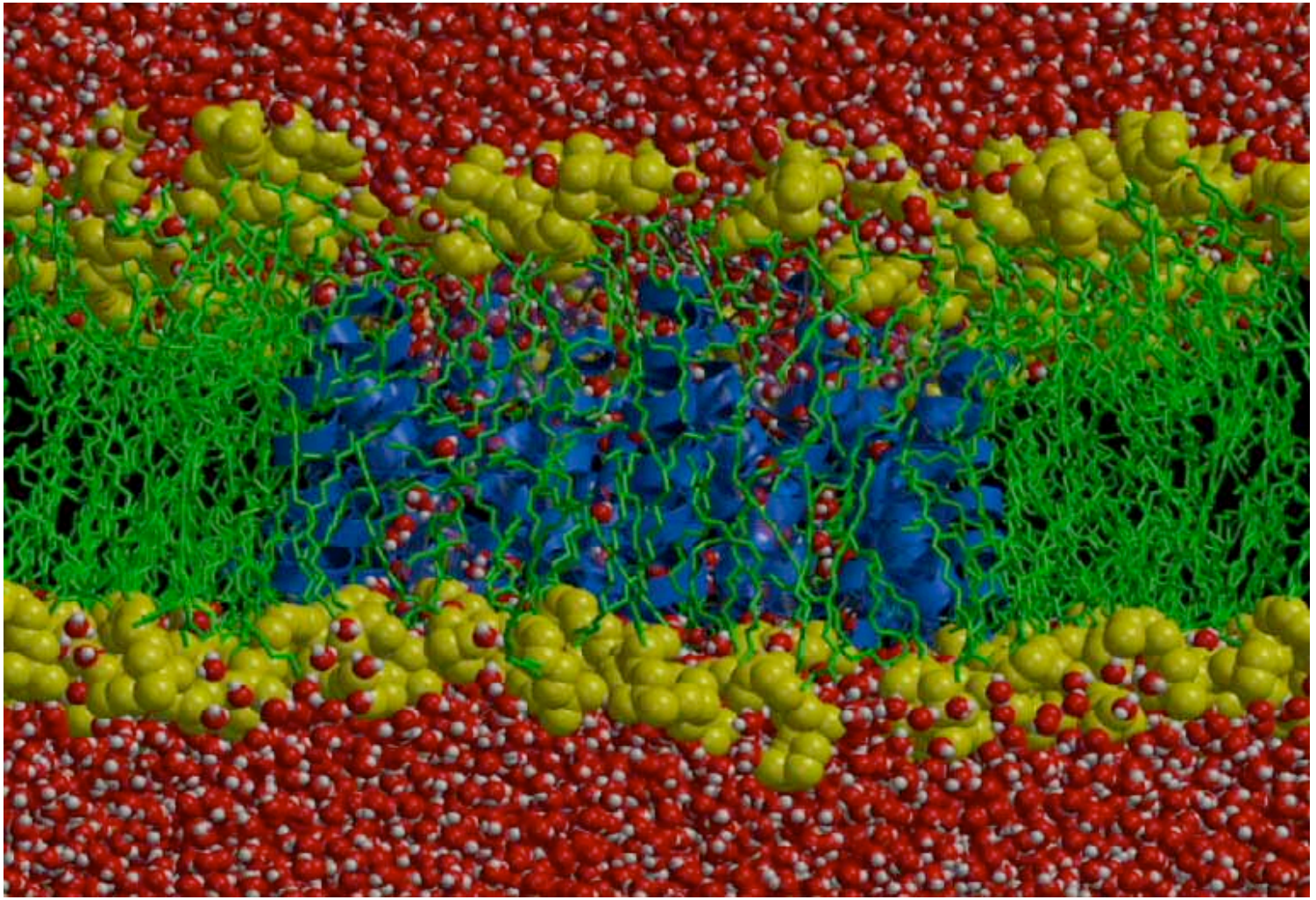
The tradeoff we can afford





Exemple : Aquaporin

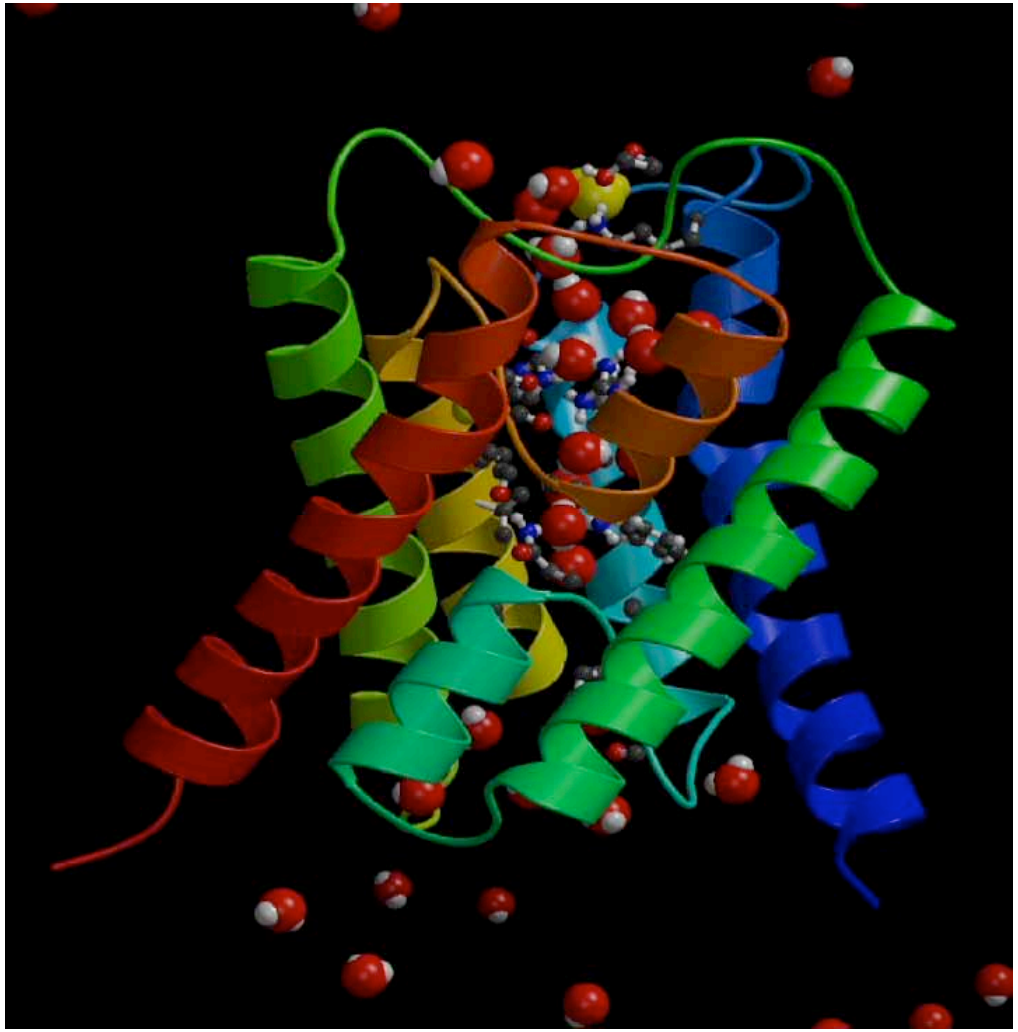
Selective translocation of water across a membrane



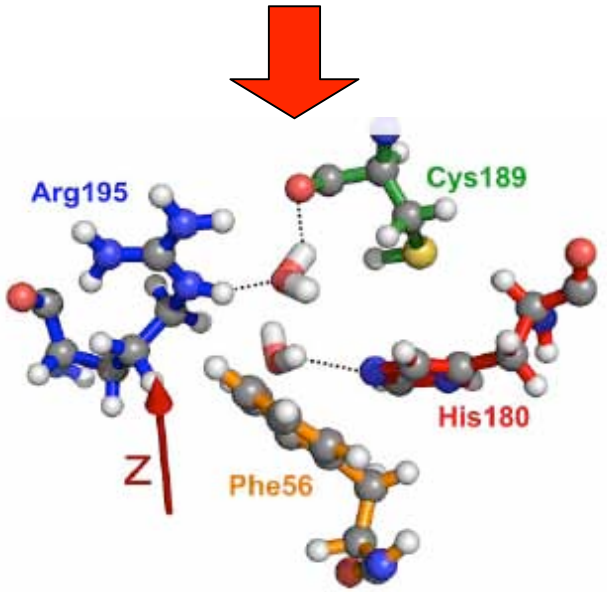
S. Hub and Bert L. de Groot. Mechanism of selectivity in aquaporins and aquaglyceroporins PNAS. 105:1198-1203 (2008)

Exemple : Aquaporin

Selective translocation of water across a membrane



Propose a model for selectivity at the atomic level



S. Hub and Bert L. de Groot. Mechanism of selectivity in aquaporins and aquaglyceroporins, PNAS 105:1198-1203 (2008)