INSTITUTE OF MOLECULAR AND
TRANSLATIONAL MEDICINE

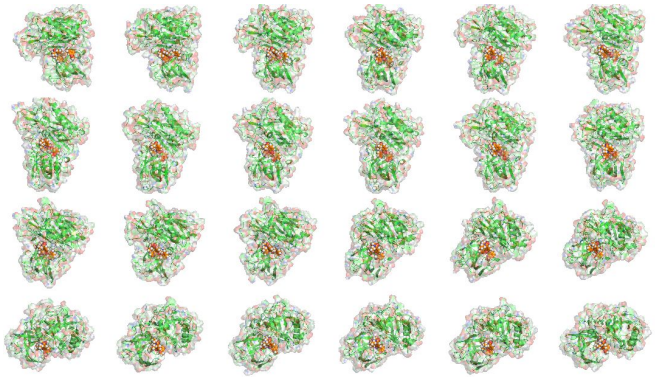
Molecular dynamics

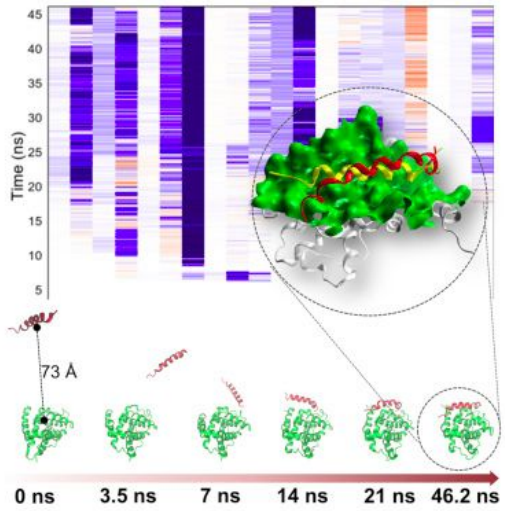
PhD student: Ivanova Aleksandra
Supervisor: Pavel Polishchuk

Molecular dynamics

1. MD simulations **mimic the physical motions of atoms present in the actual environment;**
2. The atoms and molecules are allowed to interact for a fixed period of time, giving **a view of the dynamic "evolution" of the system.**

- to explore conformational space
- to explore biological process of molecular recognition
- to investigate ligand pose stability
- to explore protein flexibility
- to estimate binding affinity of protein-ligand (protein-protein) complexes





https://doi.org/10.1016/j.str.2017.02.009

Run MD simulation (10 ps):

https://privatecloud.imtm.cz/s/xgUMAHX2iX1TJmg

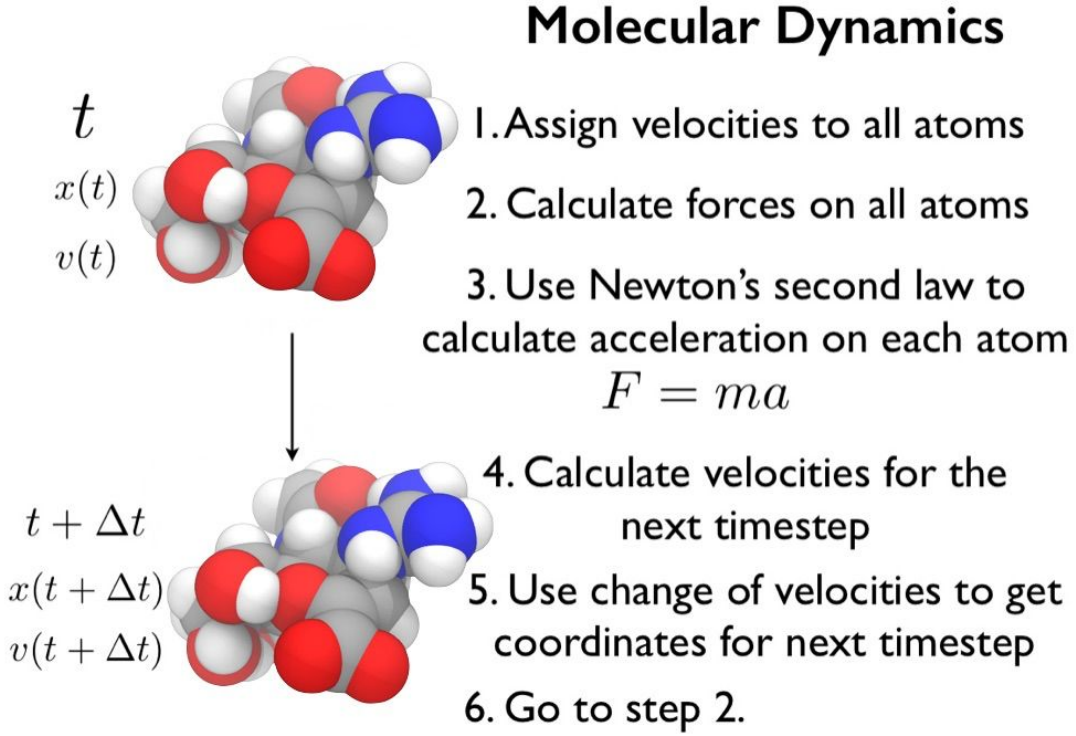
mkdir md_tutorial_student
cd md_tutorial
conda activate md

pip uninstall streamd
pip install git+https://github.com/ci-lab-cz/md-scripts.git

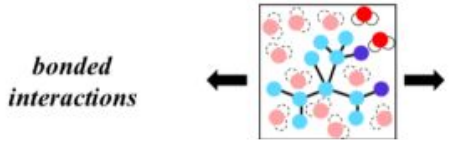
run_md -p protein_HIS.pdb -l ligand.mol --md_time 0.01 --nvt_time 10 --npt_time 10 --ncpu 8 -d mdrun

Classical Molecular Dynamics

Simulation process is based on Newton's second law



Theory: molecular interactions

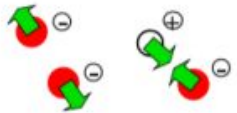


bonded interactions ← → ***non-bonded interactions***

Bond stretching $V^{bond}(\mathbf{r}^N) = \sum_{bonds\ i} \frac{1}{2} K_i^b \left[b_i(\mathbf{r}^N) - b_i^0 \right]^2$

Angle bending $V^{angle}(\mathbf{r}^N) = \sum_{angles\ i} \frac{1}{2} K_i^a \left[\theta_i(\mathbf{r}^N) - \theta_i^0 \right]^2$

Rotation around bond $V^{torsion}(\mathbf{r}^N) = \sum_{torsion\ i} K_i^\varphi \cos\left[m_i \varphi_i(\mathbf{r}^N) + \delta_i \right]$



Electrostatic interactions

van der Waals interactions

All atoms see each other through space
~ ½ N (N – 1) interactions

Goals of classical (semi-empirical) force fields

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

The computational bottleneck

Theory: force fields

All force fields were developed using different experimental data and for slightly different purposes

The most popular FFs for protein and protein-ligand simulations are:

CHARMM – proteins, lipids, nucleic acids

AMBER – peptide, protein, nucleic acids, lipid14, GAFF

OPLS - optimized to fit experimental properties of liquids, such as density and heat of vaporization

GROMOS – proteins, small molecules, also optimized for experimental parameters

Each FF works with a particular set of simulation parameters → they should not be mixed!

Implementation of calssical potential energy functions

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

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.

examples of *.mdp files

ions.mdp

```
; ions.mdp - used as input into grompp to generate ions.tpr
; Parameters describing what to do, when to stop and what to save
integrator  = steep         ; Algorithm (steep = steepest descent minimization)
emtol       = 1000.0        ; Stop minimization when the maximum force < 1000.0 kJ/mol/nm
emstep      = 0.01          ; Minimization step size
nsteps      = 50000         ; Maximum number of (minimization) steps to perform

; Parameters describing how to find the neighbors of each atom and how to calculate the interactions
cutoff-scheme   = Verlet    ; Buffered neighbor searching
ns_type         = grid      ; Method to determine neighbor list (simple, grid)
coulombtype     = cutoff    ; Treatment of long range electrostatic interactions
rcoulomb        = 1.0       ; Short-range electrostatic cut-off
rvdw            = 1.0       ; Short-range Van der Waals cut-off
pbc             = xyz       ; Periodic Boundary Conditions in all 3 dimensions
```

em.mdp

```
; minim.mdp - used as input into grompp to generate em.tpr
; Parameters describing what to do, when to stop and what to save
integrator  = steep         ; Algorithm (steep = steepest descent minimization)
emtol       = 1000.0        ; Stop minimization when the maximum force < 1000.0 kJ/mol/nm
emstep      = 0.01          ; Minimization step size
nsteps      = 50000         ; Maximum number of (minimization) steps to perform

; Parameters describing how to find the neighbors of each atom and how to calculate the interactions
cutoff-scheme   = Verlet    ; Buffered neighbor searching
ns_type         = grid      ; Method to determine neighbor list (simple, grid)
coulombtype     = PME       ; Treatment of long range electrostatic interactions
rcoulomb        = 1.0       ; Short-range electrostatic cut-off
rvdw            = 1.0       ; Short-range Van der Waals cut-off
pbc             = xyz       ; Periodic Boundary Conditions in all 3 dimensions
```

md.mdp

```
title                    = OPLS Lysozyme NPT equilibration
; Run parameters
integrator               = md        ; leap-frog integrator
nsteps                   = 5000000    ; 2 * 5000000 = 10000 ps (10 ns)
dt                       = 0.002     ; 2 fs
; Output control
nstxout                  = 0         ; suppress bulky .trr file by specifying
nstvout                  = 0         ; 0 for output frequency of nstxout,
nstfout                  = 0         ; nstvout, and nstfout
nstenergy                = 5000      ; save energies every 10.0 ps
nstlog                   = 5000      ; update log file every 10.0 ps
nstxout-compressed       = 5000      ; save compressed coordinates every 10.0 ps
compressed-x-grps        = System    ; save the whole system
; Bond parameters
continuation             = yes       ; Restarting after NPT
constraint_algorithm     = lincs     ; holonomic constraints
constraints              = h-bonds   ; bonds involving H are constrained
lincs_iter               = 1         ; accuracy of LINCS
lincs_order              = 4         ; also related to accuracy
; Neighborsearching
cutoff-scheme            = Verlet    ; Buffered neighbor searching
ns_type                  = grid      ; search neighboring grid cells
nstlist                  = 10        ; 20 fs, largely irrelevant with Verlet scheme
rcoulomb                 = 1.0       ; short-range electrostatic cutoff (in nm)
rvdw                     = 1.0       ; short-range van der Waals cutoff (in nm)
; Electrostatics
coulombtype              = PME       ; Particle Mesh Ewald for long-range electrostatics
pme_order                = 4         ; cubic interpolation
fourierspacing           = 0.16      ; grid spacing for FFT
; Temperature coupling is on
tcoupl                   = V-rescale             ; modified Berendsen thermostat
tc-grps                  = Protein Non-Protein   ; two coupling groups - more accurate
tau_t                    = 0.1     0.1           ; time constant, in ps
ref_t                    = 300     300           ; reference temperature, one for each group, in K
; Pressure coupling is on
pcoupl                   = Parrinello-Rahman     ; Pressure coupling on in NPT
pcoupltype               = isotropic             ; uniform scaling of box vectors
tau_p                    = 2.0                   ; time constant, in ps
ref_p                    = 1.0                   ; reference pressure, in bar
compressibility          = 4.5e-5                ; isothermal compressibility of water, bar^-1
; Periodic boundary conditions
pbc                      = xyz       ; 3-D PBC
; Dispersion correction
DispCorr                 = EnerPres  ; account for cut-off vdW scheme
; Velocity generation
gen_vel                  = no        ; Velocity generation is off
```

Theory: solvent and periodic boundary conditions (PBC)

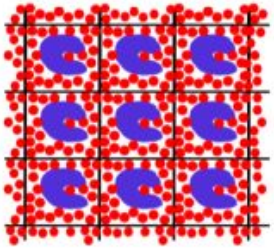
Vacuum

- Surface effects (surface tension)
- No dielectric screening

Droplets

- Still surface effects
- Only partial dielectric screening
- Evaporation of the solvent

Periodic: *rectangular system is surrounded by copies of itself*



Advantage:
- No surface effects

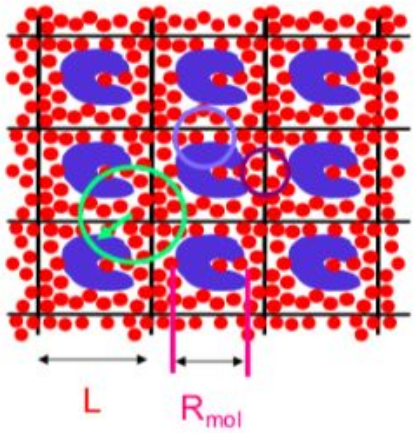
Disadvantage:
- Artificial periodicity
- High effective concentration

Probably still the best approach…

Periodic Boundary Conditions

To simulate an infinite system, **periodic boundary conditions** are often applied. This allows atoms that leave one side of the simulation box to re-enter on the opposite side.

- PBCs are often used in computer simulations and mathematical models.
- The topology of two-dimensional PBC is equal to that of a world map of some video games; the geometry of the unit cell satisfies perfect two-dimensional tiling, and when an object passes through one side of the unit cell, it reappears on the opposite side with the same velocity.



Required
(no atom sees another one twice):

$L > 2R_c$ $R_c < ½L$

Preferred
(protein does not see a copy of itself)

$L > R_c + R_{mol}$ $R_c < L - R_{mol}$

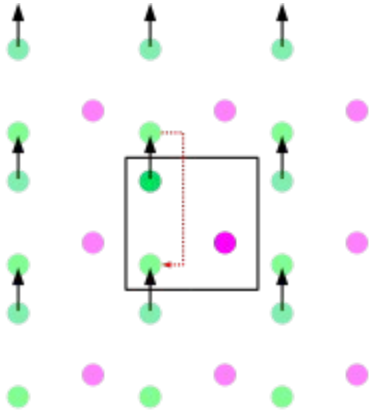
Even better
(no solvent sees two proteins)

$L > 2R_c + R_{mol}$ $R_c < ½(L - R_{mol})$

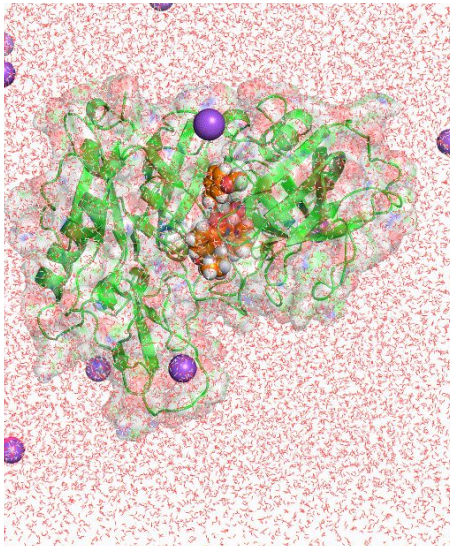
<nu mbe r>

- Cubic:

Rectangular:



Classical Molecular Dynamics



Structure Preparation

Preprocessing

Force-field
- Definitions of inter-atomic bonded and non-bonded forces (ligand and protein)

Simulation box setup
- box size/shape

Solvate system
- add HOH molecules

Neutralize system
- add NA+/CL- ions

Pre-simulation steps

Energy minimization
- 100 ps
- to ensure that the system has no steric clashes or inappropriate geometry

NVT and NPT equilibration
- 1000 ps
- equilibrate the solvent and ions around the protein

Product Simulation

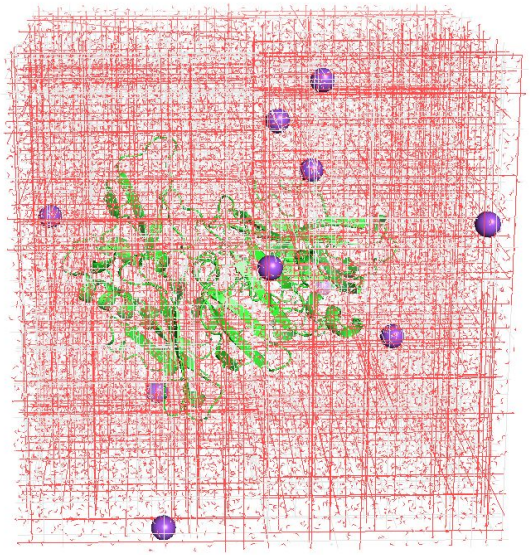
Analysis

FAST. FLEXIBLE. FREE.
GROMACS

Classical Molecular Dynamics

After we got simulated trajectory:

1) Remove **Periodic boundary conditions**

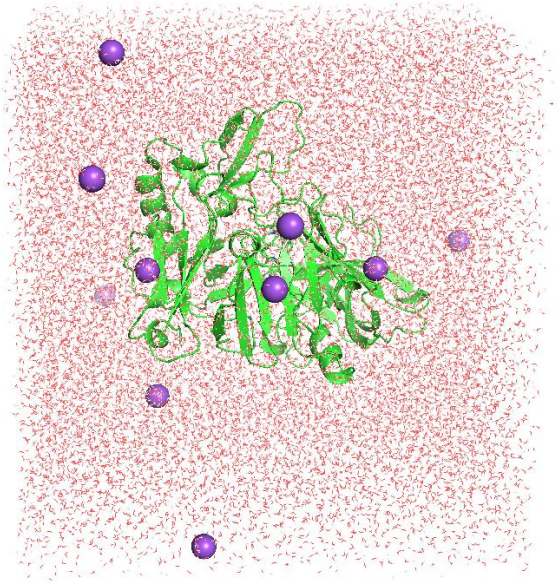


Classical Molecular Dynamics

After we got simulated trajectory:

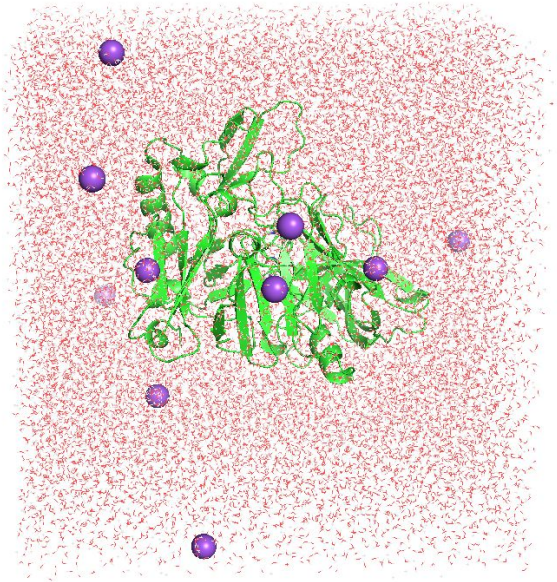
2) Center system (solvate, ions, atoms) over protein

3) Remove rotations of the system



Classical Molecular Dynamics

Now you are ready to analyze your trajectory



Metrics

1) **Root-mean-square deviation** (**RMSD**):

the RMSD is a measure of the difference between a initial conformation of the ligand/protein conformation and the analysed conformation

Root mean square deviations in structure

gmx rms, gmx rmsdist

The *root mean square deviation* ($RMSD$) of certain atoms in a molecule with respect to a reference structure can be calculated with the program gmx rms by least-square fitting the structure to the reference structure ($t_2 = 0$) and subsequently calculating the $RMSD$ ((458)).

$$RMSD(t_1, t_2) = \left[\frac{1}{M} \sum_{i=1}^{N} m_i \|\mathbf{r}_i(t_1) - \mathbf{r}_i(t_2)\|^2 \right]^{\frac{1}{2}} \quad (458)$$

where $M = \sum_{i=1}^{N} m_i$ and $\mathbf{r}_i(t)$ is the position of atom i at time t. **Note** that fitting does not have to use the same atoms as the calculation of the $RMSD$; *e.g.* a protein is usually fitted on the backbone atoms (N, C_α, C), but the $RMSD$ can be computed of the backbone or of the whole protein.

Instead of comparing the structures to the initial structure at time $t = 0$ (so for example a crystal structure), one can also calculate (458) with a structure at time $t_2 = t_1 - \tau$. This gives some insight in the mobility as a function of τ. A matrix can also be made with the $RMSD$ as a function of t_1 and t_2, which gives a nice graphical interpretation of a trajectory. If there are transitions in a trajectory, they will clearly show up in such a matrix.

Alternatively the $RMSD$ can be computed using a fit-free method with the program gmx rmsdist:

$$RMSD(t) = \left[\frac{1}{N^2} \sum_{i=1}^{N} \sum_{j=1}^{N} \|\mathbf{r}_{ij}(t) - \mathbf{r}_{ij}(0)\|^2 \right]^{\frac{1}{2}} \quad (459)$$

where the *distance* $\mathbf{r}_{ij}$ between atoms at time t is compared with the distance between the same atoms at time 0.

Metrics

2) **Radius of gyration** (**Rgyr**):

To have a rough measure for the compactness of a structure, you can calculate the *radius of gyration* with the program gmx gyrate as follows:

$$R_g = \left(\frac{\sum_i \|\mathbf{r}_i\|^2 m_i}{\sum_i m_i} \right)^{\frac{1}{2}} \tag{456}$$

where m_i is the mass of atom i and $\mathbf{r}_i$ the position of atom i with respect to the center of mass of the molecule. It is especially useful to characterize polymer solutions and proteins. The program will also provide the radius of gyration around the coordinate axis (or, optionally, principal axes) by only summing the radii components orthogonal to each axis, for instance

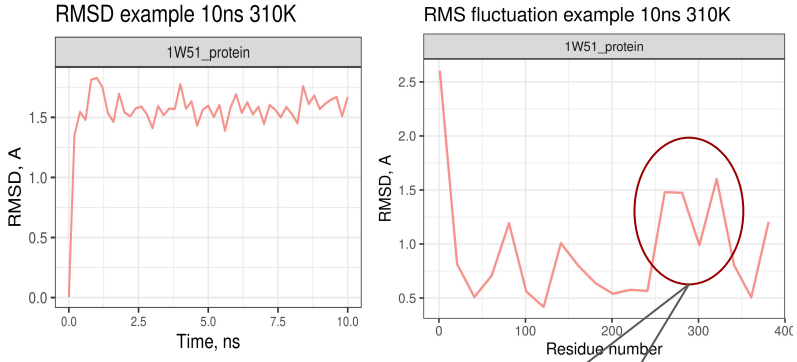
$$R_{g,x} = \left(\frac{\sum_i \left(r_{i,y}^2 + r_{i,z}^2 \right) m_i}{\sum_i m_i} \right)^{\frac{1}{2}} \tag{457}$$

Value of Rgyr:

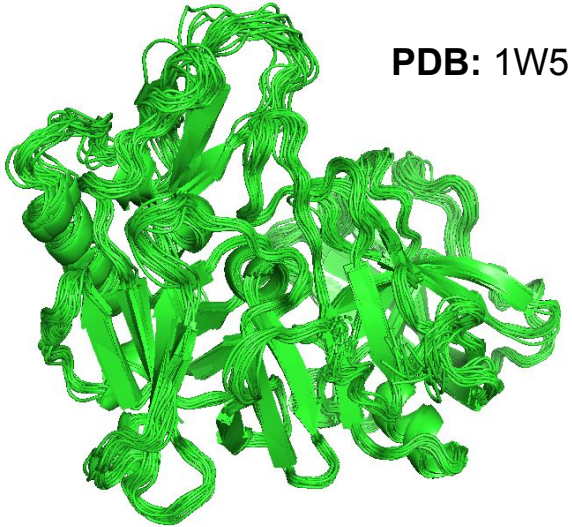
Decreasing - compression

Increasing - extension

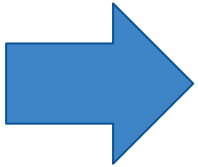
What can be done



- To explore different conformation of protein
  - To investigate internal-flexibility of protein
  - For practical use we select conformations from the most populated clusters of all conformations

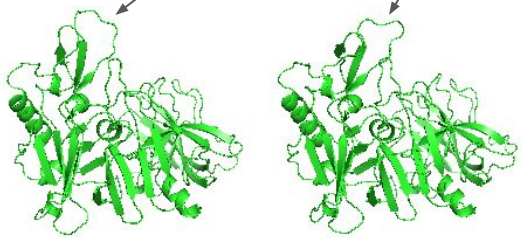


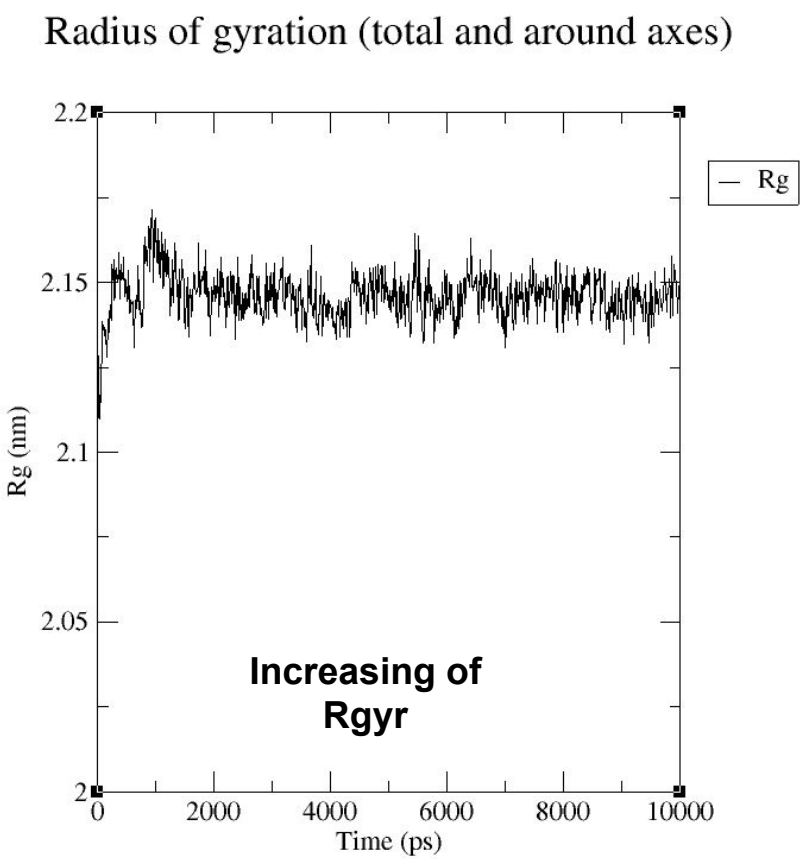
PDB: 1W51

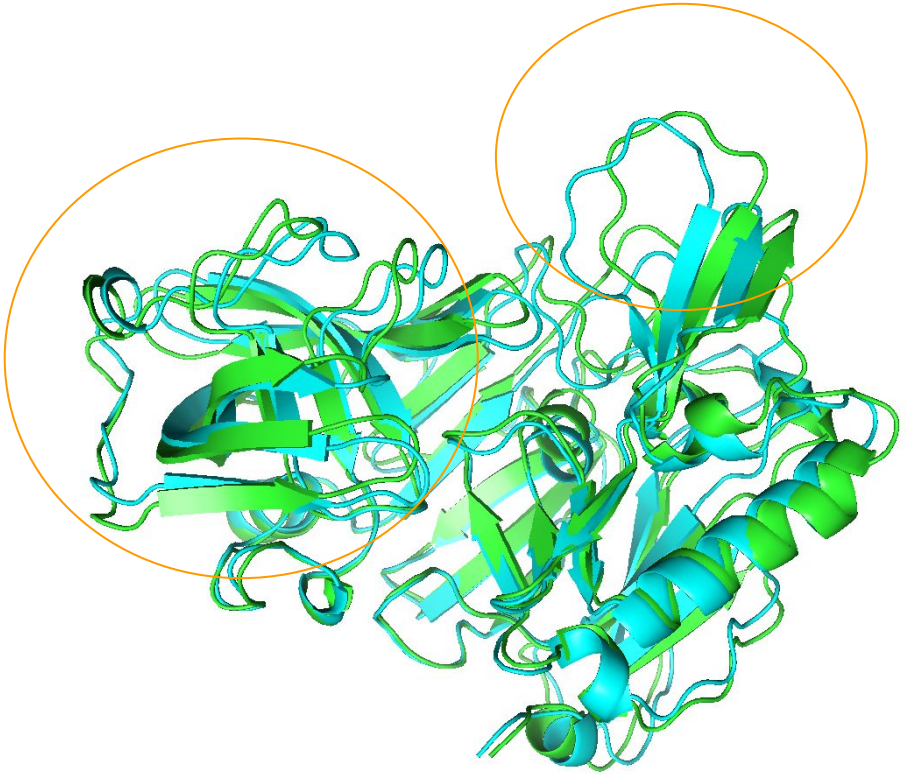


Conformation can be used:

- for further molecular docking study





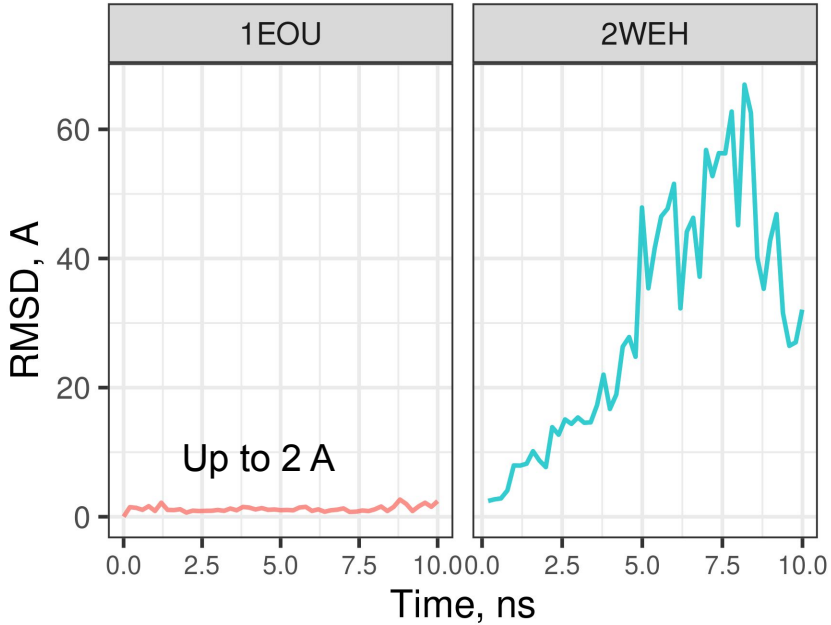


Conformer at the 1st ps
Conformer at the 1000th ps

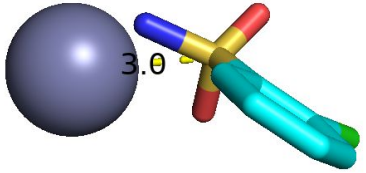
What can be done by MD

- To explore stability of ligand pose

RMSD example 10ns 310K

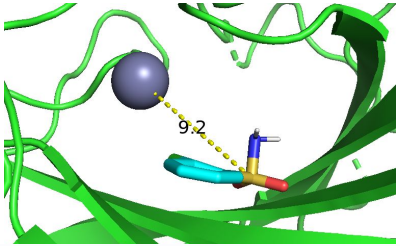


Example of incorrect pose:

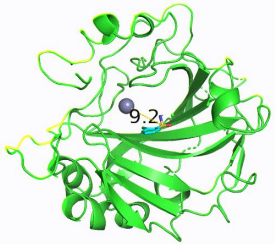


x-ray

PDB: 2WEH



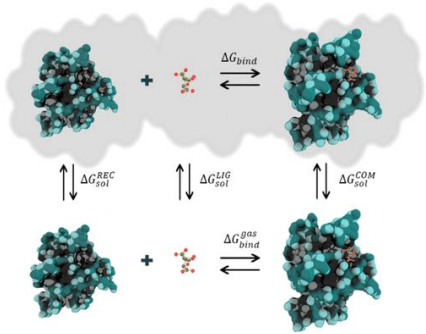
docking



MD 10ns.
Ligand goes out of the site

What can be done by MD

- to estimate binding affinity of protein-ligand complexes
  - Molecular mechanics Poisson–Boltzmann surface area (**MM/PBSA**)
  - Molecular mechanics generalized Born surface area (**MM/GBSA**)



In the MM/PBSA or MM/GBSA approach, the free energy for binding of the ligand (L) to the protein receptor (R) to form the complex (RL),

$$\Delta G_{\text{bind}} = G_{\text{RL}} - G_{\text{R}} - G_{\text{L}} \quad (4)$$

can be decomposed into contributions of different interactions and expressed as (58)

$$\Delta G_{\text{bind}} = \Delta H - T\Delta S = \Delta E_{\text{MM}} + \Delta G_{\text{sol}} - T\Delta S \quad (5)$$

in which

$$\Delta E_{\text{MM}} = \Delta E_{\text{int}} + \Delta E_{\text{ele}} + \Delta E_{\text{vdW}} \quad (6)$$

$$\Delta G_{\text{sol}} = \Delta G_{\text{PB/GB}} + \Delta G_{\text{SA}} \quad (7)$$

$$\Delta G_{\text{SA}} = \gamma \cdot \text{SASA} + b \quad (8)$$

Valdés-Tresanco, M.S., Valdés-Tresanco, M.E., Valiente, P.A. and Moreno E. ***gmx_MMPBSA: A New Tool to Perform End-State Free Energy Calculations with GROMACS.*** Journal of Chemical Theory and Computation, 2021 17 (10), 6281-6291. https://pubs.acs.org/doi/10.1021/acs.jctc.1c00645.
***MMPBSA.py*: An Efficient Program for End-State Free Energy Calculations** Bill R. Miller III, T. Dwight McGee Jr., Jason M. Swails, Nadine Homeyer, Holger Gohlke, and Adrian E. Roitberg Journal of Chemical Theory and Computation **2012** *8* (9), 3314-3321 DOI: 10.1021/ct300418h

What can be done by MD

- to estimate binding affinity of protein-ligand complexes

Total $G_{Binding}$=

- **Gas-phase molecular mechanics energy ΔE_{MM}:**
  - changes in the **internal energies** ΔE_{int} (bond, angle, and dihedral energies)
  - **electrostatic energies** ΔE_{ele}
  - **van der Waals energies** ΔE_{vdW}
- **sum of the electrostatic solvation energy G_{sol}**
  - **The polar contribution** is calculated using either the PB or GB model ($\Delta G_{PB/GB}$). Where the GB method gives an analytical expression for the polar solvation energy and is thus much faster than the PB method.
  - **nonpolar energy** is usually estimated using the solvent-accessible surface area (**SASA**)
- The change in **conformational entropy** $-T\Delta S$
  - is usually calculated by normal-mode analysis (or Interaction entropy) on a set of conformational snapshots taken from MD simulations.

In the MM/PBSA or MM/GBSA approach, the free energy for binding of the ligand (L) to the protein receptor (R) to form the complex (RL),

$$\Delta G_{bind} = G_{RL} - G_R - G_L \quad (4)$$

can be decomposed into contributions of different interactions and expressed as (58)

$$\Delta G_{bind} = \Delta H - T\Delta S = \Delta E_{MM} + \Delta G_{sol} - T\Delta S \quad (5)$$

in which

$$\Delta E_{MM} = \Delta E_{int} + \Delta E_{ele} + \Delta E_{vdW} \quad (6)$$

$$\Delta G_{sol} = \Delta G_{PB/GB} + \Delta G_{SA} \quad (7)$$

$$\Delta G_{SA} = \gamma \cdot SASA + b \quad (8)$$

GBSA/PBSA

We can perform end-state free energy calculations

$$\Delta G_{bind} = G_{RL} - G_R - G_L$$

```
Delta (Complex - Receptor - Ligand):
Energy Component       Average     SD(Prop.)         SD   SEM(Prop.)        SEM
----------------------------------------------------------------------------------
ΔBOND                     0.00          3.90       0.00         0.17       0.00
ΔANGLE                    0.00          4.77       0.00         0.21       0.00
ΔDIHED                    0.00          3.87       0.00         0.17       0.00
ΔVDWAALS               -102.13          2.26       4.27         0.10       0.19
ΔEEL                    -16.56          0.74       4.83         0.03       0.22
Δ1-4 VDW                 -0.00          1.70       0.00         0.08       0.00
Δ1-4 EEL                 -0.00          0.64       0.00         0.03       0.00
ΔEGB                     51.30          0.77       4.38         0.03       0.20
ΔESURF                  -11.85          0.02       0.36         0.00       0.02

ΔGGAS                  -118.69          2.38       5.68         0.11       0.25
ΔGSOLV                   39.45          0.77       4.34         0.03       0.19

ΔTOTAL                  -79.23          2.50       4.49         0.11       0.20
----------------------------------------------------------------------------------

Using Interaction Entropy Approximation:
ΔG binding =    -68.17 +/-    6.10
```

$$\Delta G_{bind} = \Delta H - T\Delta S = \Delta E_{MM} + \Delta G_{sol} - T\Delta S$$

$$\Delta E_{MM} = \Delta E_{int} + \Delta E_{ele} + \Delta E_{vdW}$$

$$\Delta G_{sol} = \Delta G_{PB/GB} + \Delta G_{SA}$$

$$\Delta G_{SA} = \gamma \cdot SASA + b$$

```
GENERALIZED BORN:

Complex:
Energy Component       Average     SD(Prop.)         SD   SEM(Prop.)        SEM
----------------------------------------------------------------------------------
BOND                   2858.30         46.32      46.32         2.07       2.07
ANGLE                  7239.59         66.97      66.97         2.99       2.99
DIHED                  9131.55         44.47      44.47         1.99       1.99
VDWAALS               -7387.25         51.90      51.90         2.32       2.32
EEL                  -55344.22        317.42     317.42        14.18      14.18
1-4 VDW                3275.97         24.44      24.44         1.09       1.09
1-4 EEL               36478.79         82.81      82.81         3.70       3.70
EGB                  -18524.97        275.82     275.82        12.32      12.32
ESURF                   280.64          5.77       5.77         0.26       0.26

GGAS                  -3747.27        345.71     300.04        15.45      13.40
GSOLV                -18244.33        275.88     274.96        12.33      12.28

TOTAL                -21991.60        442.29      99.56        19.76       4.45
```

Complex energy terms

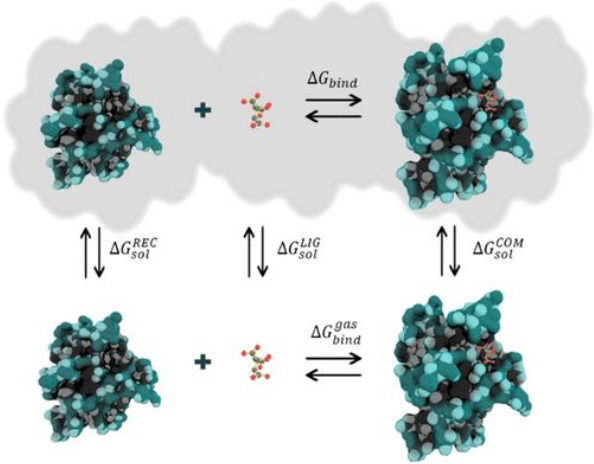
```
Receptor:
Energy Component       Average     SD(Prop.)         SD   SEM(Prop.)        SEM
----------------------------------------------------------------------------------
BOND                   2835.83         46.26      46.26         2.07       2.07
ANGLE                  7133.13         66.61      66.61         2.98       2.98
DIHED                  9079.29         44.14      44.14         1.97       1.97
VDWAALS               -7277.88         51.88      51.88         2.32       2.32
EEL                  -55330.60        315.81     315.81        14.11      14.11
1-4 VDW                3251.42         24.39      24.39         1.09       1.09
1-4 EEL               36494.62         82.78      82.78         3.70       3.70
EGB                  -18549.75        274.33     274.33        12.26      12.26
ESURF                   285.40          5.69       5.69         0.25       0.25

GGAS                  -3814.19        344.09     299.04        15.37      13.36
GSOLV                -18264.35        274.39     273.49        12.26      12.22

TOTAL                -22078.54        440.10      99.05        19.66       4.43
----------------------------------------------------------------------------------

Ligand:
Energy Component       Average     SD(Prop.)         SD   SEM(Prop.)        SEM
----------------------------------------------------------------------------------
BOND                     22.47          3.96       3.96         0.18       0.18
ANGLE                   106.46          5.13       5.13         0.23       0.23
DIHED                    52.26          4.20       4.20         0.19       0.19
VDWAALS                  -7.24          2.28       2.28         0.10       0.10
EEL                       2.93          0.87       0.87         0.04       0.04
1-4 VDW                  24.55          1.75       1.75         0.08       0.08
1-4 EEL                 -15.83          0.67       0.67         0.03       0.03
EGB                     -26.51          0.72       0.72         0.03       0.03
ESURF                     7.08          0.06       0.06         0.00       0.00

GGAS                    185.60          8.32       7.01         0.37       0.31
GSOLV                   -19.43          0.72       0.71         0.03       0.03

TOTAL                   166.17          8.35       6.96         0.37       0.31
----------------------------------------------------------------------------------
```

Protein energy terms

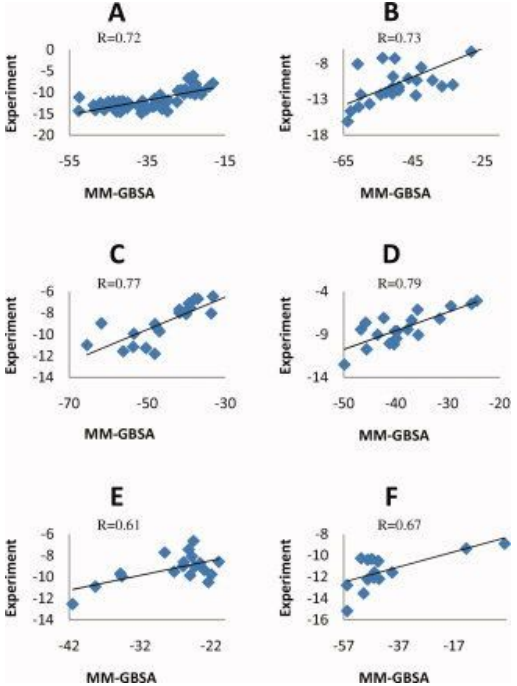
Ligand energy terms

What can be done by MD

- to estimate binding affinity of protein-ligand complexes



*Correlation between **MM-GBSA** predicted and experimental binding free energy.*



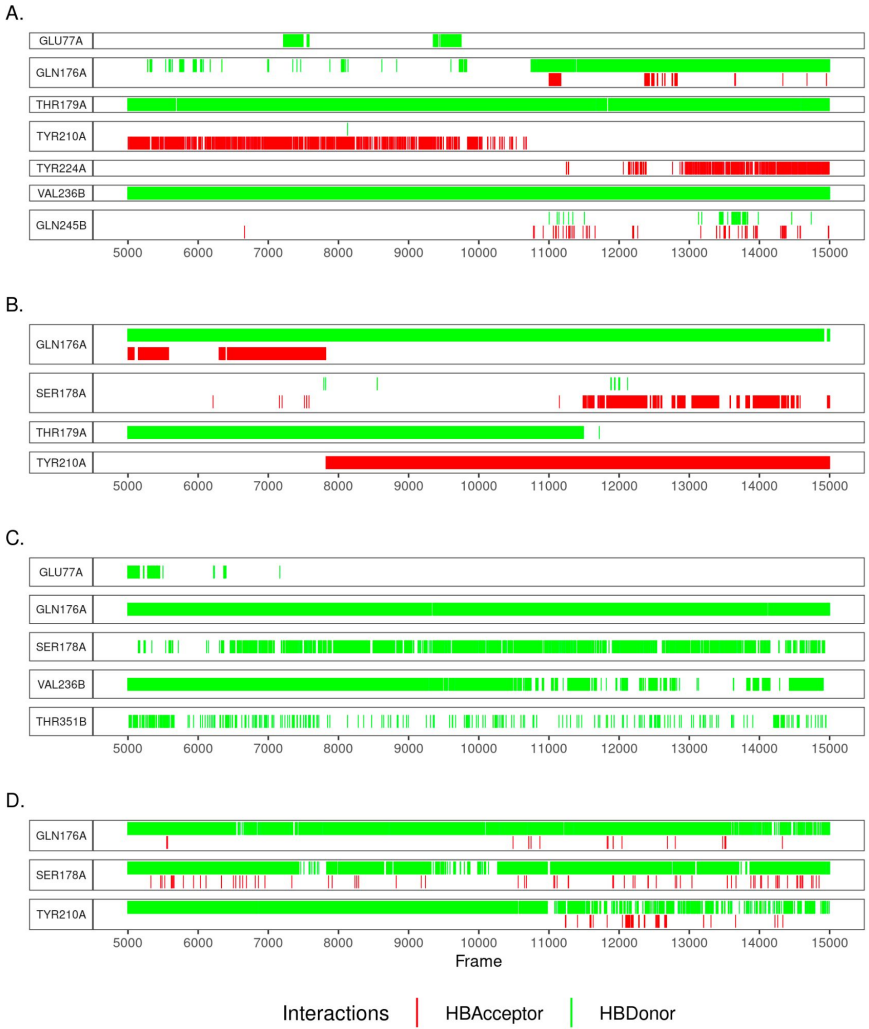
Yang T, Wu JC, Yan C, Wang Y, Luo R, Gonzales MB, Dalby KN, Ren P. Virtual screening using molecular simulations. Proteins. 2011 Jun;79(6):1940-51. doi: 10.1002/prot.23018. Epub 2011 Apr 12. PMID: 21491494; PMCID: PMC3092865.

Valdés-Tresanco, M.S., Valdés-Tresanco, M.E., Valiente, P.A. and Moreno E. ***gmx_MMPBSA: A New Tool to Perform End-State Free Energy Calculations with GROMACS.*** Journal of Chemical Theory and Computation, 2021 17 (10), 6281-6291. https://pubs.acs.org/doi/10.1021/acs.jctc.1c00645.

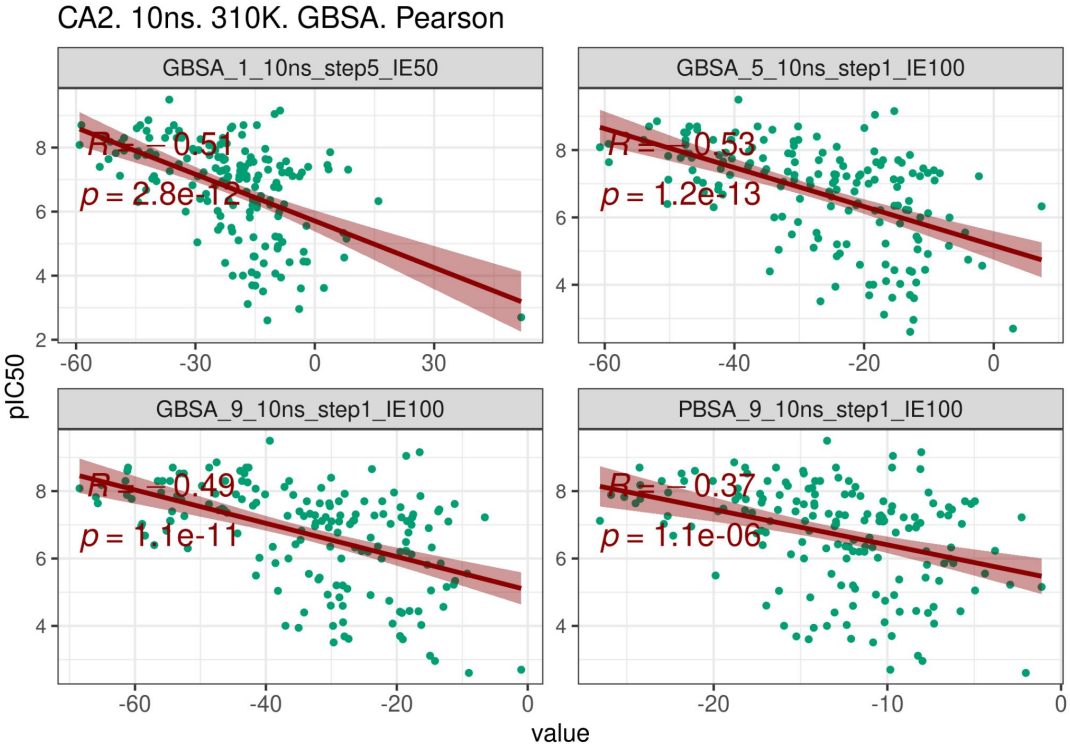
***MMPBSA.py*: An Efficient Program for End-State Free Energy Calculations** Bill R. Miller III, T. Dwight McGee Jr., Jason M. Swails, Nadine Homeyer, Holger Gohlke, and Adrian E. Roitberg Journal of Chemical Theory and Computation **2012** *8* (9), 3314-3321 DOI: 10.1021/ct300418h

What can be done by MD

- to investigate protein-ligand (protein) interaction stability



Beta Secretase



Tested parameters:

- Time of analysis
  - from 1 to 10 ns
  - from 5 to 10 ns
  - from 9 to 10 ns
- GBSA/PBSA methods

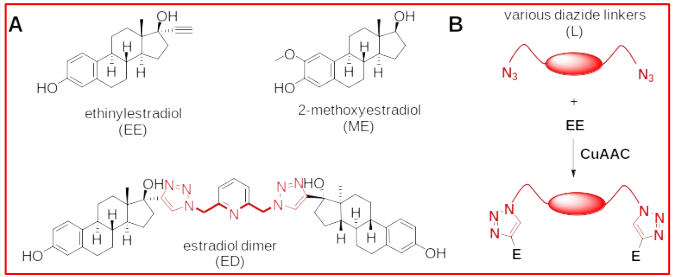
- **The best correlation:** -0.53
- Results depends on parameters but not drastically
- Further analysis may help to investigate number of the most optimal parameters

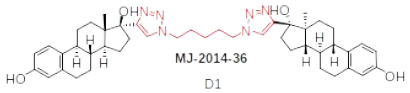
~ time of running

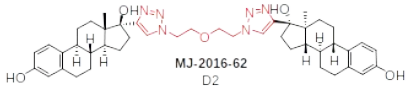
- 1 hour for 10 ns of MD
- 0.5 sec per compound for GBSA

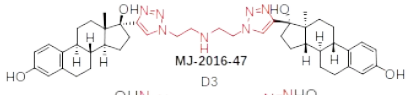
INSTITUTE OF MOLECULAR AND TRANSLATIONAL MEDICINE

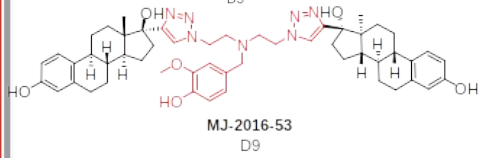
α/β-Tubulin estradiol dimers inhibitors

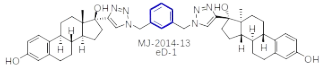


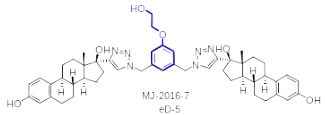


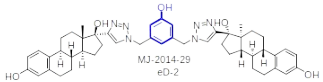


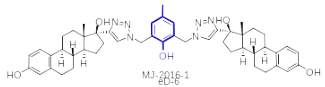












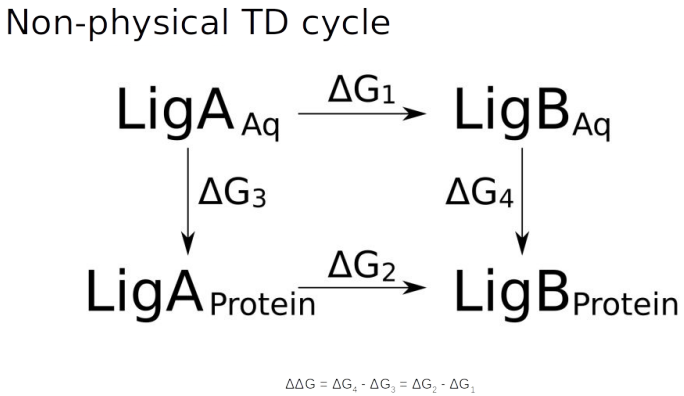
				GBSA	PBSA
ID	Vmax	Gnina	VINA	140-150 ns IE25%	
ED	4.2	7.4	-15	-60.81	-16.8
ED3	7.1	8.2	-14.7	-41.30	-2.35
ED5	2.8	8.3	-14.1	-81.59	-24.045
D1	4.1	8.2	-13.7	-70.06	-22.60
D3	7.4	8.0	-13.2	-42.25	-2.88
D9	5.9	8.2	-13.7	-48.5	-5.58
D2	2.8	8.1	-13.2	-65.72	-19.35
Pearson Corr		**0.05**	**-0.03**	**0.93**	**0.95**

Limitations of Classical Molecular Dynamics:

1) MD simulations on very large systems may require such large computer resources that they cannot easily be studied by traditional all-atom methods.
2) Simulations of processes on long timescales* (beyond about 1 microsecond) are prohibitively expensive, because they require so many time steps.
3) not accurate calculation of binding energy

*The microscopic observation of biomolecular processes such as **protein folding**, **protein interactions**, **agonist-antagonist functional recognition** and **enzyme reactions** occur on timescales ranging from microseconds to seconds.

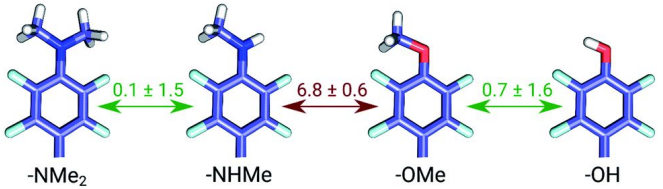
Free energy perturbation (FEP)



Free Energy Perturbation (FEP) is a computational technique in molecular dynamics simulations used to calculate the free energy difference between two different states of a molecular system. This method is widely employed in computational chemistry and drug discovery to understand thermodynamic properties, such as binding affinities or conformational changes.

Uses for predicting relative binding free energy between similar ligands

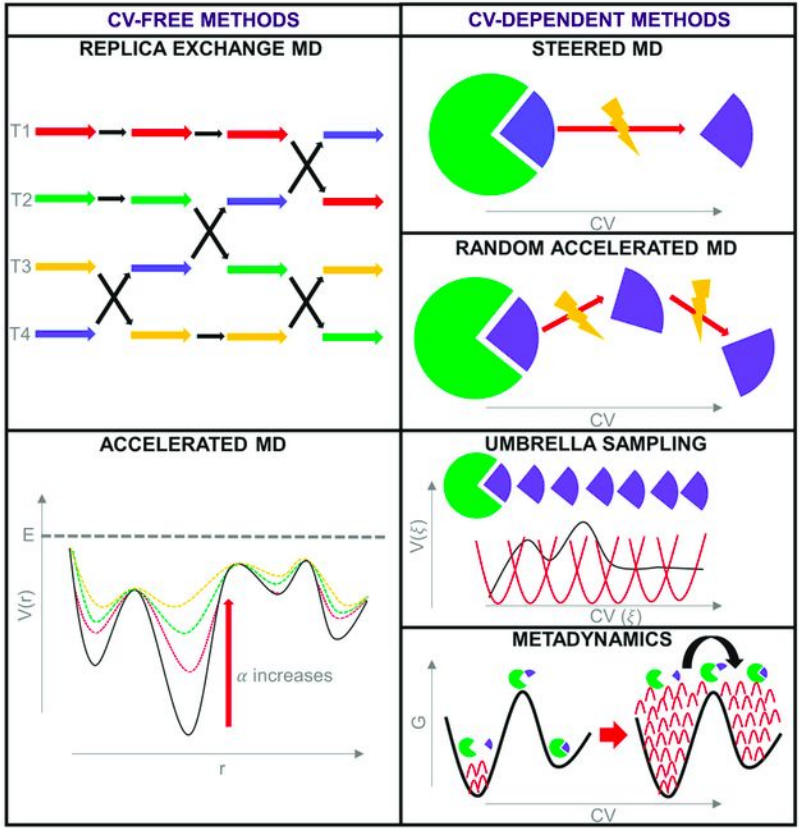
Basic Concept:
Free Energy Difference: FEP aims to calculate the free energy difference (ΔG) between two states of a system, typically denoted as **State A** and **State B**. This can include changes in conformation, solvation state, or interactions with other molecules.



perturbation

FEP calculations are **based on molecular dynamics (MD) simulations** and therefore explicitly consider both enthalpy and entropy effects of the conformational flexibility of the ligand, as well as desolvation effects within the ligand binding domain (LBD) of certain receptors. Simulations must be stable.

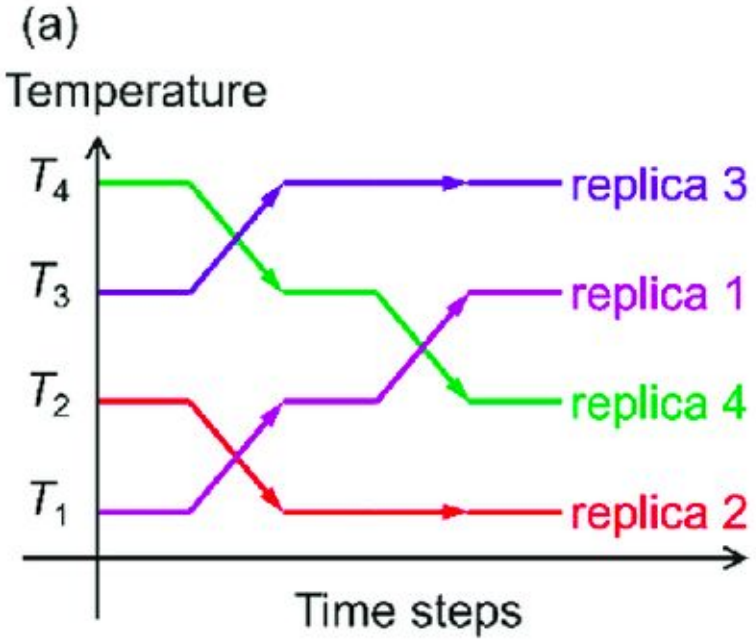
Enhanced sampling methods for molecular dynamics simulations



Enhanced sampling methods in molecular dynamics simulations are designed to overcome the limitations of traditional simulations: which may struggle to *explore rare events*, *transitions between different states*, or *high-dimensional spaces efficiently.*

These methods aim to enhance the sampling of relevant regions in the configuration space, providing more accurate and comprehensive insights into the behavior of the system.

Replica exchange method

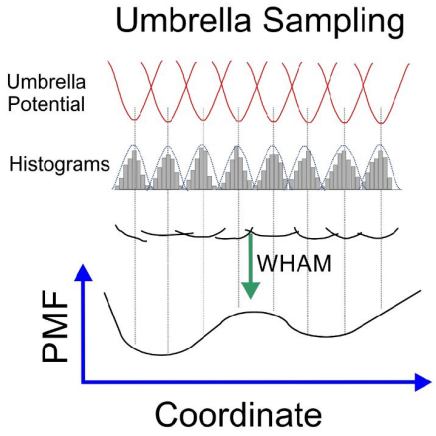


Replica Exchange Molecular Dynamics (REMD):

Idea: REMD involves running parallel simulations at different temperatures and occasionally swapping configurations between neighboring replicas.

Implementation: By exchanging configurations, REMD promotes exploration of the entire temperature space, facilitating transitions between different energy basins.

Umbrella sampling



Idea: Umbrella Sampling is used to sample a reaction coordinate by applying harmonic restraints to the system along that coordinate.

Implementation: Windows are defined along the reaction coordinate, and simulations are run independently in each window. The resulting data are then combined to obtain the free energy profile.

1. Definition of Reaction Coordinate: A reaction coordinate is chosen to describe the progress of the system throughout the simulation. This could be a distance between specific atoms, an angle, or any other coordinate that characterizes the process of interest.

2. Harmonic Restraints: Along the chosen reaction coordinate, harmonic restraints (springs) are applied to confine the system to specific values of the coordinate. These restraints prevent the system from moving too far away from the chosen values.

3. Simulation in Windows: The simulation is divided into multiple windows along the reaction coordinate. In each window, the system is simulated under the influence of the harmonic restraint. The strength of the restraint varies between windows, covering the entire range of the reaction coordinate.

4. Independent Simulations: Simulations are run independently in each window. These simulations sample different regions of the free energy landscape along the reaction coordinate. Typically, each window is equilibrated before production runs.

5. Biasing Potential: The biasing potential applied to the system is essentially a harmonic potential due to the restraints. This potential introduces a bias along the reaction coordinate, making it easier for the system to explore different states.

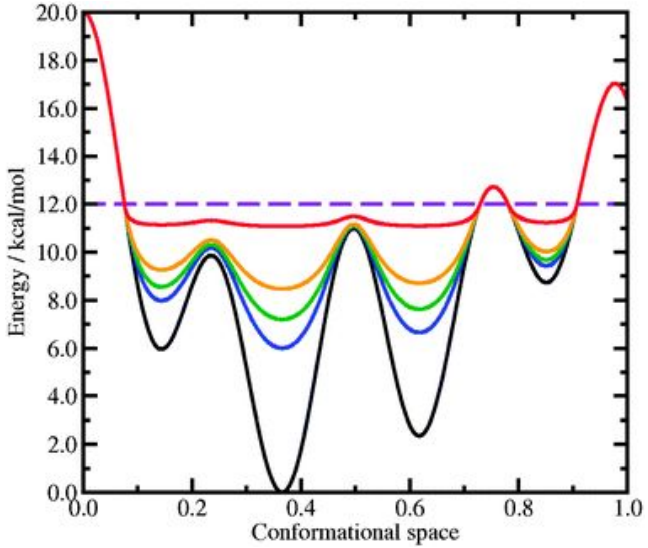
6. Sampling in Each Window: In each window, the system samples the region around the constrained value of the reaction coordinate. The biasing potential helps overcome energy barriers, allowing the system to explore different conformations.

7. Combining Data: The data from simulations in each window are combined to construct the overall free energy profile along the reaction coordinate. This is typically achieved by using the Weighted Histogram Analysis Method (WHAM) or other similar methods.

8. Free Energy Profile: The resulting free energy profile provides information about the thermodynamics of the system along the chosen reaction coordinate. It reveals the relative stability of different states and the energy barriers between them.

Accelerated Molecular Dynamics

Accelerated Molecular Dynamics (aMD) is an enhanced sampling technique used in molecular dynamics simulations to overcome energy barriers and explore conformational space more efficiently. The primary idea behind aMD is to selectively boost the potential energy of the system in regions of high energy, making it easier for the simulation to overcome barriers and sample rare events.



Potential Energy Boosting: In traditional molecular dynamics simulations, the potential energy surface may have high-energy barriers that slow down the exploration of certain regions of conformational space. Accelerated Molecular Dynamics addresses this issue by applying a biasing potential to selectively boost the potential energy in high-energy regions.

Boosting Function: A boosting function is defined to modulate the potential energy. This function is typically based on the potential energy of each atom or a collective variable. The boosting function is designed to be higher in regions of high potential energy, effectively reducing energy barriers.

Boosted Force Calculation: The boosted potential energy modifies the forces acting on the atoms. The forces experienced by the atoms are adjusted based on the boosting function, making it easier for the system to explore high-energy states.

Simulation Setup: The simulation is set up similarly to traditional molecular dynamics, with an initial configuration, force field, and integration algorithm. The only difference is the inclusion of the boosting function in the potential energy calculations.

Sampling: The simulation proceeds with the modified potential energy, allowing the system to explore regions that would be energetically unfavorable in traditional simulations. The boosted potential energy helps overcome barriers and facilitates the sampling of rare events.

Equilibration and Production Runs: The simulation typically starts with an equilibration phase to allow the system to adjust to the modified potential energy. After equilibration, production runs are performed to collect data for analysis.

Analysis: Trajectories from the simulation are analyzed to extract information about the system's behavior, such as structural changes, thermodynamic properties, and the exploration of different conformational states.

Applications: Accelerated Molecular Dynamics is applied to study various biological and chemical processes, including protein folding, ligand binding, and conformational changes. It is particularly useful for simulating rare events that occur on longer timescales.

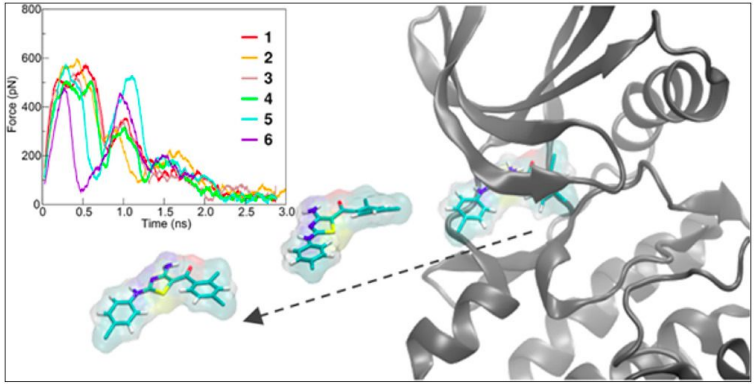
Challenges: While aMD can be a powerful tool for enhancing sampling, care must be taken in choosing the boosting function and parameters to ensure that the accelerated dynamics do not introduce artifacts. Validation against experimental data or other simulation methods is essential.

Collective variable

A **collective variable** is a descriptor of the molecular system studied predefined by the user. It must be a differentiable function of the atomic coordinates. Furthermore, its value should reflect the state of the simulated system, including metastable states

Steered molecular dynamics (SMD)

Steered Molecular Dynamics (SMD) is a simulation technique used to study the behavior of a system by applying external forces to specific atoms or groups of atoms, often to mimic a pulling or stretching action. This method is particularly useful for investigating processes such as protein unfolding, ligand binding, or other conformational changes in biological molecules



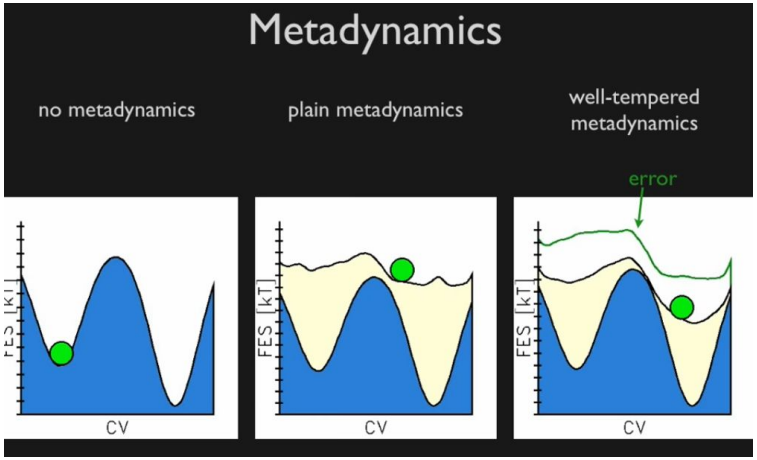
50 SMD simulations

Idea: In SMD, external forces are applied to specific atoms or groups of atoms, effectively guiding the system along a particular reaction coordinate.

Implementation: Forces are applied to mimic an experimental process (e.g., ligand binding or unfolding of a protein) to explore the associated energy landscape.

Metadynamics

Metadynamics involves **enhanced sampling over collective variables using a biased potential to force the system to leave local minima and thus sample low-probability states** (Laio and Parrinello, 2002). Metadynamics is an adaptive method, automatically biasing configurations away from those most visited.



set initial $\{\vec{r}\}$ and $\{\vec{v}\}$
set $V_{\text{bias}}(\vec{s}) := 0$

every MD step:
  compute CV values:
    $\vec{s}_t := \vec{s}(\{\vec{r}\})$

  every n MD steps:
    update bias potential:

$$V_{\text{bias}}(\vec{s}) := V_{\text{bias}}(\vec{s}) + \tau\omega \exp\left(-\frac{1}{2}\left|\frac{\vec{s}-\vec{s}_t}{\vec{\sigma}}\right|^2\right)$$

  compute atomic forces:

$$\vec{F}_i := -\frac{\partial V(\{\vec{r}\})}{\partial \vec{r}_i} - \overbrace{\frac{\partial V_{\text{bias}}(\vec{s})}{\partial \vec{s}}\bigg|_{\vec{s}_t} \frac{\partial \vec{s}(\{\vec{r}\})}{\partial \vec{r}_i}}^{\vec{F}_{\text{bias},i}}$$

  propagate $\{\vec{r}\}$ and $\{\vec{v}\}$ by Δt

Idea: Metadynamics adds a history-dependent bias potential to the potential energy surface, encouraging the system to explore regions that are rarely visited.

Implementation: Gaussian potentials are added along chosen collective variables. As the simulation progresses, these Gaussians are deposited, preventing the system from revisiting the same states.

Supervised Molecular Dynamics (SuMD)

Supervised MD (SuMD) is a computational method that enables the exploration of ligand–receptor recognition pathway in a reduced timescale.

It involves integrating machine learning models, often neural networks, with molecular dynamics simulations to improve the accuracy and efficiency of the simulations.

Prediction Step: At each simulation step, the current state of the system (positions, velocities, etc.) is input into the trained machine learning model. The model predicts the desired property or behavior based on the current state.

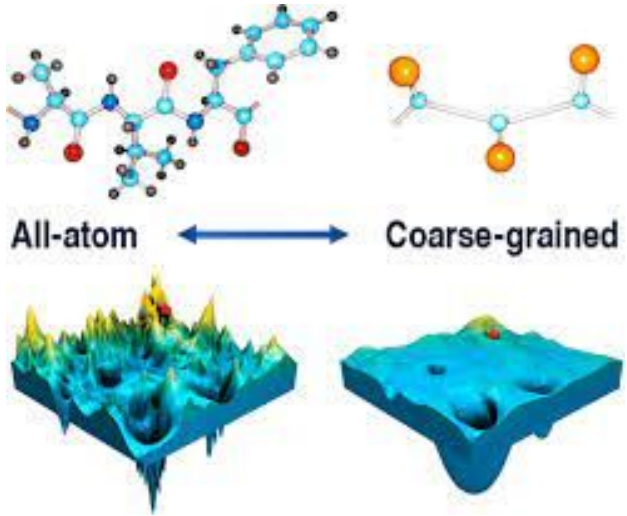
Influence on Dynamics: The predicted information is used to guide or influence the dynamics of the simulation. This can involve adjusting forces acting on particles, modifying the potential energy landscape, or altering other parameters to match the predicted behavior.

Coarse grained molecular dynamics

*Coarse-grained molecular dynamics (CGMD) i*s a simulation technique that simplifies the representation of molecular systems by grouping multiple atoms into a single interaction site. This approach reduces the computational cost and allows for the simulation of larger and longer time-scale processes compared to all-atom simulations.

Coarse-grained models are widely used for molecular modeling of biomolecules at various granularity levels.

- Allows us to model extremely large systems



Take-home message

1) **Molecular dynamics (MD)** is a computer simulation method for analyzing the physical movements of atoms and molecules.

2) Molecular dynamics gives us more accurate information about molecular interaction than docking

3) Molecular dynamics allows us:
- to explore dynamic evolution of the system (flexibility of protein, protein-ligand complex, etc.)
- to calculate binding free energy
- to explore conformation space
- to explore protein-ligand (protein-protein) interaction profile

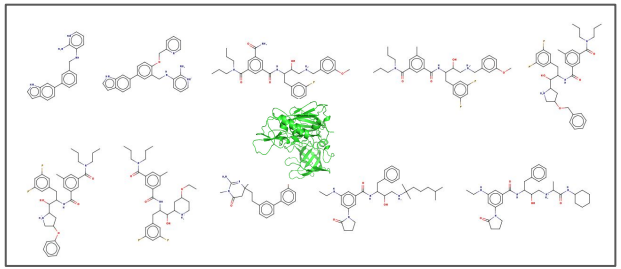
4) The accuracy of the simulation depends on the choice of an appropriate force field, and validation against experimental data is crucial.

Sources:

- **GROMACS documentation -** http://manual.gromacs.org/documentation/2016/index.html
- **GROMACS reference manual** - http://manual.gromacs.org/documentation/2016/manual-2016.pdf
- **J. Lemkul tutorials** – http://www.mdtutorials.com/gmx/index.html

High-throughput MD

Aim: to implement an easy to use tool to run whole pipeline of Molecular Dynamics simulation automatically



Perform **High-Throughput Molecular Dynamics**

Rank compounds by:

- Stability of previously obtained pose of ligand
- Calculated Binding Energy
- Stability of protein-ligand contacts

Preprocessing

Structure Preparation
- add missing atoms
- protonation
- etc

Force-field
- Definitions of inter-atomic bonded and non-bonded forces

Simulation box setup
- box size/shape

Solvate system
- add HOH molecules

Neutralize system
- add NA+/CL- ions

Pre-simulation steps

Energy minimization

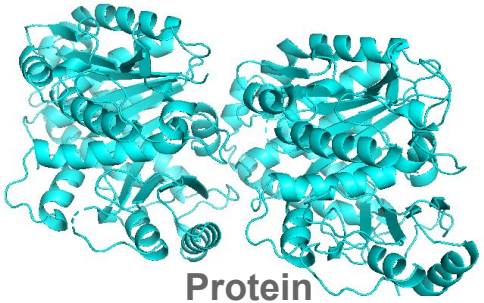
NVT and NPT equilibration
- 1000 ps

Product Simulation

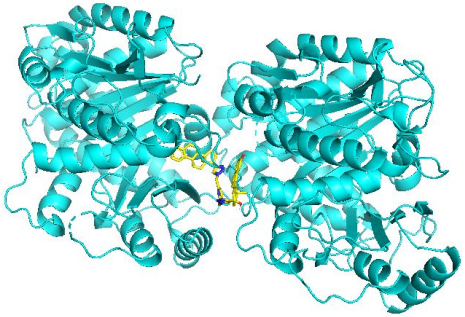
Analysis

Main features of the tool:

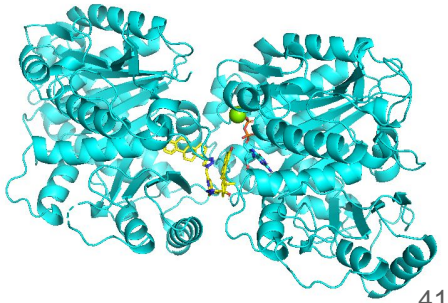
- **User control of simulation time**
  - from 10 ps to 1 μs
- **Default preset optimal parameters to run Molecular Dynamics**
  - can be easily modified
  - useful as teaching source
- **Support of modeling of different molecular systems**
  - protein only, protein-ligand, protein-ligand-cofactors
- **Support of simulation of boron-containing molecules (Gaussian-based calculations)**
- **Support of simulation with parameterized by MCPBPY metal atoms**
- **Automatic analysis of simulation:**
  - RMSD plots for both protein and ligand objects
  - Plot of flexibility of each amino acids (RMSF)
- **Support of analysis by additional instruments:**
  - ***ProLIF***: Ligand-Protein interactions (time-dependent function, stability analysis)
  - ***MM(PB)GBSA:*** Calculation of Binding Energy



Protein



Protein-Ligand



Protein-Ligand-Cofactor

RCSB PDB Deposit Search Visualize Analyze Download Learn More Documentation Careers MyPDB Contact us

198,528 Structures from the PDB
1,000,361 Computed Structure Models (CSM)

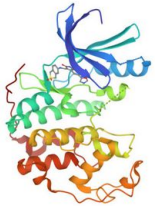
3D Structures | Enter search term(s), Entry ID(s), or sequence | Include CSM

Advanced Search | Browse Annotations | Help

PDB-101 wwPDB EMDataResource Nucleic Acid Database wwPDB Foundation

Structure Summary | 3D View | Annotations | Experiment | Sequence | Genome | Ligands | Versions

Biological Assembly 1



3D View: Structure | 1D-3D View | Electron Density | Validation Report | Ligand Interaction

Global Symmetry: Asymmetric - C1
Global Stoichiometry: Monomer - A1

Find Similar Assemblies

1KE7

Display Files | Download Files

CYCLIN-DEPENDENT KINASE 2 (CDK2) COMPLEXED WITH 3-{[(2,2-DIOXIDO-1,3-DIHYDRO-2-BENZOTHIEN-5-YL)AMINO]METHYLENE}-5-(1,3-OXAZOL-5-YL)-1,3-DIHYDRO-2H-INDOL-2-ONE

PDB DOI: 10.2210/pdb1KE7/pdb

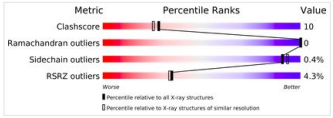
Classification: TRANSFERASE
Organism(s): Homo sapiens
Expression System: Spodoptera frugiperda
Mutation(s): No

Deposited: 2001-11-14 **Released:** 2002-05-14
Deposition Author(s): Bramson, H.N., Corona, J., Davis, S.T., Dickerson, S.H., Edelstein, M., Frye, S.V., Gampe, R.T., Hassell, A.H., Shewchuk, L.M., Kuyper, L.F.

Experimental Data Snapshot

Method: X-RAY DIFFRACTION
Resolution: 2.00 Å
R-Value Free: 0.235
R-Value Work: 0.191

wwPDB Validation | 3D Report | Full Report



Macromolecules

Find similar proteins by: Sequence (by identity cutoff) | 3D Structure

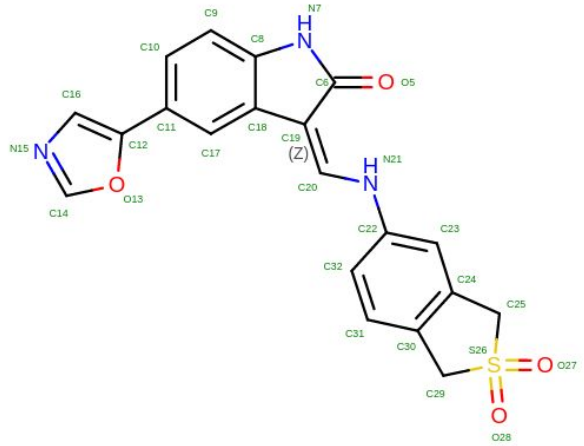
Entity ID: 1

Molecule	Chains	Sequence Length	Organism	Details	Image
Cell division protein kinase 2	A	298	Homo sapiens	**Mutation(s):** 0 **Gene Names:** CDK2, CDK N2 **EC:** 2.7.1.37 (PDB Primary Data), 2.7.11.22 (UniProt)	

UniProt & NIH Common Fund Data Resources

Find proteins for P24941 *(Homo sapiens)* Explore P24941 Go to UniProtKB: P24941

1KE7: Ligand LS3



Small Molecules

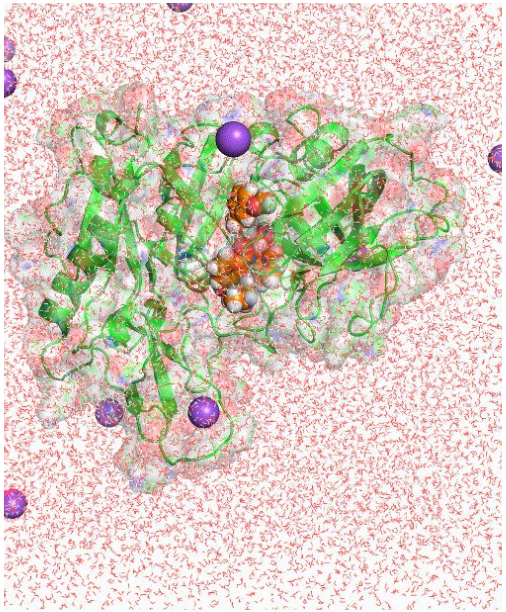
Ligands 1 Unique

ID	Chains	Name / Formula / InChI Key	2D Diagram	3D Interactions
LS3 Query on LS3 Download Ideal Coordinates CCD File Download Instance Coordinates	B [auth A]	3-{[(2,2-DIOXIDO-1,3-DIHYDRO-2-BENZOTHIEN-5-YL)AMINO]METHYLENE}-5-(1,3-OXAZOL-5-YL)-1,3-DIHYDRO-2H-INDOL-2-ONE $C_{20}H_{15}N_3O_4S$ FTQYGMLRLRXBPT-IDUWFGFVSA-N		Ligand Interaction

Binding Affinity Annotations

ID	Source	Binding Affinity
LS3	BindingDB: 1KE7	IC50: 8.9 (nM) from 1 assay(s)
	Binding MOAD: 1KE7	IC50: 8.9 (nM) from 1 assay(s)
	PDBBind: 1KE7	IC50: 8.9 (nM) from 1 assay(s)

Classical Molecular Dynamics



Structure Preparation

Preprocessing

Force-field
- Definitions of inter-atomic bonded and no-bonded forces (ligand and protein)

Simulation box setup
- box size/shape

Solvate system
- add HOH molecules

Neutralize system
- add NA+/CL- ions

Pre-simulation steps

Energy minimization
- 100 ps
- to ensure that the system has no steric clashes or inappropriate geometry

NVT and NPT equilibration
- 1000 ps
- equilibrate the solvent and ions around the protein

Product Simulation

Analysis

FAST. FLEXIBLE. FREE.
GROMACS

Protein preparation

https://www.cgl.ucsf.edu/chimera/

about projects people publications
resources visit us search

UCSF CHIMERA

an Extensible Molecular Modeling System

UCSF Chimera is a program for the interactive visualization and analysis of molecular structures and related data, including density maps, trajectories, and sequence alignments. It is available free of charge for noncommercial use. Commercial users, please see Chimera commercial licensing.

We encourage Chimera users to try ChimeraX for much better performance with large structures, as well as other major advantages and completely new features. ChimeraX includes a significant subset of Chimera features (with more to come, see the missing features list) and is under active development. Users may choose to use both programs, and it is fine to have both installed.

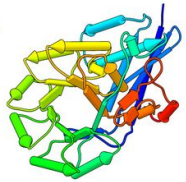
Chimera is no longer under active development, and is only updated for critical maintenance. Chimera development was supported by a grant from the National Institutes of Health (P41-GM103311) that ended in 2018.

Feature Highlight

Pipes and Planks

The PipesAndPlanks tool shows protein helices as "pipes" (cylinders) and strands as "planks" (rectangular boxes), with connectors for the intervening coil. Adjustable settings include pipe radius, plank width, colors, and whether to include arrowheads to show chain N→C directionality (see image how-to).

(More features...)



Gallery Sample

Peroxiredoxin Wreath

Peroxiredoxins are enzymes that help cells cope with stressors such as high levels of reactive oxygen species. The image shows a decameric peroxiredoxin from human red blood cells (Protein Data Bank entry 1qmv), styled as a holiday wreath.

See also the RBVI holiday card gallery.

(More samples...)

Quick Links

Documentation
Getting Started
User's Guide
Command Index
Tutorials and Videos
Guide to Volume Data
Release Notes
Download
What's New in Daily Builds
Map of Download Locations
Galleries
Image Gallery
Animation Gallery
Publications and Talks
Related Databases and Software
Citing Chimera
Contact Us

Recent Citations

Imprinted antibody responses against SARS-CoV-2 Omicron sublineages. Park YJ, Pinto D *et al. Science*. 2022 Nov 11;378(6620):619-627.

Bending forces and nucleotide state jointly regulate F-actin structure. Reynolds MJ, Hachicho C *et al. Nature*. 2022 Nov 10;611(7935):380-386.

Bestrophin-2 and glutamine synthetase form a complex for glutamate release. Owji AP, Yu K *et al. Nature*. 2022 Nov 3;611(7934):180-187.

Cryo-EM structures of human m6A writer complexes. Su S, Li S *et al. Cell Res*. 2022 Nov;32(11):982-994.

Algorithmic design of 3D wireframe RNA polyhedra. Elonen A, Natarajan AK *et al. ACS Nano*. 2022 Oct 25;16(10):16608-16616.

Previously featured citations...

Chimera Search

Go
Google™ Search

News

September 27, 2022

Website downtime: The RBVI website (Chimera, ChimeraX, etc.) and RBVI-hosted web services will be down for maintenance from Tue, Sep 27 9pm PDT, through Wed, possibly extending to Thu, Sep 29 5pm PDT.

December 20, 2021

The RBVI wishes you a safe and happy holiday season! See our 2021 card and the gallery of previous cards back to 1985.

December 17, 2021

Chimera production release 1.16 is now available. This will be the last release to support Windows 7. See the release notes for what's new.

Previous news...

Upcoming Events

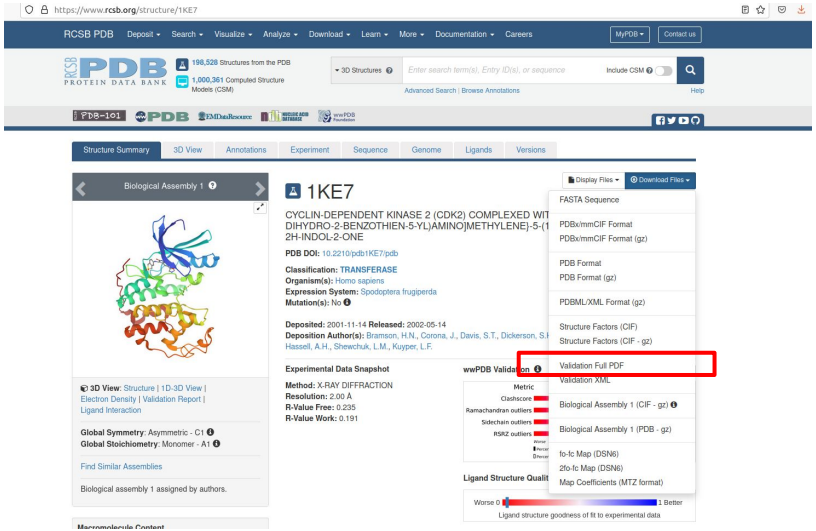
About RBVI | Projects | People | Publications | Resources | Visit Us

https://www.cgl.ucsf.edu/chimera/

Protein preparation

1. Download structure from PDB (https://www.rcsb.org/) using PDBID

   Download Files -> PDB Format



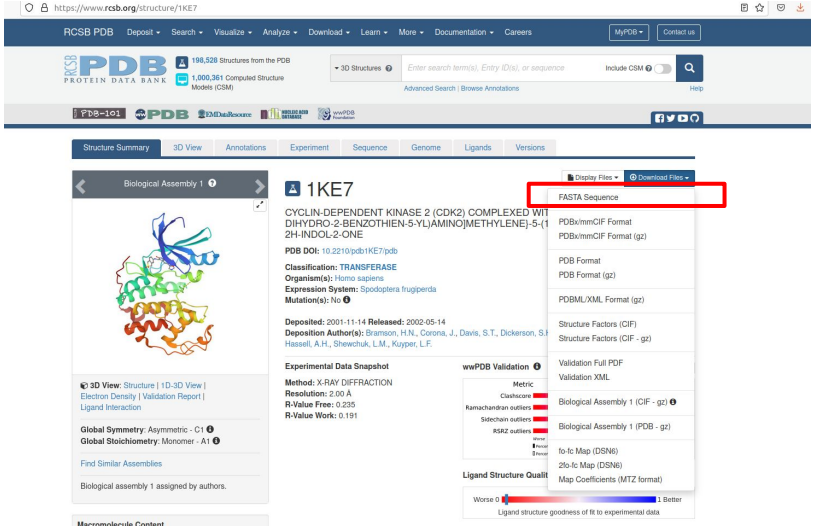
Protein preparation

1. Download structure from PDB (https://www.rcsb.org/) using PDBID

   Download Files -> PDB Format

2. Download sequence from PDB or from UniProt

   Download Files -> Fasta Sequence



Protein preparation

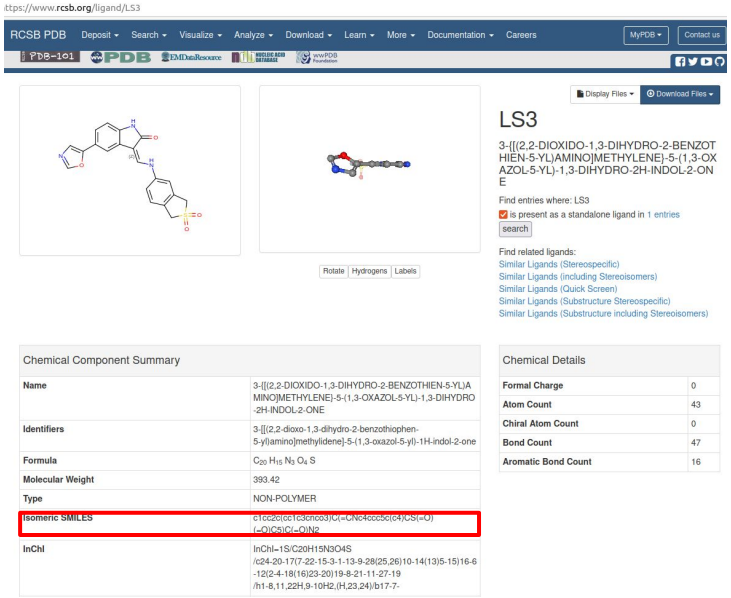
1. Download structure from PDB (https://www.rcsb.org/) using PDBID

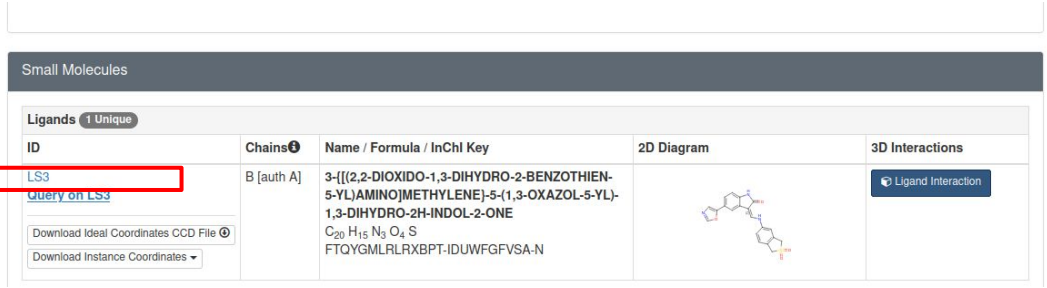
   Download Files -> PDB Format

2. Download sequence from PDB or from UniProt

   Download Files -> Fasta Sequence

3. *Save smiles of ligand into file (ligand.smi)*

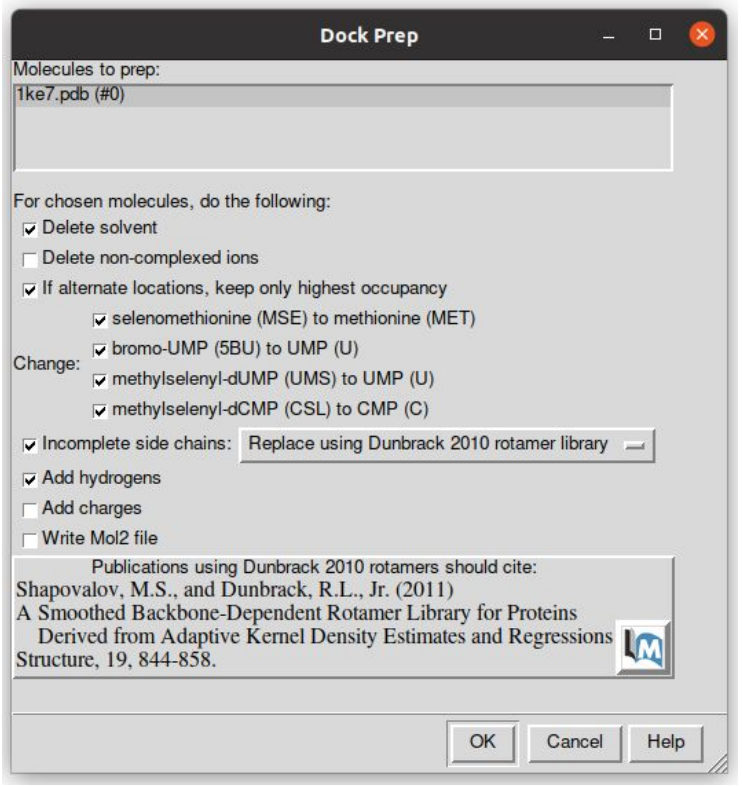




Protein preparation

4. Open Fasta and PDB in **Chimera**
    a. Dock Prepare

Structure Editing -> Dock Prep



Protein preparation

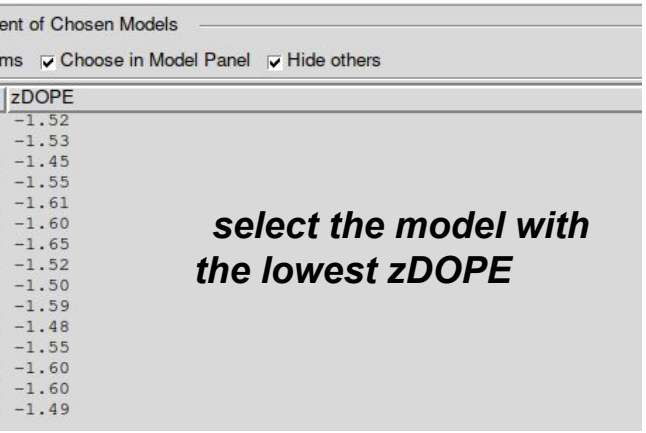
4. Open Fasta and PDB in **Chimera**

   b. fill missing loops by Modeller

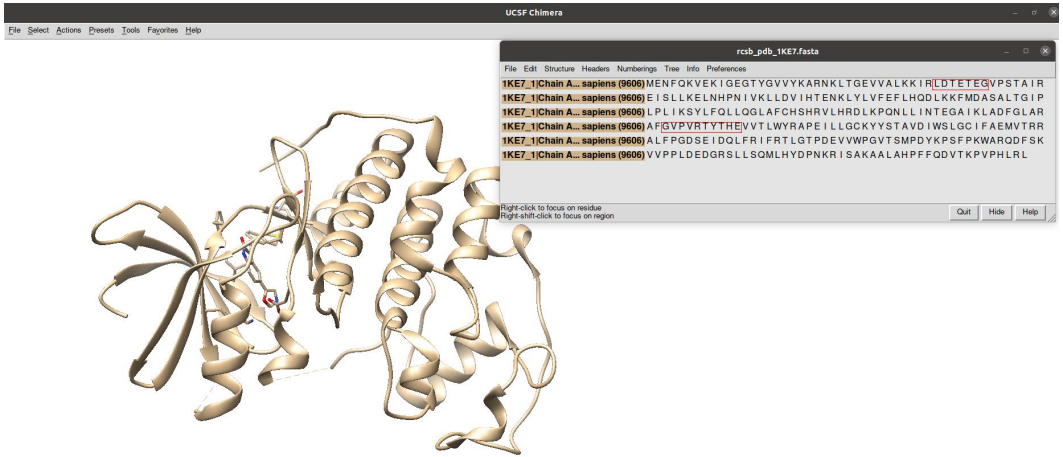
      Tools -> Sequence -> Sequence

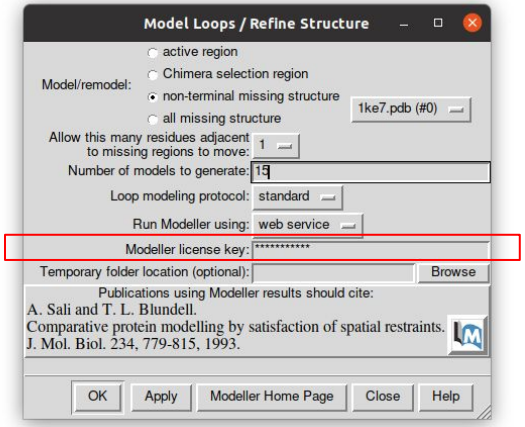
      Sequence -> Structure -> Modeller (loops/refinement)

   c. add Hs to selected model



select the model with the lowest zDOPE





Protein preparation

https://ambermd.org/Questions/HIS.html

AMBER Histidine residues

Histidine (HIS in normal pdb files) is really one of three possible residues:

HID: Histidine with hydrogen on the delta nitrogen

HIE: Histidine with hydrogen on the epsilon nitrogen

HIP: Histidine with hydrogens on both nitrogens; this is positively charged.

It is up to the user to inspect the environment of each histidine and identify the type that is appropriate.

4. Open Fasta and PDB in Chimera

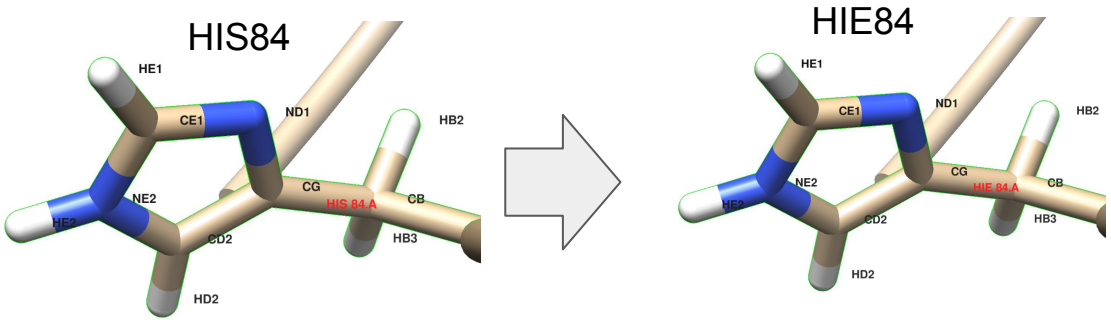
   d. Set HIS protonated state

Tools -> General Controls -> Command line

Put in the Command line:

setattr r type HID :HIS@HD1,DD1,TD1,HND
setattr r type HIP :HID@HE2,DE2,TE2
setattr r type HIE :HIS@HE2

check visually active site



Chimera preparation

4. Open Fasta and PDB in **Chimera**

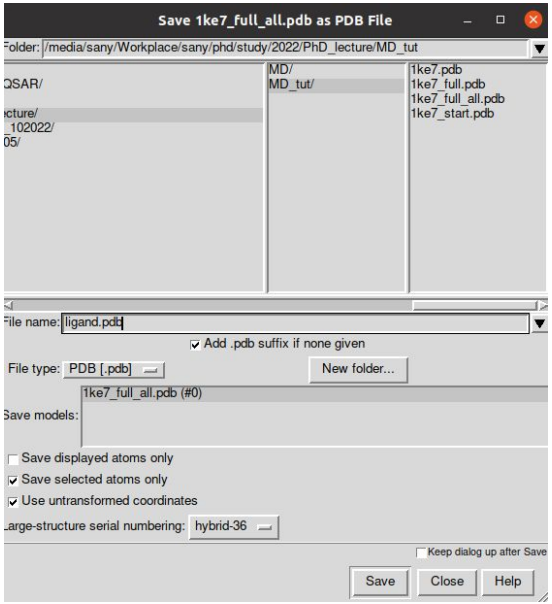
   e. Save protein only (plus ions, metals) and ligand only in *pdb* formats separately

Select -> Residue -> Ligand_id

File -> Save PDB -> Save selected only ligand.pdb

Select -> Invert (all models)

File -> Save PDB -> Save selected only protein_prepared.pdb



Ligand preparation

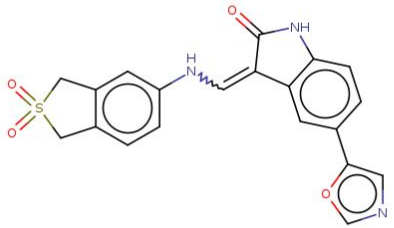
5. Save protonated smiles (pH 7.4) by Marvin into *ligand_74.smi*

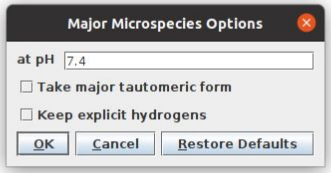
   Open Marvin Sketch

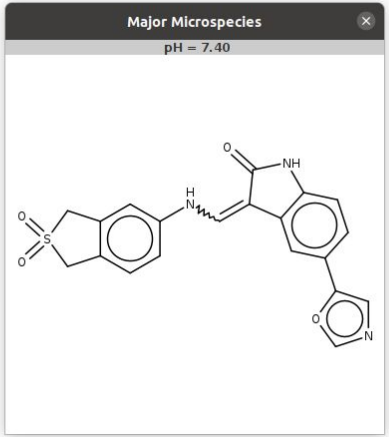
   File -> Open -> ligand.smi

   Calculations -> Protonation -> Major Microspecies -> 7.4 -> Ctrl+L (copy as Smiles)

   Save to ligand_74.smi







Ligand preparation

6. Transform pdb of ligand into mol format (add bonds information from smiles)

   Script https://github.com/ci-lab-cz/md-scripts/blob/master/scripts/pdb2mol.py

   Run in Bash/Shell:

   python md-scripts/scripts/pdb2mol.py -i ligand.pdb --smiles ligand_74.smi -o ligand.mol

Protein Ligand preparation

Input Files for MD:

protein_prepared.pdb

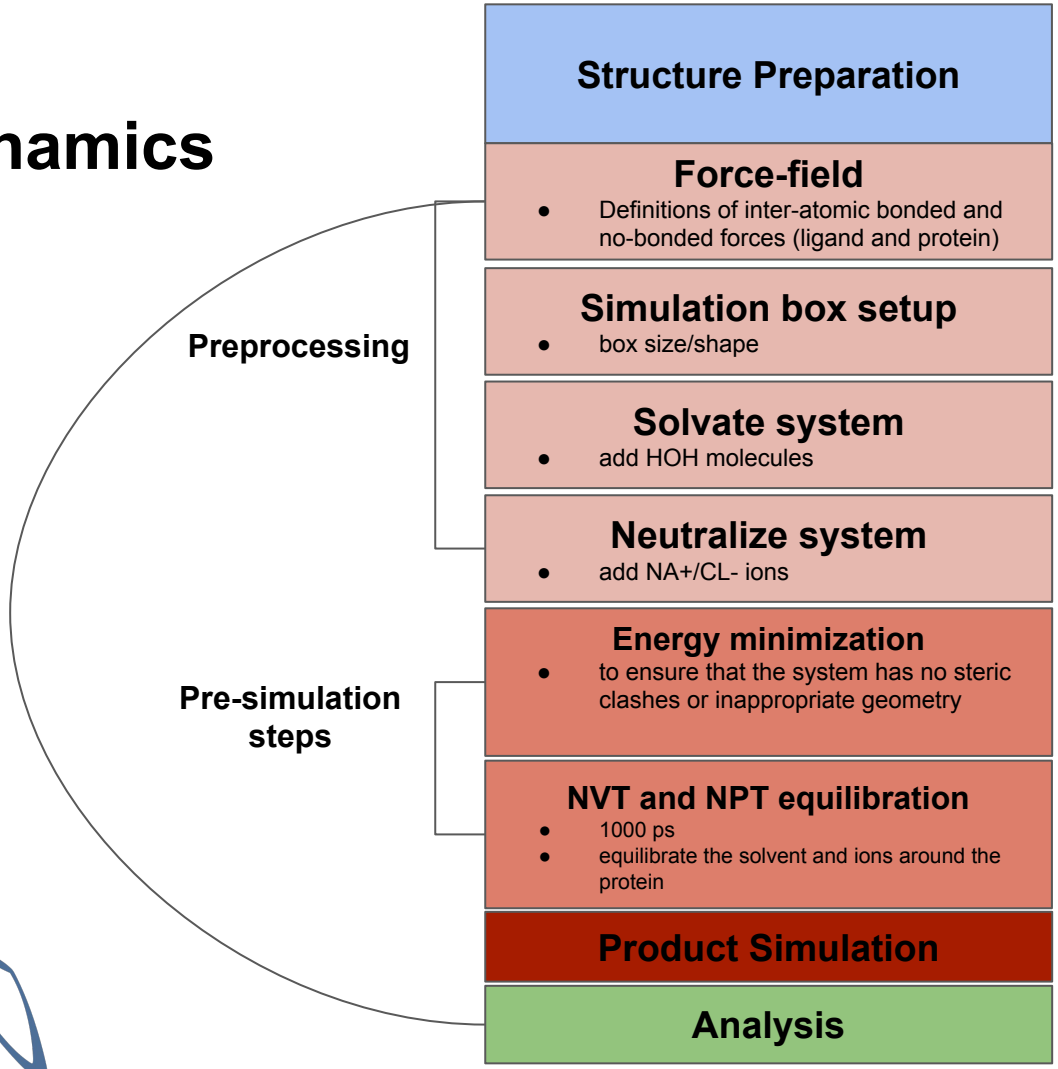
- *no missing non-terminal atoms/residues*
- *removed non-protein residues*
- *added all hydrogens*
- *set HID/HIE/HIP states of HIS*

ligand.mol

- *protonated at 7.4 pH*
- *added all hydrogens*

Classical Molecular Dynamics

https://github.com/ci-lab-cz/md-scripts



FAST. FLEXIBLE. FREE.
GROMACS

Practice: force fields

Prepare the protein topology with pdb2gmx

gmx pdb2gmx

Synopsis

```
gmx pdb2gmx [-f [<.gro/.g96/...>]] [-o [<.gro/.g96/...>]] [-p [<.top>]]
            [-i [<.itp>]] [-n [<.ndx>]] [-q [<.gro/.g96/...>]]
            [-chainsep <enum>] [-merge <enum>] [-ff <string>]
            [-water <enum>] [-[no]inter] [-[no]ss] [-[no]ter]
            [-[no]lys] [-[no]arg] [-[no]asp] [-[no]glu] [-[no]gln]
            [-[no]his] [-angle <real>] [-dist <real>] [-[no]una]
            [-[no]ignh] [-[no]missing] [-[no]v] [-posrefc <real>]
            [-vsite <enum>] [-[no]heavyh] [-[no]deuterate]
            [-[no]chargegrp] [-[no]cmap] [-[no]renum] [-[no]rtpres]
```

Description

`gmx pdb2gmx` reads a .pdb (or .gro) file, reads some database files, adds hydrogens to the molecules and generates coordinates in GROMACS (GROMOS), or optionally .pdb, format and a topology in GROMACS format. These files can subsequently be processed to generate a run input file.

New files:
protein.gro: a molecular structure in Gromos87 format.
topol.top: molecular topology
posre.itp: position restraints

gmx pdb2gmx -f protein.pdb -o protein.gro -water tip3p -ignh <<< “AMBER99SB-ILDN”

https://manual.gromacs.org/documentation/current/onlinehelp/gmx-pdb2gmx.html

Practice: force fields

Prepare the protein topology with pdb2gmx

gmx pdb2gmx

Synopsis

```
gmx pdb2gmx [-f [<.gro/.g96/...>]] [-o [<.gro/.g96/...>]] [-p [<.top>]]
            [-i [<.itp>]] [-n [<.ndx>]] [-q [<.gro/.g96/...>]]
            [-chainsep <enum>] [-merge <enum>] [-ff <string>]
            [-water <enum>] [-[no]inter] [-[no]ss] [-[no]ter]
            [-[no]lys] [-[no]arg] [-[no]asp] [-[no]glu] [-[no]gln]
            [-[no]his] [-angle <real>] [-dist <real>] [-[no]una]
            [-[no]ignh] [-[no]missing] [-[no]v] [-posrefc <real>]
            [-vsite <enum>] [-[no]heavyh] [-[no]deuterate]
            [-[no]chargegrp] [-[no]cmap] [-[no]renum] [-[no]rtpres]
```

pdb2gmx does not work on ligand

Description

`gmx pdb2gmx` reads a .pdb (or .gro) file, reads some database files, adds hydrogens to the molecules and generates coordinates in GROMACS (GROMOS), or optionally .pdb, format and a topology in GROMACS format. These files can subsequently be processed to generate a run input file.

gmx pdb2gmx -f protein.pdb -o protein.gro -water tip3p -ignh <<< “AMBER99SB-ILDN”

https://manual.gromacs.org/documentation/current/onlinehelp/gmx-pdb2gmx.html

Automation tools for ligand topology

AMBER	**Antechamber**	Parametrizes molecules using GAFF
	acpype	A Python interface to Antechamber, writes GROMACS topologies
CHARMM	**CGenFF**	The official CHARMM General Force Field server
GROMOS87/ GROMOS96	**PRODRG 2.5**	An automated server for topology generation
	ATB	A newer server for topology generation, uses GROMOS96 54A7
OPLS-AA	**TopolBuild**	Converts a Tripos .mol2 file into a topology
	TopolGen	A Perl script to convert an all-atom .pdb file to a topology
	LigParGen	A server from the Jorgensen group to produce OPLS topologies

Practice: water model

```
Select the Water Model:
 1: TIP3P     TIP 3-point, recommended
 2: TIP4P     TIP 4-point
 3: TIP4P-Ew  TIP 4-point optimized with Ewald
 4: TIP5P     TIP 5-point (see http://redmine.gromacs.org/issues/1348 for issues)
 5: SPC       simple point charge
 6: SPC/E     extended simple point charge
 7: None
```

TIP3P - 3-site rigid water molecule with charges and Lennard-Jones parameters assigned to each of the 3 atoms.

	SPC	TIP3P	SPC/E
r(OH), Å	1.0	0.9572	1.0
HOH, deg	109.47	104.52	109.47
A, 10^3 kcal Å^{12}/mol	629.4	582.0	629.4
B, kcal Å^6/mol	625.5	595.0	625.5
q(O)	−0.82	−0.834	−0.8476
q(H)	+0.41	+0.417	+0.4238

gmx pdb2gmx -f protein.pdb -o protein.gro -water tip3p -ignh <<< “AMBER99SB-ILDN”

https://manual.gromacs.org/documentation/current/onlinehelp/gmx-pdb2gmx.html

Ligand Force fields

Prepare the ligand topology using external tools

```
antechamber -i ligand.pdb -fi pdb -o lig.mol2 -fo mol2 -c bcc -pf y -s 2
parmchk2 -i lig.mol2 -f mol2 -o lig.frcmod
python pmed_amb2gmx.py -p lig.prmtop -x lig.inpcrd -o UNL
```

Build the topology for the protein-ligand system

Insert a line that says #include lig.itp" into topol.top after the position restraint file is included

Make adjustments is in the [molecules] directive

New files:

lig.mol2 *- file of ligand with charges and bonds*
lig.frcmod *- Parameter modification file*
posre_ligand.itp *- positional restraints*
ligand.gro *- ligand in Gromos87 format*
ligand.inpcrd - *Amber coordinate file of ligand*
ligand.prmtop *- Amber topology file*
ligand.top *- Gromacs topology file*

Changed files: topol.top

Solvation

To perform simulation we should add water molecules to mimic real environment.

There are two steps to defining the box and filling it with solvent:

1) Define the box dimensions using the editconf module
2) Fill the box with water using the solvate module.

We will use a simple cubic box as the unit cell.
First define the box using editconf:

```
gmx editconf -f protein.gro -o protein_newbox.gro -c -d 1.0 -bt cubic
```

a solute-box distance of 1.0 nm will mean that there are at least 2.0 nm between any two periodic images of a protein. This distance should be sufficient for just about any cutoff scheme commonly used in simulations.

New files:
protein_newbox.gro: protein+box system

Solvation

To perform simulation we should add water molecules to mimic real environment.

On the previous step we have defined a box, then we should fill it with solvent (water). Solvation is accomplished using solvate command:

gmx solvate -cp protein_newbox.gro -cs spc216.gro -o protein_solv.gro -p topol.top

-cp - the configuration of the protein (-cp)

-cs - the configuration of the solvent (-cs)

spc216.gro is a generic equilibrated 3-point solvent model (suitable for SPC, SPC/E, or TIP3P water, since they are all three-point water models)

New files:
protein_solv.gro: *protein + solvent system*
Changed files:
topol.top

Adding ions

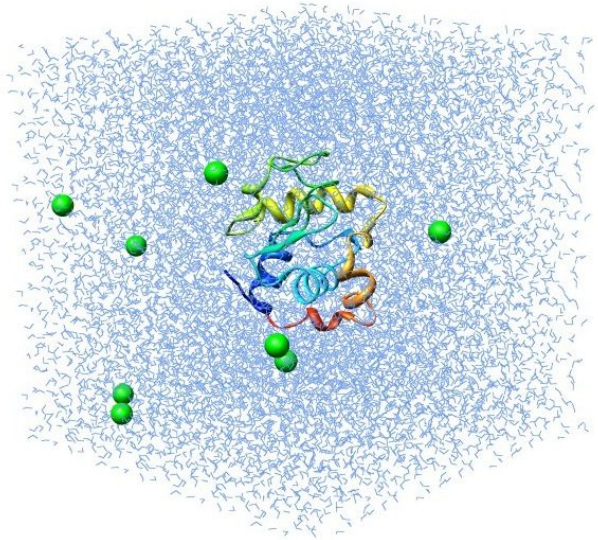
We need to add ions of CL- or NA+ to neutralize our protein:

```
gmx grompp -f ions.mdp -c protein_solv.gro -p topol.top -o ions.tpr
```

```
gmx genion -s ions.tpr -o protein_solv_ions.gro -p topol.top -pname NA -nname CL -neutral
```

New files:
protein_solv_ions.gro: *protein + solvent + ions system*
ions.tpr: *portable binary run input file. This file contains the starting structure, the molecular topology and all the simulation parameters.*

Changed files: topol.top



Solvated protein with ions

Energy minimization

Before we can begin dynamics, we must ensure that the system has no steric clashes or inappropriate geometry (equilibrium state). **The structure is relaxed through a process called energy minimization (EM)**.

Prepare files:

```
gmx grompp -f minim.mdp -c protein_solv_ions.gro -p topol.top -o  em.tpr
```

Run the minimization:

```
gmx mdrun -v -deffnm em –s em.tpr
```

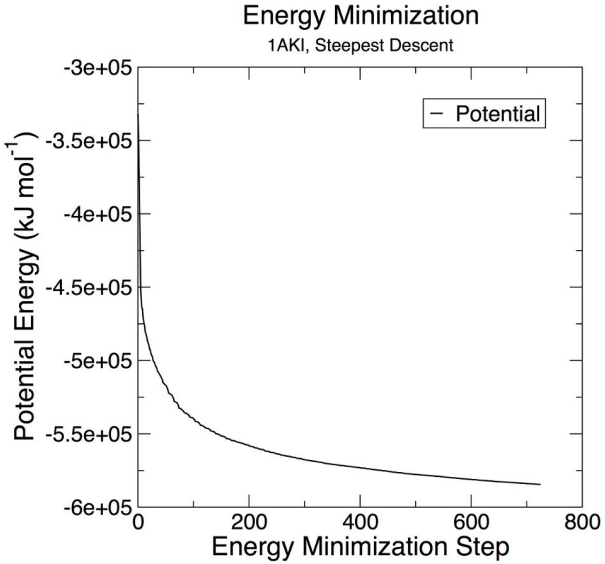
New files:
em.log: ASCII-text log file of the EM process
em.edr: Binary energy file
em.tpr: Contains the starting structure, the molecular topology and all the simulation parameters.
em.trr: Binary full-precision trajectory
em.gro: Energy-minimized structure

Energy minimization

How to do an analysis:

The em.edr file contains all of the energy terms that GROMACS collects during EM. You can analyze any .edr file using the GROMACS energy module:

```
gmx energy -f em.edr -o potential.xvg << "Potential"
```



Epot should be negative, and (for a simple protein in water) on the order of 10^5-10^6

maximum force, **Fmax**, "emtol = 1000.0" should be no greater than target 1000 kJ mol-1 nm-1.

Equilibration

Energy Minimization ensured that we have a reasonable starting structure, but to begin real dynamics, we must equilibrate the solvent and ions around the protein.

Why do we need equilibration?

If we were to attempt unrestrained dynamics at this point, the system may collapse.

The reason is that the solvent is mostly optimized within itself, and not necessarily with the solute. It needs to be brought to the temperature we wish to simulate and establish the proper orientation around the solute (the protein). After we achieved the correct temperature (based on kinetic energies), we will apply pressure to the system until it reaches the proper density.

Controlling the system

Thermodynamic system has a number of state variables which describe its macroscopic state such as

- Particle number, volume, temperature, pressure, total energy

They are not all independent, but connected by equations of state

Example: Ideal gas of non-interacting point particles

$$p = p(V,T) \qquad p \cdot V = N \cdot k_B \cdot T$$
$$E = E(V,T) \qquad p = \mathcal{S} k_B T$$
$$E = \frac{3}{2} N k_B T$$

https://web.mst.edu/~vojtat/class_5403/MolecularDynamics.pdf

Restrain the system

To equilibrate the solvent and ions around the protein we need to apply position restraints on the heavy atoms of the protein.

The purpose of **posre.itp** is to apply a position restraining force on the heavy atoms of the protein (anything that is not a hydrogen). Movement is permitted, but only after overcoming a substantial energy penalty.

NVT equilibration

Canonical ensemble (NVT)

- Particle number N
- Volume V
- Temperature T

} External parameters

- Total energy E
- Pressure P

} Observables to be calculated

In MD simulation: some state variables are external parameters, others are observables to be calculated

Requires a **thermostat**, an algorithm that adds and removes energy to keep the temperature constant

- Velocity rescaling based on equipartition theorem

$$\langle E_{kin} \rangle = \frac{1}{2} \sum_i m_i \langle \vec{v}_i^{\,2} \rangle = \frac{3}{2} N k_B T$$

- Berendsen thermostat, Anderson thermostat

https://web.mst.edu/~vojtat/class_5403/MolecularDynamics.pdf

NVT equilibration

to run NVT equlibration:

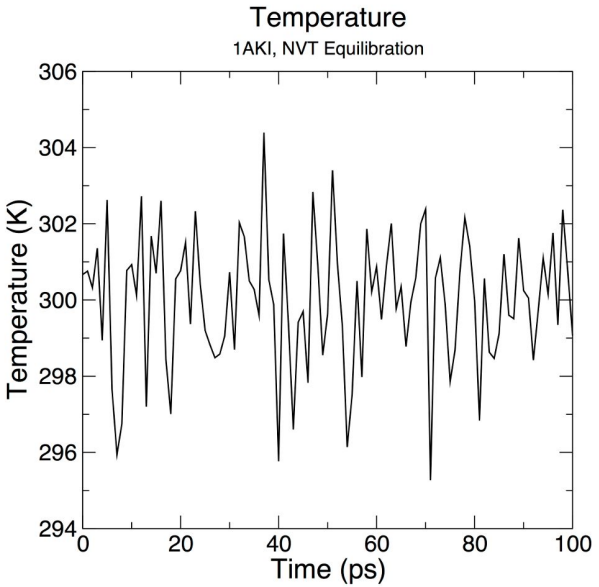
gmx grompp -f nvt.mdp -c em.gro -r em.gro -p topol.top -o nvt.tpr

gmx mdrun -deffnm nvt -s nvt.tpr

An analysis:

gmx energy -f nvt.edr -o temperature.xvg <<< “Temperature”

New files:
nvt.log: *ASCII-text log file of the equilibration process*
nvt.edr: *Binary energy file*
nvt.trr: *Binary full-precision trajectory*
nvt.gro: *NVT-minimized structure*



NPT equilibration

Isothermal–isobaric ensemble (NPT)

- Particle number N
- Pressure P
- Temperature T

} External parameters

- Total energy E
- Volume V

} Observables to be calculated

Requires a **barostat** in addition to the thermostat, an algorithm that changes volume to keep the pressure constant

https://web.mst.edu/~vojtat/class_5403/MolecularDynamics.pdf

NPT equilibration

Note that we are now including the -t flag to include the checkpoint file from the NVT equilibration

The coordinate file (-c) is the final output of the NVT simulation.

to run NPT equlibration:

gmx grompp -f npt.mdp -c nvt.gro -r nvt.gro -t nvt.cpt -p topol.top -o npt.tpr gmx mdrun -deffnm npt -s npt.tpr

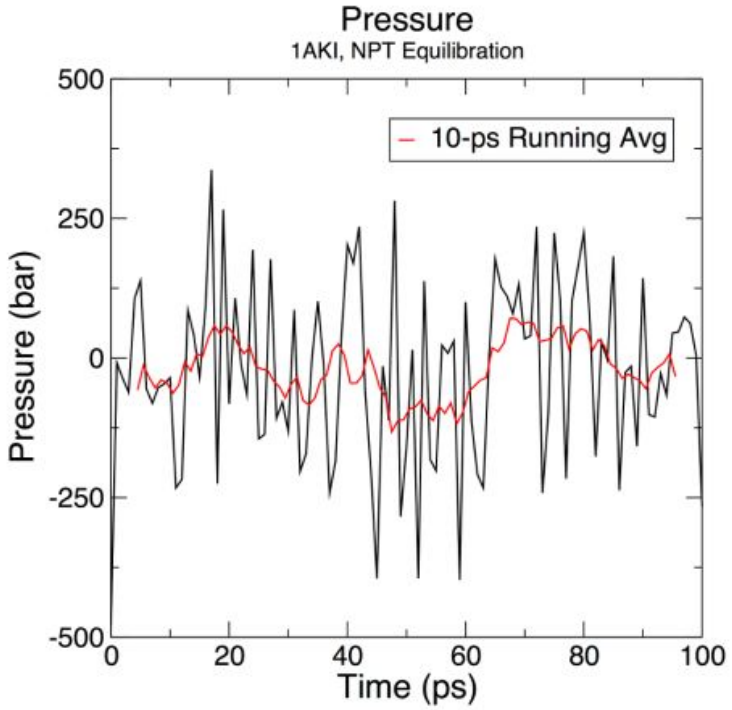
An analysis:

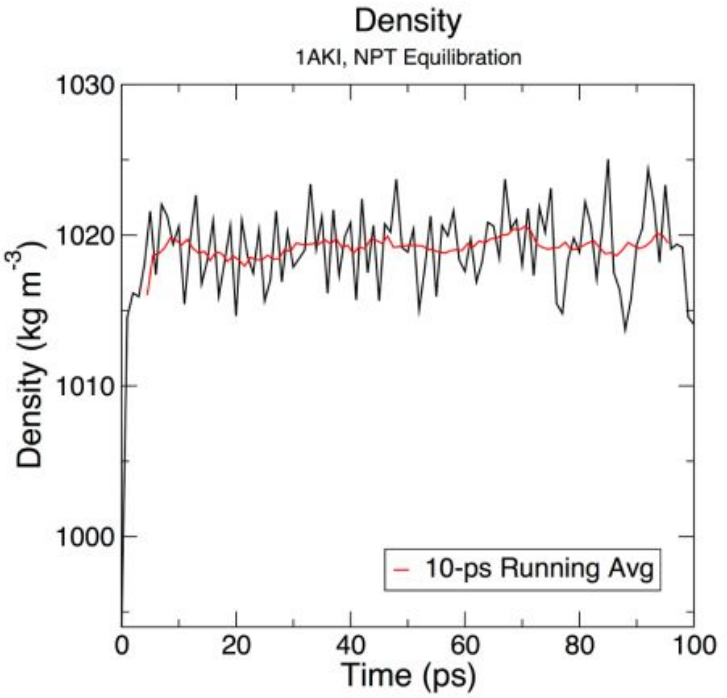
gmx energy -f npt.edr -o pressure.xvg <<< ”Pressure”

gmx energy -f npt.edr -o density.xvg <<< “Density”

New files:
nvt.log: *ASCII-text log file of the equilibration process*
nvt.edr: *Binary energy file*
nvt.trr: *Binary full-precision trajectory*
nvt.gro: *NVT-minimized structure*

NPT equilibration



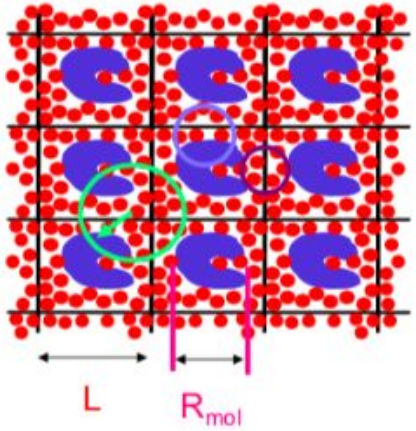


In our tool we run 1000 ps NPT equilibration by default

Periodic Boundary Conditions

Periodic boundary conditions (PBCs) are a set of boundary conditions which are often chosen for approximating a large (infinite) system by using a small part called a unit cell.

- PBCs are often used in computer simulations and mathematical models.
- The topology of two-dimensional PBC is equal to that of a world map of some video games; the geometry of the unit cell satisfies perfect two-dimensional tiling, and when an object passes through one side of the unit cell, it reappears on the opposite side with the same velocity.



Required
(no atom sees another one twice):

$L > 2R_c$ $R_c < ½L$

Preferred
(protein does not see a copy of itself)

$L > R_c + R_{mol}$ $R_c < L - R_{mol}$

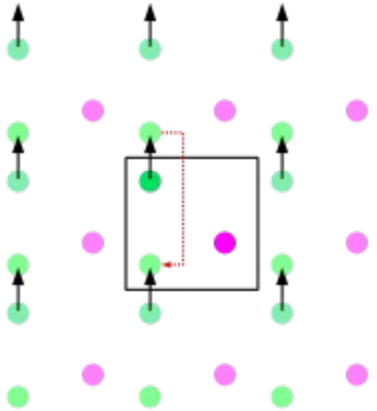
Even better
(no solvent sees two proteins)

$L > 2R_c + R_{mol}$ $R_c < ½(L - R_{mol})$

<nu mbe r>

- Cubic:

Rectangular:



Production MD

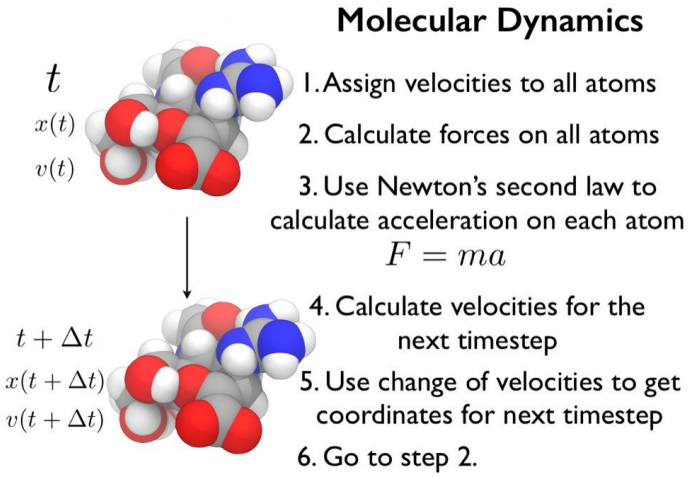
the system is now well-equilibrated at the desired temperature and pressure. We are now ready to release the position restraints and run production MD

We will run MD simulation

gmx grompp -f md.mdp -c npt.gro -t npt.cpt -p topol.top -o md_0_1.tpr

gmx mdrun -deffnm md_0_1 -s md_0_1.tpr

New files:
md_0_1.tpr: portable binary run input file. It contains the starting structure of the simulation, the molecular topology and all the simulation parameters.
md_0_1.log: *ASCII-text log file of the equilibration process*
md_0_1.cpt: *portable checkpoint file. The complete state of the simulation is stored in the checkpoint file, including extended thermostat/barostat variables, random number states and NMR time averaged data*
md_0_1.edr*: Binary energy file*
md_0_1.xtc: *Binary full-precision trajectory*
md_0_1.gro*:* starting structure of the simulation in



Analysis of calculated MD simulation

Remove PBC:

gmx trjconv -s md_out.tpr -f md_out.xtc -pbc nojump -o md_out_noj_noPBC.xtc <<< "System"

Center system:

gmx trjconv -s md_out.tpr -f md_out_noj_noPBC.xtc -o md_centermolsnoPBC.xtc -pbc mol -center -n index.ndx <<< "Protein_UNL System"

Alignment of all frames (Remove rotations and translations):

gmx trjconv -s md_out.tpr -f md_centermolsnoPBC.xtc -fit rot+trans -o md_fit.xtc -n index.ndx <<< "Protein_UNL System"

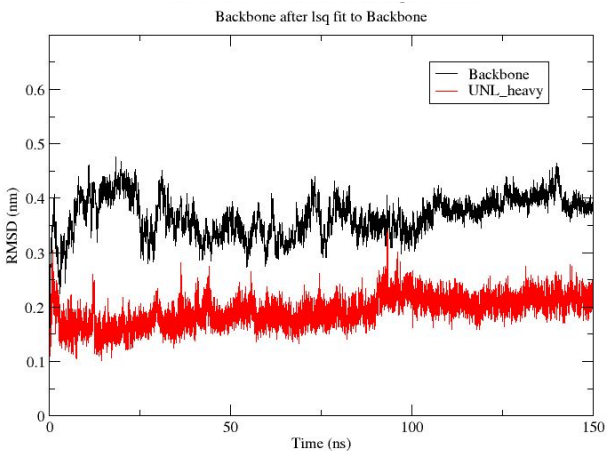
Analysis of calculated MD simulation

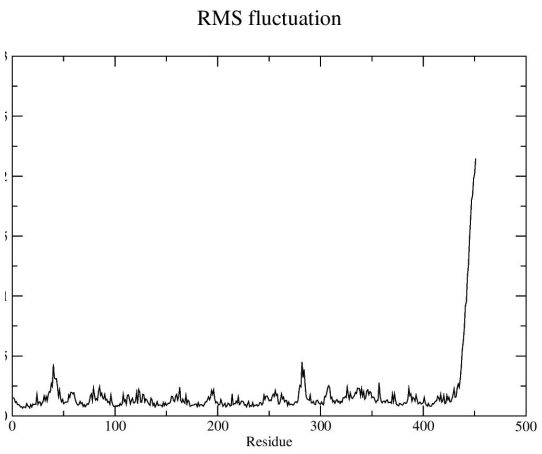
RMSD:

gmx rms -s md_0_1.tpr -f md_0_1_noPBC.xtc -o rmsd.xvg -tu ns

RMSF:

gmx rmsf -s md_0_1.tpr -f md_0_1_noPBC.xtc -o rmsf.xvg -oq rmsf.pdb -res





xmgrace

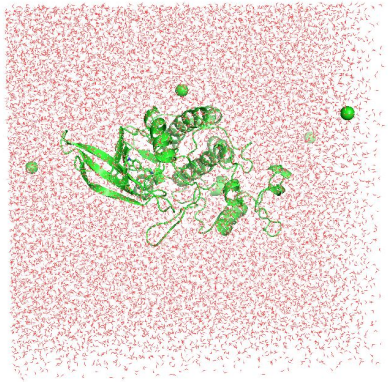
MD simulation by one command:

Output files:

md_out.tpr: portable binary run input file. It contains the starting structure of the simulation, the molecular topology and all the simulation parameters.
md*_out.*log: *ASCII-text log file of the equilibration process*
md*_out.*cpt: *portable checkpoint file. The complete state of the simulation is stored in the checkpoint file, including extended thermostat/barostat variables, random number states and NMR time averaged data*
***md*_out.*edr*:** *Binary energy file*
md*_out.*xtc: *Binary full-precision trajectory*
***md*_out.*gro*:** starting structure of the simulation

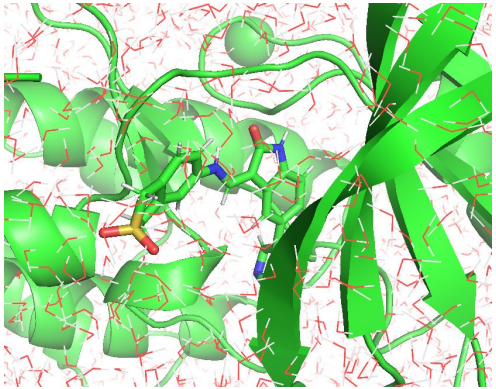
md_fit.xtc: *fitted trajectory (removed PBC and the rotation and translation, centered) to use for rmsd and energy calculation analysis*
md_short_forcheck.xtc: *fitted short trajectory (each 100 step is skipped)*
frame.pdb: a frame from the trajectory to provide topology

MD trajectory analysis files:
rmsd.xvg - rmsd of Backbone of protein
rmsd_UNL.xvg - rmsd of the heavy atoms of ligand
gyrate.xvg - radius of gyration of the protein
rmsf.xvg - root mean square fluctuation of each amino acids



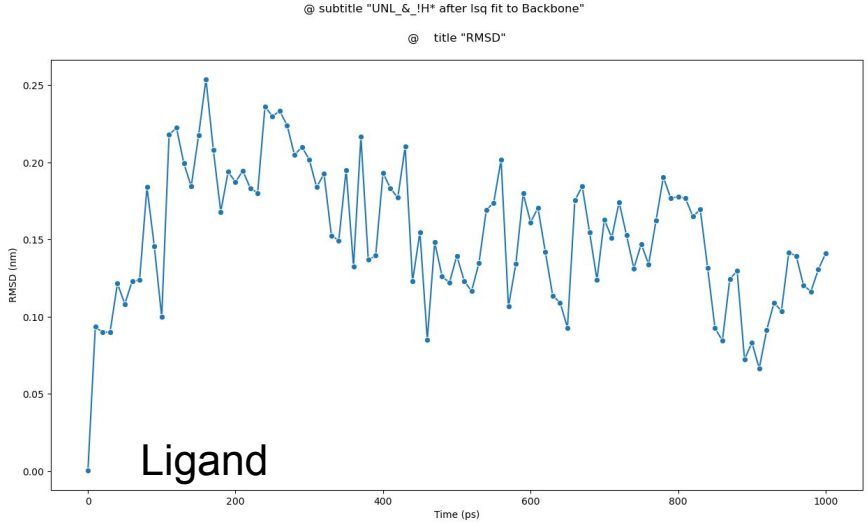
Check your own MD trajectory

frame.pdb - a frame from the trajectory to provide topology
md_fit.xtc - your fitted MD trajectory
md_short_forcheck.xtc - short part of the simulation



Analysis of calculated MD simulation

```
(md) [anikonenko1@cn299.karolina MD_tutorial]$ ls mdrun/md_files/md_run/protein_HIS_ligand_1/
all.itp              em.tpr      ligand_1.itp           mdout.mdp              npt.edr  nvt.gro        pressure.png       rmsf.pdb                                                  topol.top
all_ligand_resid.txt em.trr      md_centermolsnoPBC.xtc md_out_noj_noPBC.xtc   npt.gro  nvt.log        pressure.xvg       rmsf.png
complex.gro          frame.pdb   md_fit.xtc             md_out.tpr             npt.log  nvt.mdp        rmsd_ligand_1.png  rmsf.xvg
density.png          gyrate.png  md.mdp                 md_out.xtc             npt.mdp  nvt.tpr        rmsd_ligand_1.xvg  solv.gro
density.xvg          gyrate.xvg  md_out.cpt             md_short_forcheck.xtc  npt.tpr  nvt.trr        rmsd.png           solv_ions.gro
em.edr               index.ndx   md_out.edr             minim.mdp              npt.trr  posre.itp      rmsd_xtal.png      streamd_bash_protein_HIS_ligand__26-11-2023-20-10-17.log
em.gro               ions.mdp    md_out.gro             newbox.gro             nvt.cpt  potential.png  rmsd_xtal.xvg      temperature.png
em.log               ions.tpr    md_out.log             npt.cpt                nvt.edr  potential.xvg  rmsd.xvg           temperature.xvg
```

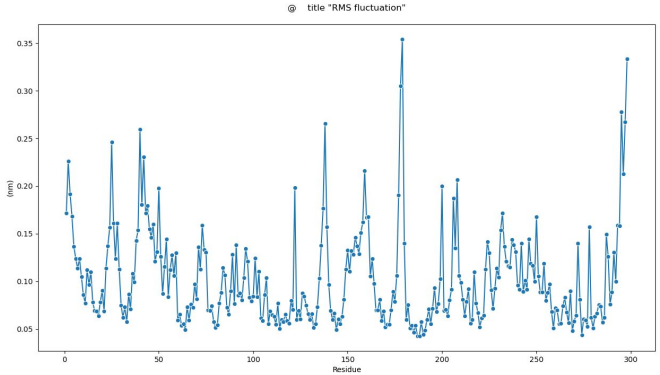


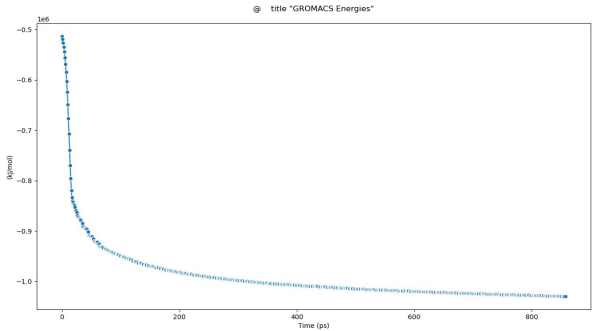
1 ns

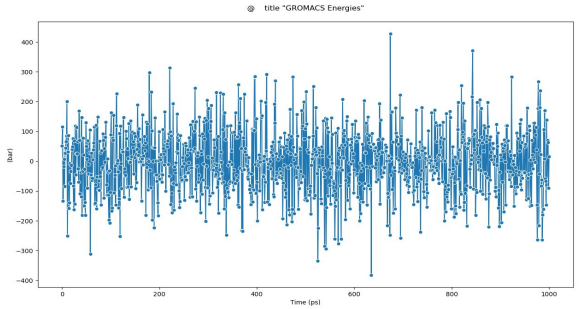
Analysis of the calculated MD simulation

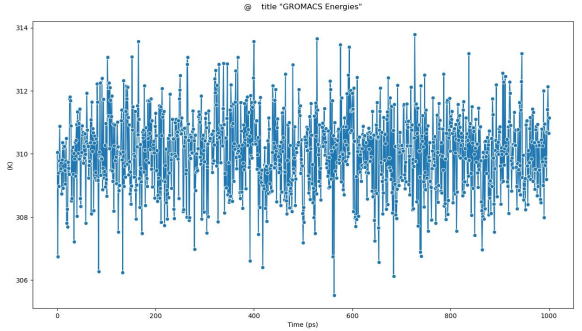
```
# This file was created Mon Nov 28 18:33:21 2022
# Created by:
#         :-) GROMACS - gmx rms, 2021.4-EasyBuild-4.5.1-PLUMED-2.7.3 (-:
#
# Executable:   /apps/all/GROMACS/2021.4-foss-2020b-PLUMED-2.7.3/bin/gmx
# Data prefix:  /apps/all/GROMACS/2021.4-foss-2020b-PLUMED-2.7.3
# Working dir:  /mnt/proj1/dd-22-84/MD_tutorial/ligand_01ns
# Command line:
#   gmx rms -s md_out.tpr -f md_fit.xtc -o rmsd.xvg -n index.ndx -tu ps
# gmx rms is part of G R O M A C S:
#
# God Rules Over Mankind, Animals, Cosmos and Such
#
@    title "RMSD"
@    xaxis  label "Time (ps)"
@    yaxis  label "RMSD (nm)"
@TYPE xy
@ subtitle "Backbone after lsq fit to Backbone"
    0.0000000    0.0004955
   10.0000000    0.0872569
   20.0000000    0.0738152
   30.0000000    0.1004452
   40.0000000    0.1048534
   50.0000000    0.0945487
   60.0000000    0.0887801
   70.0000000    0.0993429
   80.0000000    0.1069565
   90.0000000    0.1168087
  100.0000000    0.1420962
```

```
# This file was created Mon Nov 28 18:33:21 2022
# Created by:
#         :-) GROMACS - gmx rms, 2021.4-EasyBuild-4.5.1-PLUMED-2.7.3 (-:
#
# Executable:   /apps/all/GROMACS/2021.4-foss-2020b-PLUMED-2.7.3/bin/gmx
# Data prefix:  /apps/all/GROMACS/2021.4-foss-2020b-PLUMED-2.7.3
# Working dir:  /mnt/proj1/dd-22-84/MD_tutorial/ligand_01ns
# Command line:
#   gmx rms -s md_out.tpr -f md_fit.xtc -o rmsd_UNL.xvg -n index.ndx -tu ps
# gmx rms is part of G R O M A C S:
#
# God Rules Over Mankind, Animals, Cosmos and Such
#
@    title "RMSD"
@    xaxis  label "Time (ps)"
@    yaxis  label "RMSD (nm)"
@TYPE xy
@ subtitle "UNL_&_!H* after lsq fit to Backbone"
    0.0000000    0.0005219
   10.0000000    0.0578194
   20.0000000    0.0515443
   30.0000000    0.1366709
   40.0000000    0.1673483
   50.0000000    0.1613055
   60.0000000    0.1436562
   70.0000000    0.1480425
   80.0000000    0.1522363
   90.0000000    0.1496438
  100.0000000    0.0998629
```







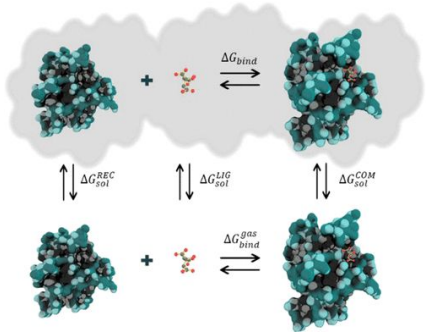


MMPBSA / MMGBSA

End-state free energy calculations
with GROMACS files

What can be done by MD

- to estimate binding affinity of protein-ligand complexes
  - Molecular mechanics Poisson–Boltzmann surface area (**MM/PBSA**)
  - Molecular mechanics generalized Born surface area (**MM/GBSA**)



In the MM/PBSA or MM/GBSA approach, the free energy for binding of the ligand (L) to the protein receptor (R) to form the complex (RL),

$$\Delta G_{\text{bind}} = G_{\text{RL}} - G_{\text{R}} - G_{\text{L}} \quad (4)$$

can be decomposed into contributions of different interactions and expressed as (58)

$$\Delta G_{\text{bind}} = \Delta H - T\Delta S = \Delta E_{\text{MM}} + \Delta G_{\text{sol}} - T\Delta S \quad (5)$$

in which

$$\Delta E_{\text{MM}} = \Delta E_{\text{int}} + \Delta E_{\text{ele}} + \Delta E_{\text{vdW}} \quad (6)$$

$$\Delta G_{\text{sol}} = \Delta G_{\text{PB/GB}} + \Delta G_{\text{SA}} \quad (7)$$

$$\Delta G_{\text{SA}} = \gamma \cdot \text{SASA} + b \quad (8)$$

Valdés-Tresanco, M.S., Valdés-Tresanco, M.E., Valiente, P.A. and Moreno E. ***gmx_MMPBSA: A New Tool to Perform End-State Free Energy Calculations with GROMACS.*** Journal of Chemical Theory and Computation, 2021 17 (10), 6281-6291. https://pubs.acs.org/doi/10.1021/acs.jctc.1c00645.
***MMPBSA.py*: An Efficient Program for End-State Free Energy Calculations** Bill R. Miller III, T. Dwight McGee Jr., Jason M. Swails, Nadine Homeyer, Holger Gohlke, and Adrian E. Roitberg Journal of Chemical Theory and Computation **2012** *8* (9), 3314-3321 DOI: 10.1021/ct300418h

What can be done by MD

- to estimate binding affinity of protein-ligand complexes

Total $G_{Binding}$=

- **Gas-phase molecular mechanics energy ΔE_{MM}:**
  - changes in the **internal energies** ΔE_{int} (bond, angle, and dihedral energies)
  - **electrostatic energies** ΔE_{ele}
  - **van der Waals energies** ΔE_{vdW}
- **sum of the electrostatic solvation energy G_{sol}**
  - **The polar contribution** is calculated using either the PB or GB model ($\Delta G_{PB/GB}$). Where the GB method gives an analytical expression for the polar solvation energy and is thus much faster than the PB method.
  - **nonpolar energy** is usually estimated using the solvent-accessible surface area (**SASA**)
- The change in **conformational entropy** $-T\Delta S$
  - is usually calculated by normal-mode analysis (or Interaction entropy) on a set of conformational snapshots taken from MD simulations.

In the MM/PBSA or MM/GBSA approach, the free energy for binding of the ligand (L) to the protein receptor (R) to form the complex (RL),

$$\Delta G_{bind} = G_{RL} - G_R - G_L \quad (4)$$

can be decomposed into contributions of different interactions and expressed as (58)

$$\Delta G_{bind} = \Delta H - T\Delta S = \Delta E_{MM} + \Delta G_{sol} - T\Delta S \quad (5)$$

in which

$$\Delta E_{MM} = \Delta E_{int} + \Delta E_{ele} + \Delta E_{vdW} \quad (6)$$

$$\Delta G_{sol} = \Delta G_{PB/GB} + \Delta G_{SA} \quad (7)$$

$$\Delta G_{SA} = \gamma \cdot SASA + b \quad (8)$$

MMPBSA.in

```
Sample input file for PB/GB calculation
#This input file is meant to show only that gmx_MMPBSA works. Although, we tried to use the input
files as recommended in the
#Amber manual, some parameters have been changed to perform more expensive calculations in a
reasonable amount of time. Feel free to change the parameters
#according to what is better for your system.
&general
sys_name="PB_GB_IE",
startframe=1, interval=1, verbose=2, PBRadii=3,
interaction_entropy=1, ie_segment=100, temperature=310
/
&gb
igb=5, saltcon=0.150,
/
&pb
istrng=0.15, fillratio=4.0, radiopt=0, indi=1, exdi=80.0
/
```

run_gbsa -i mdrun/md_files/md_run/protein_HIS_ligand_1/

```
EDISPER              -1407.08         18.20         18.20          5.49          5.49

GGAS                 -3184.42        101.56         74.86         30.62         22.57
GSOLV                -2452.28         57.05         52.79         17.20         15.92

TOTAL                -5636.70        116.49         31.95         35.12          9.63

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
Ligand:
Energy Component       Average     SD(Prop.)          SD   SEM(Prop.)         SEM
--------------------------------------------------------------------------------
BOND                     11.40          2.28          2.28          0.69          0.69
ANGLE                    45.46          2.79          2.79          0.84          0.84
DIHED                    23.24          2.06          2.06          0.62          0.62
VDWAALS                  -3.31          0.62          0.62          0.19          0.19
EEL                      95.11          0.99          0.99          0.30          0.30
1-4 VDW                   7.97          0.54          0.54          0.16          0.16
1-4 EEL                -225.15          1.58          1.58          0.48          0.48
EPB                     -34.82          0.85          0.85          0.26          0.26
ENPOLAR                  40.24          0.17          0.17          0.05          0.05
EDISPER                 -42.92          0.18          0.18          0.06          0.06

GGAS                    -45.28          4.62          3.37          1.39          1.02
GSOLV                   -37.50          0.89          0.95          0.27          0.29

TOTAL                   -82.78          4.71          2.96          1.42          0.89

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
Delta (Complex - Receptor - Ligand):
Energy Component       Average     SD(Prop.)          SD   SEM(Prop.)         SEM
--------------------------------------------------------------------------------
ΔBOND                    -0.00          2.08          0.00          0.63          0.00
ΔANGLE                   -0.00          2.09          0.00          0.63          0.00
ΔDIHED                    0.00          1.45          0.00          0.44          0.00
ΔVDWAALS                -46.46          0.60          2.50          0.18          0.75
ΔEEL                    -41.29          0.26          7.77          0.08          2.34
Δ1-4 VDW                 -0.00          0.30          0.00          0.09          0.00
Δ1-4 EEL                  0.00          1.33          0.00          0.40          0.00
ΔEPB                     64.17          0.71          4.26          0.22          1.28
ΔENPOLAR                -32.09          0.07          0.75          0.02          0.23
ΔEDISPER                 57.51          0.07          0.85          0.02          0.26

ΔGGAS                   -87.75          0.66          7.62          0.20          2.30
ΔGSOLV                   89.59          0.72          4.38          0.22          1.32

ΔTOTAL                    1.84          0.98          6.93          0.29          2.09
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
Using Interaction Entropy Approximation:
ΔG binding =      4.51 +/-     6.93
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
```

GENERALIZED BORN:
POISSON BOLTZMANN:

PBSA

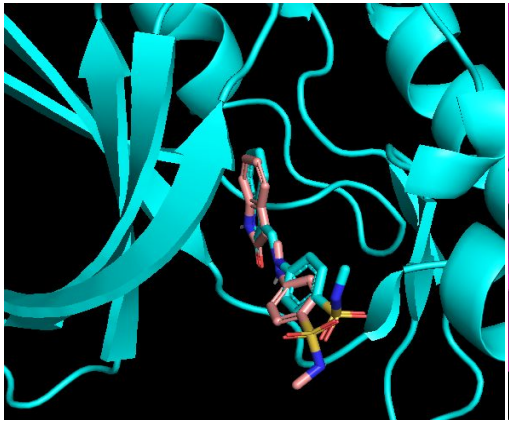
	A	B	C
	fname	**ΔG_binding**	**ΔG_binding_+/-**
	protein_HIS_igand_1	-35.81	0.74

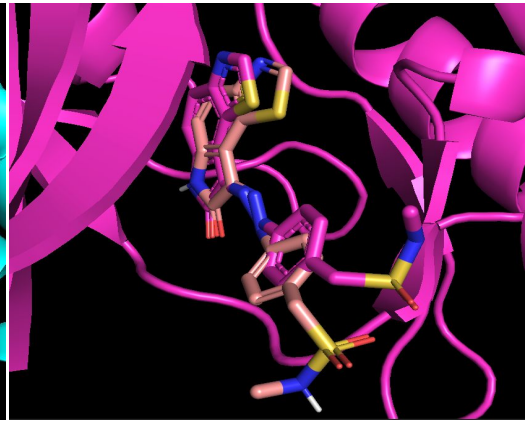
GBSA

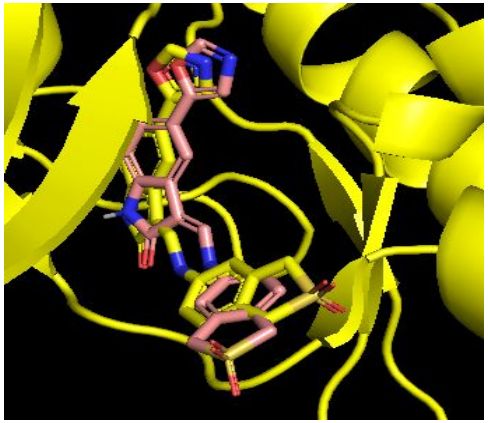
fname	**ΔG_binding**	**ΔG_binding_+/-**
protein_HIS_ligand_1	-24.31	4.1

MMPBSA Energy and **MMGBSA Energy** cannot be compared within the different methods. But you can rank your molecules by energies obtained from each method separately.

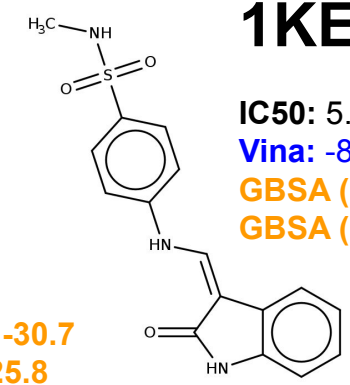
CYCLIN-DEPENDENT KINASE 2 (CDK2)





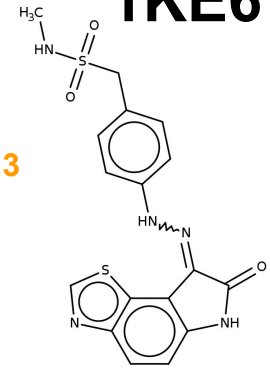


1KE5



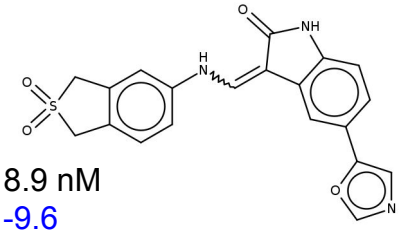
IC50: 560 nM
Vina: -8.7
GBSA (0.1 ns): -30.7
GBSA (1 ns): -25.8

1KE6



IC50: 5.7 nM
Vina: -8.9
GBSA (0.1ns): -23.3
GBSA (1ns): -21.5

1KE7



IC50: 8.9 nM
Vina: -9.6
GBSA (0.1ns): -19.3
GBSA (1ns): -27.1

Thank you for your attention!