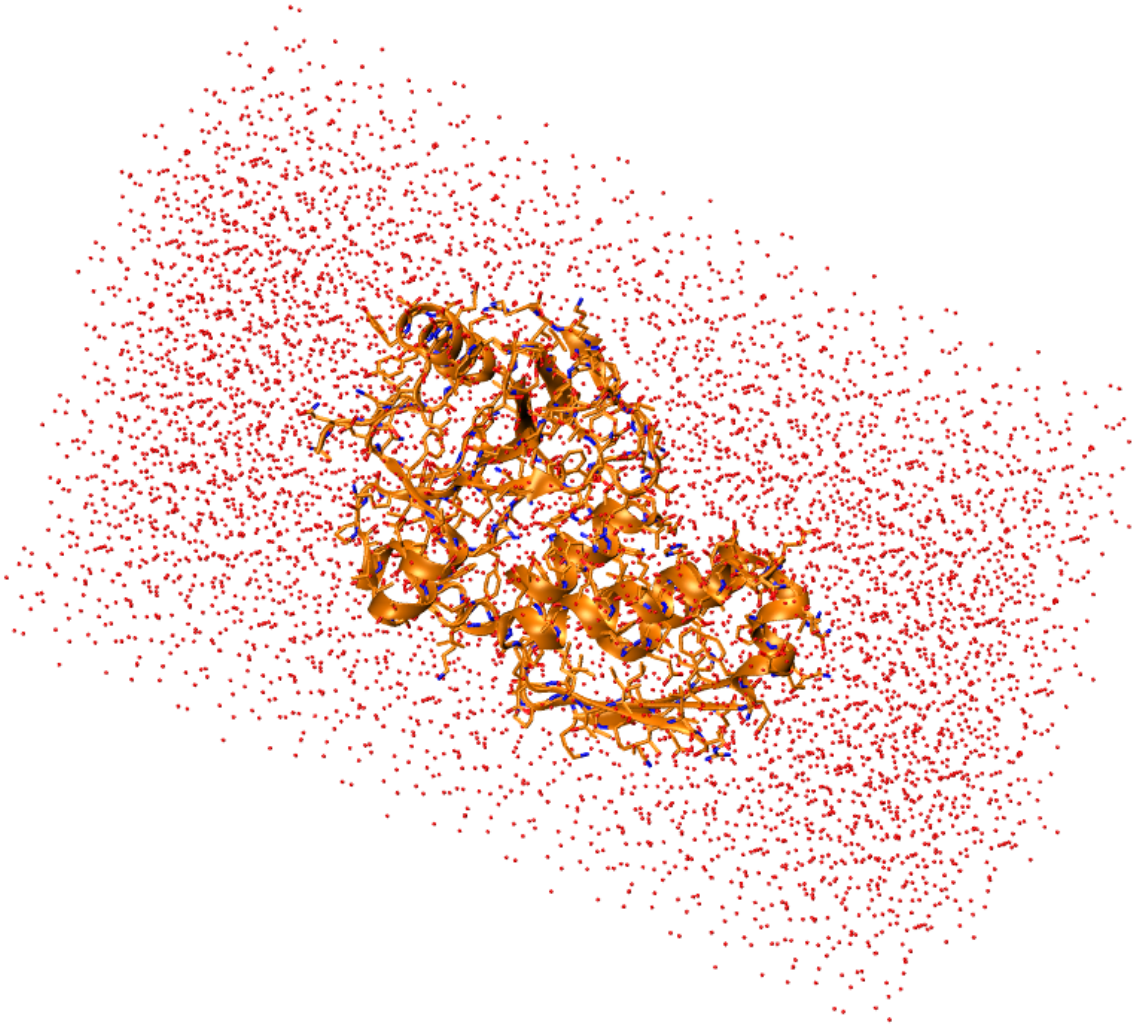
Wintersemester 2016/2017
Biomolecular Engineering/Nanobiophysics Module

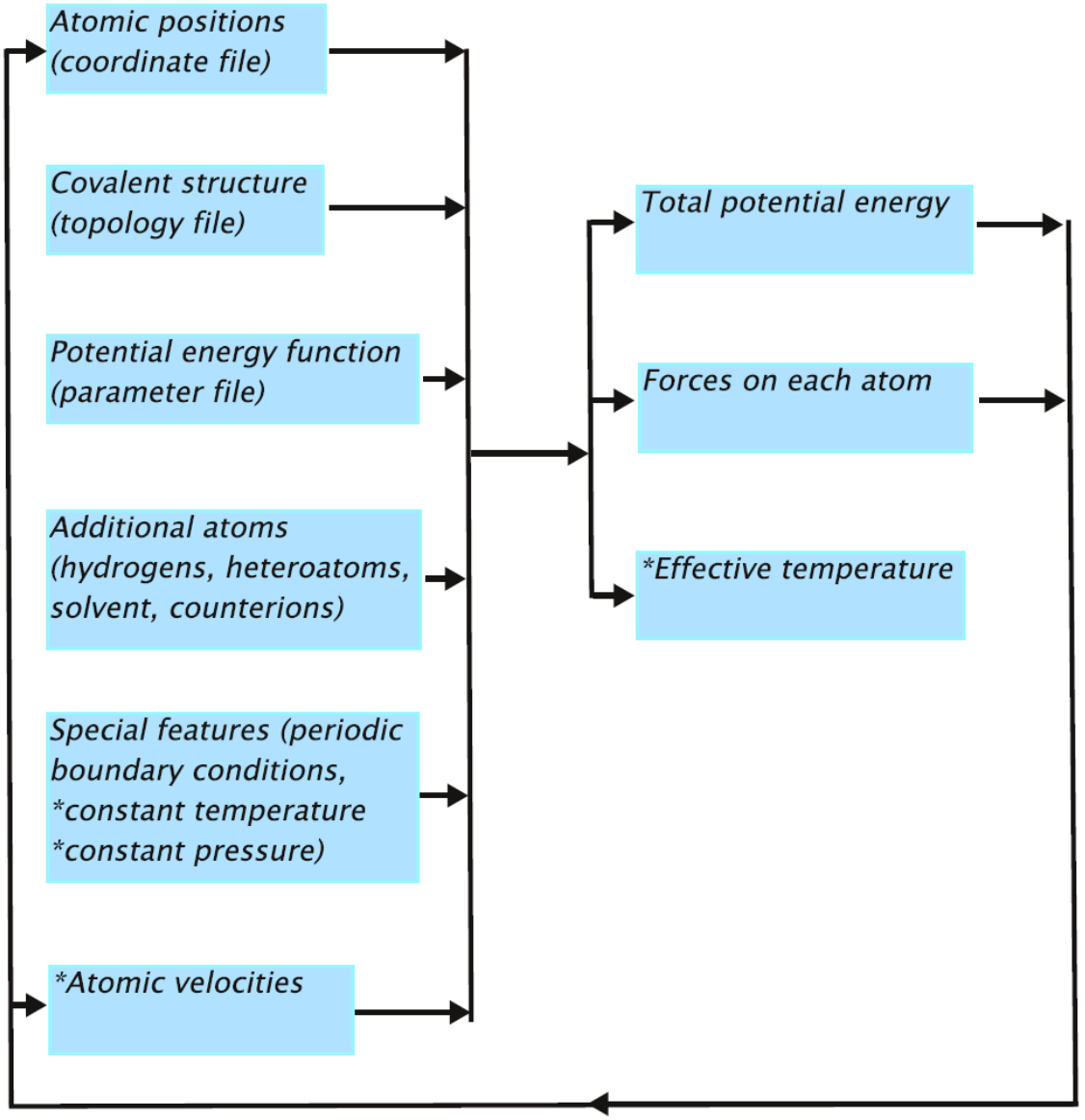
LECTURE 2: MOLECULAR DYNAMICS



LECTURE 2: OUTLINE
MOLECULAR DYNAMICS

- MD pipeline and principles
- Minimization
- MD with fixed physico-chemical parameters
- Energy calculations
- Biased MD
- Special cases of MD
- MD software
- Case study: fluorinated amino acids

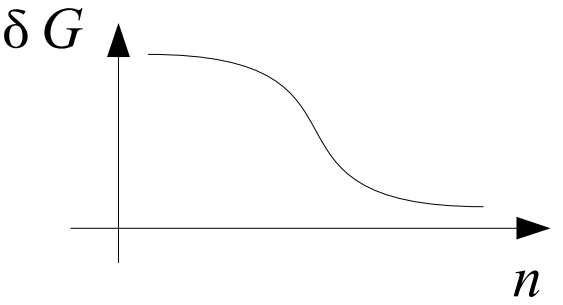
MOLECULAR DYNAMICS PIPELINE



$$V(\vec{r}) = \sum_{bonds} K_r (r - r_{eq})^2 + \sum_{angles} K_\theta (\theta - \theta_{eq})^2 + \sum_{dihedrals} \frac{V_n}{2} (1 + \cos[n\phi - \gamma]) + \sum_{i<j}^{atoms} \left(\frac{A_{ij}}{R_{ij}^{12}} - \frac{B_{ij}}{R_{ij}^{6}} \right) + \sum_{i<j}^{atoms} \frac{q_i q_j}{\epsilon R_{ij}}$$

MINIMIZATION

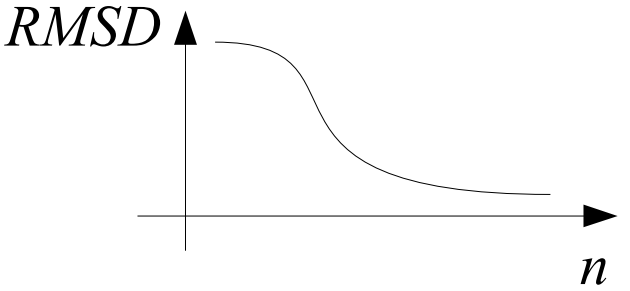
- **ΔG of the system decreases iteratively**



- **Criteria:** $\left|\delta G_n - \delta G_{n-1}\right| < \epsilon_1$

$$RMSD_{n,n-1} < \epsilon_2$$

$$RMSD = \sum_{i=1}^{N} \sqrt{(\vec{r}_n - \vec{r}_{ref})^2}$$



- **Algorithms:**

  - **Steepest descent** $\vec{r}_n = \vec{r}_{n-1} + \lambda_n \frac{\vec{F}_n}{F_n} ; \lambda_n > 0$

  - **Conjugate gradient** $\vec{r}_n = \vec{r}_{n-1} + \lambda_n (\vec{F}_n + \frac{F_n^2}{F_{n-1}^2} \frac{\vec{F}_{n-1}}{F_{n-1}})$

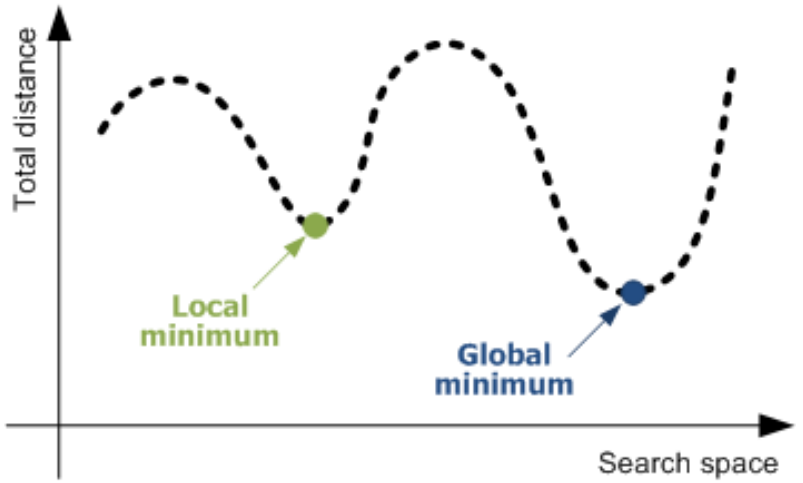
  - **Their combination (first SD, then CG)**

MINIMIZATION vs MD

$$V(\vec{r}) = \sum_{bonds} K_r (r - r_{eq})^2 + \sum_{angles} K_\theta (\theta - \theta_{eq})^2 + \sum_{dihedrals} \frac{V_n}{2}(1 + \cos[n\varphi - \gamma]) + \sum_{i<j}^{atoms} \left(\frac{A_{ij}}{R_{ij}^{12}} - \frac{B_{ij}}{R_{ij}^{6}}\right) + \sum_{i<j}^{atoms} \frac{q_i q_j}{\epsilon R_{ij}}$$

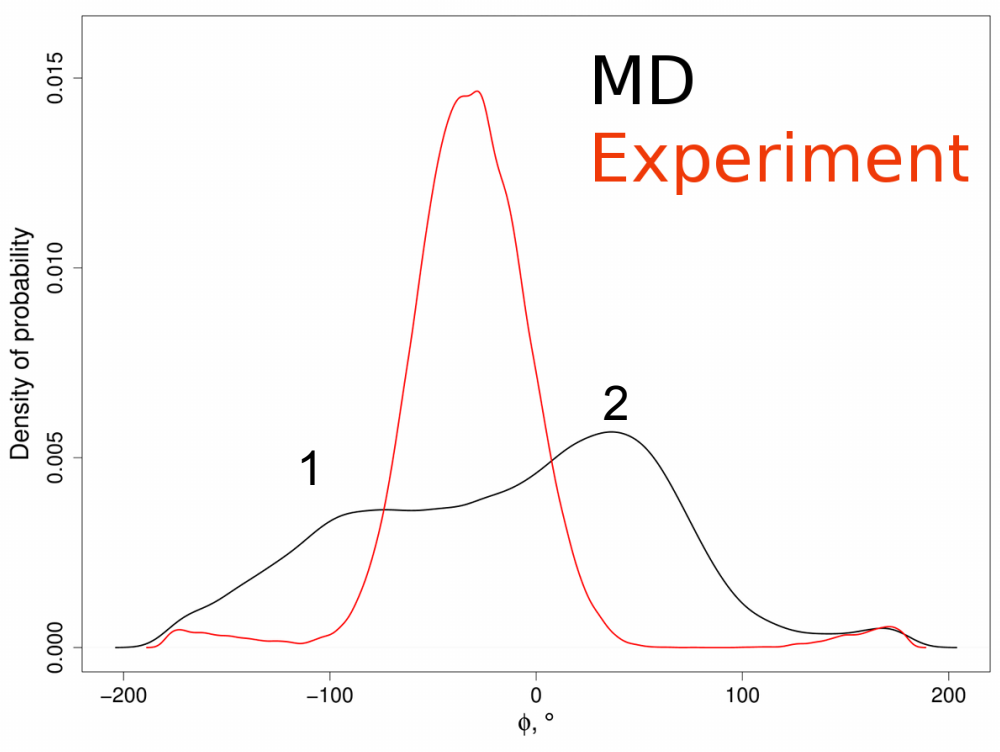
$V(\vec{r}) \neq f(t)$ – *minimization*

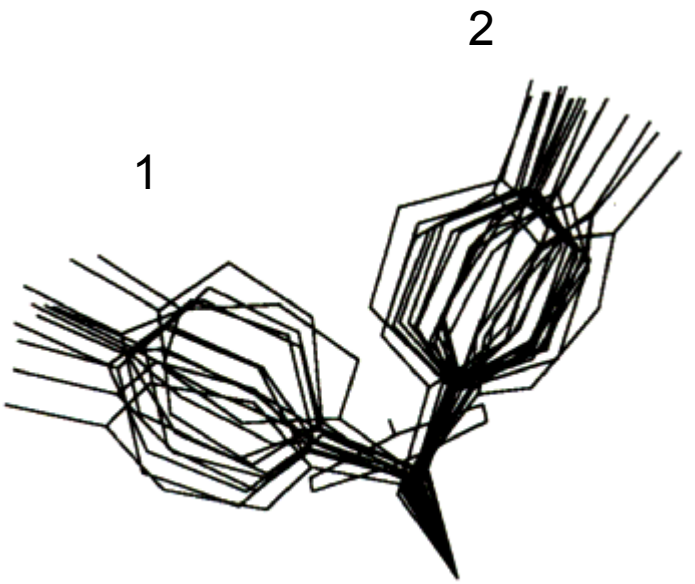
$V(\vec{r}) = f(t)$ – *molecular dynamics*



- **Minimization: one minimum**
- **MD: *potentially* several minima (sampling challenge!)**
- **Experiment: mixture**

MD vs EXPERIMENT





- **Two rotamers in X-ray:**

  - High resolution (TYRA, TYRB records in the PDB)

  - Low resolution: average structure (TYR)

- **For some properties MD resolution could be higher then experimental**

MD: VERLET INTEGRATION

- ***Verlet integration*** **is a numerical method frequently used to integrate Newton's equations of motion:**

$$\vec{r}(t+\delta t)=\vec{r}(t)+\vec{v}(t)\delta t+\vec{a}(t)\frac{\delta t^2}{2}+\vec{b}(t)\frac{\delta t^3}{6}+O(\delta t^4)$$

$$\vec{r}(t-\delta t)=\vec{r}(t)-\vec{v}(t)\delta t+\vec{a}(t)\frac{\delta t^2}{2}-\vec{b}(t)\frac{\delta t^3}{6}+O(\delta t^4)$$

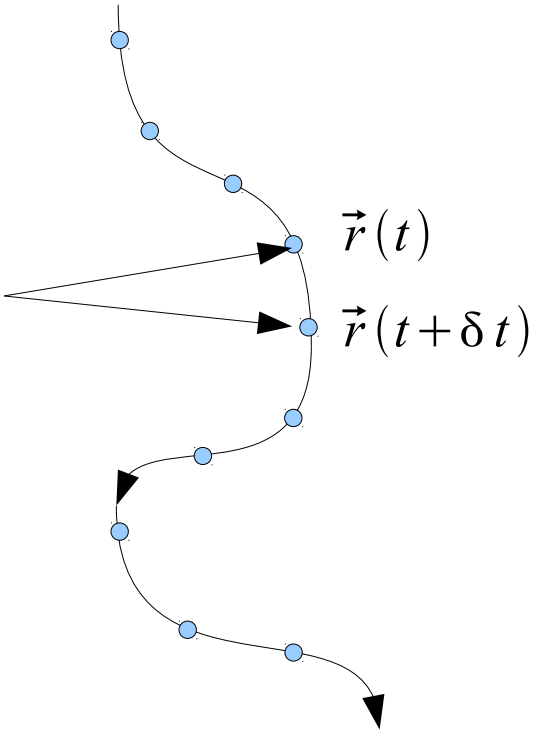
$$\vec{r}(t+\delta t)=2\vec{r}(t)-\vec{r}(t-\delta t)+\vec{a}(t)\delta t^2+O(\delta t^4)$$

$$\vec{a}(t)=\frac{1}{m}\vec{F}(\vec{r}(t))$$

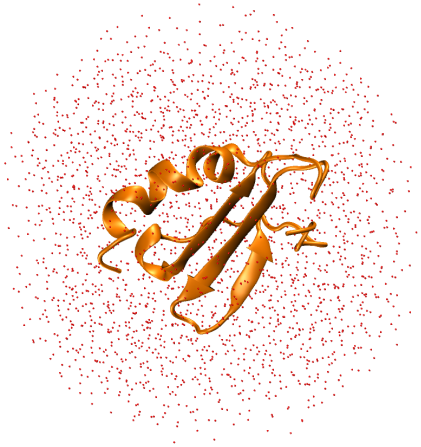
$$\vec{r}(t+\delta t)=2\vec{r}(t)-\vec{r}(t-\delta t)+\frac{1}{m}\vec{F}(\vec{r}(t))\delta t^2+O(\delta t^4)$$

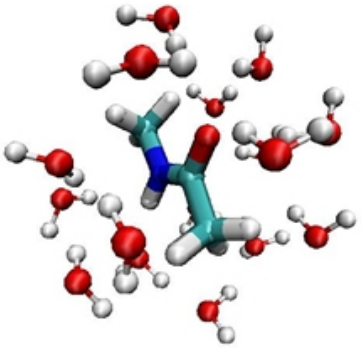
$$\vec{v}(t)=\frac{\vec{r}(t+\delta t)-r(t+\delta t)}{2\delta t}$$

$$\vec{r}(t+\delta t)=\vec{r}(t)+\vec{v}(t+\frac{1}{2}\delta t)\delta t\,; \qquad \vec{v}(t+\frac{1}{2}\delta t)=\vec{v}(t-\frac{1}{2}\delta t)+\vec{a}(t)\delta t$$



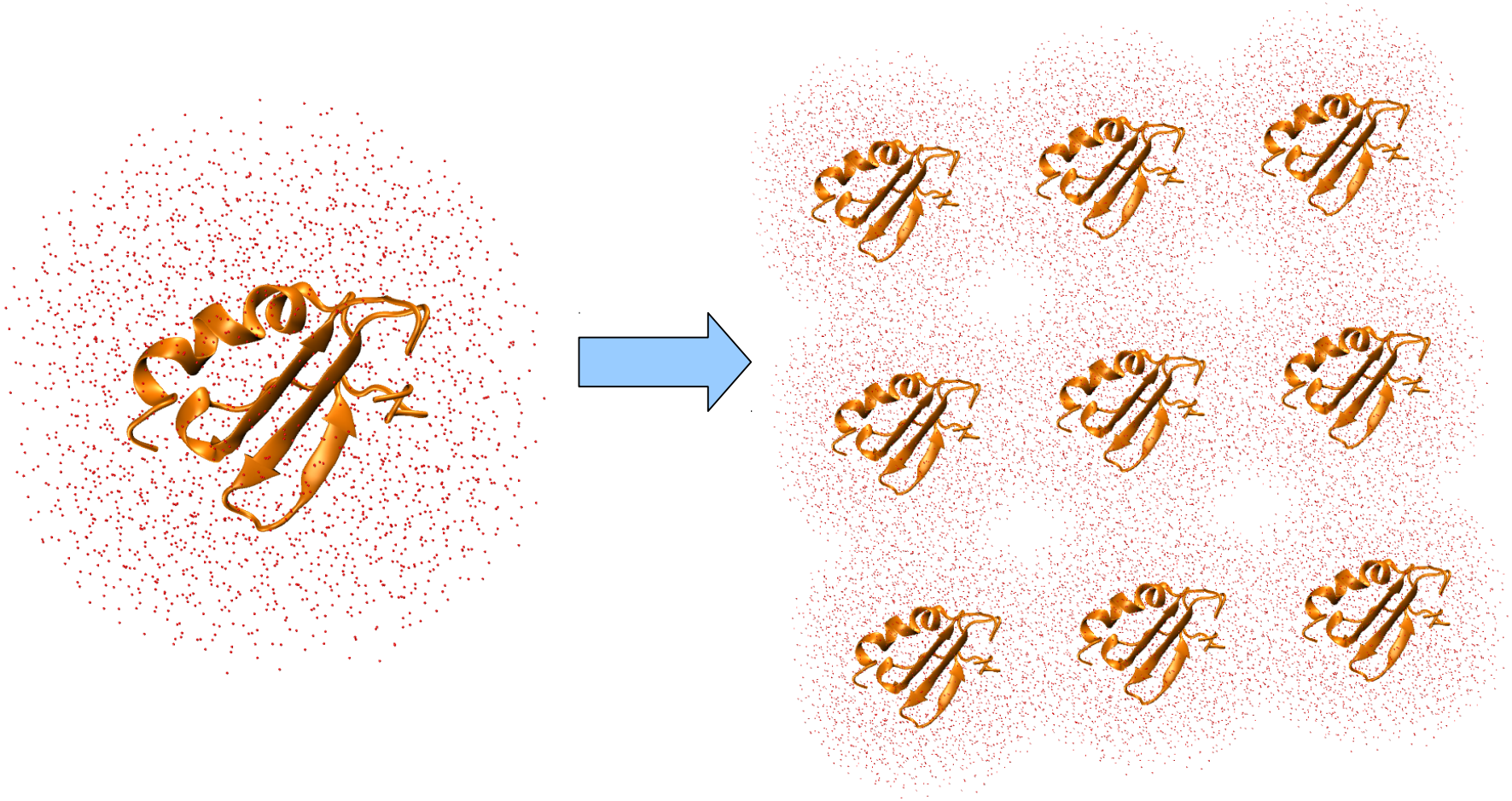
SOLVENT IN MD





- *In vacuo* – no solvent
- Implicit solvent – continuous solvent with averaged macroscopic properties
- Explicit solvent – each solvent molecule is given explicitly

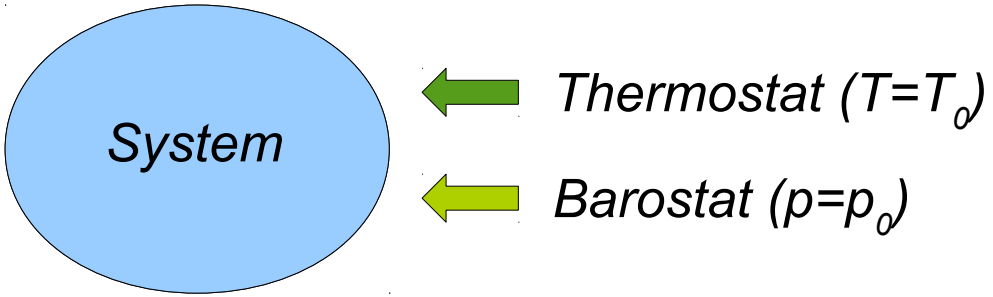
PERIODIC BOUNDARY CONDITIONS



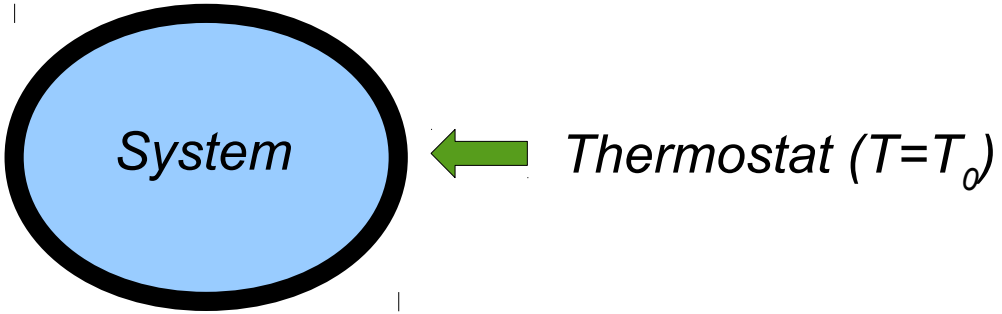
- PBC: each cell unit is repeated in space infinite number of times to provide space continuity

STATISTICAL ENSEMBLES IN MD

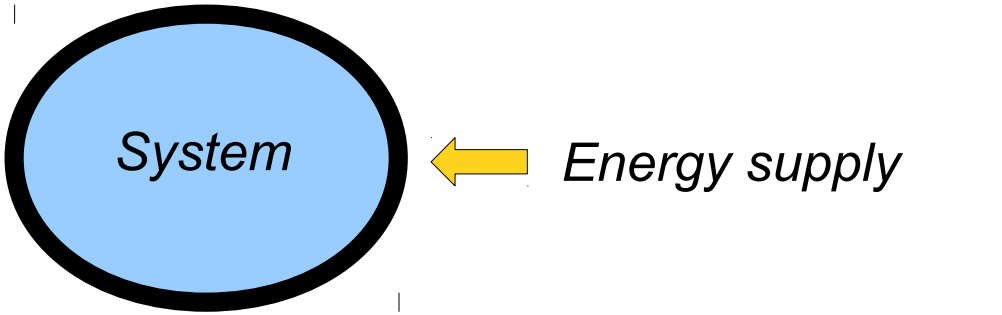
- **NTP: constant number of particles, temperature, pressure**



- **NTV: constant number of particles, temperature, volume**



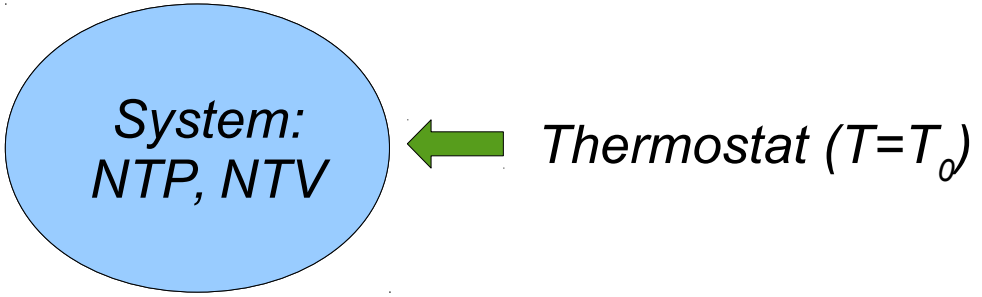
- **NVE: constant number of particles, volume, full energy of the system**



TEMPERATURE COUPLING

- **Weak (Berendsen) coupling**

$$\frac{dT}{dt} = -\alpha(T - T_0)\,;\, \alpha > 0$$

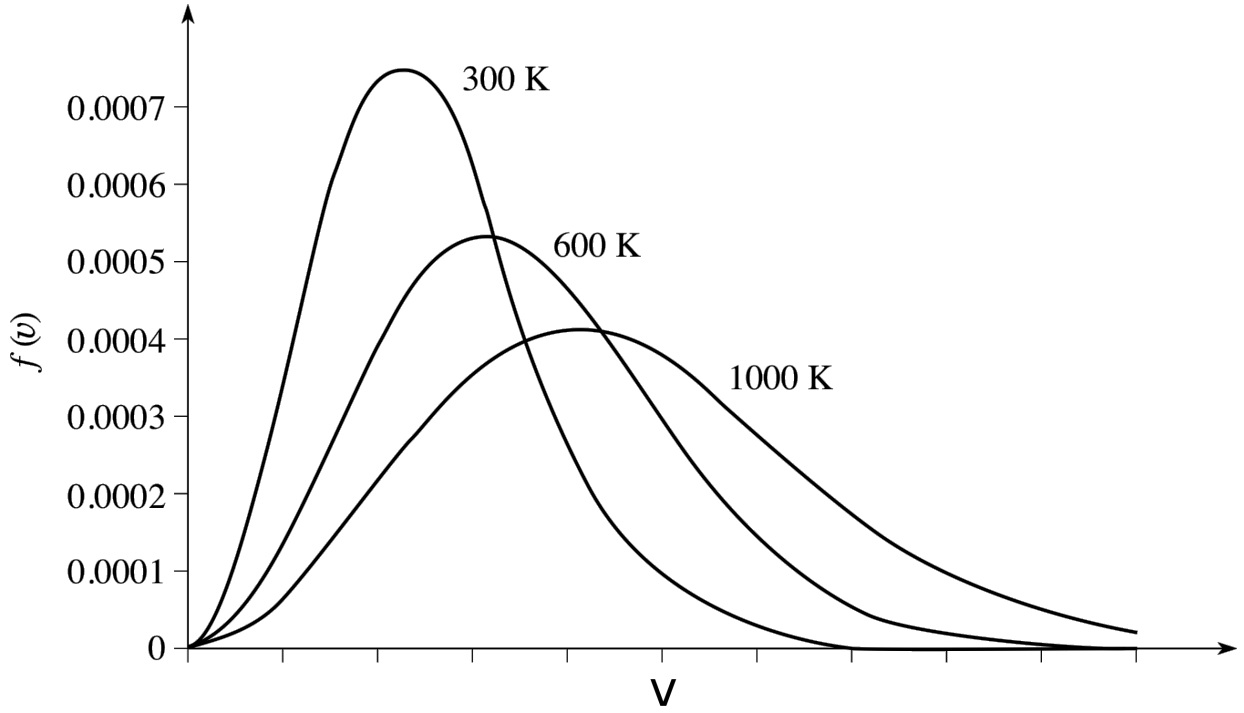


- **Andersen coupling**

Imaginary collisions randomize velocities

- **Langevin coupling**

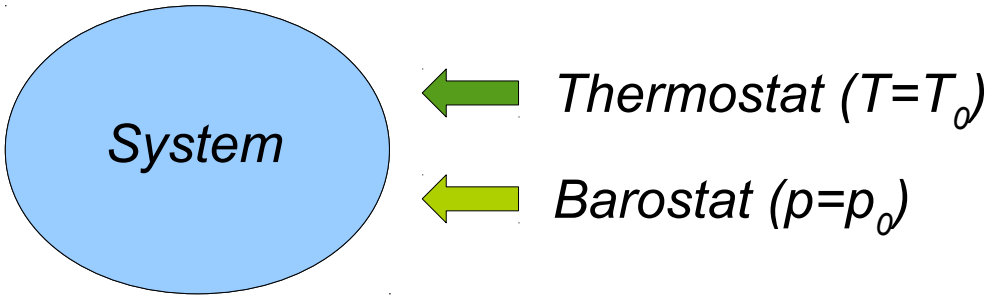
More frequent than physical collisions randomize velocities from M-B distribution



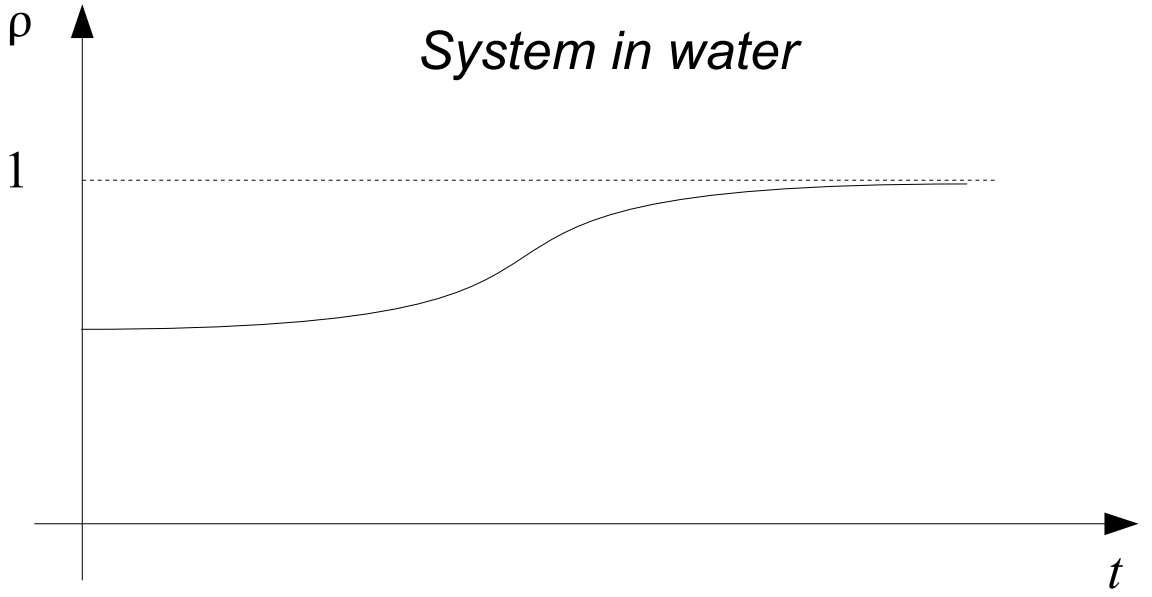
PRESSURE COUPLING

➢ Berendsen coupling

$$\frac{dp}{dt} = -\alpha(p - p_0); \alpha > 0$$

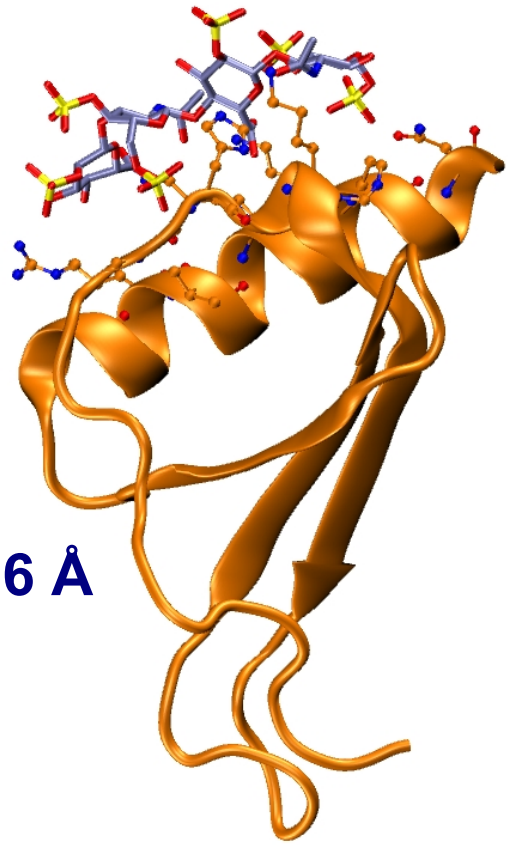


➢ Density convergence

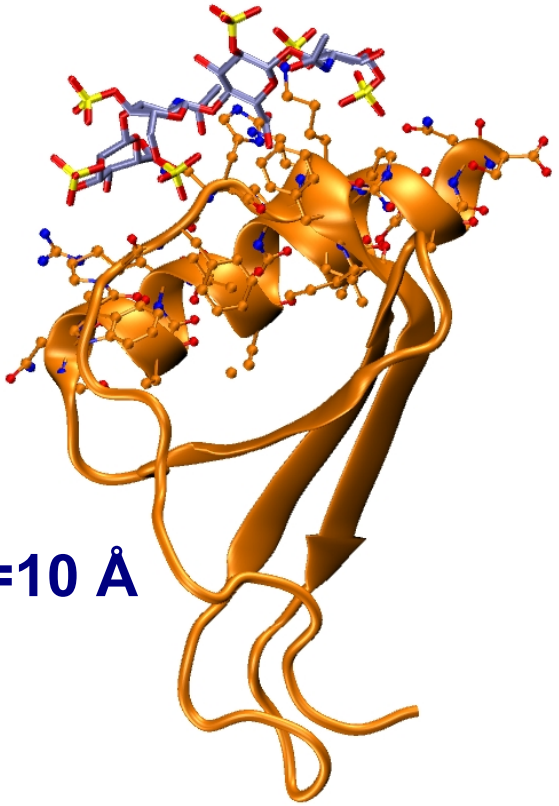


CUT-OFF

cut_off=6 Å



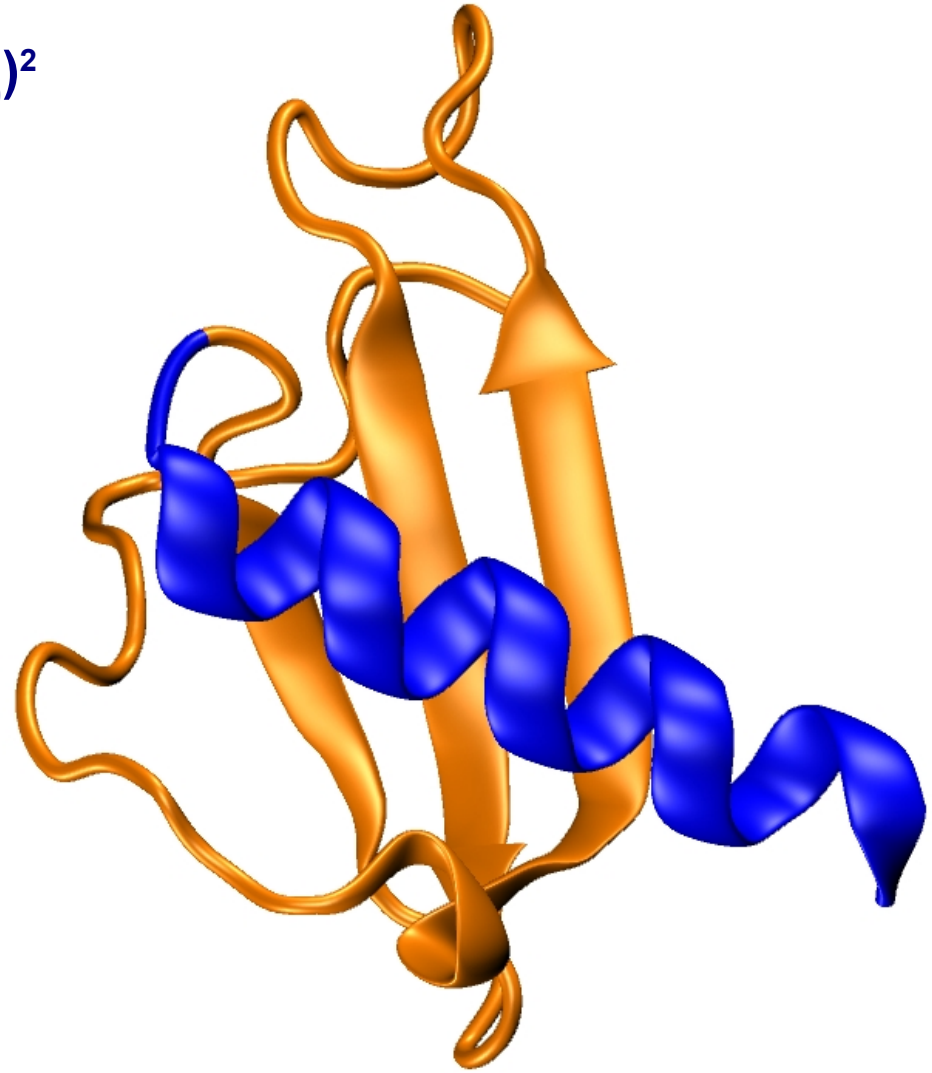
cut_off=10 Å



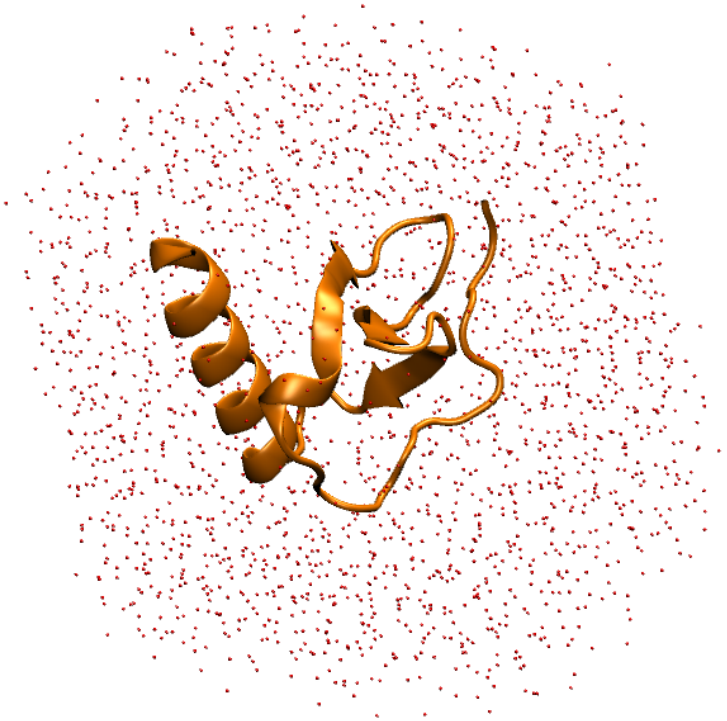
Number of atomic pairs ~ $(cut\_off)^6$

RESTRAINTS IN MD

- Harmonic constraints: $F \sim (X - X_0)^2$
  - Bonds
  - Angles
  - Dihedrals
  - Distances between atoms
  - CM distances
- Freezing (fixing atoms)

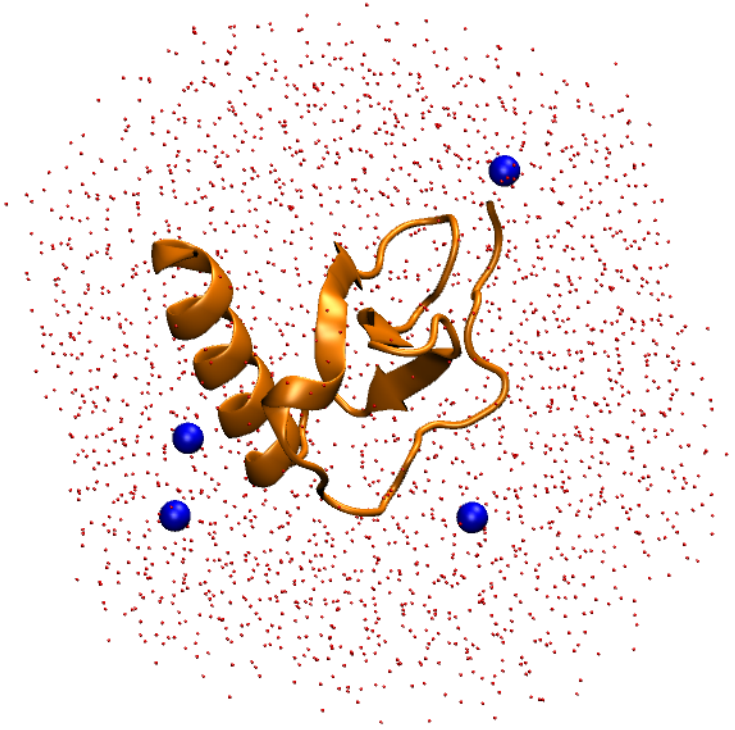


COUNTERIONS



q≠0

Electrostatics wrong!



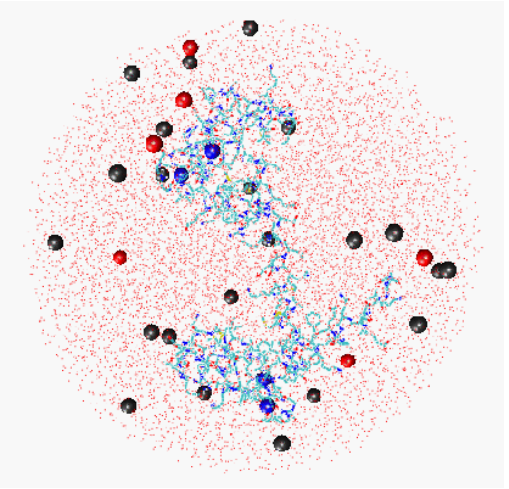
q=0, Na^+/Cl^- added

IONIC STRENGTH AND pH IN MD

- **Molarity calculations:**

  ~ 10^3 water molecules

  ~ 1mM salt concentration

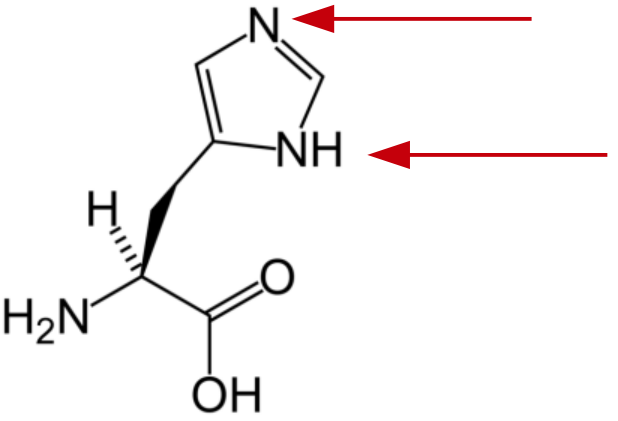


  ~ 10 cationic and anionic atoms added (Na^+ and Cl^-, f.i.)

- **Protonation state is manually adjusted (depending on pK_a) but only ONCE:**

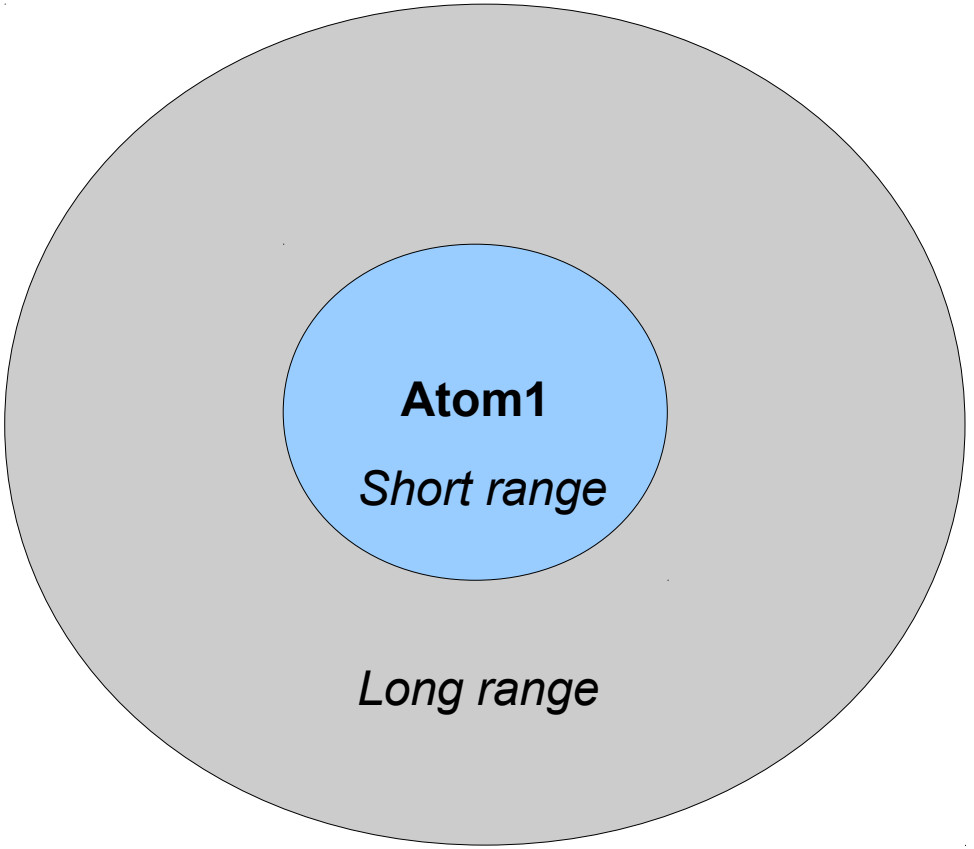
  - Asp, Glu, Lys, Arg – two residue libraries

  - His – three residue libraries



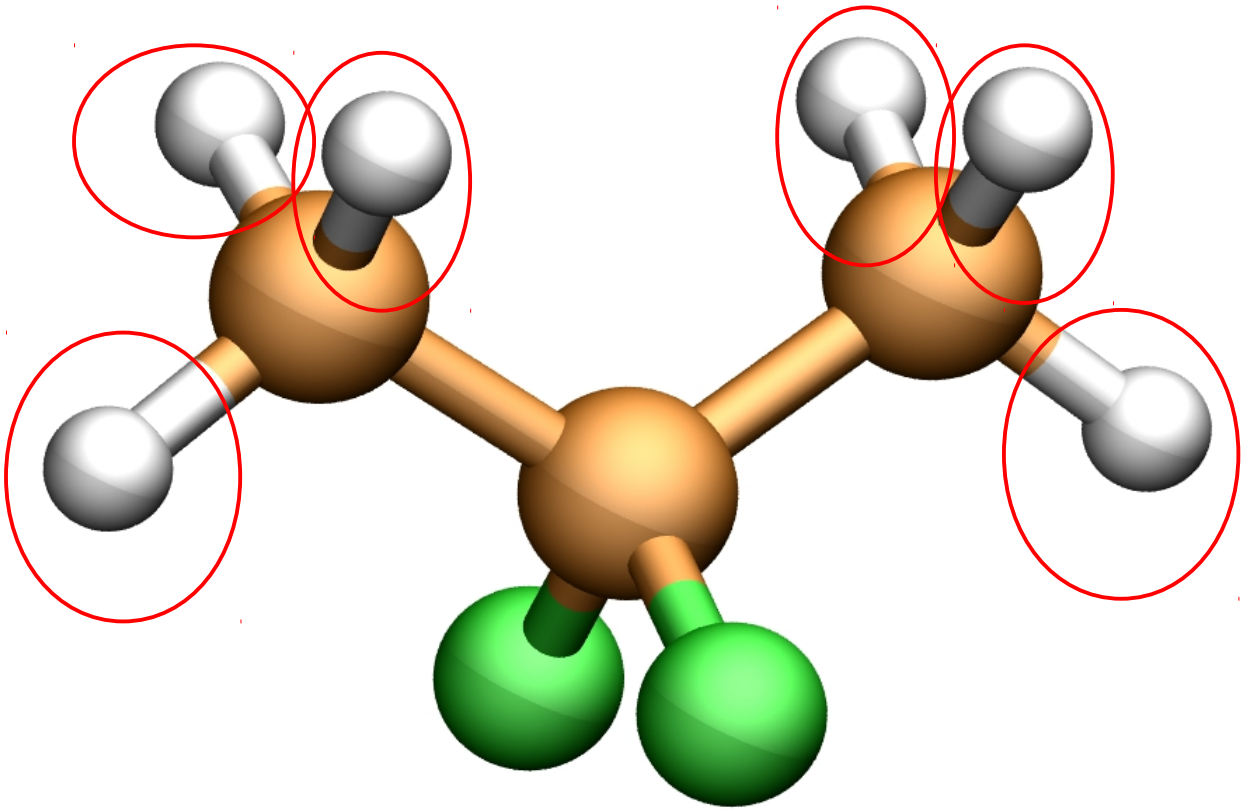
PARTICLE MESH EWALD METHOD

- **Particle Mesh Ewald method is utilized for electrostatic calculations in PBC. PME uses (Ewald) summation, which replaces the summation of interaction energies in real space with an equivalent summation in Fourier space. This increases speed and helps summation's convergence.**

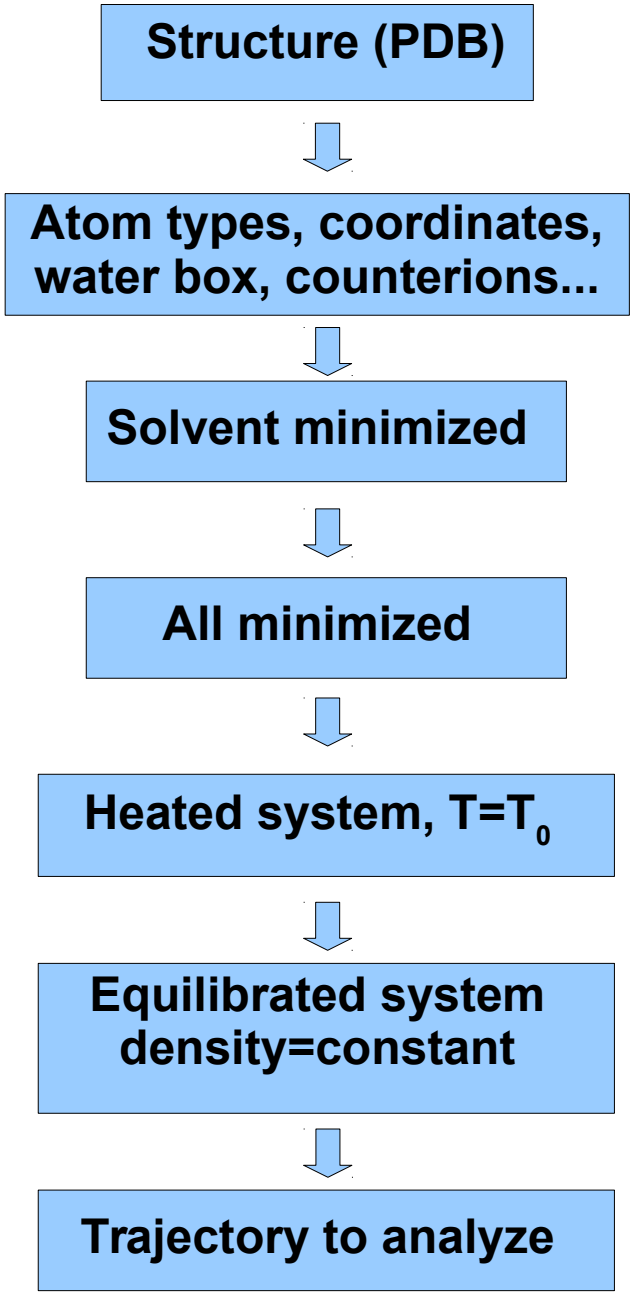


SHAKE ALGORITHM

- SHAKE algorithm constraints all motions of covalently bound hydrogens since these motions are the fastest (only dynamics!)



EXAMPLE: STANDARD MD RUN



Application of the force field

First minimization with solute fixed

Second minimization of the whole system

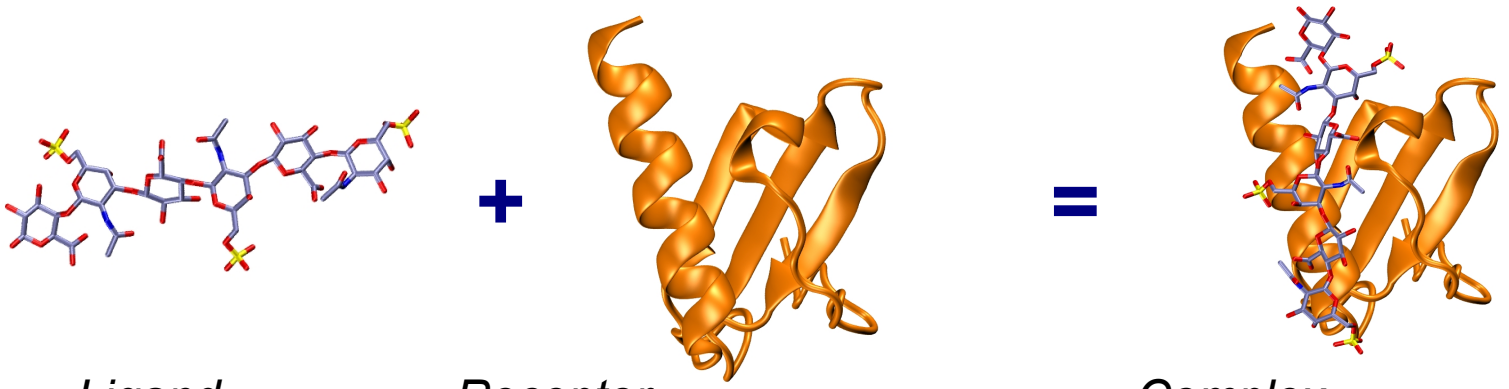
MD: heating from 0 to T_0

MD: putting the constant pressure

MD simulation in NTP ensemble

ENERGY CALCULATIONS

$$\delta G = \delta H - T\delta S; \quad \delta G_{binding} = \delta G_{complex} - (\delta G_{receptor} + \delta G_{ligand})$$



Ligand *Receptor* *Complex*

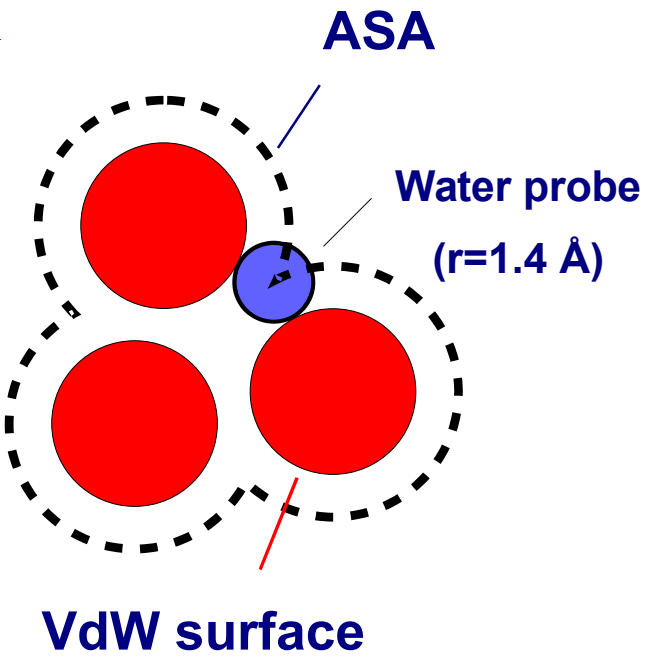
- **MM-PBSA (Molecular Mechanics Poisson Boltzmann Surface Area)**
- **Free energy perturbation**
- **Biased molecular dynamics**
- **Normal mode analysis for entropy calculations**

MM-PBSA

- **Molecular Mechanics-Poisson-Boltzmann Surface Area**

$$\delta G = \delta G_{vac} + \delta G_{solv}\,;\ \delta G_{vac} = \delta G_{ele} + \delta G_{vdw}$$

Molecular Mechanics (force field)



$$\delta G_{solv} = \delta G_{el} + \delta G_{nonel}$$

$$\delta G_{nonel} \sim ASA$$

$$\delta G_{el} = \frac{1}{2}\int (\rho(r)\phi(r))$$

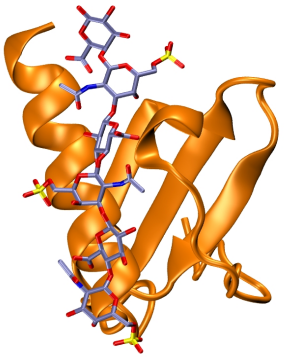
$$\varphi(r) = -4\pi\rho(r) + \kappa^2 \epsilon(r)\,\varphi(r) \quad - \quad Poisson-Boltzmann\ equation$$

$$G_{el,GB} = \frac{1}{8\pi}\left(\frac{1}{\epsilon_0} - \frac{1}{\epsilon}\right)\sum_{i<j}^{N} \frac{q_i q_j}{f_{GB}} \quad - \quad Generalized\ Born\ approximation$$

$$f_{GB} = \sqrt{r_{ij}^2 + a_{ij}^2 e^{-D}} \qquad D = \left(\frac{r_{ij}}{2a_{ij}}\right)^2,\ a_{ij} = \sqrt{a_i a_j}$$

MM-PBSA

MD trajectory



Frames extraction

N frames

Minimization in implicit solvent

MM-PBSA energies: mean and standard deviation

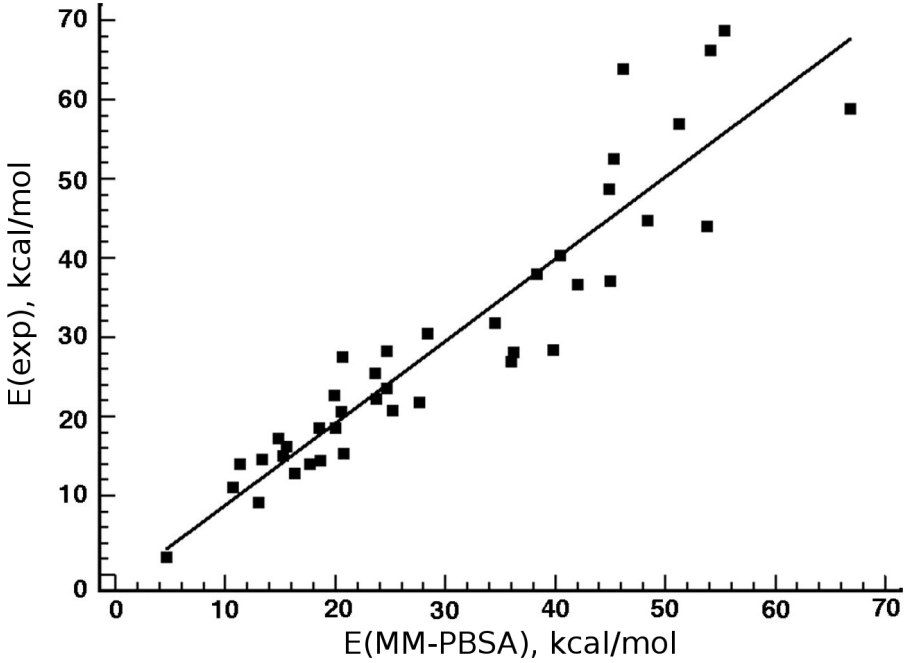
- **Coordinates from a the same SINGLE trajectory are used for Complex, Receptor and Ligand**
- **Strong dependence on particular frames**
- $\delta G_{binding} = \delta G_{complex} - (\delta G_{receptor} + \delta G_{ligand})$

MM-PBSA

- Components: Ele, VDW, reaction field, solvation, total
- Works well for complexes of bigger molecules
- Neglection of explicit solvent molecules
- r~0.4-0.9 depending on a predominant interaction type/frames choice
- Reliable relative but not absolute energies
- Comparable values for similar systems

Differences (Complex - Receptor - Ligand):

Energy Component	Average	Std. Dev.	Std. Err. of Mean
BOND	0.0000	0.0001	0.0000
ANGLE	-0.0000	0.0001	0.0000
DIHED	-0.0000	0.0001	0.0000
VDWAALS	-47.6566	5.0198	0.5612
EEL	-577.7110	69.5639	7.7775
1-4 VDW	0.0000	0.0001	0.0000
1-4 EEL	-0.0000	0.0000	0.0000
EPB	580.9776	66.2979	7.4123
ECAVITY	-5.8694	0.2482	0.0278
DELTA G gas	-625.3676	68.8867	7.7018
DELTA G solv	575.1082	66.3250	7.4154
DELTA G binding =	-50.2594 +/-	8.4920	0.9494



MM-PBSA per residue decomposition

- Contribution of each individual residue to binding
- Fast screening of possible mutants
- Alanine scanning

Energy Decomposition Analysis (All units kcal/mol): Generalized Born solvent

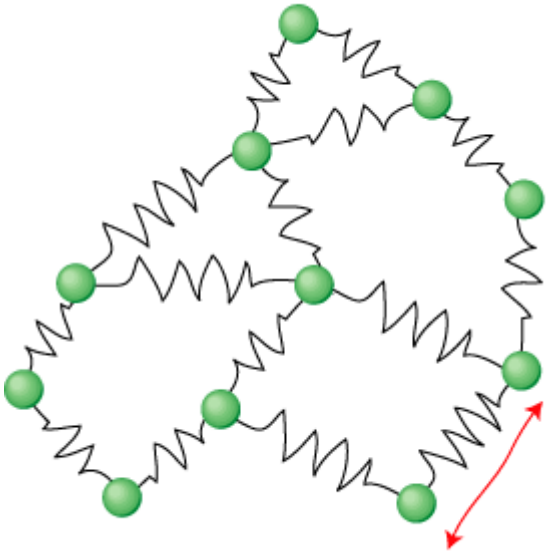
DELTAS:
Total Energy Decomposition:

Residue	Location	Internal	van der Waals	Electrostatic	Polar Solvation	Non-Polar Solv.	TOTAL
ILE 217	R ILE 217	0.000 +/- 4.933	-0.004 +/- 1.893	0.059 +/- 1.796	-0.048 +/- 1.107	0.000 +/- 0.085	0.008 +/- 5.690
VAL 218	R VAL 218	0.000 +/- 3.830	-0.008 +/- 1.862	-0.056 +/- 3.153	0.111 +/- 2.434	0.000 +/- 0.006	0.047 +/- 5.831
ARG 219	R ARG 219	0.000 +/- 5.914	-0.037 +/- 2.388	-4.663 +/- 14.675	4.807 +/- 11.967	-0.001 +/- 0.169	0.107 +/- 19.982
SER 220	R SER 220	0.000 +/- 2.907	-0.052 +/- 1.536	-0.038 +/- 6.548	0.191 +/- 5.590	-0.001 +/- 0.069	0.100 +/- 9.216
CYX 221	R CYX 221	0.000 +/- 2.618	-0.053 +/- 1.337	-0.055 +/- 2.409	0.218 +/- 1.711	0.000 +/- 0.008	0.111 +/- 4.168
LYS 222	R LYS 222	0.000 +/- 4.188	-0.064 +/- 2.145	-10.357 +/- 8.479	10.493 +/- 6.943	0.000 +/- 0.094	0.072 +/- 11.927
CYX 223	R CYX 223	0.000 +/- 2.960	-0.060 +/- 1.571	0.146 +/- 1.609	-0.095 +/- 0.654	0.000 +/- 0.000	-0.009 +/- 3.774
SER 224	R SER 224	0.000 +/- 3.123	-0.019 +/- 2.079	11.115 +/- 8.194	-10.999 +/- 7.466	0.000 +/- 0.036	0.097 +/- 11.703
ALA 225	R ALA 225	0.000 +/- 3.207	-1.240 +/- 1.304	-30.420 +/- 7.670	35.837 +/- 7.577	-0.337 +/- 0.221	3.840 +/- 11.325
VAL 226	R VAL 226	0.000 +/- 3.875	-2.888 +/- 1.996	0.285 +/- 2.212	0.382 +/- 2.203	-0.569 +/- 0.090	-2.790 +/- 5.362
LYS 227	R LYS 227	0.000 +/- 4.877	-0.474 +/- 1.885	-22.334 +/- 12.328	23.168 +/- 13.005	-0.011 +/- 0.206	0.349 +/- 18.668
PHE 228	R PHE 228	0.000 +/- 4.286	-4.959 +/- 1.641	-2.484 +/- 2.101	3.943 +/- 1.951	-0.849 +/- 0.083	-4.348 +/- 5.412
PRO 229	R PRO 229	0.000 +/- 3.396	-1.093 +/- 1.151	-6.375 +/- 1.701	4.238 +/- 1.192	-0.180 +/- 0.048	-3.410 +/- 4.144
GLN 230	R GLN 230	0.000 +/- 4.028	-0.689 +/- 1.402	2.560 +/- 2.847	-0.989 +/- 1.394	-0.007 +/- 0.044	0.874 +/- 5.314
LEU 231	R LEU 231	0.000 +/- 4.078	-1.093 +/- 2.054	-0.185 +/- 1.777	0.327 +/- 0.594	-0.086 +/- 0.053	-1.036 +/- 4.936
CYX 232	R CYX 232	0.000 +/- 2.802	-0.048 +/- 1.387	0.081 +/- 1.366	-0.197 +/- 0.590	0.000 +/- 0.000	-0.163 +/- 3.463
LYS 233	R LYS 233	0.000 +/- 4.619	-0.014 +/- 2.263	-12.038 +/- 11.108	11.972 +/- 10.286	0.000 +/- 0.045	-0.080 +/- 15.989
PHE 234	R PHE 234	0.000 +/- 4.331	-0.021 +/- 1.933	0.100 +/- 1.838	-0.113 +/- 1.039	0.000 +/- 0.039	-0.034 +/- 5.192
CYX 235	R CYX 235	0.000 +/- 2.724	-0.003 +/- 1.324	-0.005 +/- 3.402	0.003 +/- 2.498	0.000 +/- 0.066	-0.005 +/- 5.195
ASP 236	R ASP 236	0.000 +/- 3.692	-0.002 +/- 1.608	11.143 +/- 13.493	-10.994 +/- 12.507	0.000 +/- 0.152	0.146 +/- 18.834
VAL 237	R VAL 237	0.000 +/- 3.260	-0.001 +/- 1.807	0.080 +/- 2.382	-0.078 +/- 1.820	0.000 +/- 0.084	0.001 +/- 4.784

ENTROPY CALCULATIONS

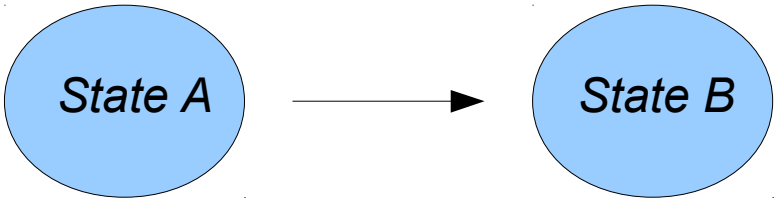
$$\delta S = \delta S_{translational} + \delta S_{rotational} + \delta S_{vibrational}$$

- Translational and rotational entropies correspond to the number of freedom degrees
- Vibrational entropy is derived from the trajectory:
  - Vibrations are represented in matrix
  - Matrix is diagonalized (normal modes)
  - Eigenvectors and eigenvalues are calculated
  - Entropy is derived related to the cut-off for eigenvalues
- Very expensive computation, applicable for relatively small molecules



FREE ENERGY PERTURBATION

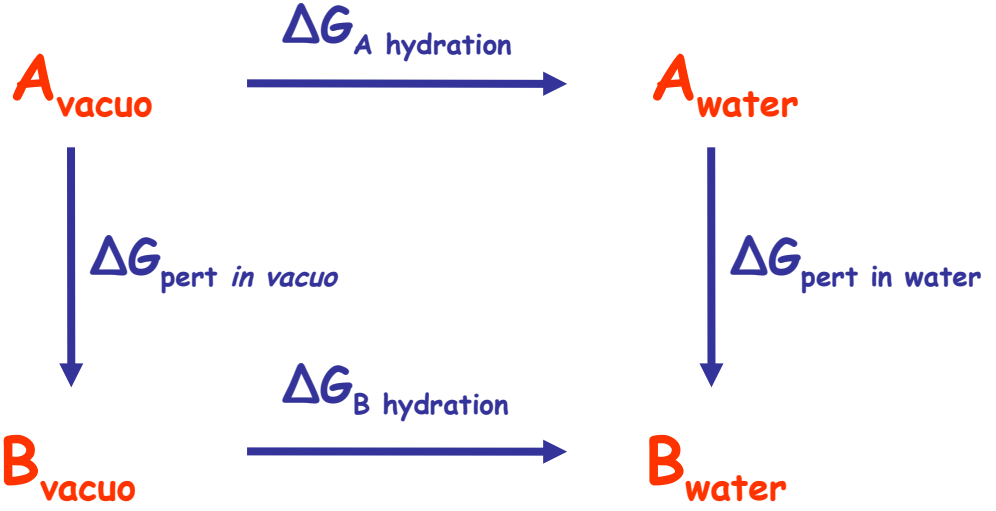
$$\delta G = G_A - G_B$$



$$\delta G = G(\lambda = 1) - G(\lambda = 0) = \int_0^1 \left(\frac{\delta V}{\delta \lambda}\right)_\lambda d\lambda - \textit{Thermodynamic integration}$$

$$\delta G = \sum_{i=1}^{n} w_i \left(\frac{\delta V}{\delta \lambda}\right)_{\lambda_i}$$

Example:



- **Electrostatics**
- **VDW**

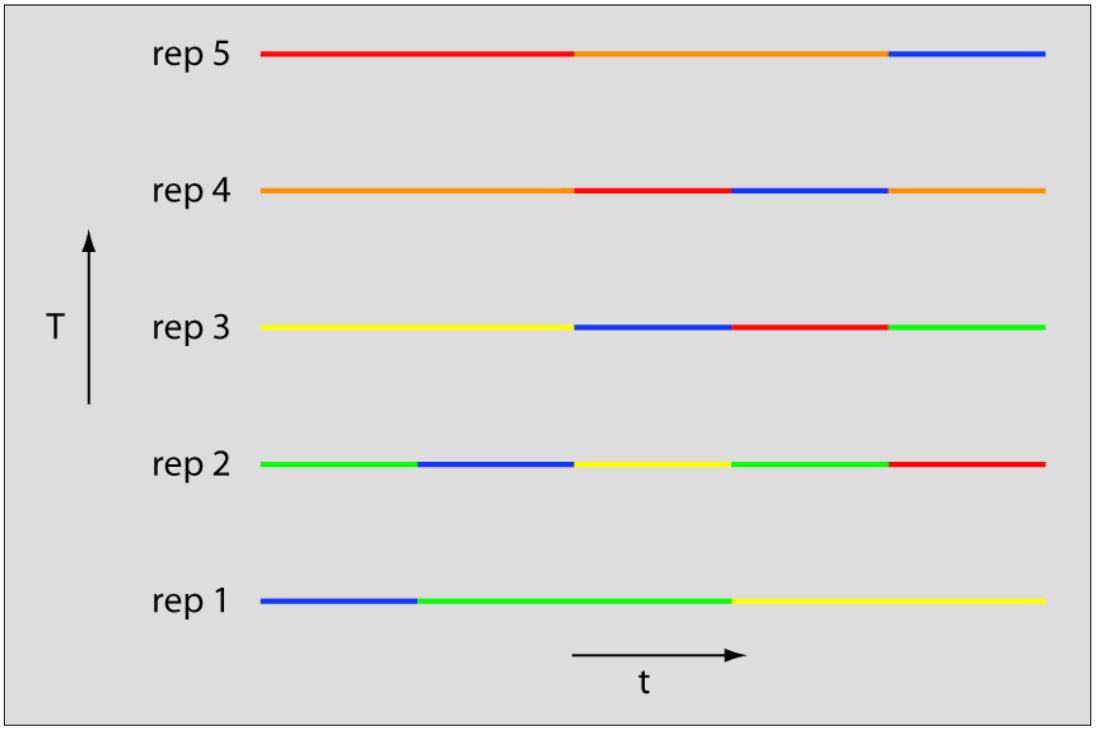
$$\Delta G_{\text{A hydration}} - \Delta G_{\text{B hydration}} = \Delta G_{\text{pert in vacuo}} - \Delta G_{\text{pert in water}}$$

REPLICA EXCHANGE MD

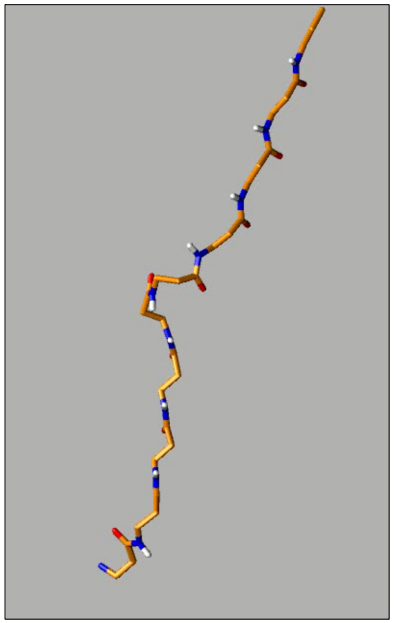
- **RE (replica exchange) MD: states $_i$ and $_j$ by T_i and T_j:**

$$p = min\left(1, \frac{\exp -\left(\frac{\delta G_j}{k T_i} - \frac{\delta G_i}{k T_j}\right)}{\exp -\left(\frac{\delta G_i}{k T_i} - \frac{\delta G_j}{k T_j}\right)}\right) = min\left(1, \exp\left(\delta G_i - \delta G_j\right)\left(\frac{1}{k T_i} - \frac{1}{k T_j}\right)\right)$$

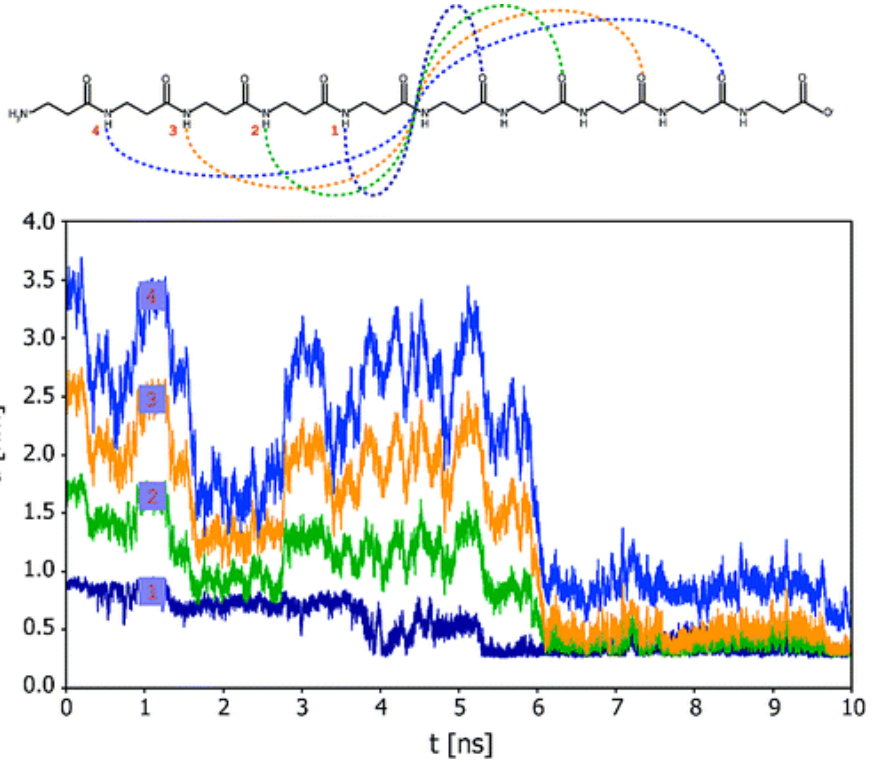
- **N replicas of the same structure**
- **Implicit solvent**
- **More effective sampling**
- **Computationally very expensive**
- **Principal component analysis**

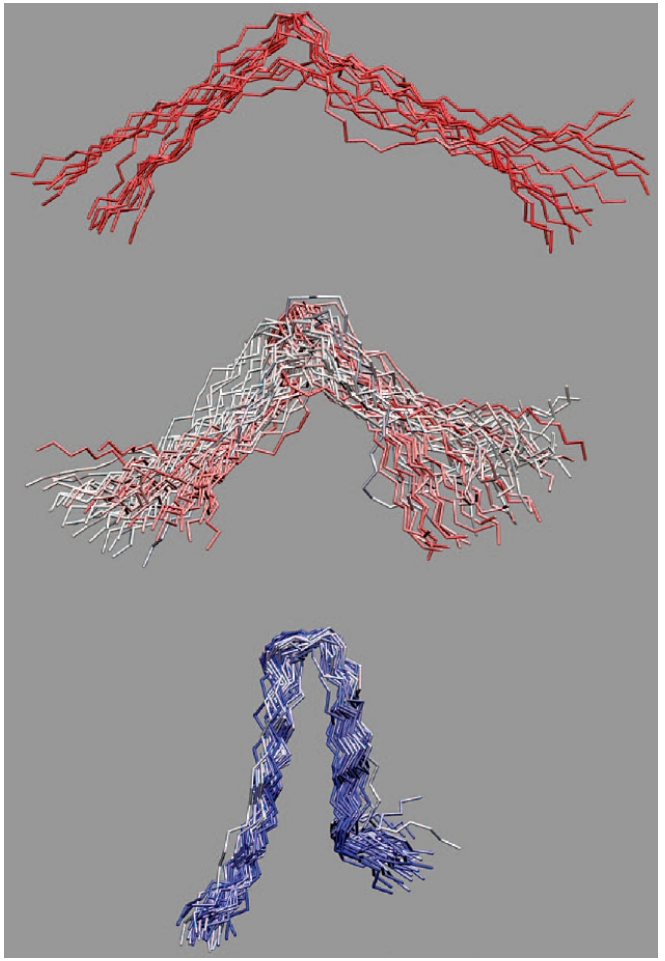


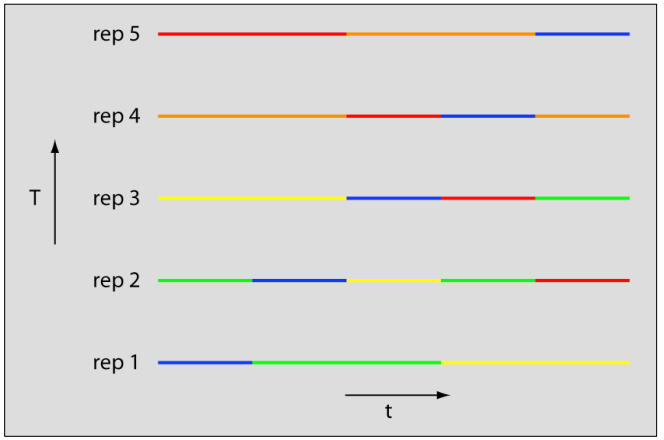
REMD: EXAMPLE 1



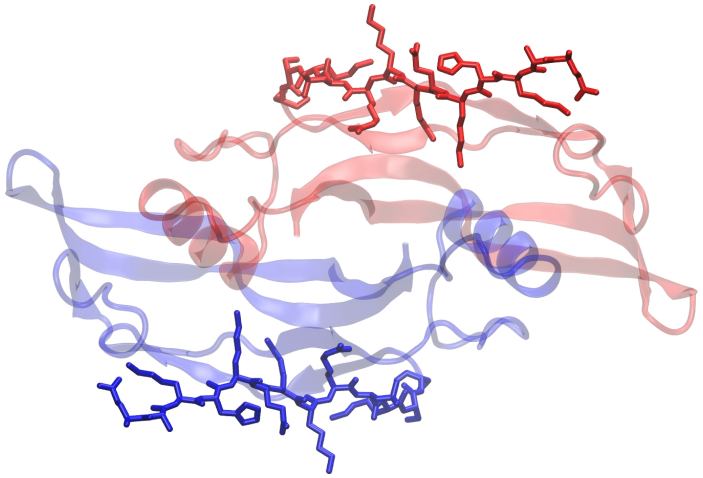
β-hairpins



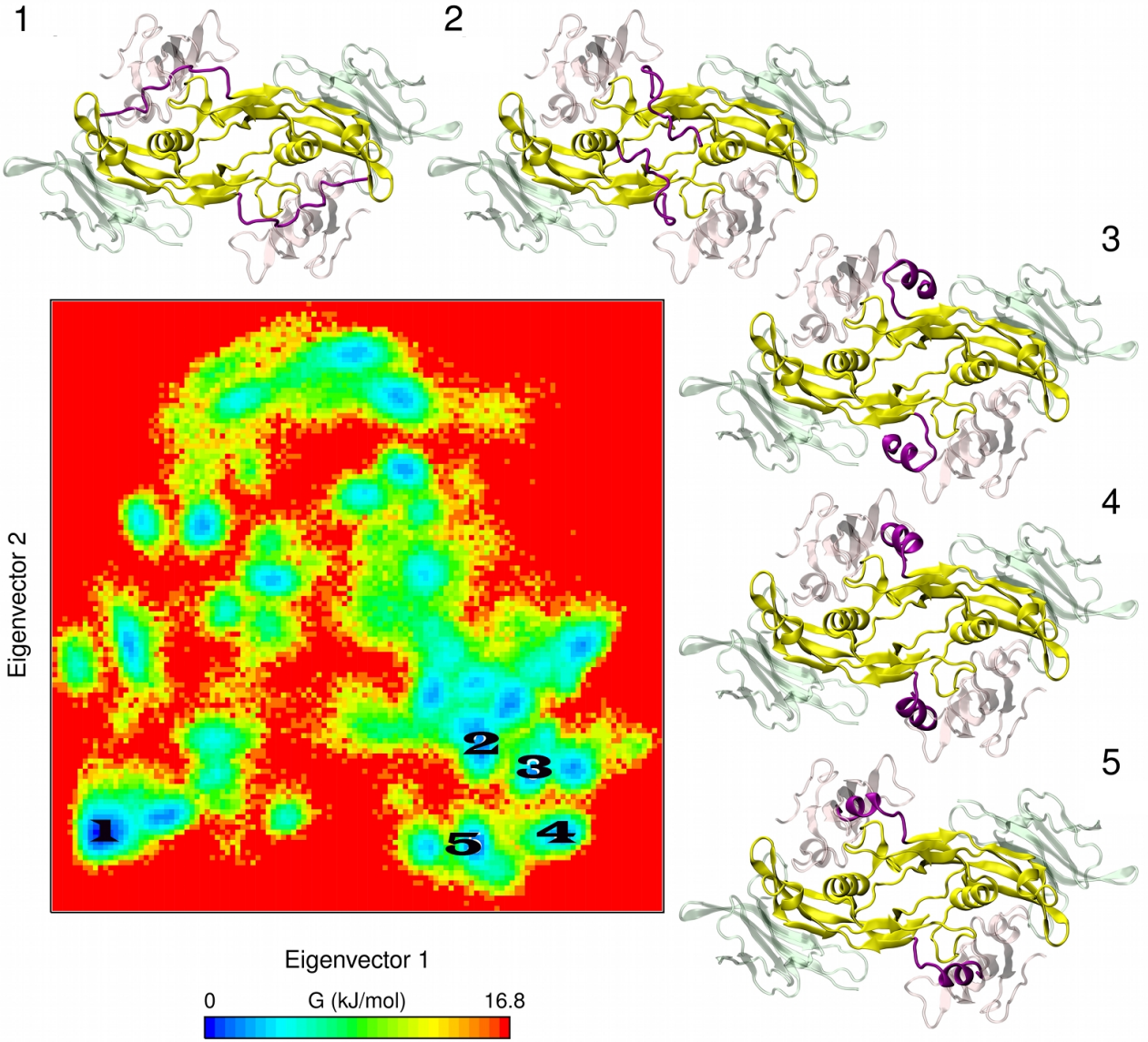




REMD: EXAMPLE 2



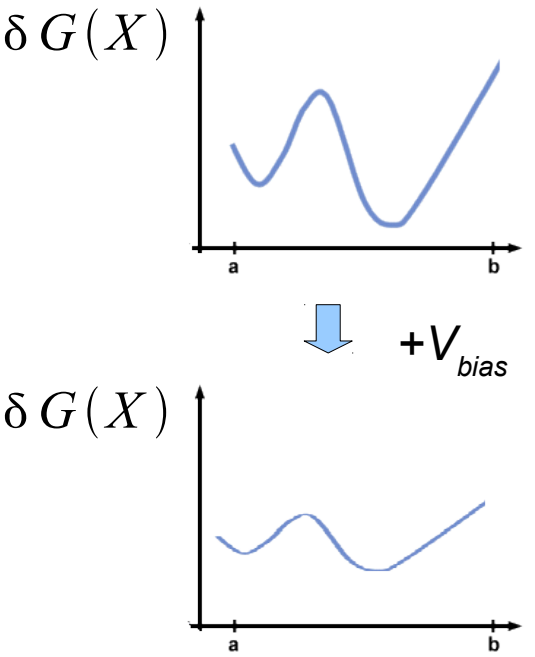
- BMP-2 dimer
- 13 aa free, rest frozen
- 12 replicas
- 100 ns

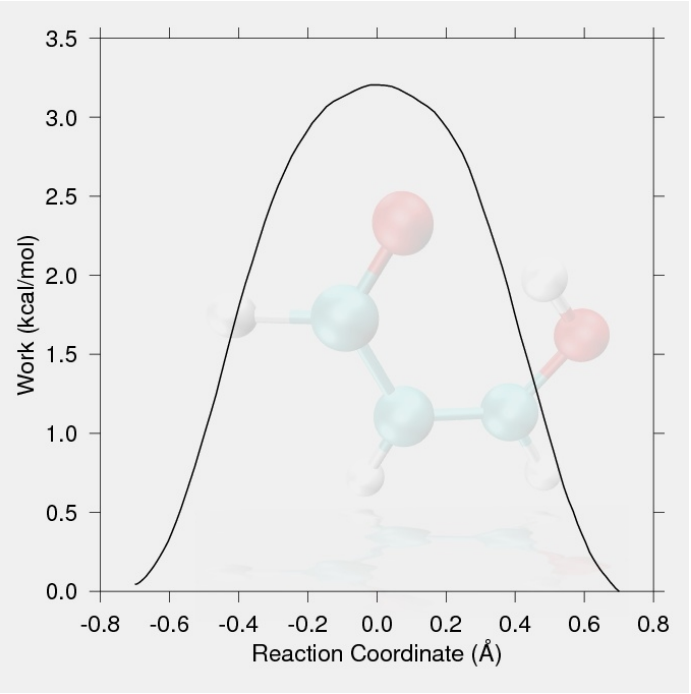


BIASED MOLECULAR DYNAMICS

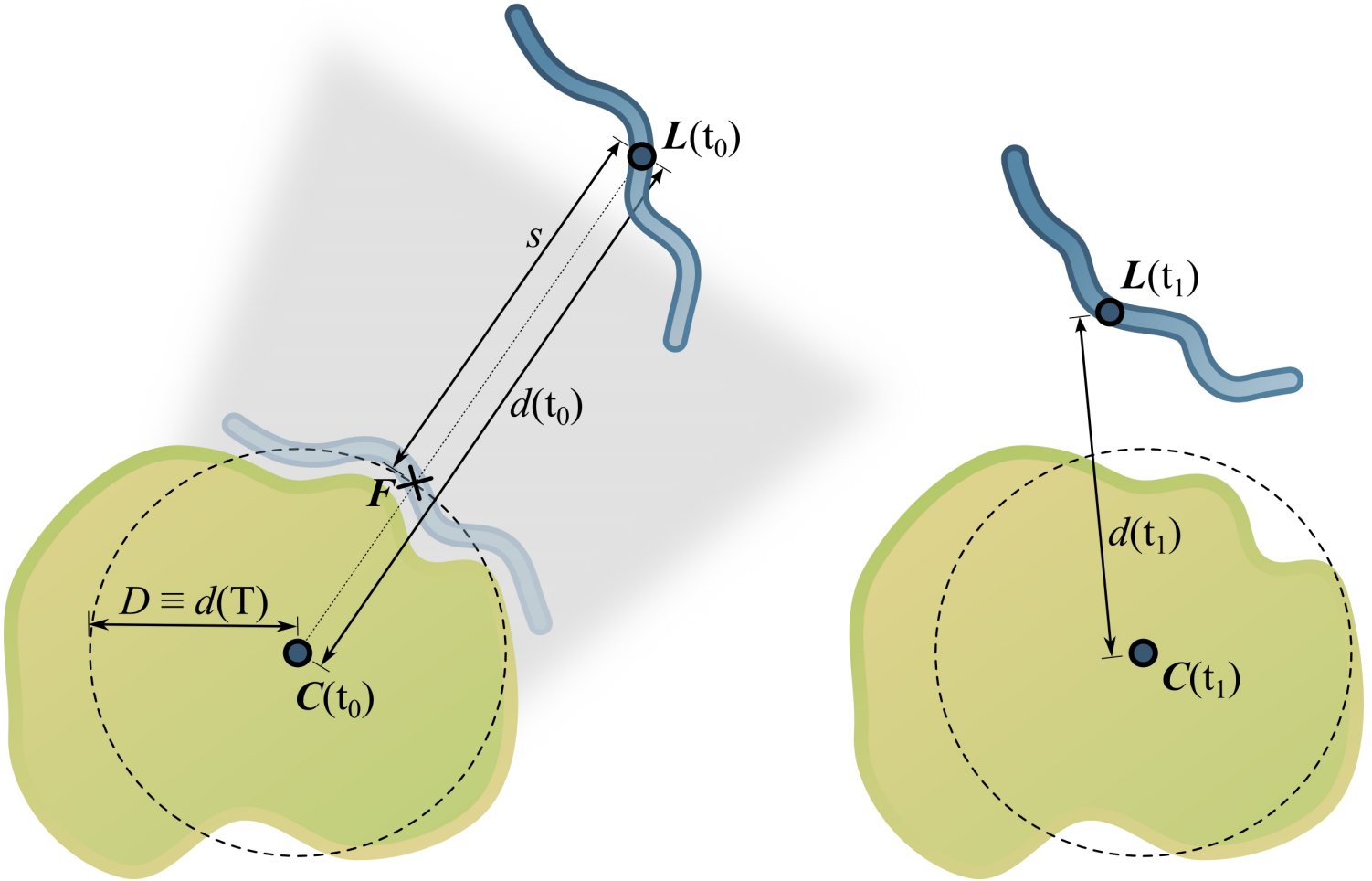
- Additional potential is applied to the system
- Speeding up the transition between A and B or search for global minimum
- Examples:
  - Targeted MD (X changes from X_1 to X_2), ~ (100 trajectories)

$$\delta G(X_1) - \delta G(X_2) = <Work\ of\ the\ applied\ force>$$





TARGETED MD-BASED DOCKING: EXAMPLE



- Dynamic Molecular Docking method
- Explicit solvent
- Fully flexible receptor and ligand
- MM-PBSA scoring for an ensemble of structures

ACCELERATED MOLECULAR DYNAMICS

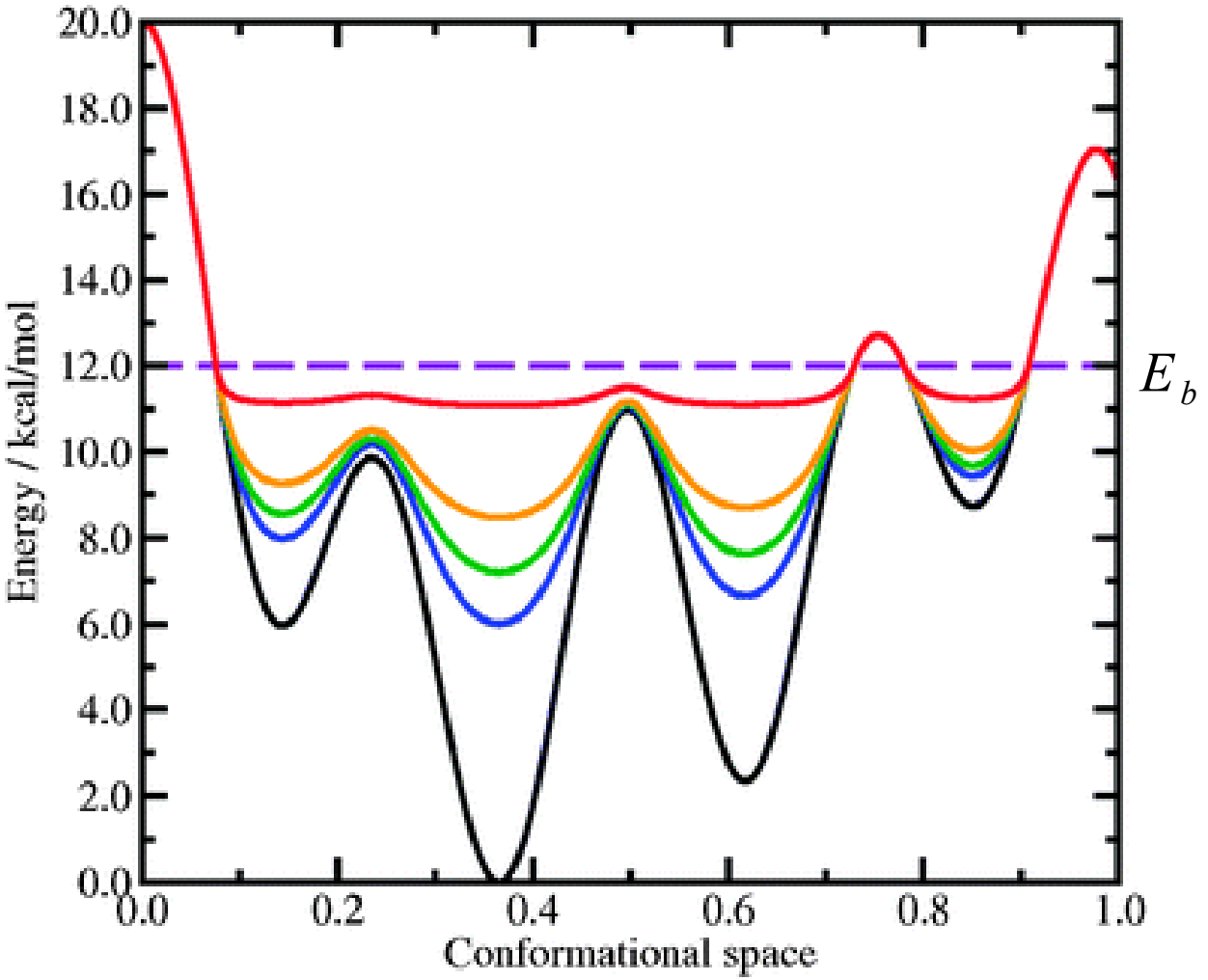
$$V(\vec{r}) = \sum_{bonds} K_r (r - r_{eq})^2 + \sum_{angles} K_\theta (\theta - \theta_{eq})^2 + \sum_{dihedrals} \frac{V_n}{2}(1 + \cos[n\phi - \gamma]) + \sum_{i<j}^{atoms} \left(\frac{A_{ij}}{R_{ij}^{12}} - \frac{B_{ij}}{R_{ij}^{6}}\right) + \sum_{i<j}^{atoms} \frac{q_i q_j}{\epsilon R_{ij}}$$

$$V_a(\vec{r}) = V(\vec{r}), V(\vec{r}) \geq E_b$$

$$V_a(\vec{r}) = V(\vec{r}) + \delta V(\vec{r}), V(\vec{r}) \leq E_b$$

$$\delta V(\vec{r}) = \frac{(E_b - V(\vec{r}))^2}{E_b - V(\vec{r}) + \alpha}$$

- **Inverted AMD**
- **Selective AMD**
- **Replica Exchange AMD**
- **Adaptive AMD**
- ***Ab initio* AMD**



BROWNIAN MD

- **Langevin dynamics: degrees of freedom are replaced with stochasticity**

$$M\frac{d^2\vec{r}}{dt^2} = -d^2\frac{V(\vec{r})}{dt^2} - \gamma M d\frac{\vec{r}}{dt} + \sqrt{2\gamma kTM}\,R(t);\quad <R(t)> = 0;\; <R(t)R(t')> = \delta(t-t')$$

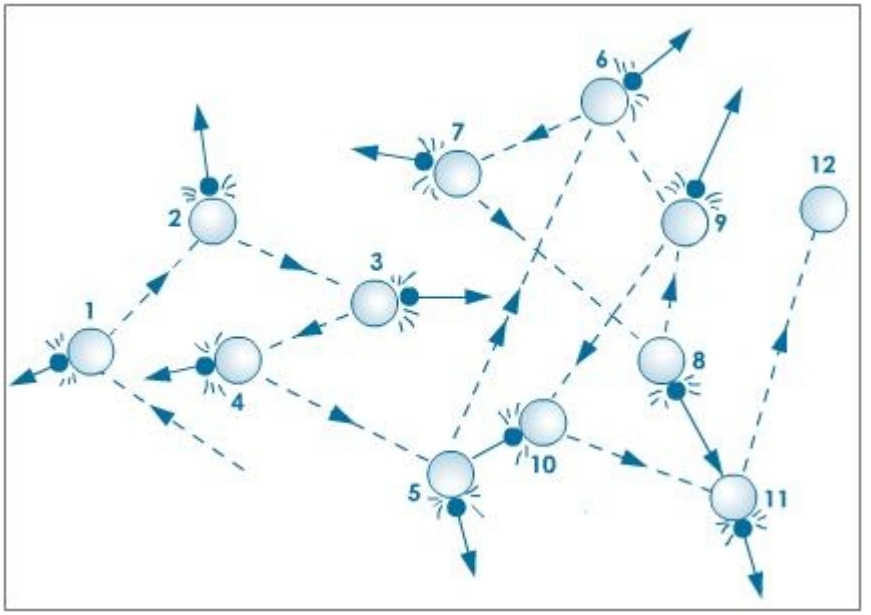
- **Browninan dynamics:**

$$0 = -d^2\frac{V(\vec{r})}{dt^2} - \gamma M d\frac{\vec{r}}{dt} + \sqrt{2\gamma kTM}\,R(t)$$

- **Longer time scales than normal MD: μs**

- **Bigger molecules**

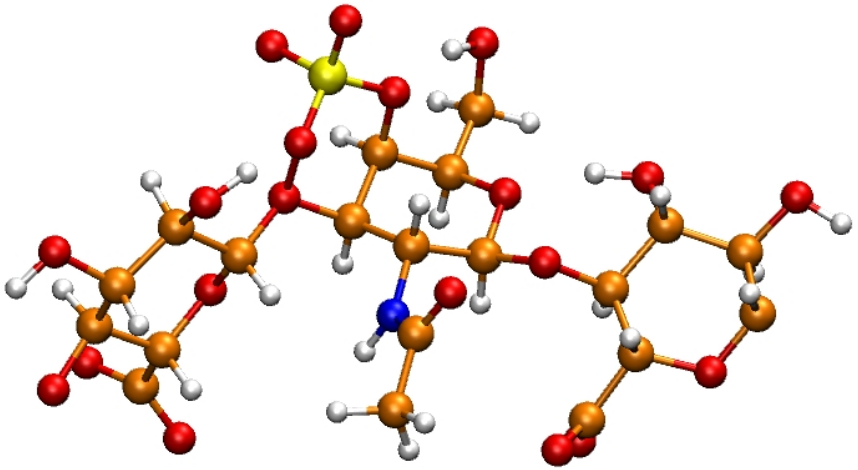
- **Implicit solvent**



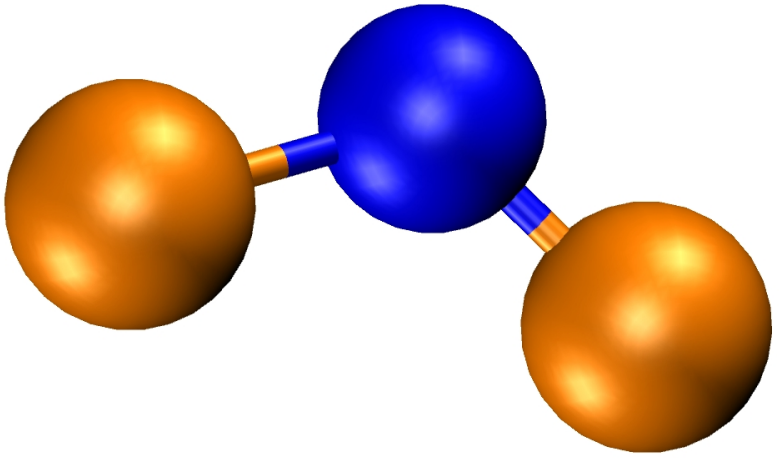
COARSE-GRAINED MD

MD (Force field)

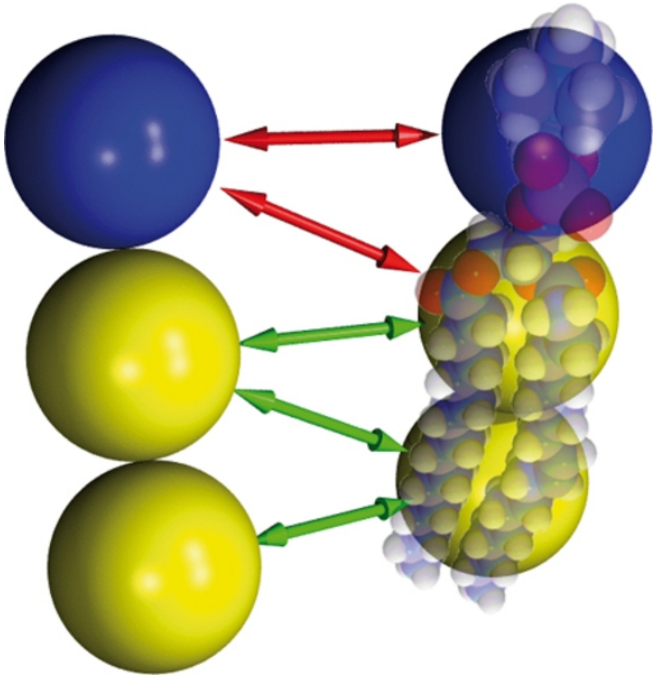
All-atomic

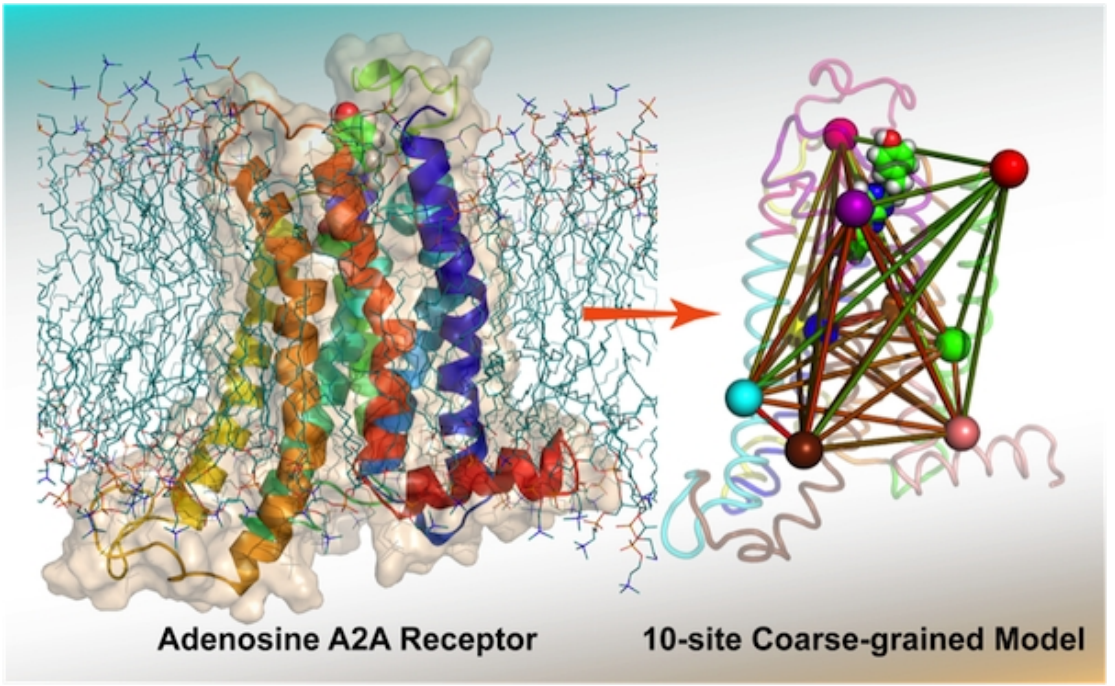


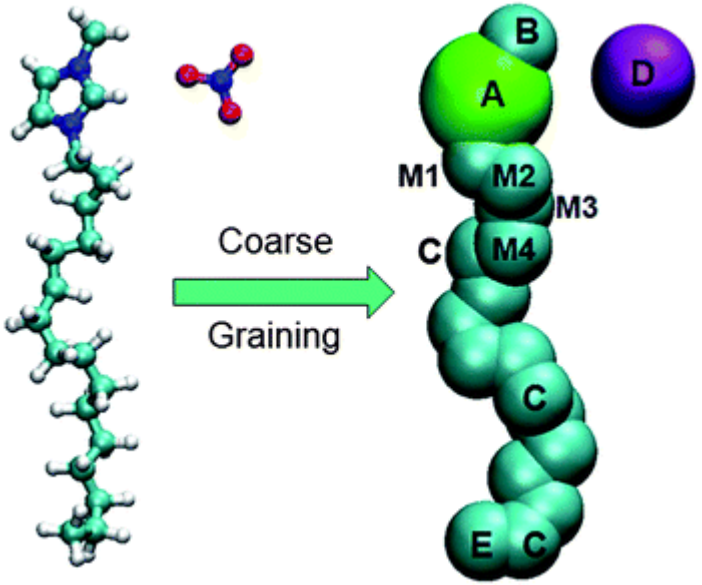
Coarse-grained

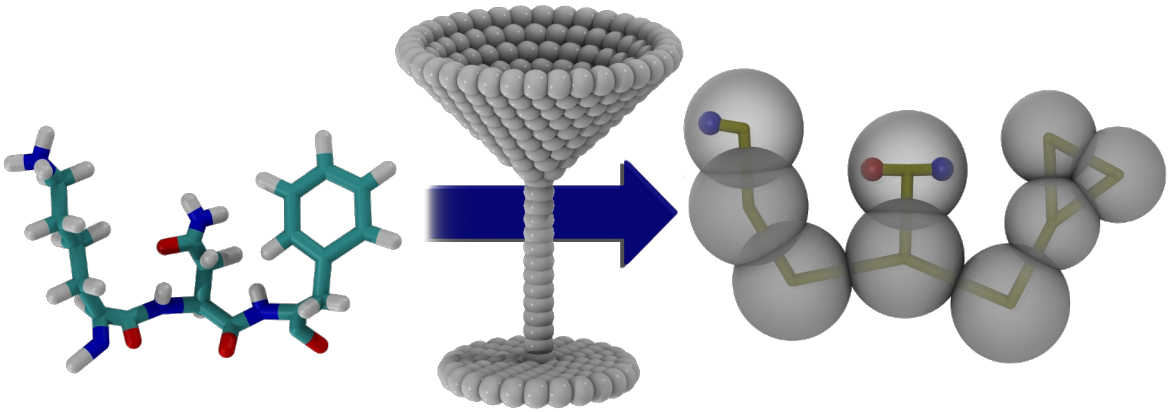


COARSE-GRAINED EXAMPLES

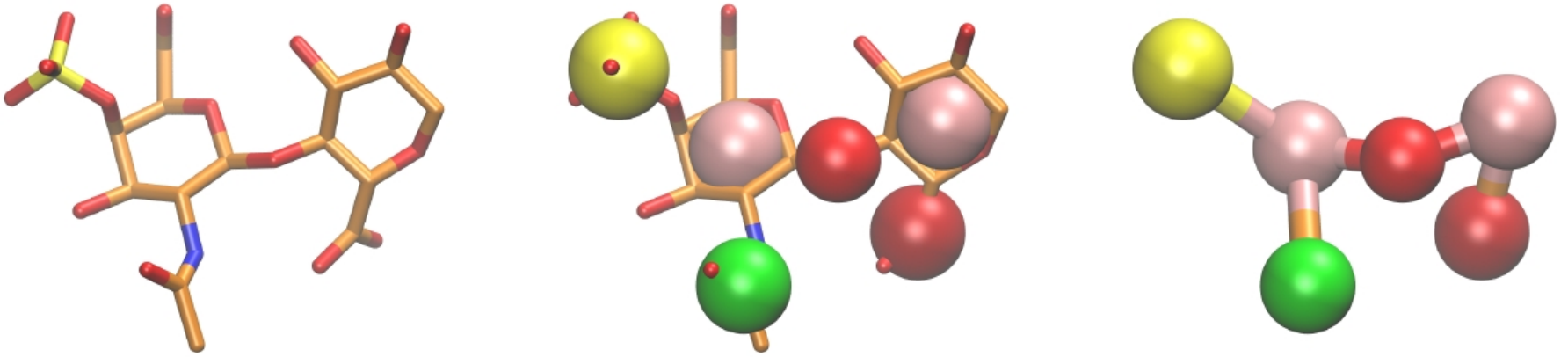








COARSE-GRAINED EXAMPLE FOR GAGs



All-atomic **Coarse-grained**

- 1 pseudoatom = several real atoms
- New atomic types
- New atomic charges
- New parameters: bond, angle, dihedral, L-J
- Global and local properties of CG should reproduce AA

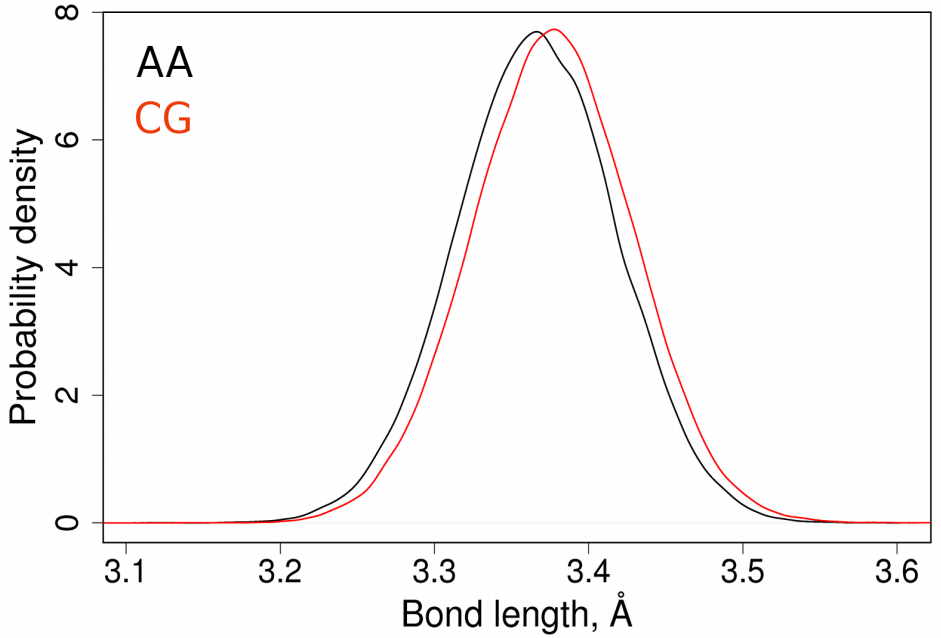
CG MD: PARAMETERS DERIVATION

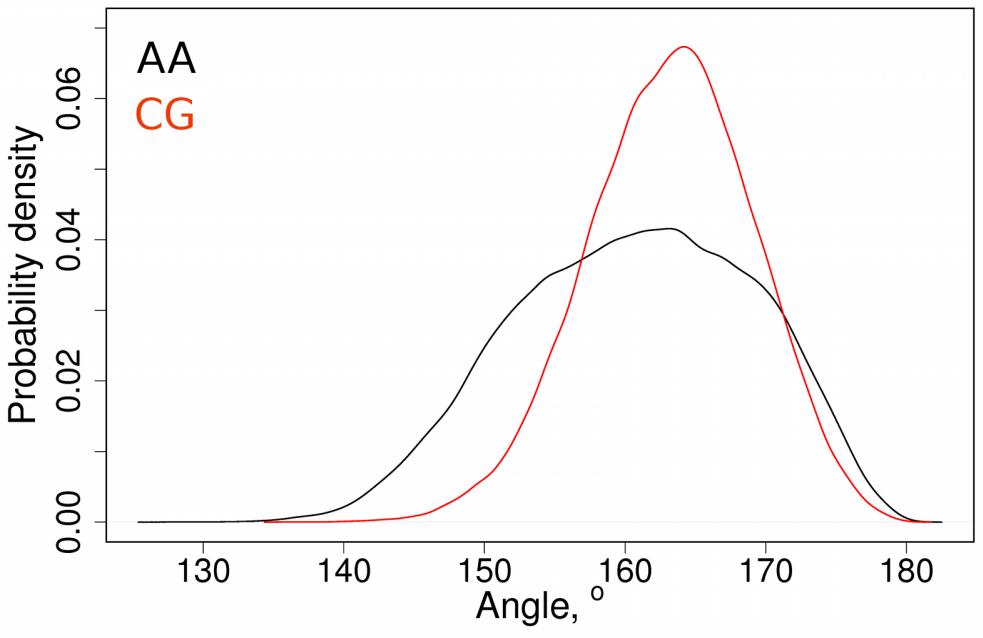
- **Charges for pseudoatoms: empirical or AA-consistent**
- **Bonds and angles for pseudoatoms: harmonic approximation**

$$\Delta G = K(X - X_0)^2 \; ; \; \rho = Ae^{\frac{-\Delta G}{kT}} = \frac{e^{-\frac{(X-X_0)^2}{2\sigma^2}}}{\sqrt{2\pi}\sigma}$$

$$K = \frac{kT}{2\sigma^2}; T = 300K$$

$$\Delta G = K(\alpha - \alpha_0)^2 \; ; \; \rho = Ae^{\frac{-\Delta G}{kT}} = \frac{e^{-\frac{(\alpha-\alpha_0)^2}{2\sigma^2}}}{\sqrt{2\pi}\sigma}$$



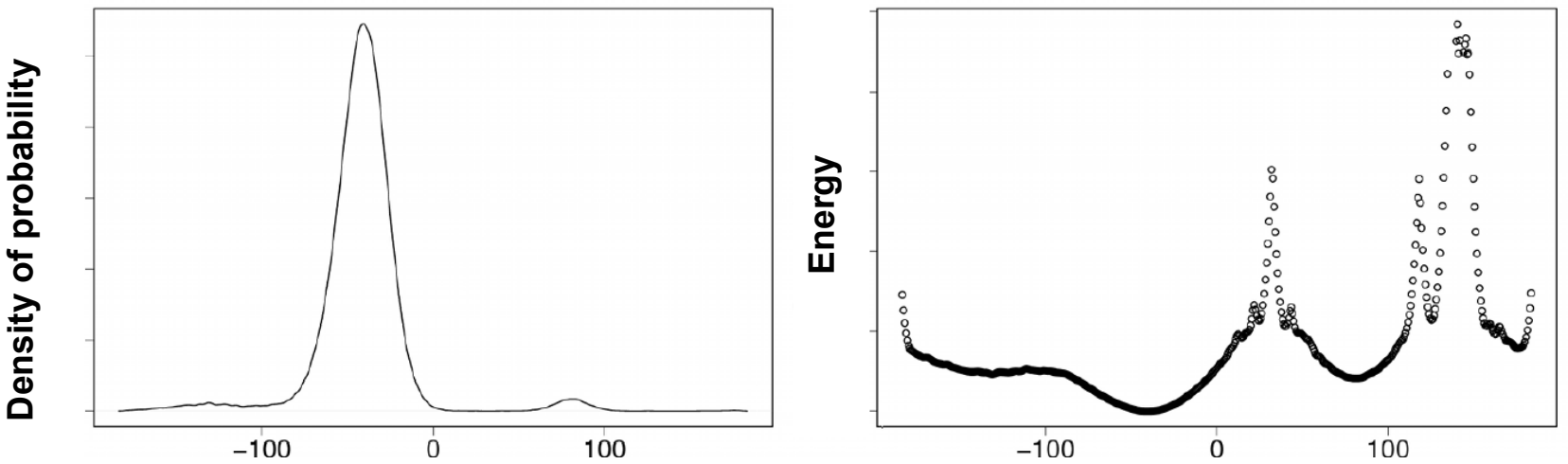


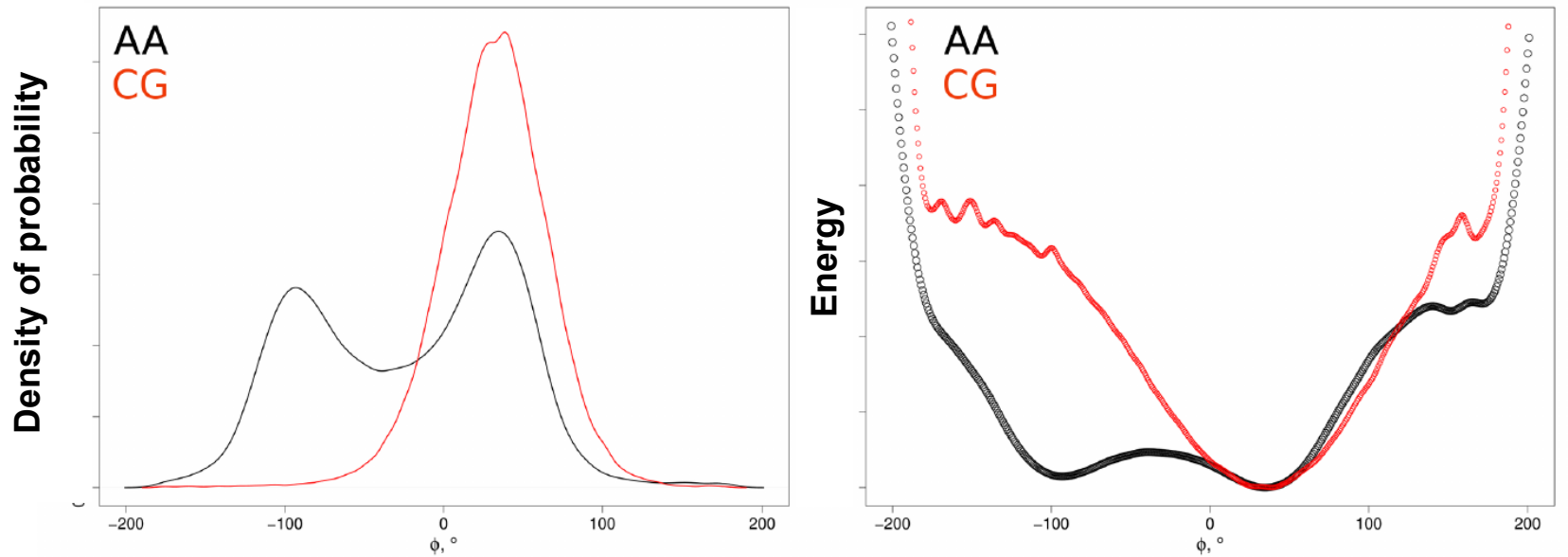
CG MD: PARAMETERS DERIVATION

- **Dihedral angles for pseudoatoms**

$$\Delta G = \frac{A_{barrier}}{2}(1 + cos(N\phi - \rho_0))$$

$A_{barrier}$: potential barrier height; N: periodicity; ρ_0: phase





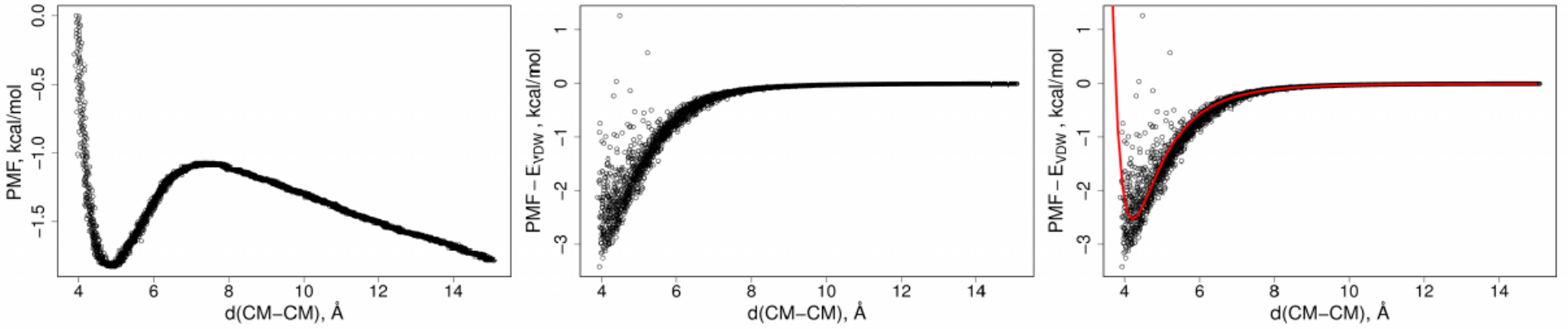
CG MD: PARAMETERS DERIVATION

➢ Lennard-Jones parameters for pseudoatoms

$$\Delta G = \frac{A}{r^{12}} - \frac{B}{r^{6}}$$

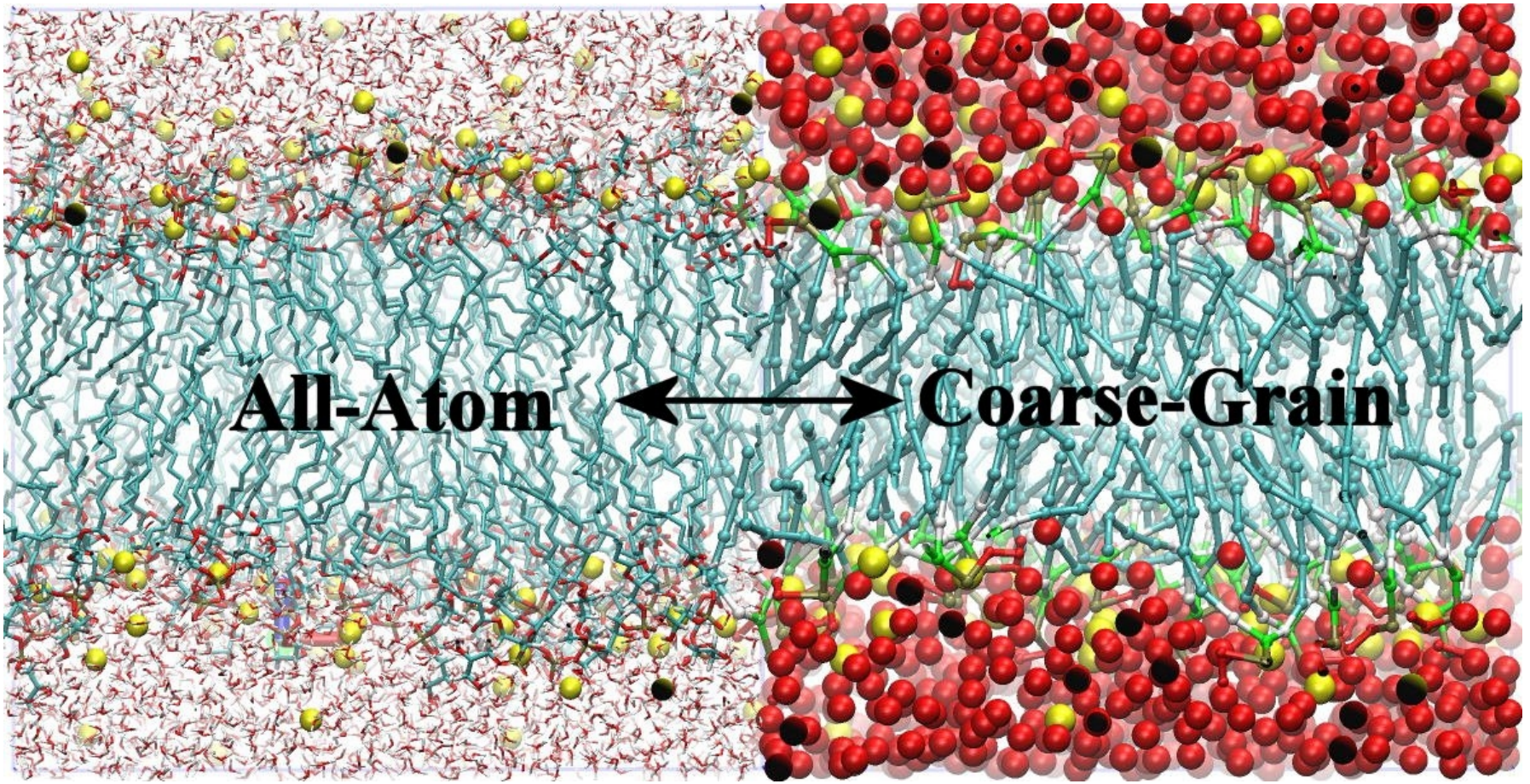
- AA molecules with 0 partial charges
- SMD: r(CM-CM) $\in [r_0 - 0.5; 15]$, *in vacuo*, 100 kcal/(mol Å), 1000 steps, 10 ns
- PMF is corrected by E_{VDW} for unbound molecules
- A and B are extracted by interpolation

Example: N-sulfate (HNH-SO_3)



CG MD: APPLICATION

- Multimacromolecular complexes
- Membranes
- Self-organizing systems/folding



MD: TIME/LENGTH SCALES

- From ps (10^{-12}) to ~10 μs (10^{-5} s)

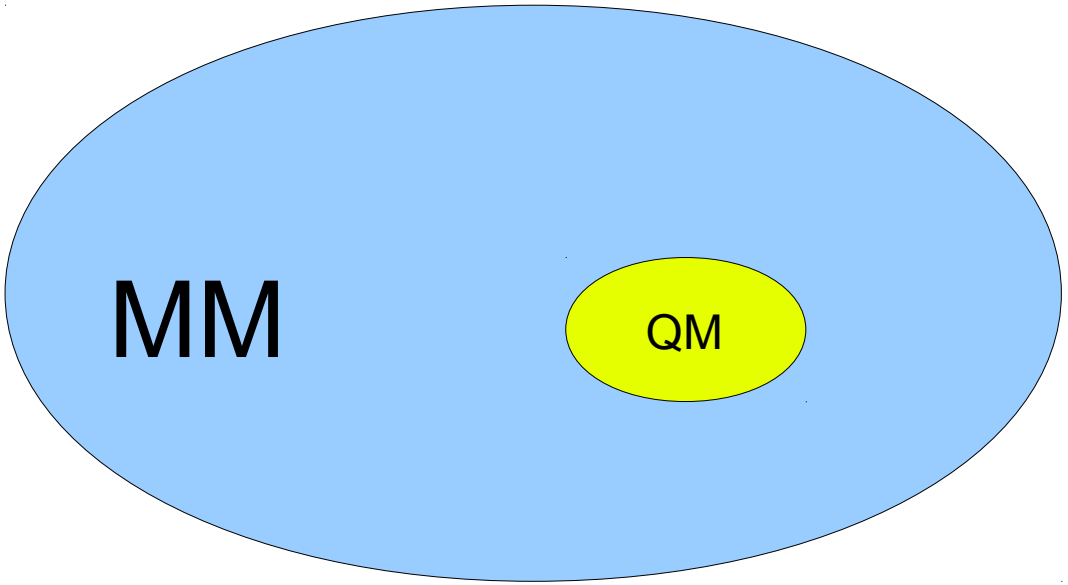
- Able to study
  - Mobility of molecules / equilibrated complexes
  - Atomic details of interactions (H-bonds, vdW contacts, salt-bridges)
  - Thermodynamics of equilibrated systems
  - Folding of small proteins

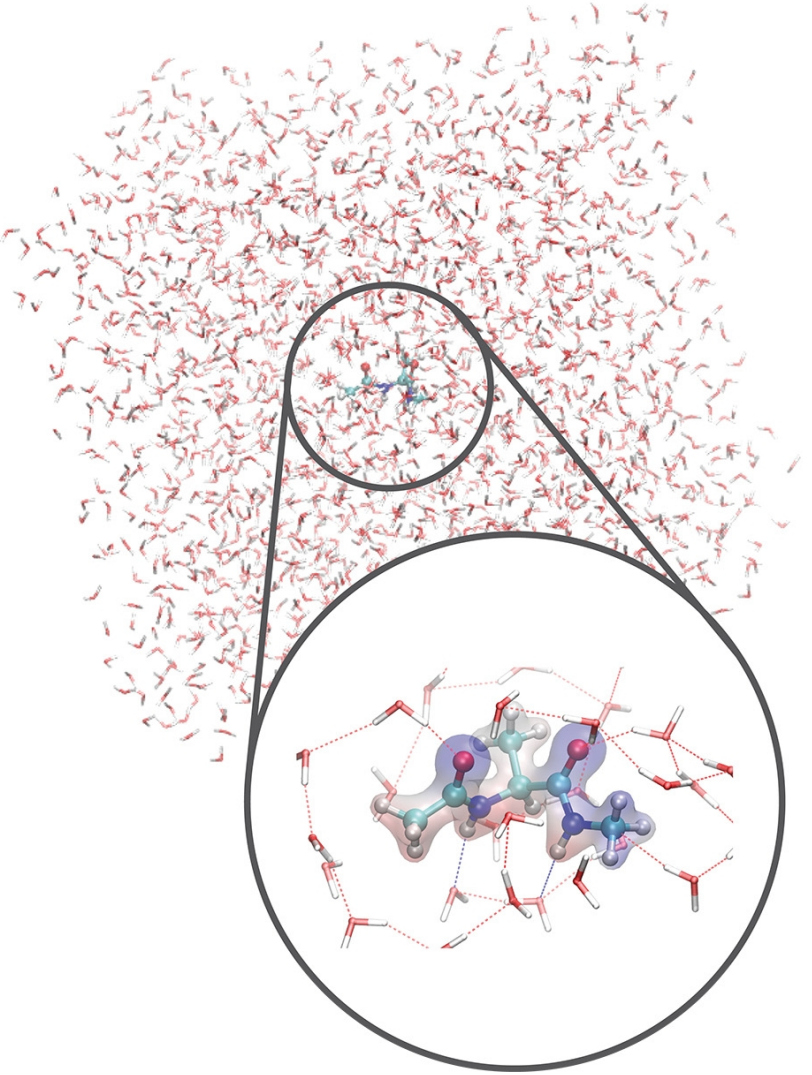
- Unable to study :(
  - Kinetics of molecular association (in general)
  - Directly catalytical reactions

- 1 ns of MD takes from hours to days
- Length: from Å to ~ 1000 Å

QM/MM

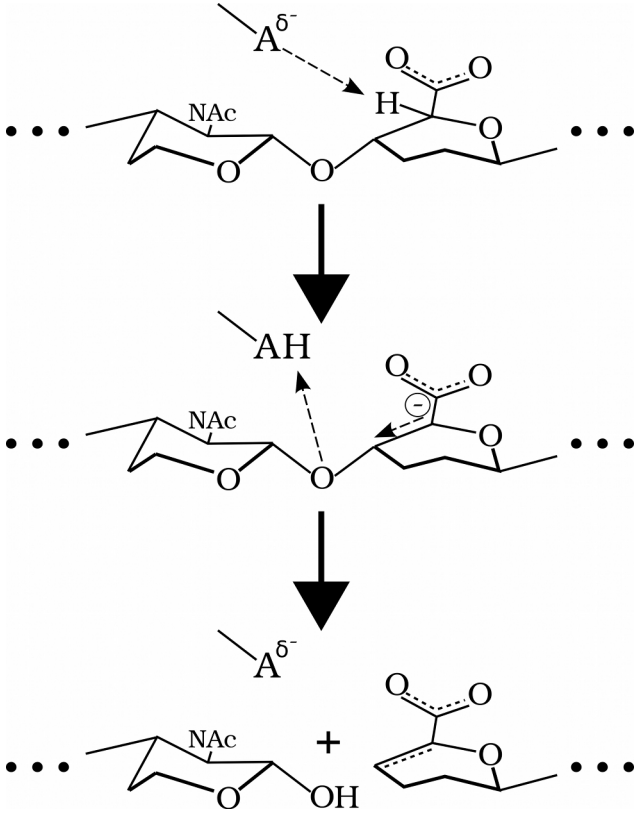
- Quantum effects are indispensable (f.i.: enzymatic reactions, conformational analysis)
- System is divided into QM and MM part
- Challenge for atoms in the QM/MM border
- Kinetics are accessible
- Computationally expensive

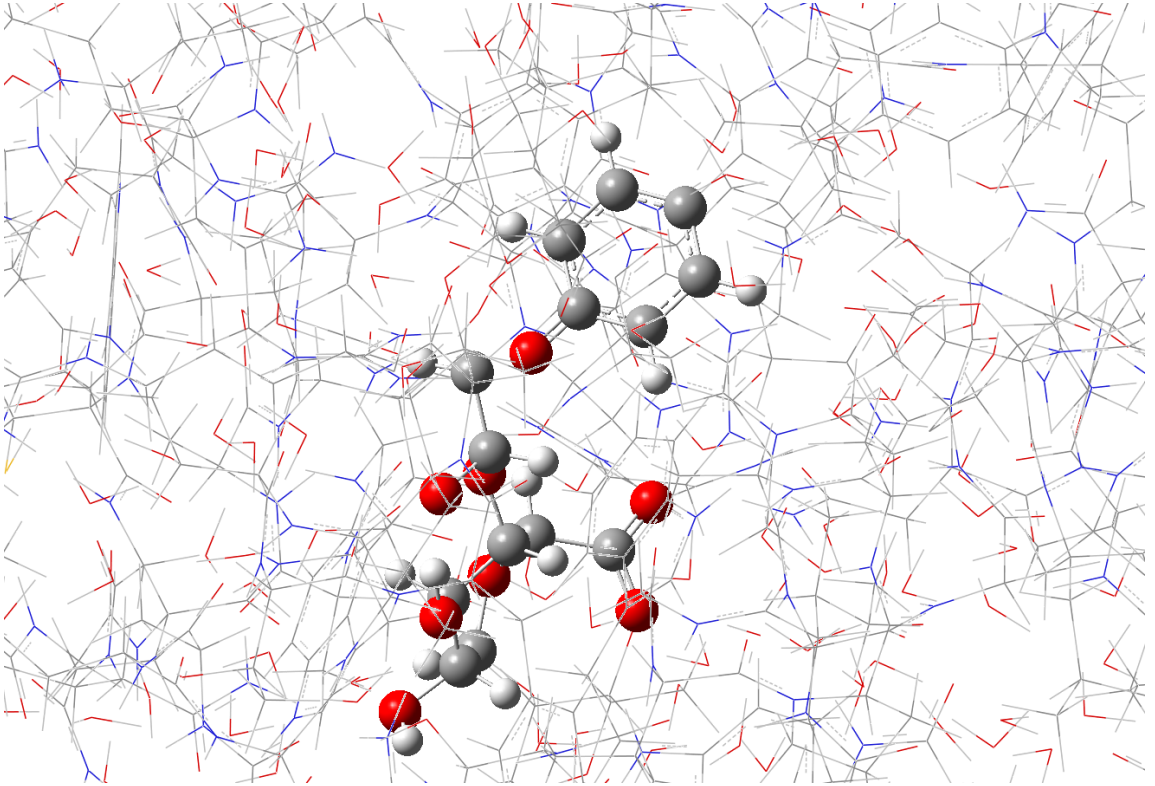




QM/MM: EXAMPLE

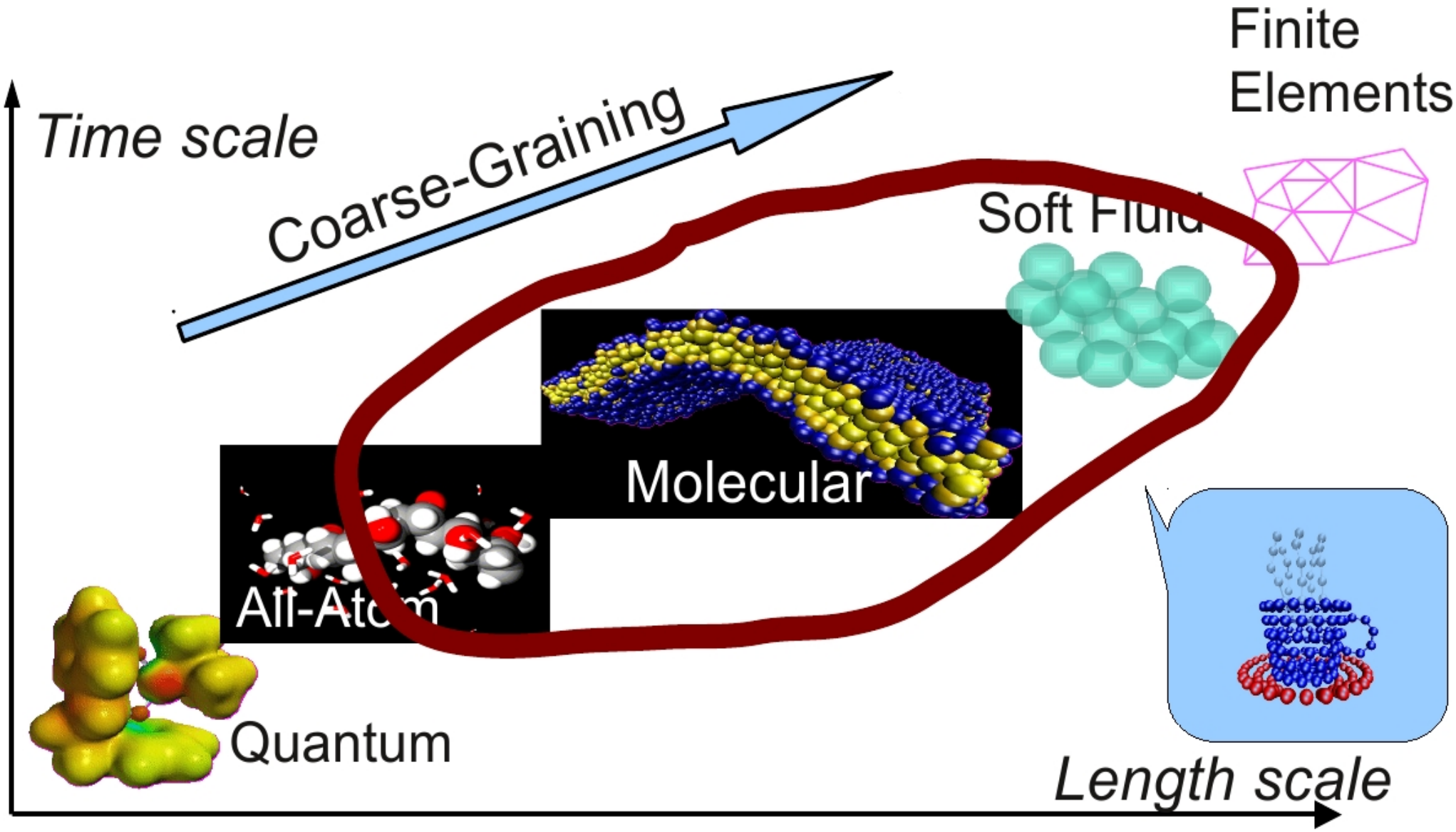
Reaction of proton transfer in two steps





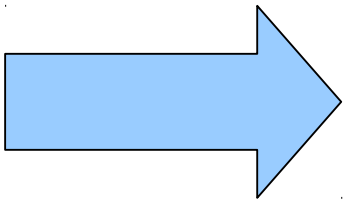
- **Structures of substrates and products**
- **Energetic properties of transitional complex**
- **Consideration of protein environment**

MD: TIME/LENGTH SCALES



FORCE FIELDS IN MD

- AMBER
- GROMOS
- CHARMM
- ...



Proteins

DNA

Sugars

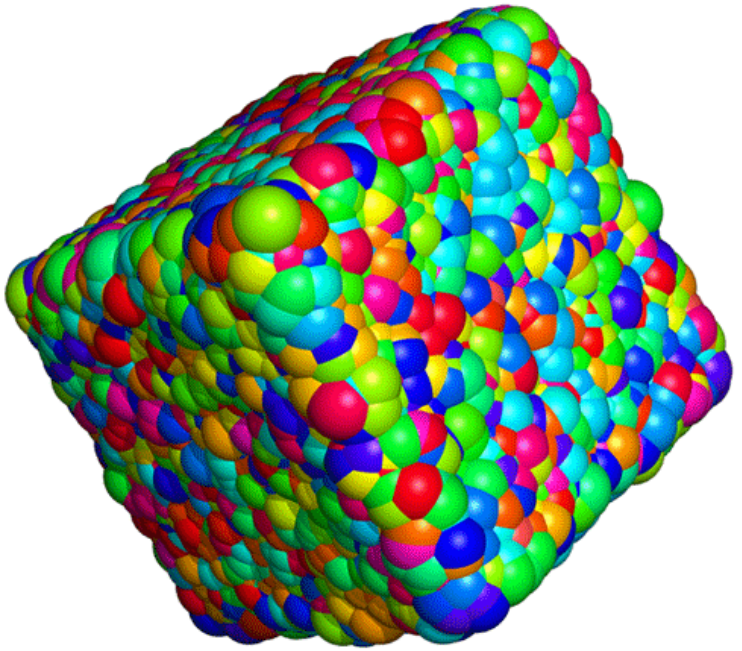
Lipids

Small molecules

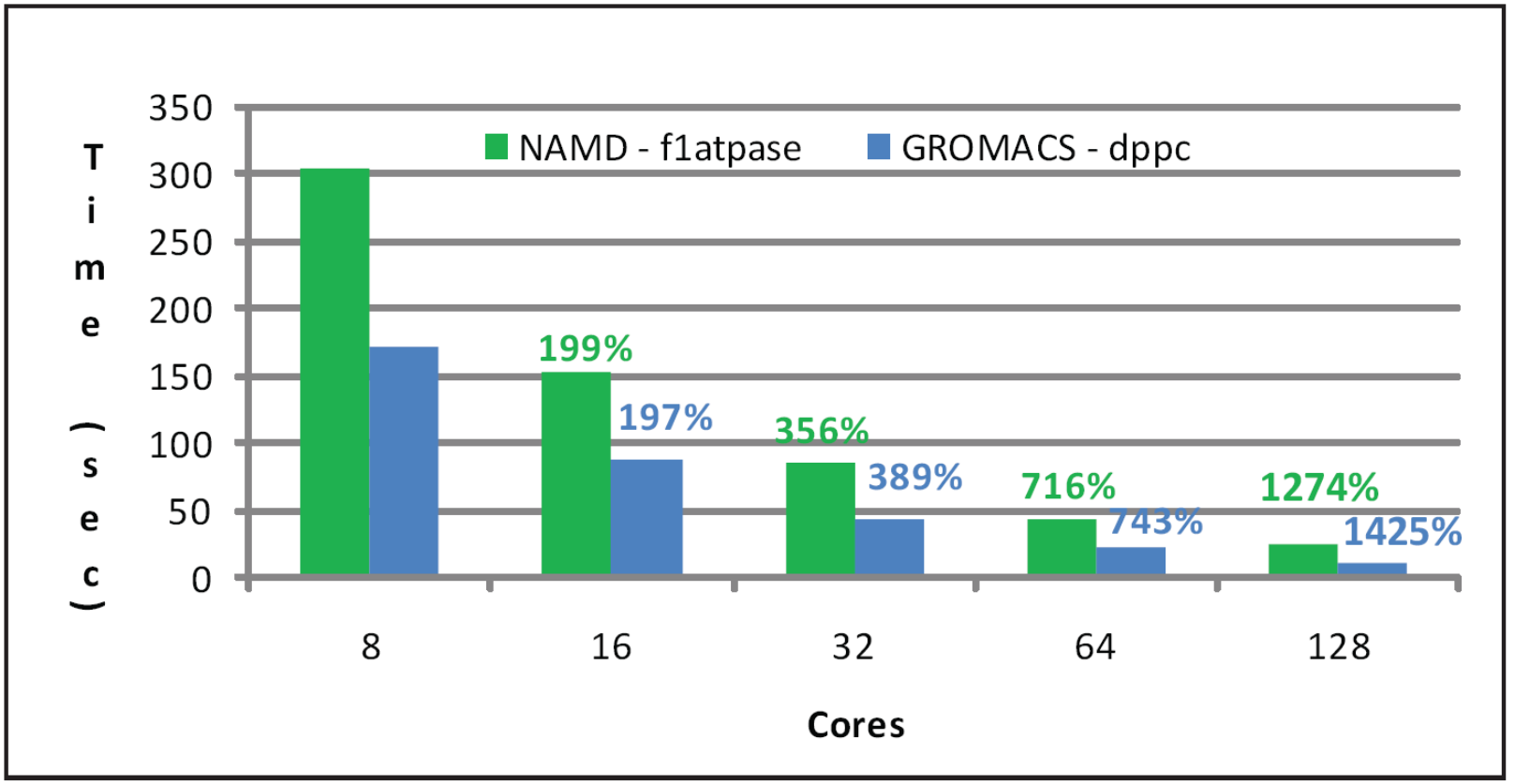
MD SOFTWARE

- GROMACS (Open Source)
- AMBER
- CHARMM
- NAMD (Open Source)
- TINKER

...



MD SCALING, CALCULATION TIME



- **Scalability: how much you speed up the parallel calculations, when increase number of cores used**

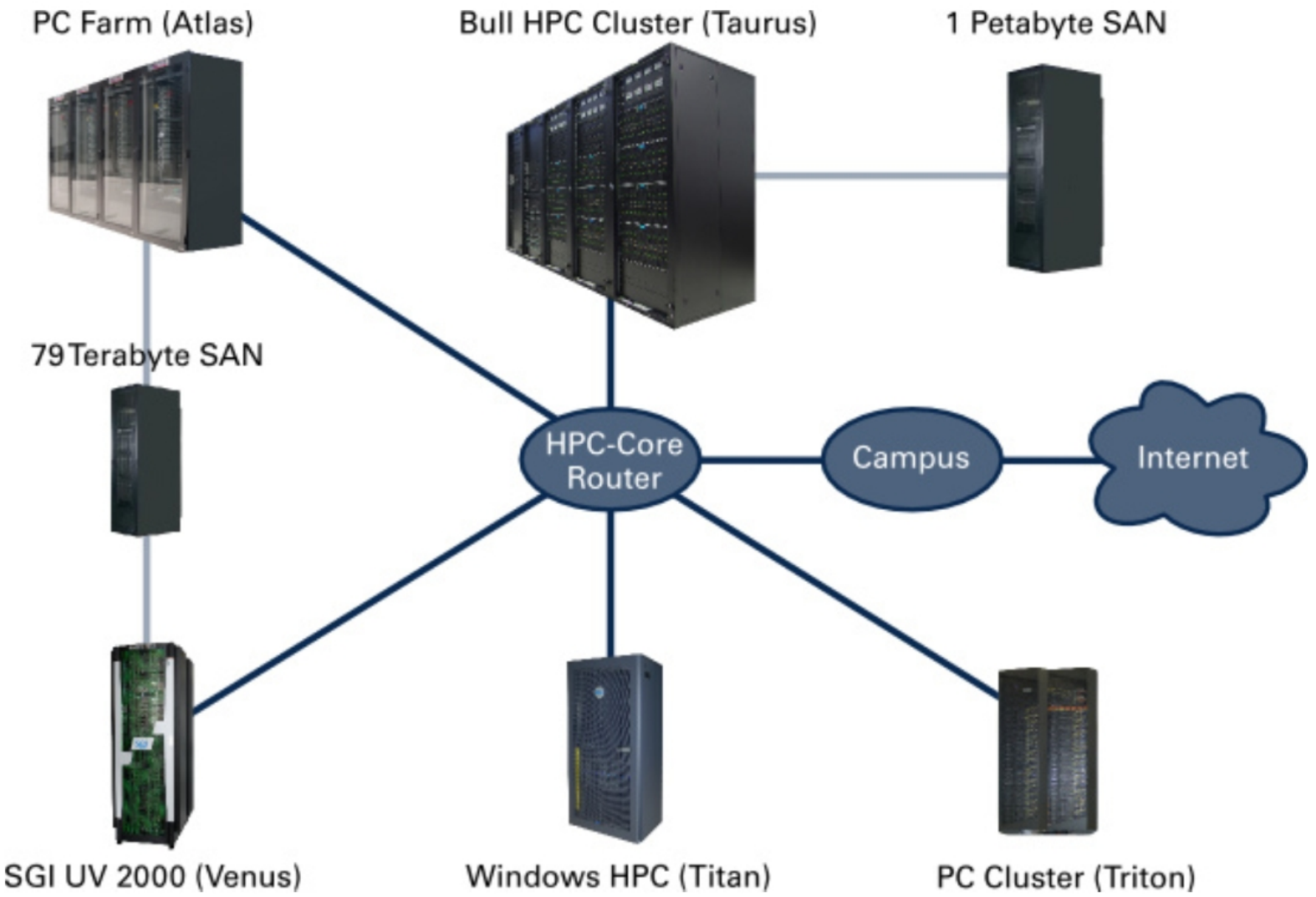
$$Scalability = \frac{t_{1\,core}}{t_{n\,cores}}$$

Example: 10ns, ~10000 atoms, AMBER 10, 32 cores, 2 days

DRESDEN SUPERCOMPUTER CENTER

http://tu-dresden.de/die_tu_dresden/zentrale_einrichtungen/zih/hpc

HIGH PERFORMANCE COMPUTING (HPC) AT ZIH



DRESDEN SUPERCOMPUTER CENTER

Bull HPC-Cluster (Taurus)

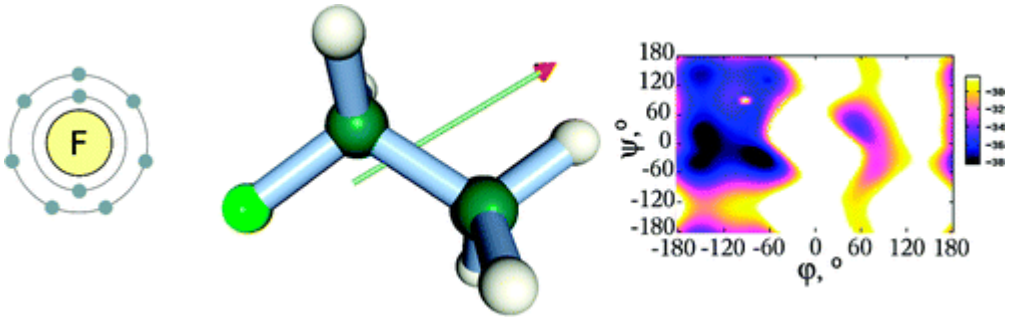
- **For highly parallel HPC applications**
- Island 1: 4320 cores Intel E5-2690 (Sandy Bridge) 2.90GHz
- Island 2: 704 cores Intel E5-2450 (Sandy Bridge) 2.10GHz + 88 NVidia Tesla K20x GPUs
- Island 3: 2160 cores Intel X5660 (Westmere) 2.80GHz
- SMP nodes with 1TB main memory
- 1 PB SAN disk storage
- Bullx Linux 6.3, batch system Slurm
- 137 TFlop/s total peak performance (w/o GPUs)

Megware PC-Farm (Atlas)

- **For capacity computing**
- 5888 cores AMD Opteron 6274 2,2 GHz
- 13 TB memory
- 92 nodes with 64 cores each and 64 to 512 GB memory
- 79 TB SAN disk storage
- SuSE Linux Enterprise Server 11, batch system LSF
- 51,8 TFlop/s peak performance

CASE STUDY: FLUORINATED AMINO ACIDS

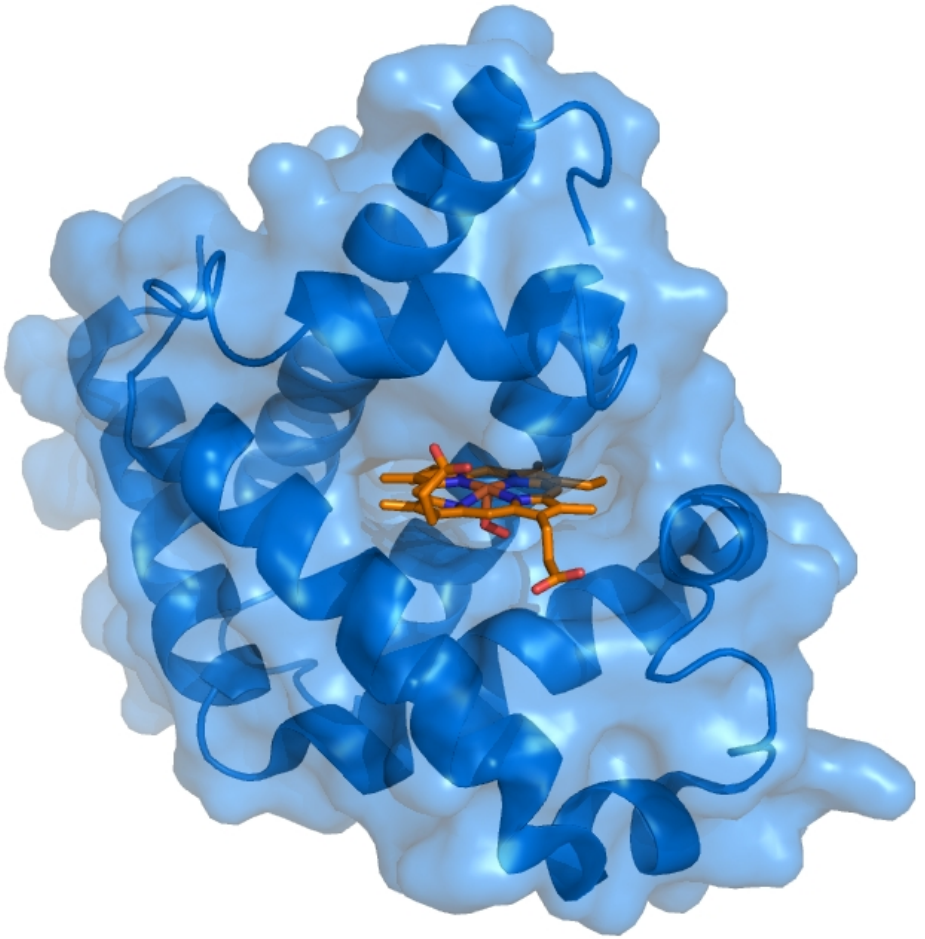
- **Aim: theoretical analysis of fluorinated amino acids in protein environments**
- **Object: fluorinated amino acids**
- **Experimental data: calorimetry, HPLC**



NON-NATURAL AMINO ACIDS

Peptides/proteins engineering

Incorporation of non-natural amino acids



Improved properties not feasible within natural proteins

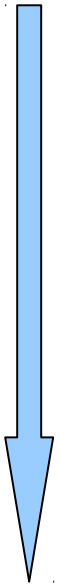
FLUORINE UNIQUE PROPERTIES

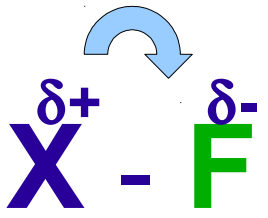
F: $1S^2\ 2S^2\ 2P^5$

Electronegativity (χ)

$$\chi_A - \chi_B = (\mathrm{eV})^{-1/2}\sqrt{E_d(AB) - [E_d(AA) + E_d(BB)]/2}$$

H	2.20
C	2.55
S	2.58
N	3.04
O	3.44
F	3.98

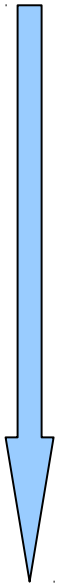




High χ and small r_a

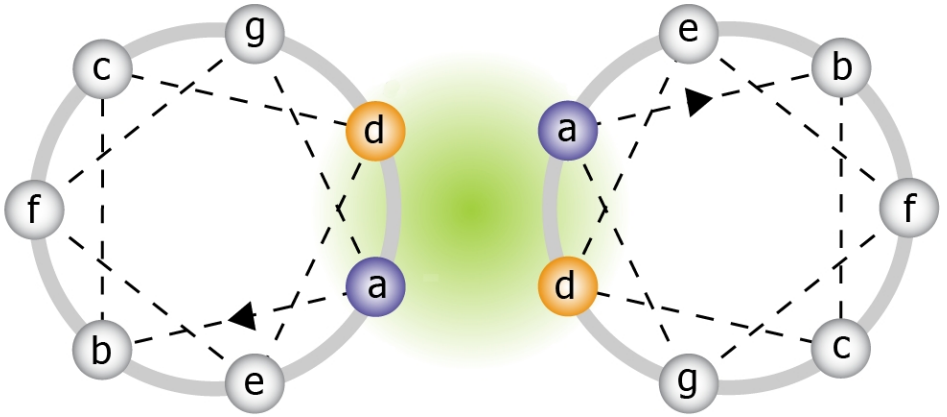
Atomic radius, Å

H	1.20
F	1.47
O	1.52
N	1.55
C	1.70
S	1.80



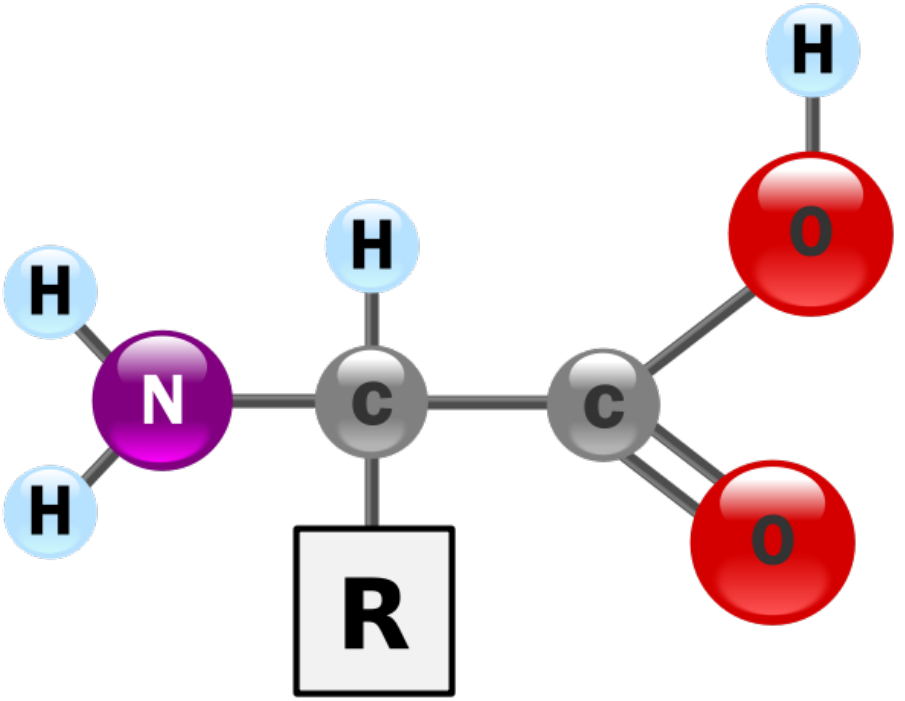
GOALS

- **Theoretical characterization of fluorinated amino acids**
- **Study of fluorinated amino acids in protein environments**



FLUORINATED AMINO ACIDS

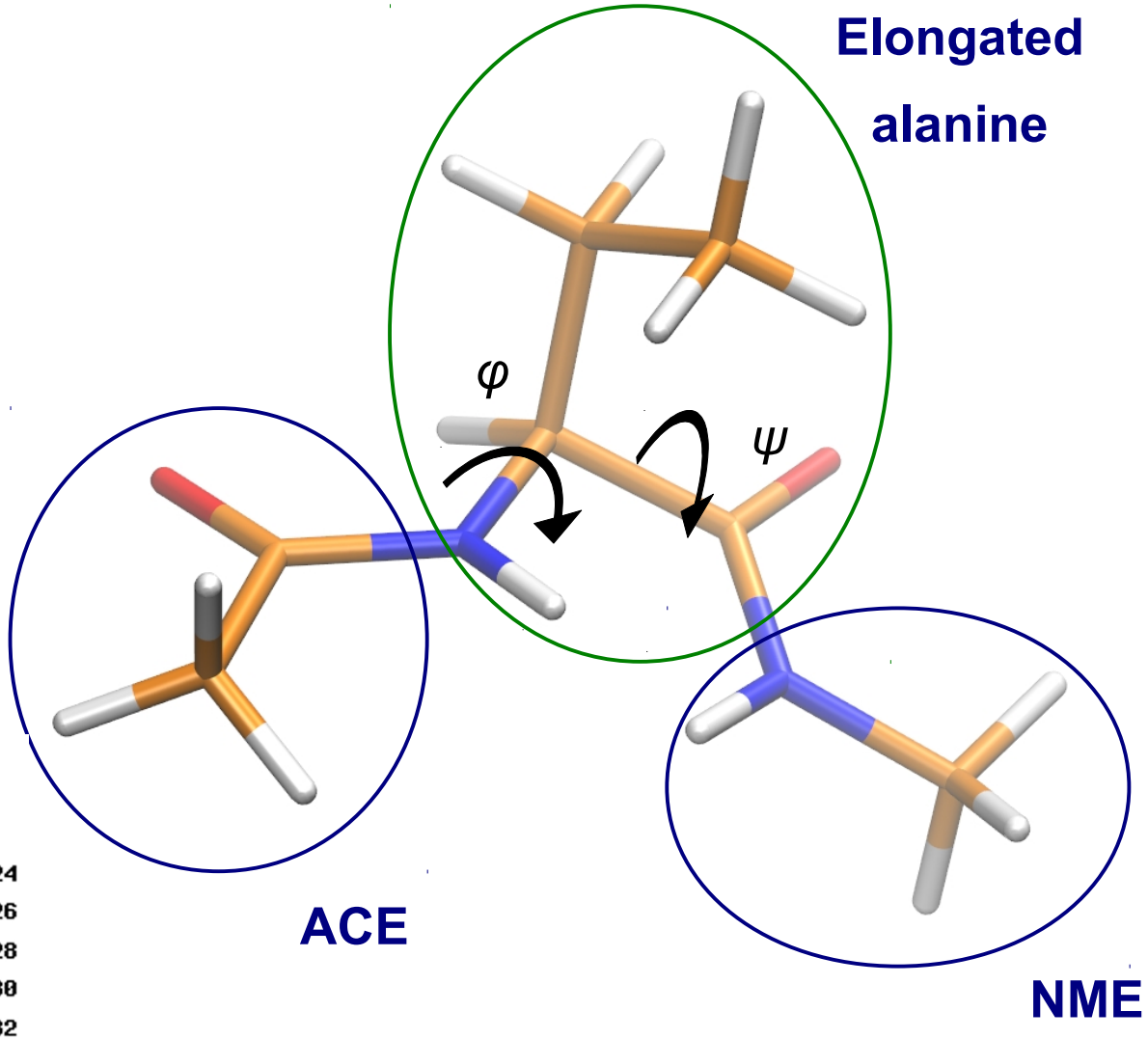
AA	R-group
Eal	$-CH_2-CH_3$
MfeGly	$\mathbf{-CH_2-CFH_2}$
DfeGly	$\mathbf{-CH_2-CF_2H}$
TfeGly	$\mathbf{-CH_2-CF_3}$
DfpGly	$\mathbf{-CH_2-CF_2CH_3}$

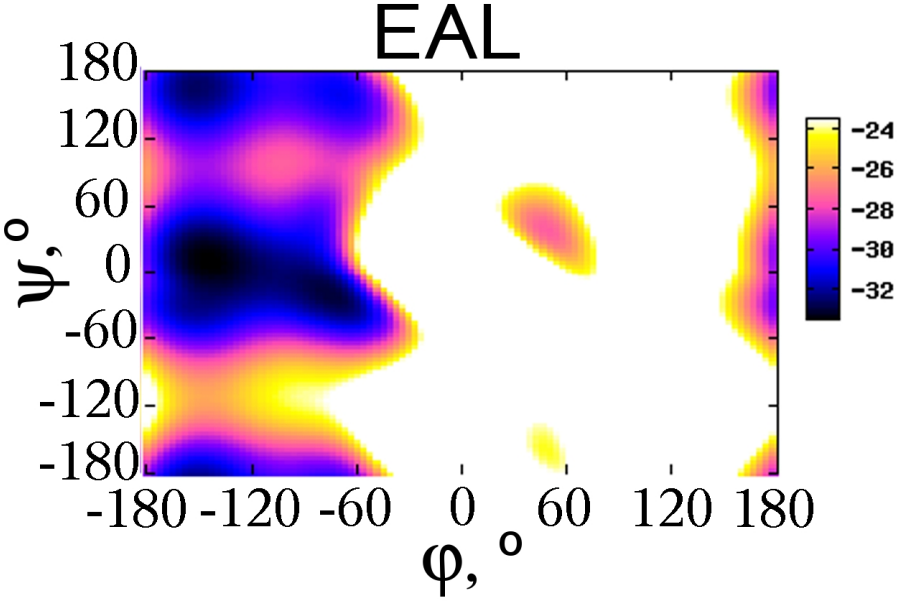


Non-natural residue libraries for AMBER ff (RESP charges, 6-31G*, α/Ext).

SINGLE AA ANALYSIS EXAMPLE

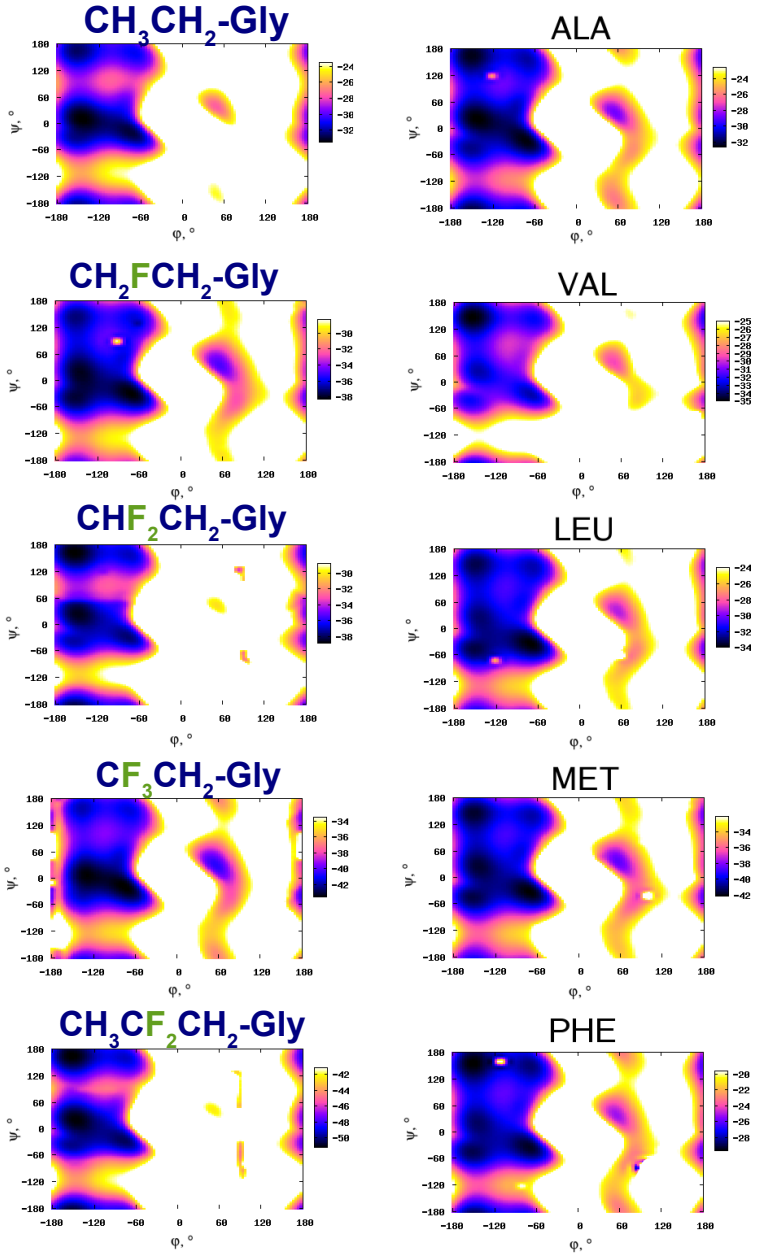
- Dipeptide: ACE-Xxx-NME
- φ,ψ ∈ [-180;180], step 5º
- Minimization with fixed (φ_i ,ψ_j)
- 36*36=1296 points with $E_{i,j}$



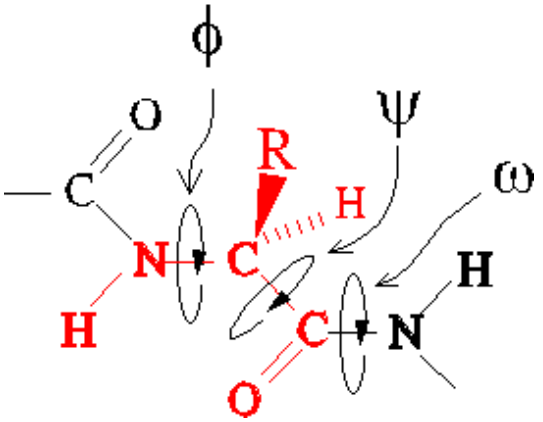


$$P(\phi,\psi)=e^{\frac{-(E(\phi,\psi)-E_{min})}{kT}}$$

SECONDARY STRUCTURE PROPERTIES



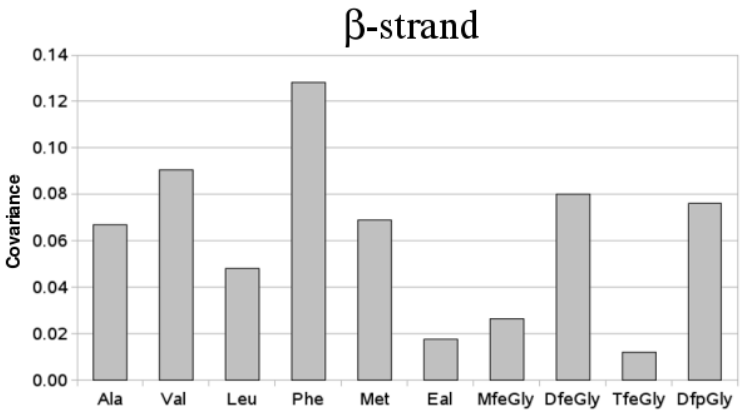
Ace-Xxx-Nme

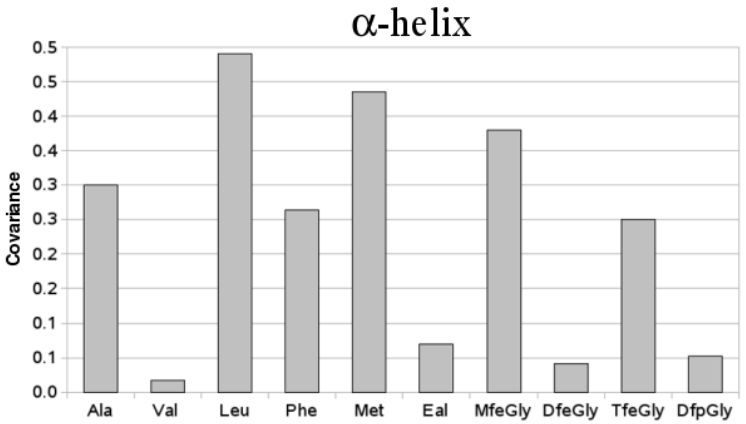


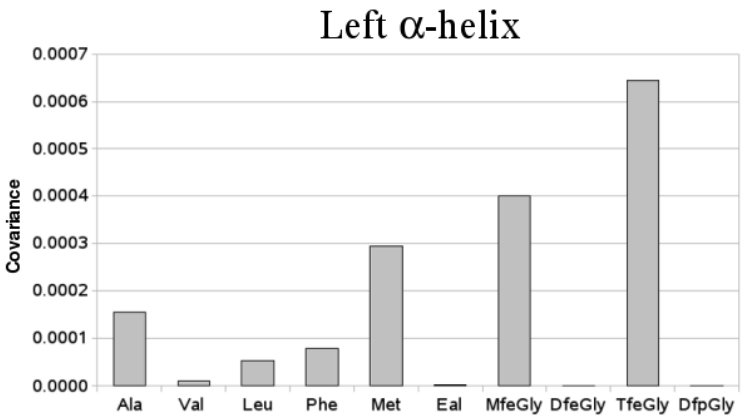
$$Cov = \sum_{\phi,\psi=-180}^{180} w_{\phi,\psi} \exp \frac{-\left(E_{\phi,\psi} - E_{\min}\right)}{RT}$$

PDB

$$w_{\phi_0,\psi_0} = \sum_{\phi \in [\phi_0-5;\phi_0+5),\psi \in [\psi_0-5;\psi_0+5)} \frac{N_{\phi,\psi}}{N_{total}}$$

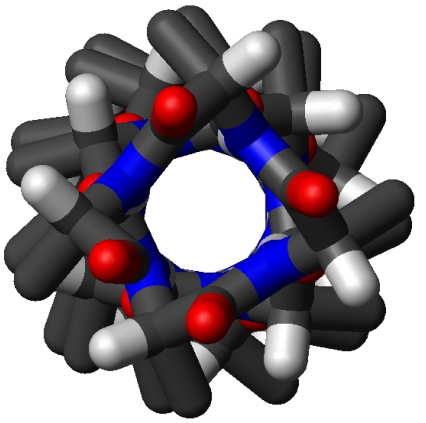






CH_2FCH_2-Gly

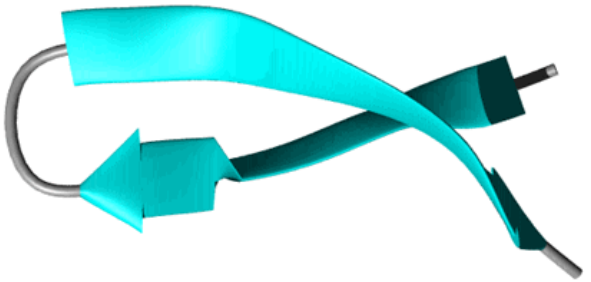
CF_3CH_2-Gly



α-helix

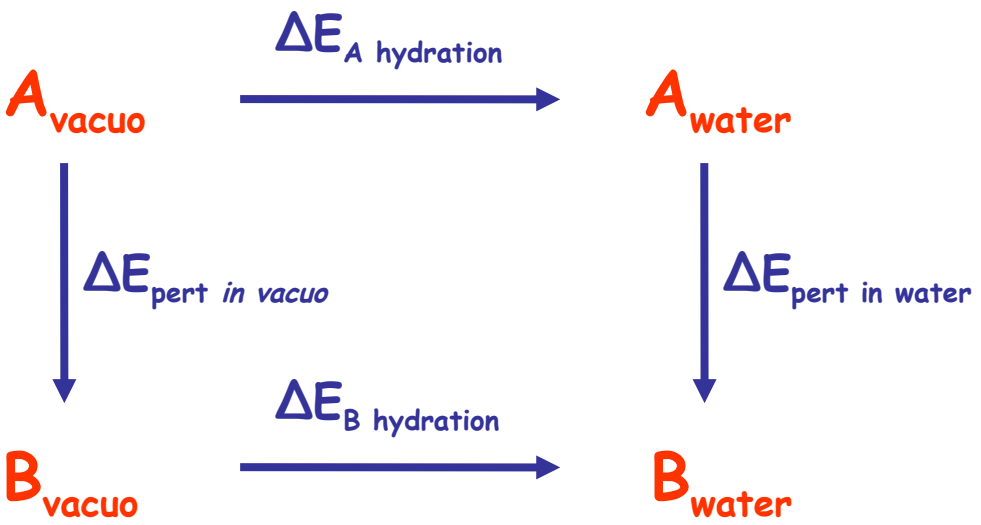
CHF_2CH_2-Gly

$CH_3CF_2CH_2$-Gly



β-strand

HYDRATION ENERGY CALCULATIONS



Free energy perturbation:

$$\Delta E \approx \Delta V = \int_{\lambda=0}^{\lambda=1} \frac{dV}{d\lambda} d\lambda$$

$$\Delta E_{A\ hydration} - \Delta E_{B\ hydration} = \Delta E_{pert\ in\ vacuo} - \Delta E_{pert\ in\ water}$$

CH_2FCH_2-Gly

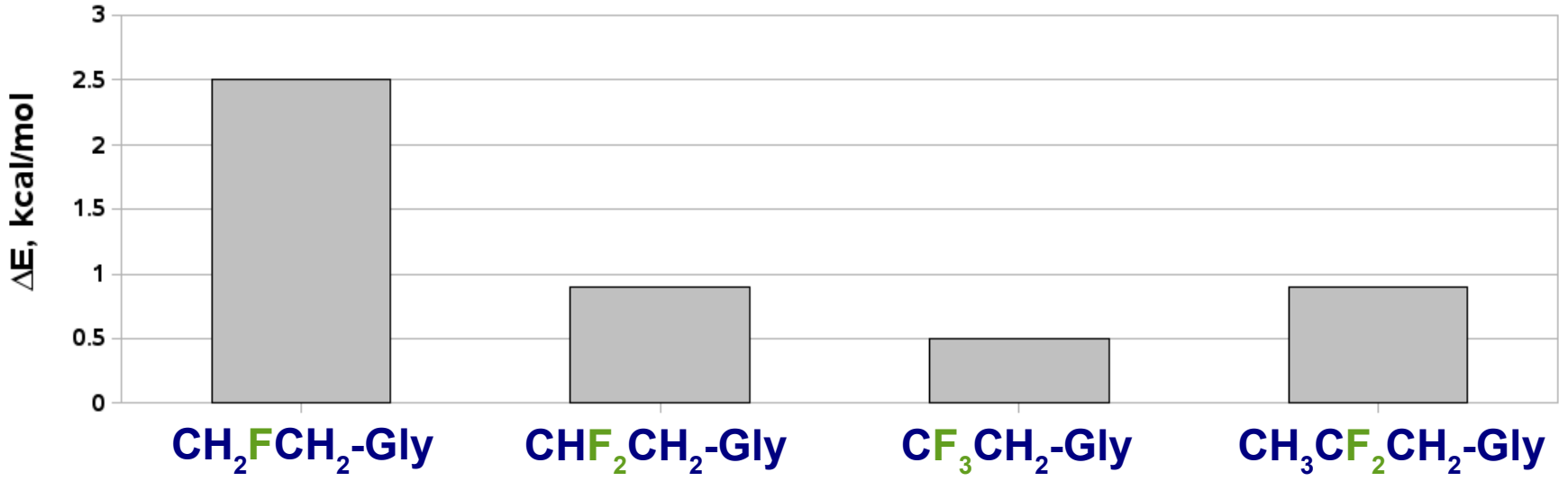
CHF_2CH_2-Gly

CF_3CH_2-Gly

$\longrightarrow$ CH_3CH_2-Gly

$CH_3CF_2CH_2$-Gly $\longrightarrow$ $CH_3CH_2CH_2$-Gly

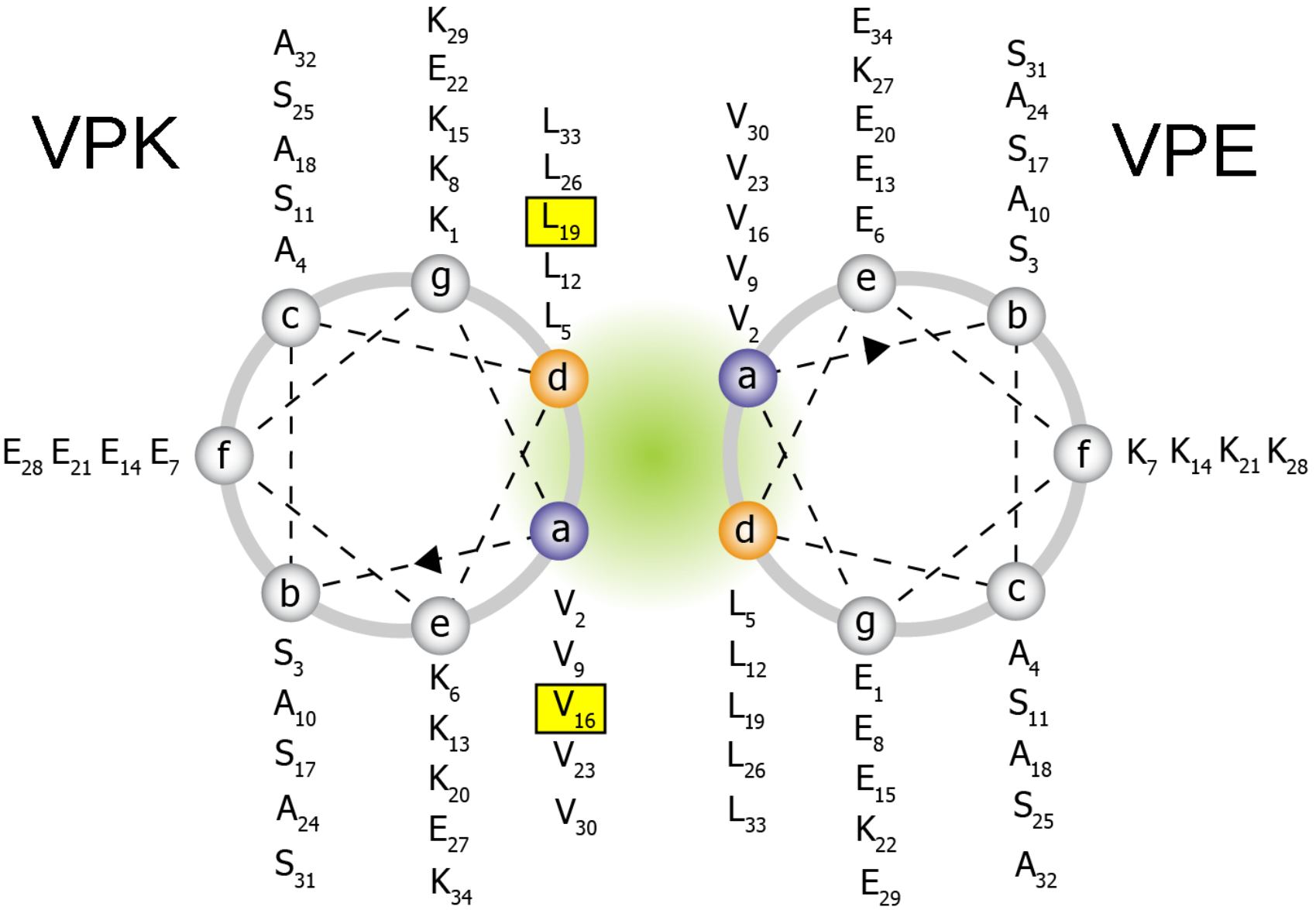
HYDRATION ENERGY CHANGE DUE TO DEFLUORINATION



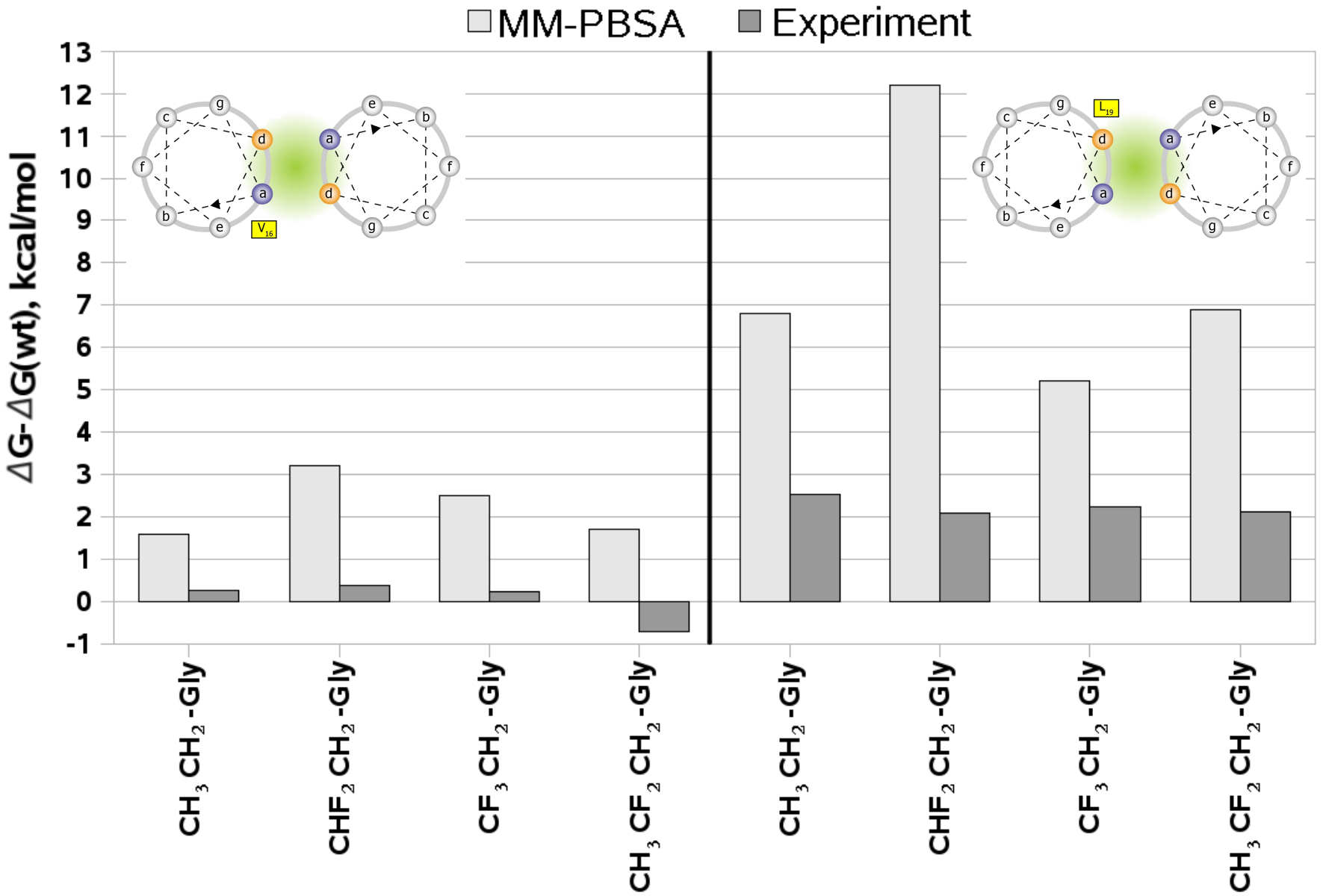
Electrostatics **Non-polar solvation**

Hydrophobicity

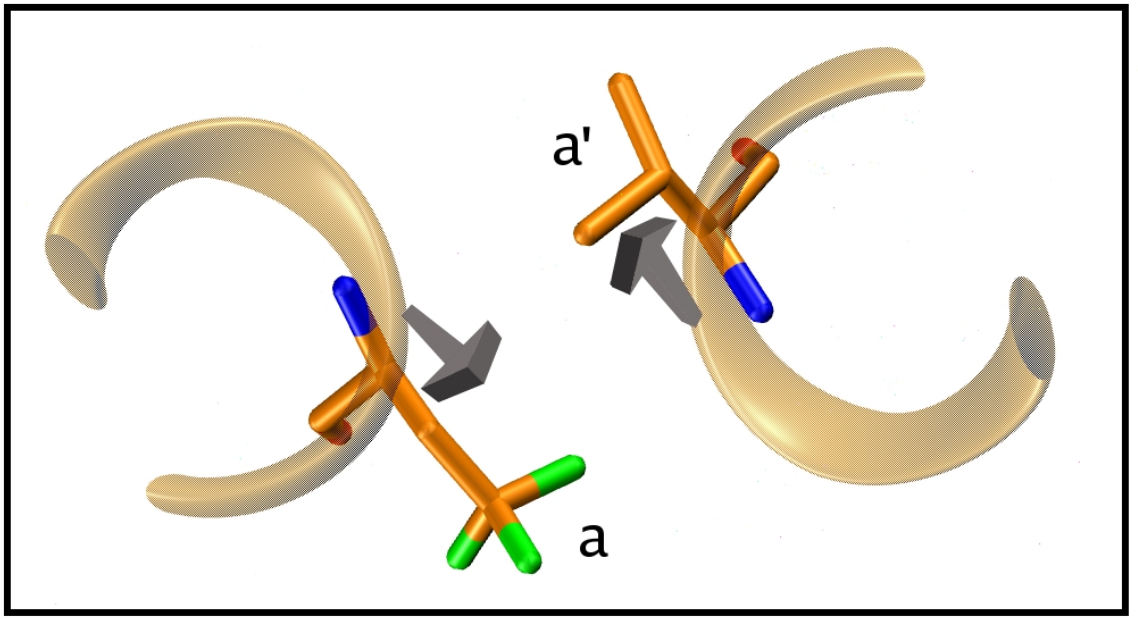
PARALLEL COILED COIL SYSTEM

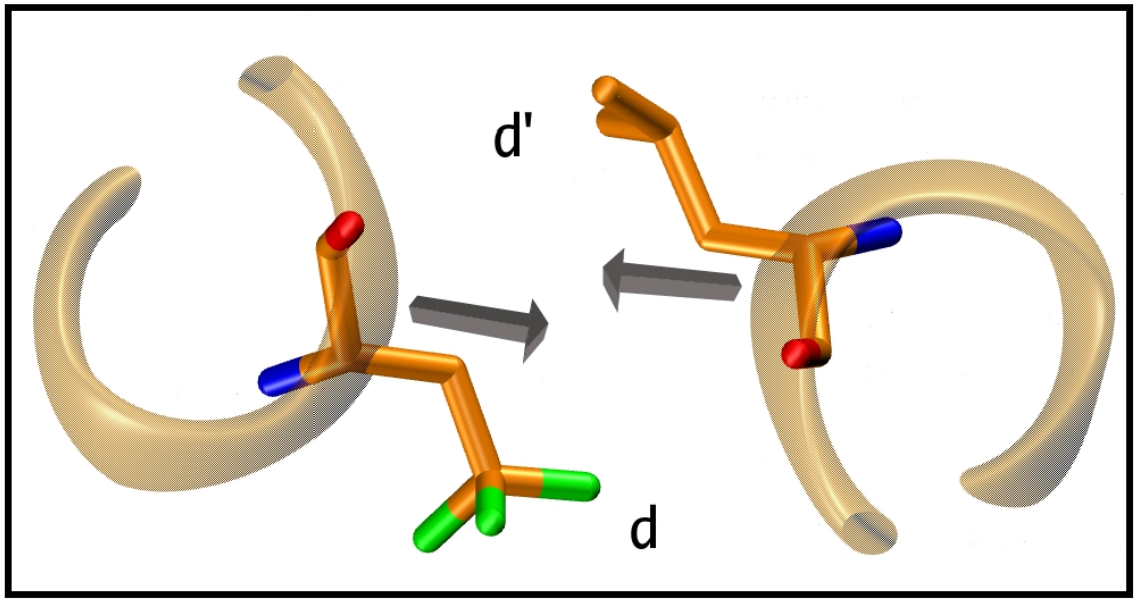


MD vs CD: COILED COIL UNFOLDING



PACKING AT a16 AND d19 POSITIONS





$d(C_\beta\text{-}C_{\beta'})_{a16} - d(C_\beta\text{-}C_{\beta'})_{d19} \sim 1Å$

Fluorine substitution effect:

1. polarity
2. steric demand
3. packing

LECTURE 2: OUTLINE
MOLECULAR DYNAMICS

- MD pipeline and principles
- Minimization
- MD with fixed physico-chemical parameters
- Energy calculations
- Biased MD
- Special cases of MD
- MD software
- Case study: fluorinated amino acids