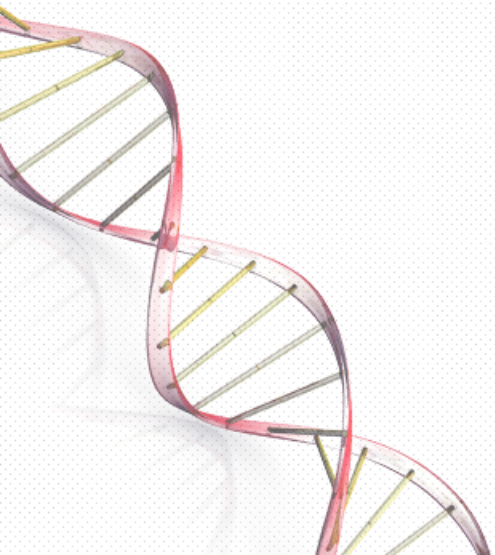
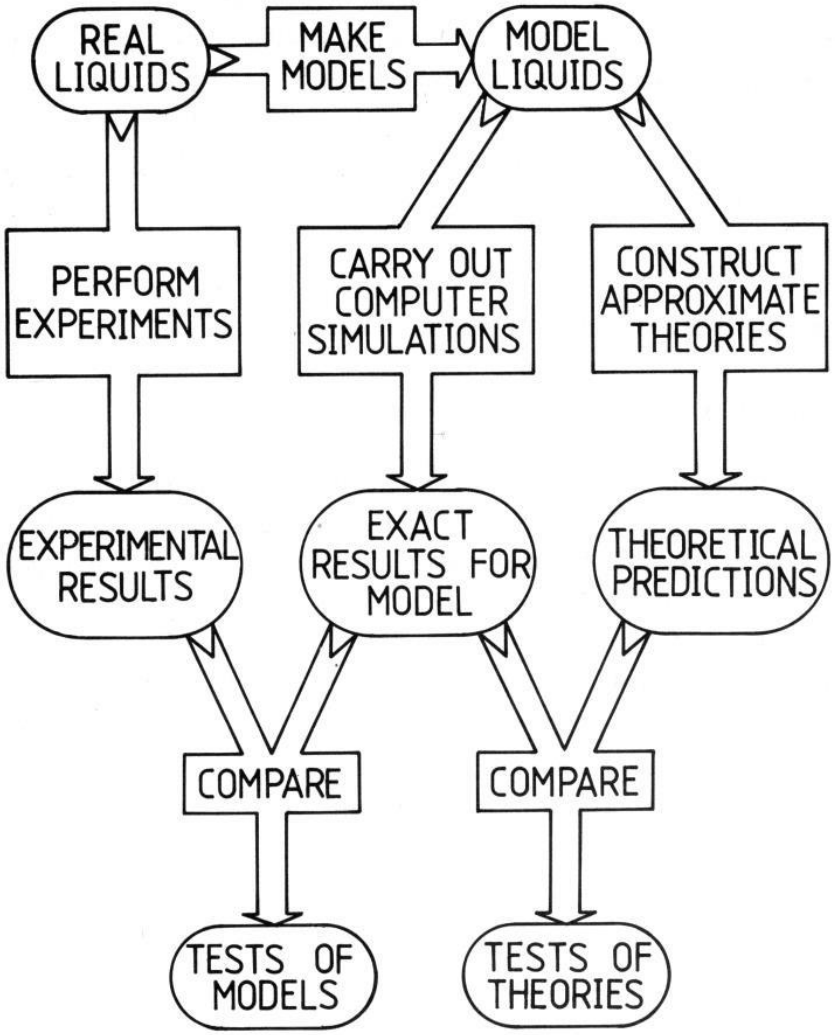
Analiza trajektorii dynamiki molekularnej

- Semestr Letni 2018



Magdalena Mozolewska

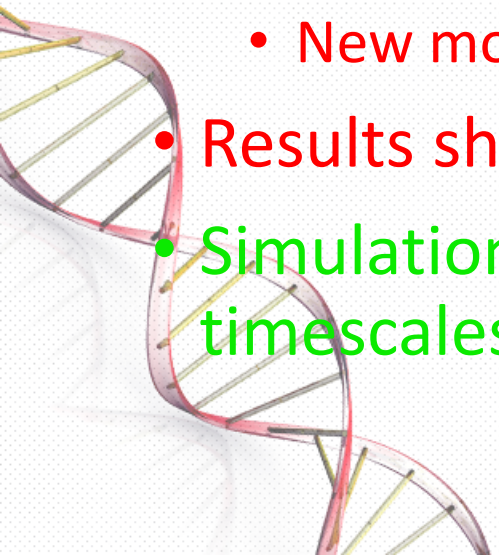
Schemat interakcji teoria-obliczenia-eksperyment



M.P. Allen D.J Tildesley, *Computer Simulation of Liquids*, Oxford Univ. Press 1987

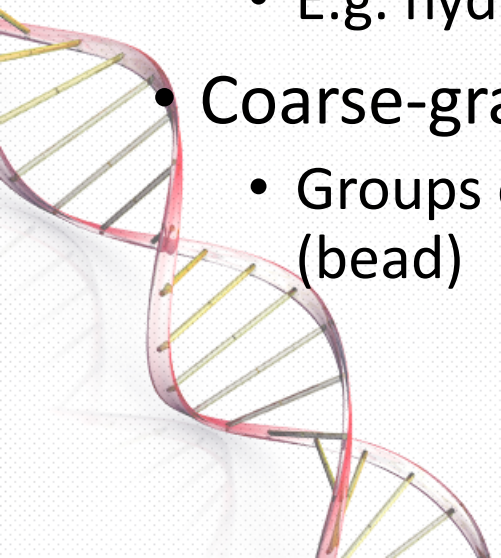
Force field

- Set of formulas, parameters and weights
- Potential (conformational) energy defined based on the positions (coordinates) of atoms (interaction sites) in internal or cartesian coordinates
- Atom is the smallest interacting object
- Usage limited to the parametrized objects
  - New molecules have to be parametrized first
- Results should be validated
- Simulation of relatively large systems in reasonable timescales possible



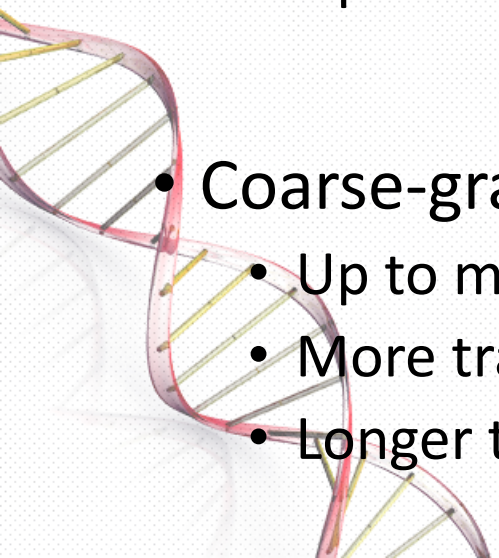
Force field

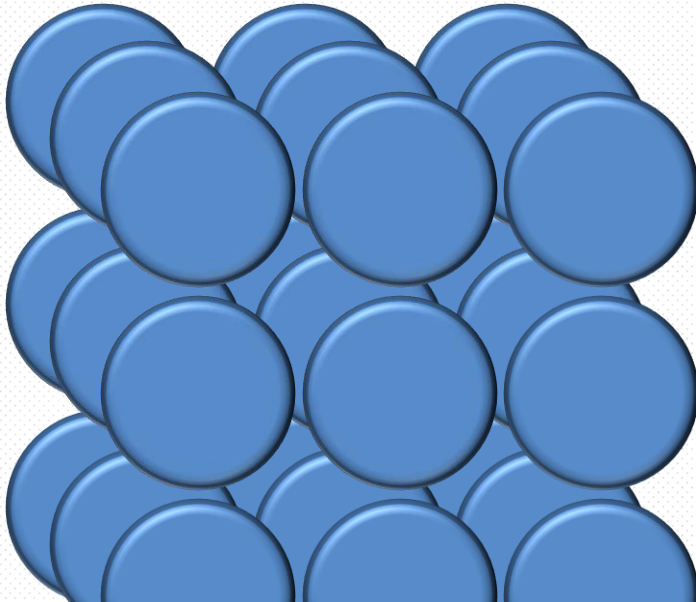
- All-atom force fields
  - Every atom is parametrized and interacting
  - Some bonds (e.g. with hydrogen) can be frozen
  - Time step of 0.1 to 0.2 fs
  - Most versatile
- United-atom force fields
  - E.g. hydrogens unified with heavy atoms
- Coarse-grained force fields
  - Groups of atoms simplified to single interacting site (bead)



Computational methods

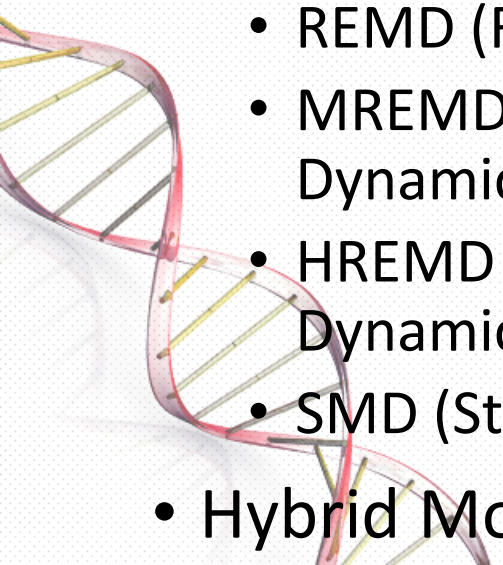
- Quantum mechanics
  - Up to hundreds of atoms
- DFT and semiempirical
  - Up to thousands of atoms
- All-atom force fields
  - Up to millions of atoms
- Coarse-grained force fields
  - Up to millions of atoms
  - More trajectories = better sampling
  - Longer time scales





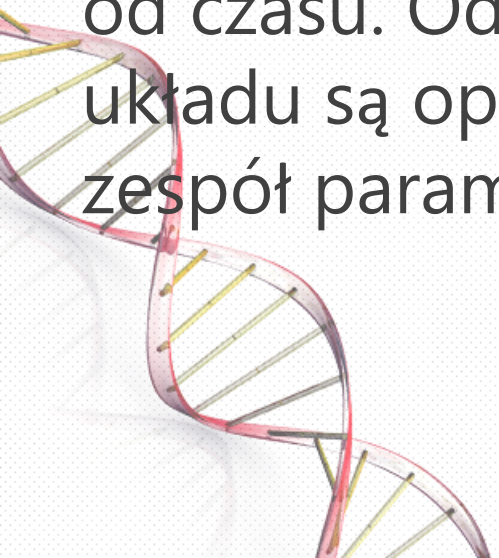
Simulation types

- Single-point energy calculation
- Energy minimization
- Monte Carlo
  - MCM (Monte Carlo-with-Minimization)
  - CSA (Conformational Space Annealing)
  - CFMC (Conformational Family Monte Carlo)
- Molecular dynamics
  - REMD (Replica Exchange Molecular Dynamics)
  - MREMD (Multiplexed Replica Exchange Molecular Dynamics)
  - HREMD (Hamiltonian Replica Exchange Molecular Dynamics)
  - SMD (Steered Molecular Dynamics)
- Hybrid Monte Carlo



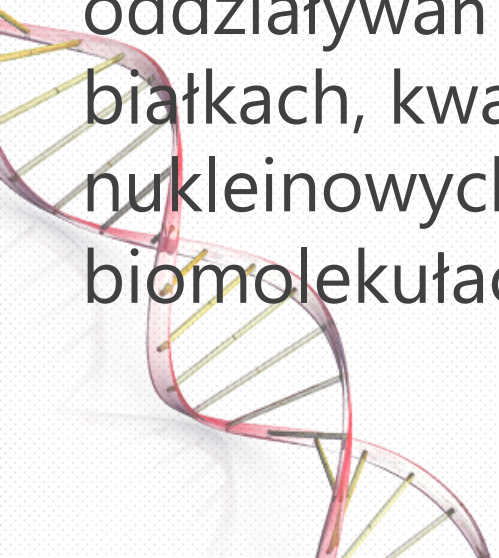
Dynamika molekularna (MD)

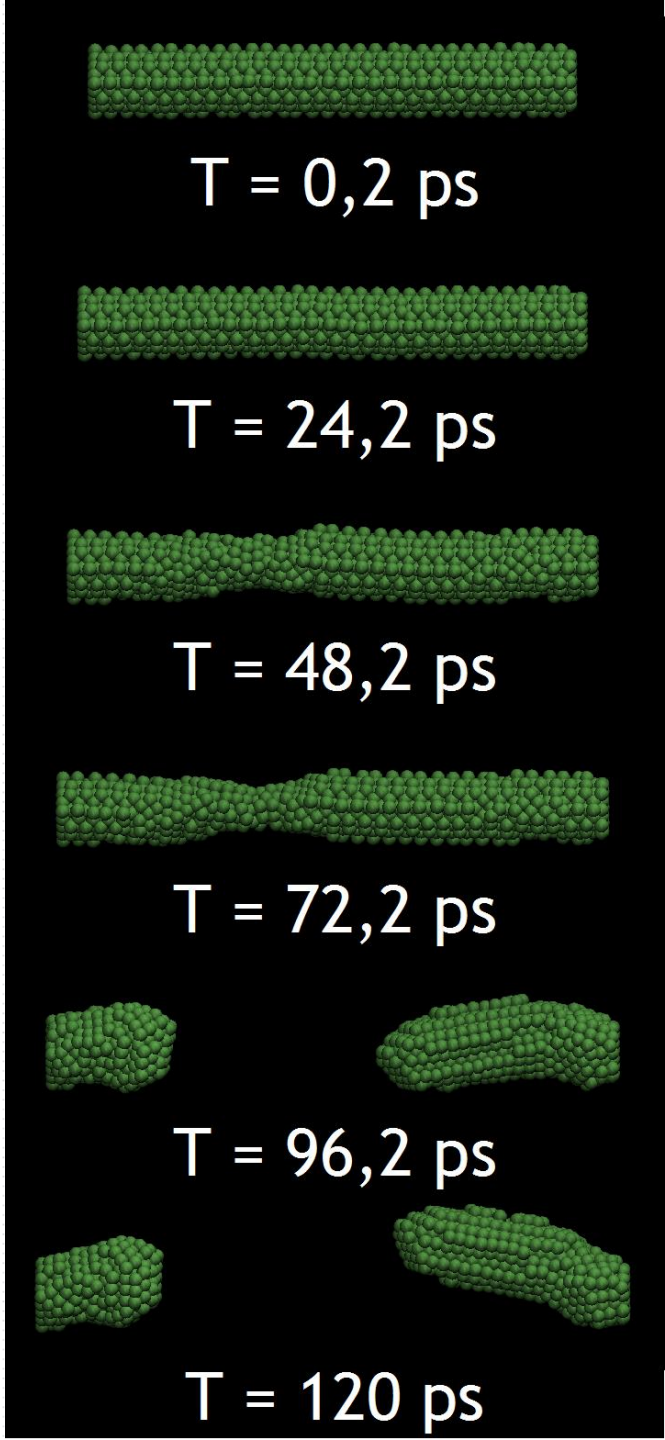
- numeryczne rozwiązywanie i komputerowa symulacja przestrzeni fazowej dla modelu układu molekuł. Elementarne poprzez całkowanie równań ruchu Newtona lub kompleksowo z uwzględnieniem licznych oddziaływań w celu uzyskania informacji o właściwościach zależnych od czasu. Oddziaływania między elementami układu są opisywane przez pewną funkcję oraz zespół parametrów dla tej funkcji.



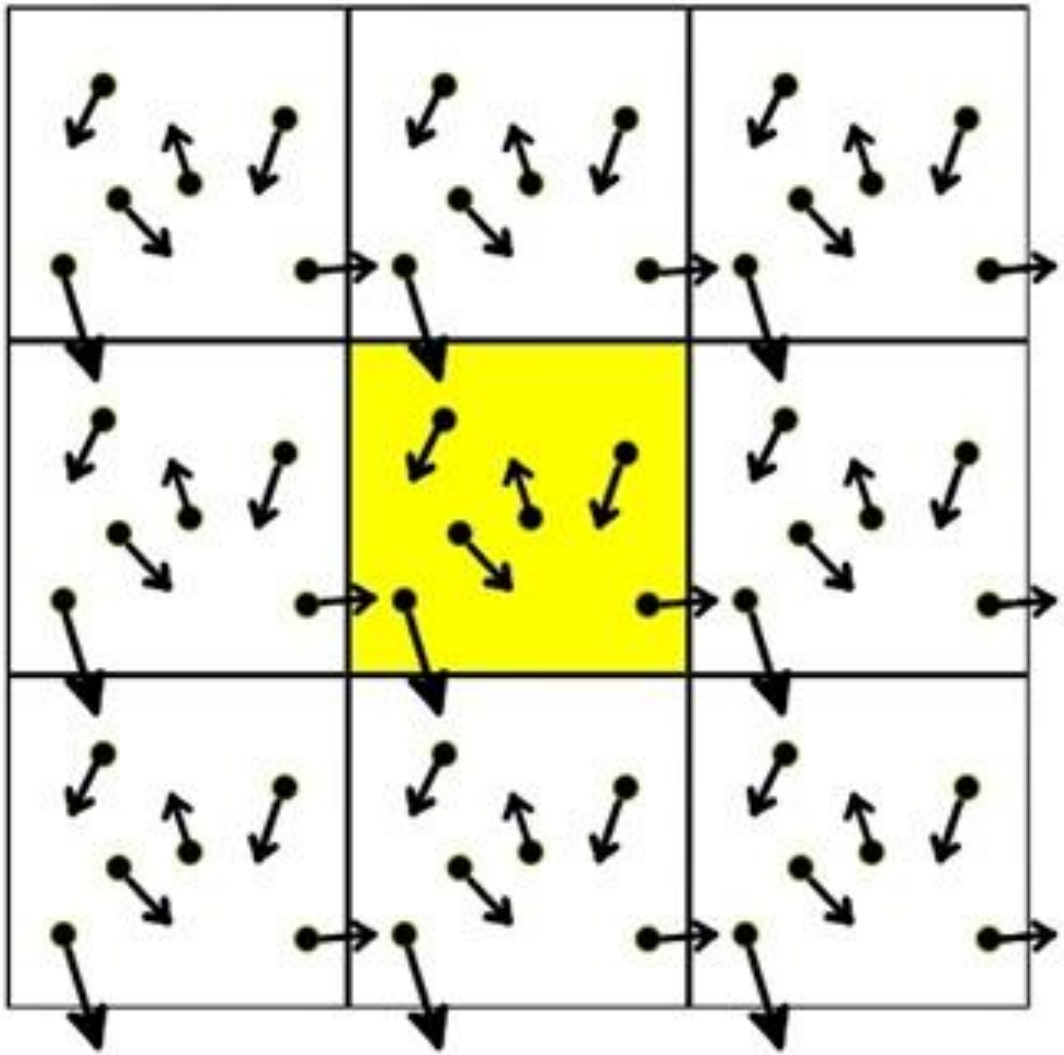
Dynamika molekularna

- Dynamika molekularna znajduje zastosowanie między innymi w biochemii jako narzędzie do poznawania struktury i oddziaływań w białkach, kwasach nukleinowych i innych biomolekułach.



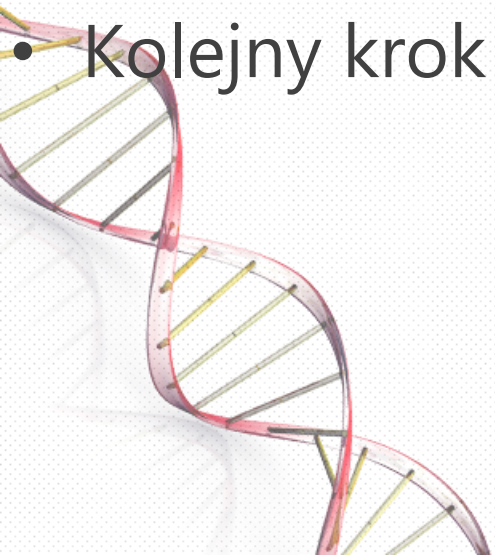


Dynamika molekularna



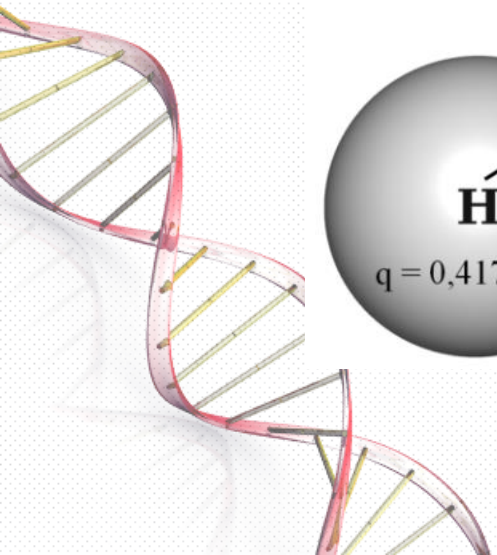
Dynamika molekularna - przebieg

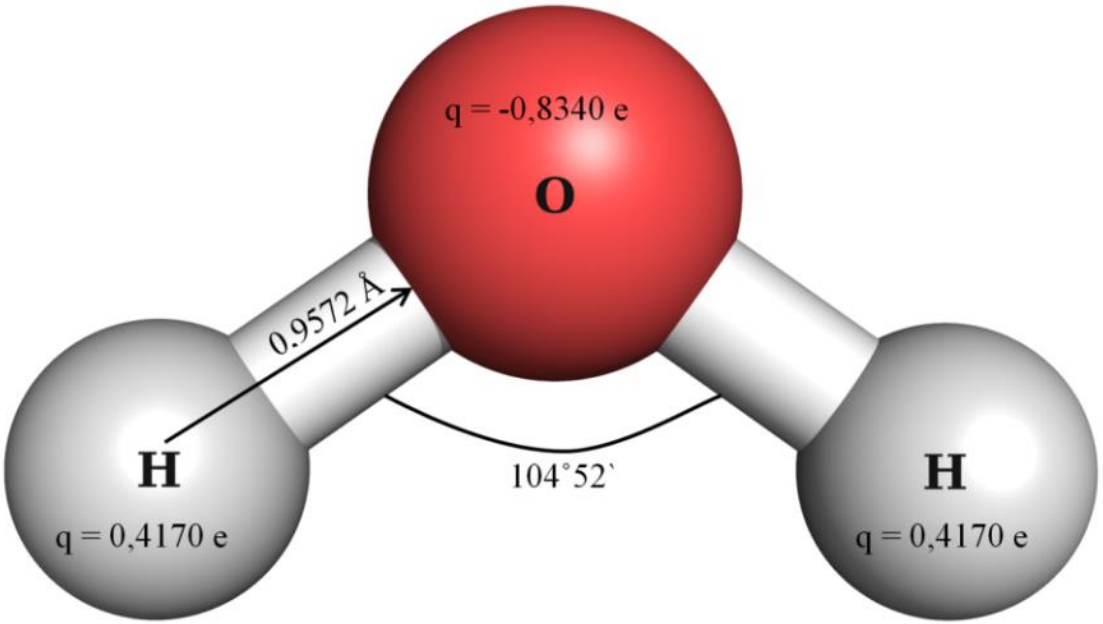
- Definiowanie położeń oraz prędkości atomów
- Wyznaczanie potencjału oddziaływania pomiędzy atomami oraz rozwiązywanie równania ruchu dla wszystkich atomów w układzie
- Obliczanie wielkości fizycznych oraz współrzędnych atomów
- Kolejny krok czasowy



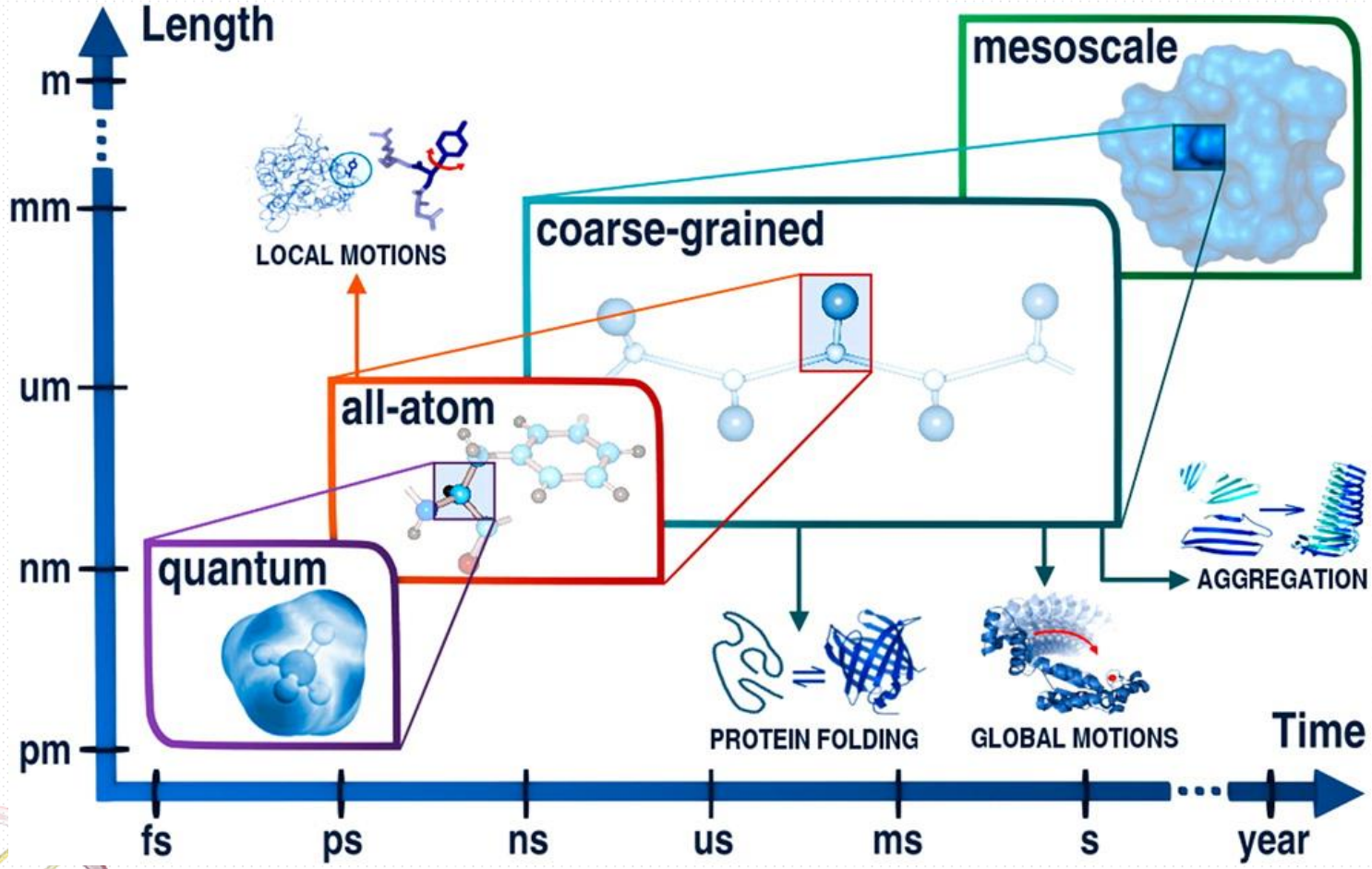
Dynamika molekularna - środowisko

- najlepsze przybliżenie warunków fizjologicznych
- w roztworze wodnym (z udziałem lub bez udziału jonów)



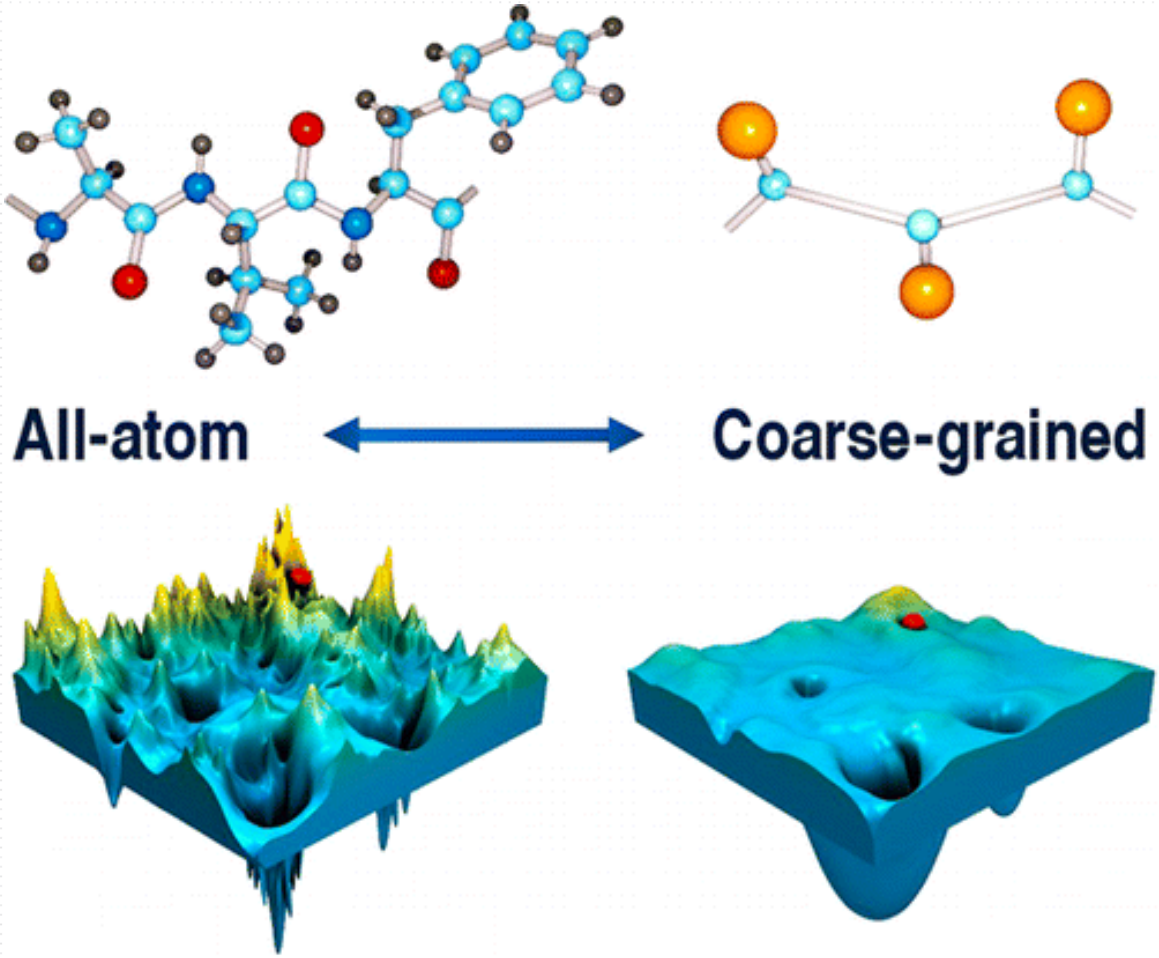


Skale czasowe symulacji

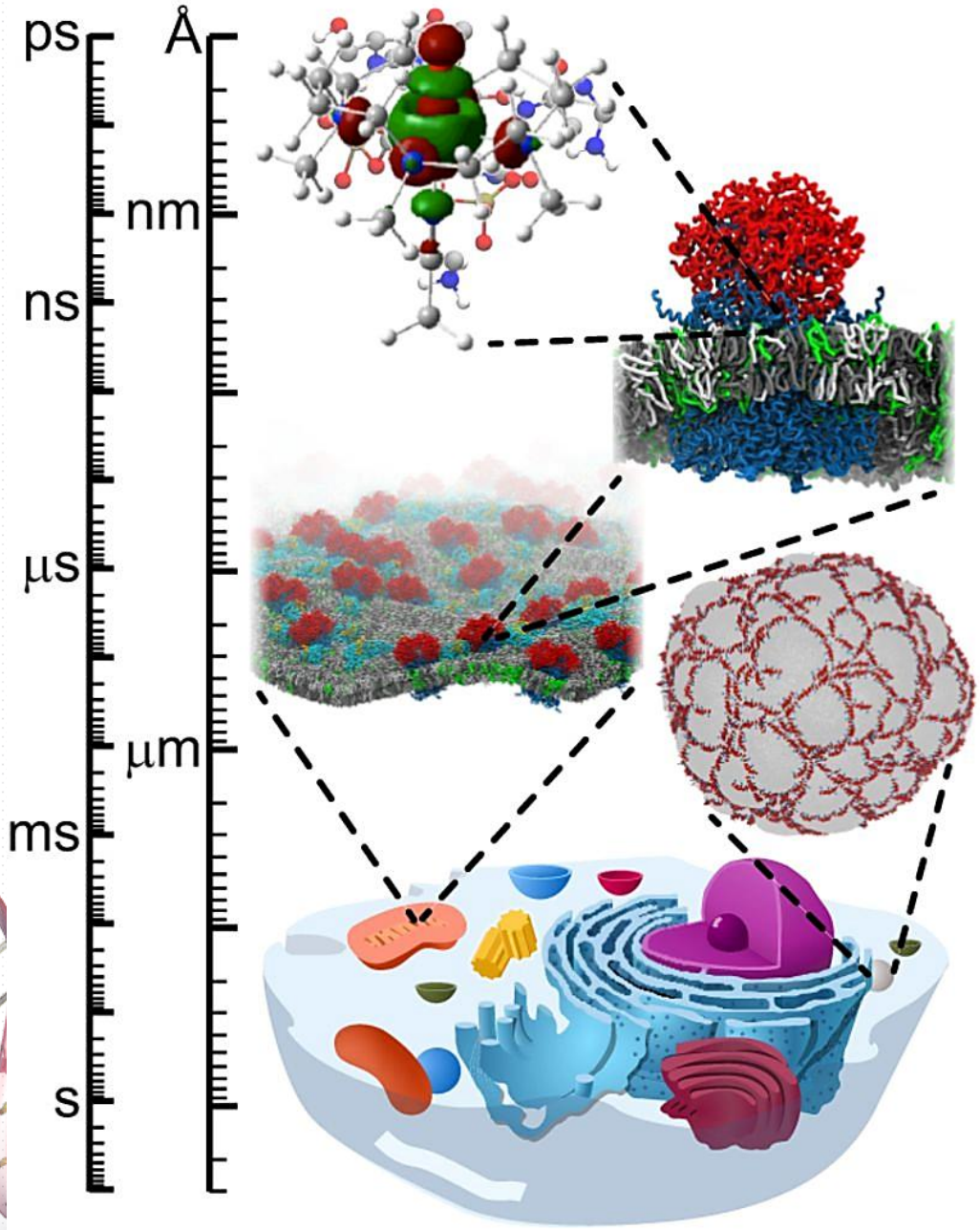


Kmiecik S. et al., *Chem. Rev.*, 2016, 116 (14), pp 7898–7936

II. Coarse-grained force fields



Kmiecik S. et al., *Chem. Rev.*, 2016, 116 (14), pp 7898–7936



Quantum
- atoms, electrons and electron clouds included
- explicit solvent
- quantum mechanics

All-atom
- all or most atoms present
- explicit solvent
- molecular dynamics

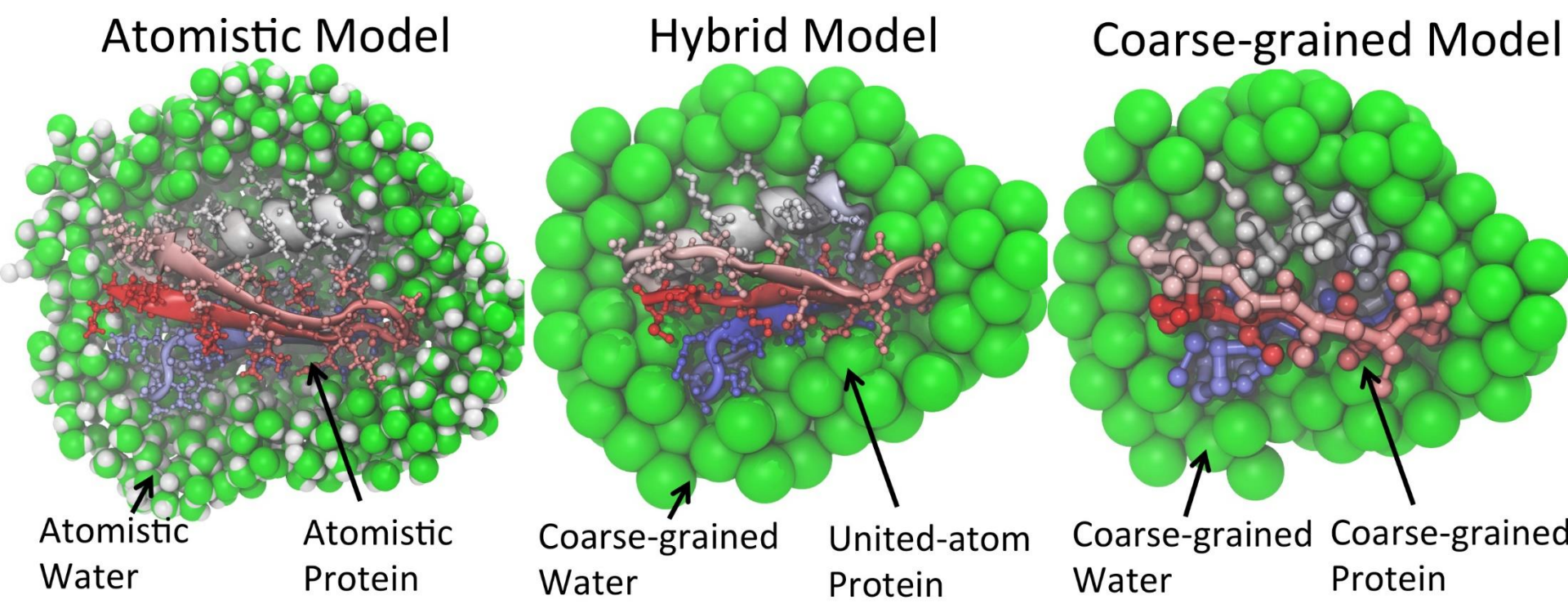
Coarse-grained
- beads comprising a few atoms
- explicit or implicit solvent
- molecular dynamics

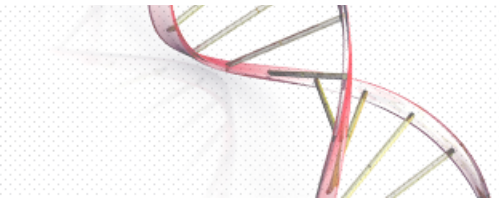
Supra-coarse-grained
- interaction sites comprising many atoms, protein parts or proteins
- implicit solvent
- stochastic dynamics

Continuum
- materials as a continuous mass
- implicit solvent
- continuum mechanics

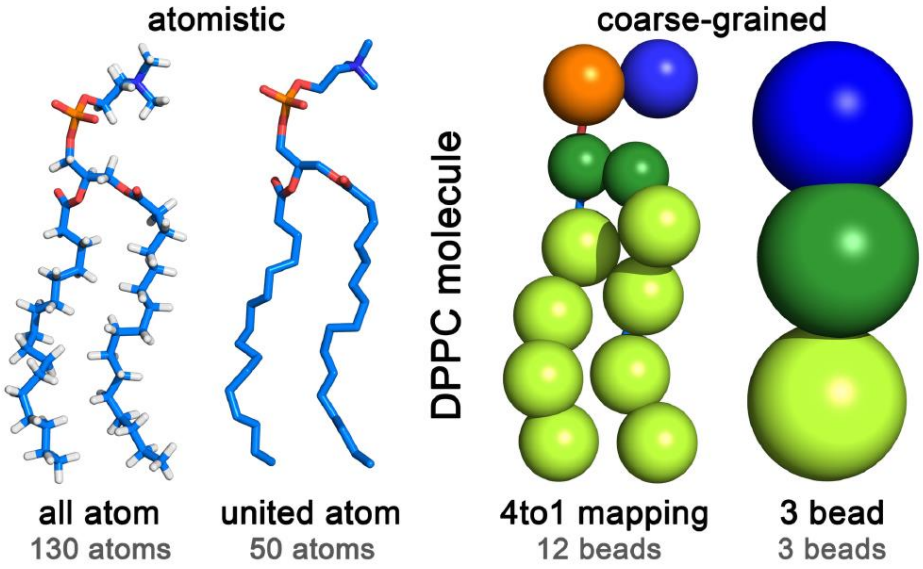
Ingólfsson et al. *J Cell Sci* 2016 129: 257-268

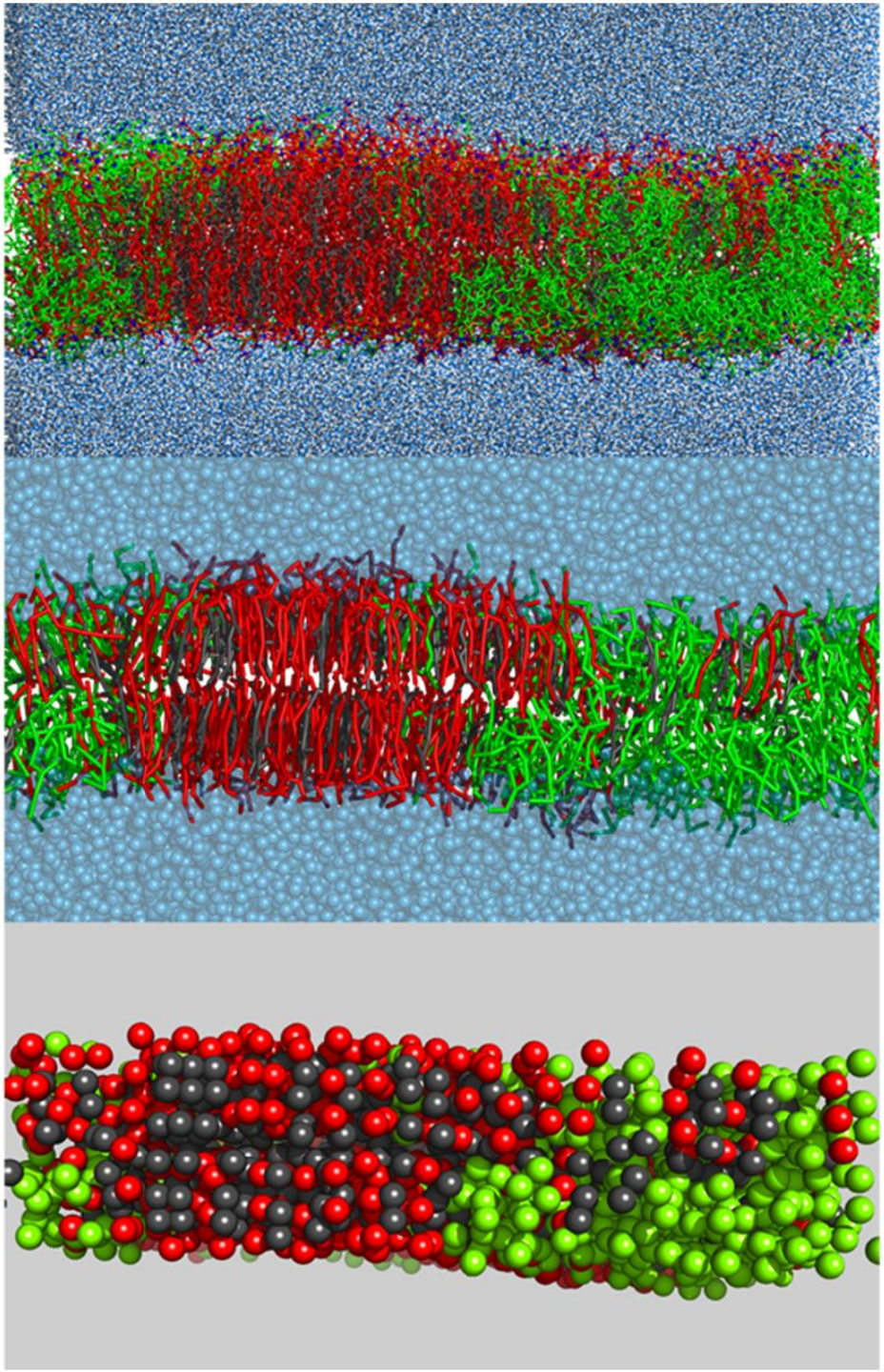
All-atom and/or coarse-grained





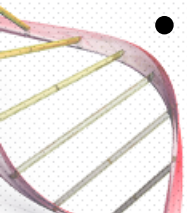
Levels of coarse-graining

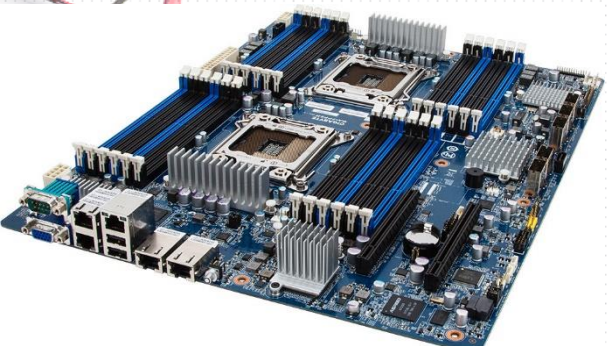


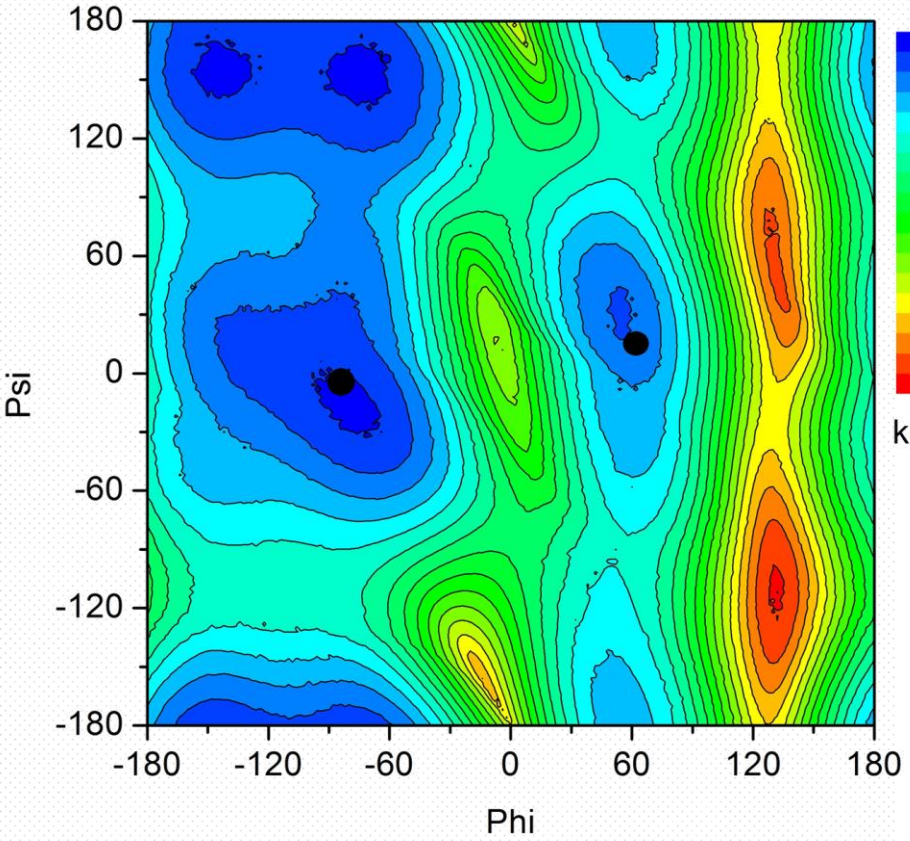


Anton supercomputer

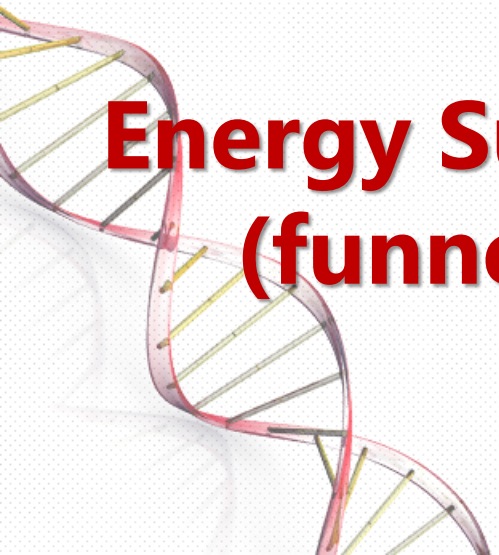
- Specialized computer to do MD simulations
- Use special ASICs (application-specific integrated circuits)
- 512-nodes of Anton can achieve 17 000 ns per day for a 23 558 atoms.
- 2*20 core CPUs can achieve 41 ns per day.
- 2*36 core CPUs can achieve 80 ns per day.
- 1 GPU (GTX 1080Ti) can achieve 625 ns per day.

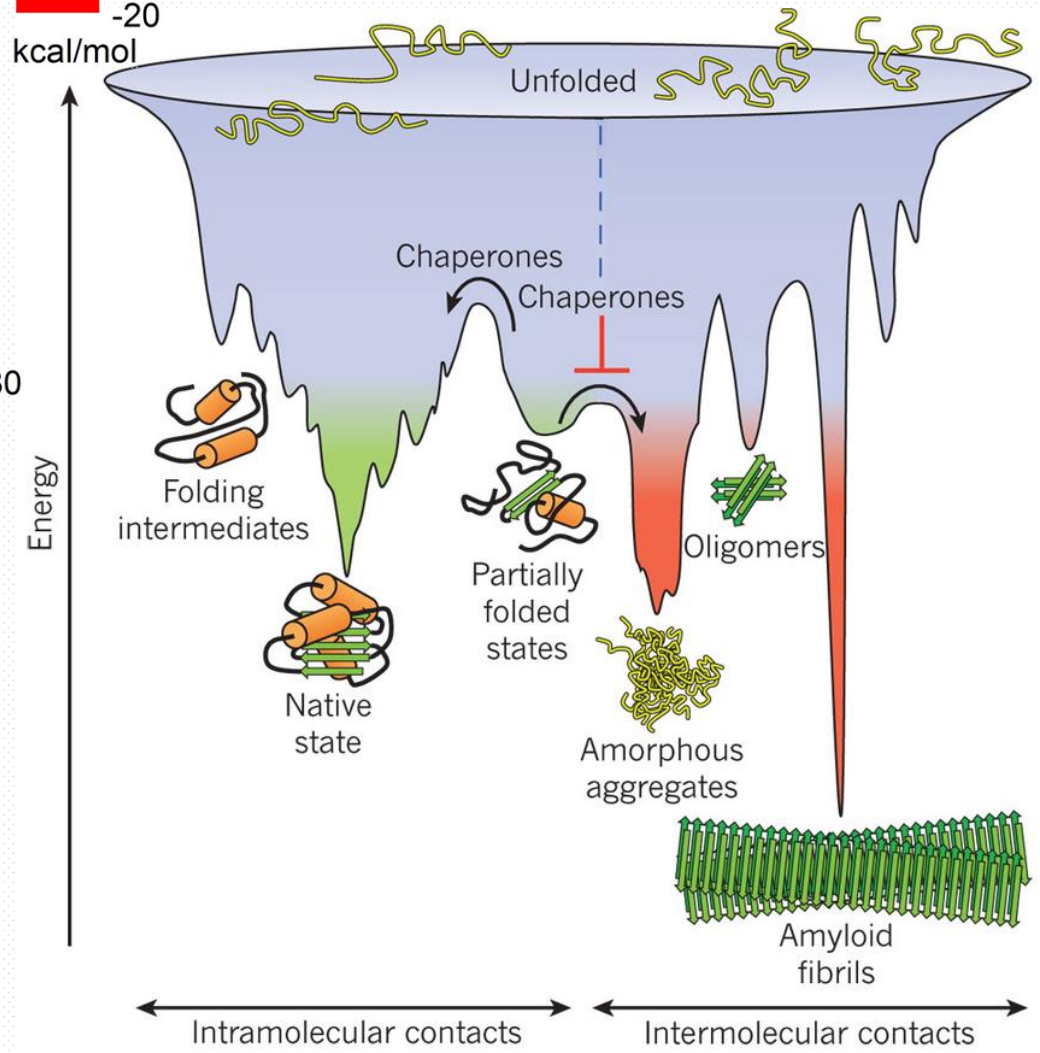






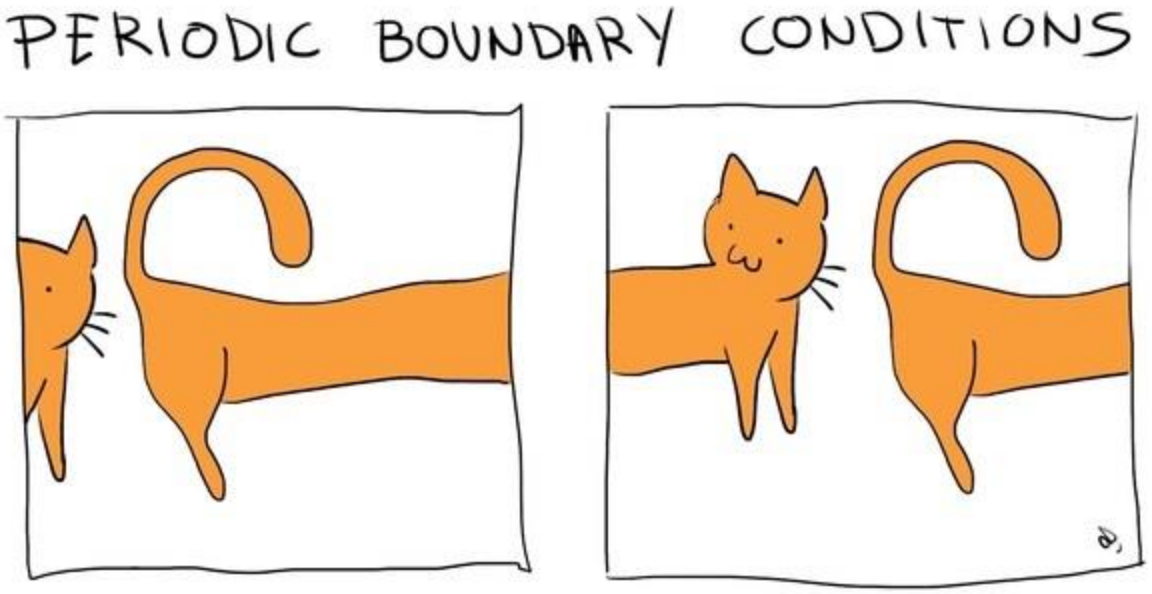
Energy Surface (funnel?)



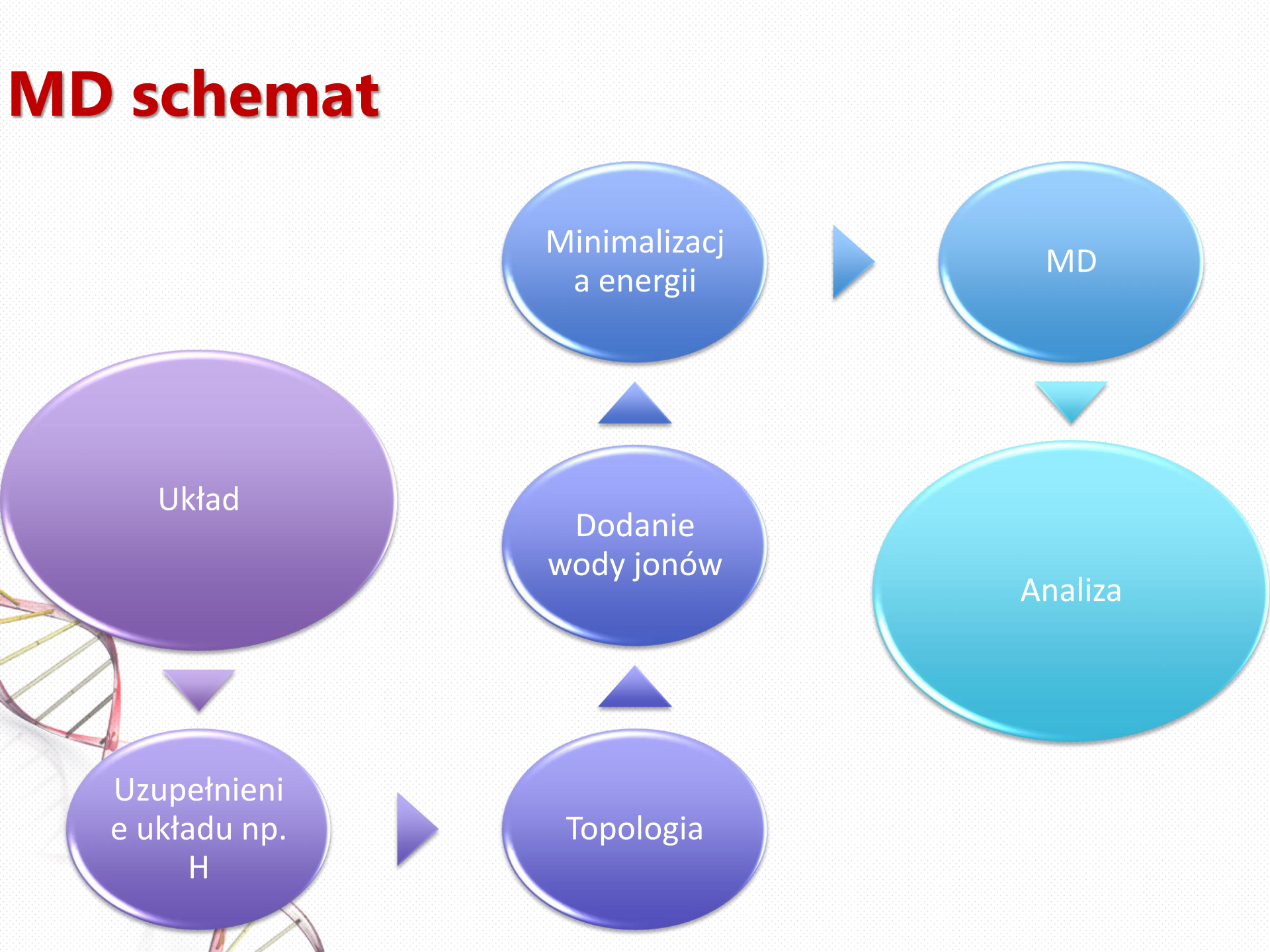


Types of MD simulations

- In vacuum
- In implicit solvent
- In explicit solvent with box
- In periodic boundary conditions



MD schemat

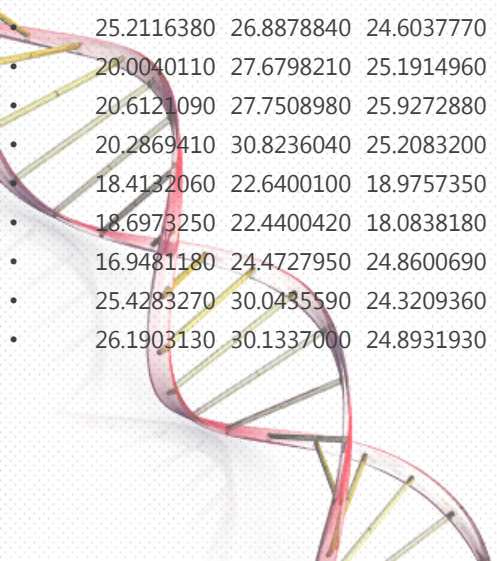

Układ
Uzupełnienie układu np. H
Topologia
Dodanie wody jonów
Minimalizacja energii
MD
Analiza

Topologia

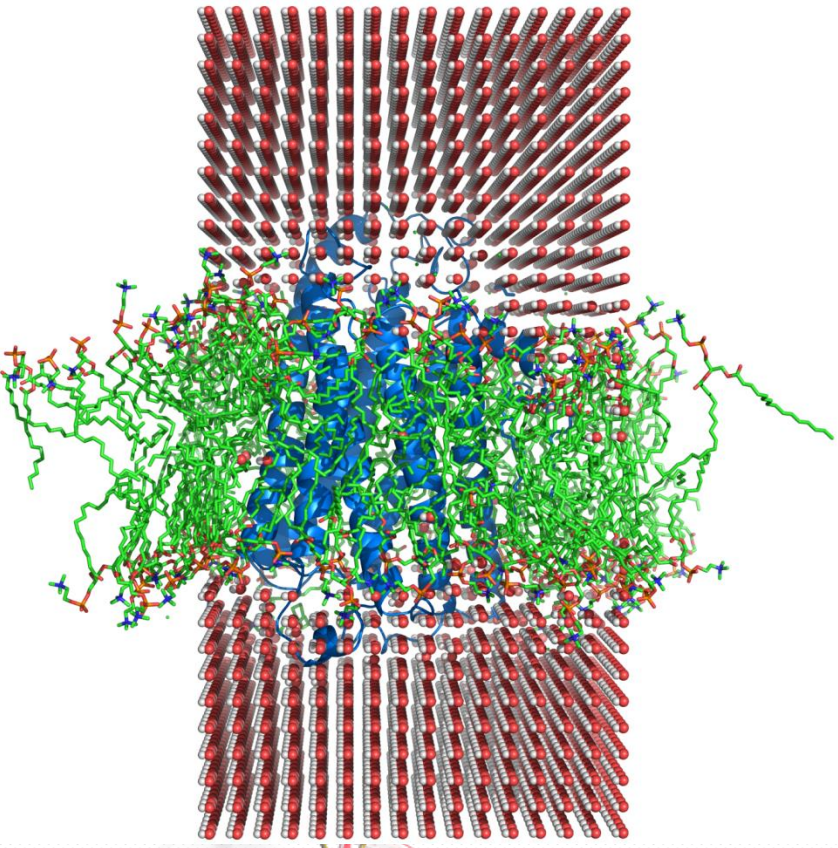
- %VERSION VERSION_STAMP = V0001.000 DATE = 06/30/15 11:44:23
- %FLAG TITLE
- %FORMAT(20a4)
- ACE
- %FLAG POINTERS
- %FORMAT(10I8)
- 1912 9 1902 9 25 11 43 24 0 0
- 2619 633 9 11 24 13 21 20 10 1
- 0 0 0 0 0 0 0 1 10 0
- 0
- %FLAG ATOM_NAME
- %FORMAT(20a4)
- HH31CH3 HH32HH33C O N H CA HA CB HB1 HB2 HB3 C O N H CH3 HH31
- HH32HH33O H1 H2 O H1 H2 O H1 H2 O H1 H2 O H1 H2 O H1 H2
- O H1 H2 O H1 H2 O H1 H2 O H1 H2 O H1 H2 O H1 H2 O H1
- H2 O H1 H2 O H1 H2 O H1 H2 O H1 H2 O H1 H2 O H1 H2 O
- H1 H2 O H1 H2 O H1 H2 O H1 H2 O H1 H2 O H1 H2 O H1 H2
- O H1 H2 O H1 H2 O H1 H2 O H1 H2 O H1 H2 O H1 H2 O H1
- H2 O H1 H2 O H1 H2 O H1 H2 O H1 H2 O H1 H2 O H1 H2 O
- H1 H2 O H1 H2 O H1 H2 O H1 H2 O H1 H2 O H1 H2 O H1 H2

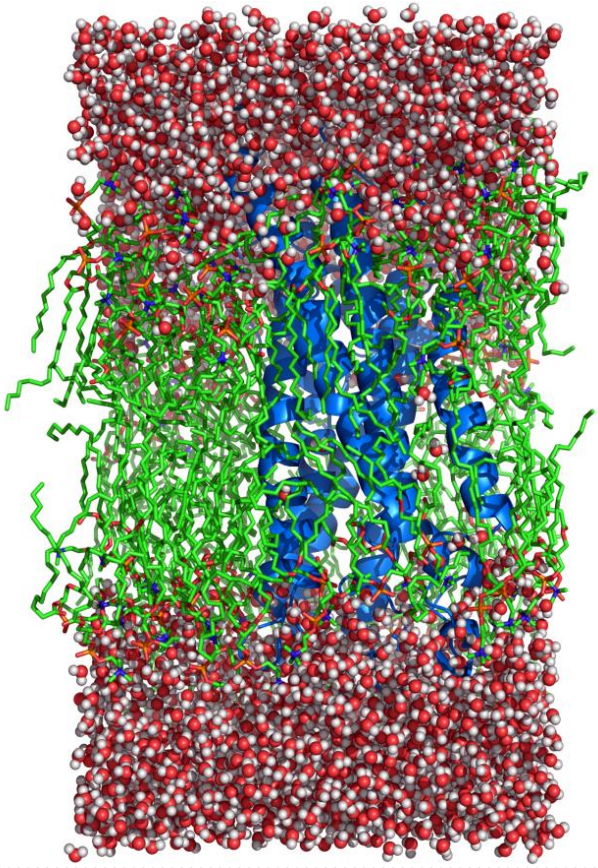
Koordynaty

- ACE
- 1912
- 13.6813322 13.1481714 15.2733473 13.6813322 14.2381714 15.2733488
- 13.1675952 14.6020204 16.1631727 13.1675902 14.6020234 14.3835287
- 15.1087512 14.7889664 15.2733457 16.0719112 14.0255774 15.2733421
- 15.2367066 16.1178202 15.2733455 14.4144513 16.7043316 15.2733474
- 16.5345934 16.7620967 15.2733444 17.0889272 16.4637102 16.1631639
- 17.3426356 16.3690139 14.0412007 16.8045928 16.6695344 13.1421471
- 18.3118152 16.8671068 14.0675580 17.4898714 15.2890438 14.0319637
- 16.3940071 18.2775899 15.2733501 15.2819757 18.8008742 15.2733549
- 17.5273845 18.9830547 15.2733512 18.4183327 18.5073334 15.2733483
- 17.5273863 20.4320551 15.2733549 16.4999074 20.7959063 15.2733591
- 18.0411296 20.7959028 16.1631770 18.0411212 20.7959068 14.3835300
- 27.8556240 25.5899380 22.4329930 27.7750290 24.6753990 22.1621640
- 27.4420660 26.0866450 21.7269640 25.9874280 27.3678400 24.3139200
- 25.2116380 26.8878840 24.6037770 26.6387210 26.6884050 24.1395360
- 20.0040110 27.6798210 25.1914960 19.1724990 27.4246920 25.5911470
- 20.6121090 27.7508980 25.9272880 19.9841370 31.7138300 25.3872980
- 20.2869410 30.8236040 25.2083200 20.7284960 32.2708520 25.1595230
- 18.4132060 22.6400100 18.9757350 17.5248760 22.9816070 18.8737230
- 18.6973250 22.4400420 18.0838180 16.3765070 25.2405750 24.8597050
- 16.9481180 24.4727950 24.8600690 16.3766610 25.5394510 23.9503660
- 25.4283270 30.0435590 24.3209360 25.3591550 29.1025530 24.1598490
- 26.1903130 30.1337000 24.8931930 17.6189940 20.6490110 26.6465980

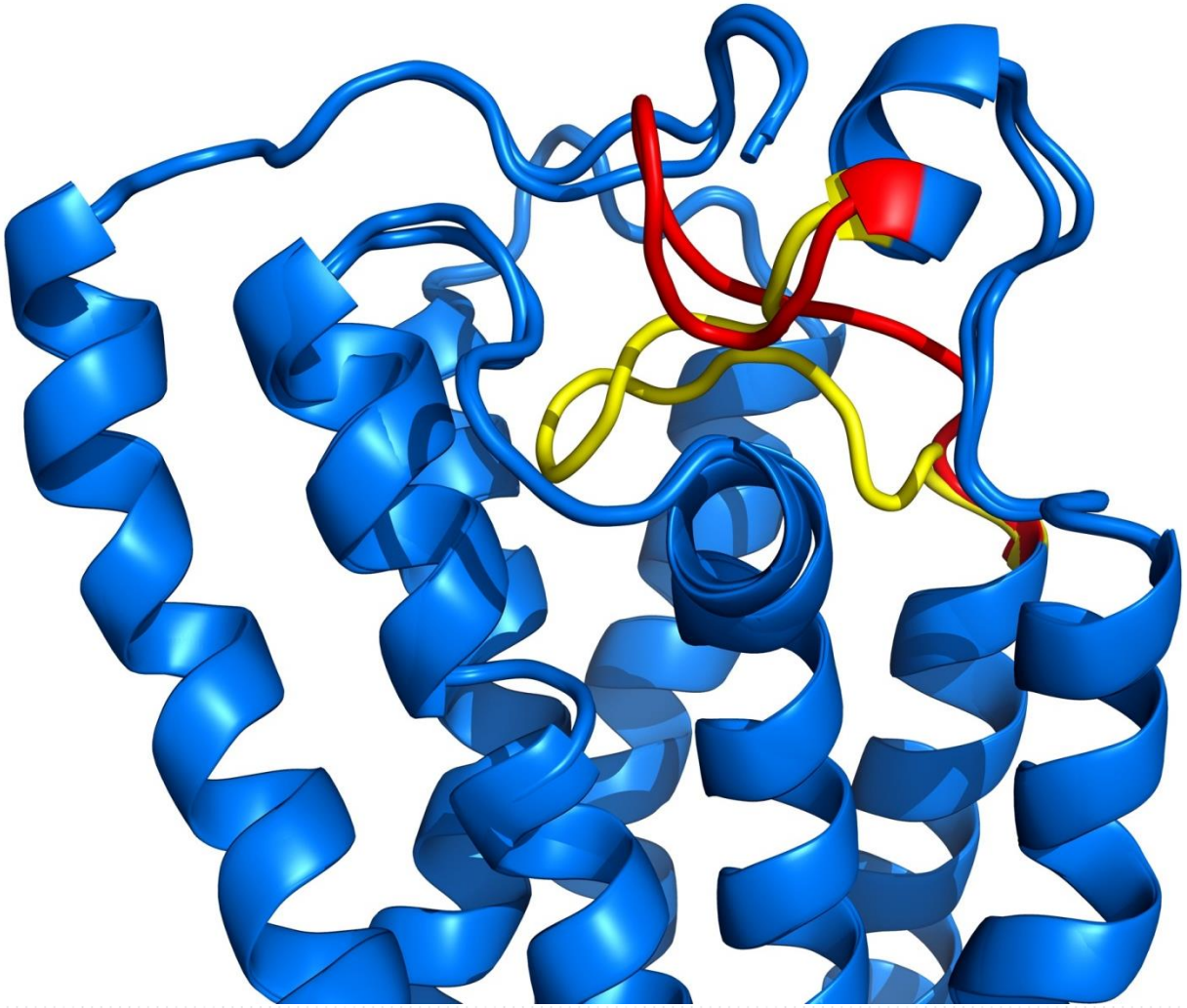


Ekwilibracja wody

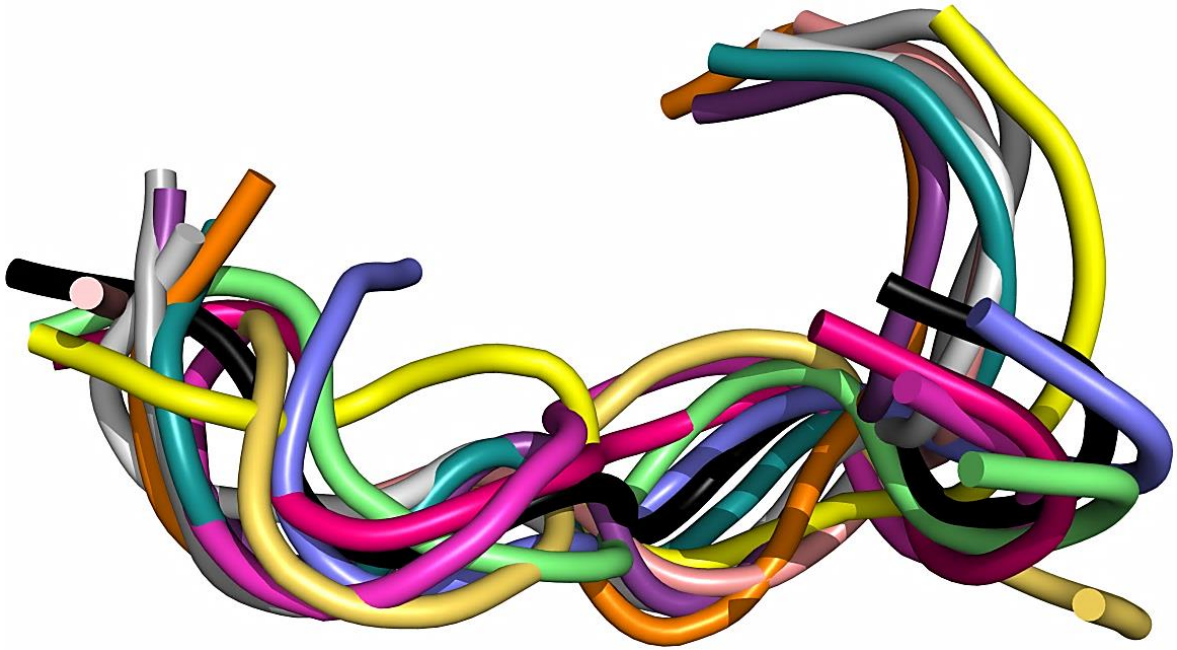




Labilne pętle

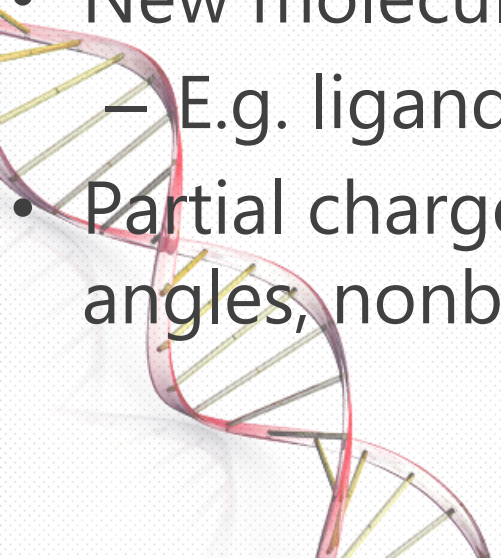


Labilny ligand



Parametrization of the nonstandard molecules and residues

- Antechamber
- GAFF
- R.E.D. server
- Nonstandard amino-acid residues
  - Based on existing residues
- New molecules
  - E.g. ligands
- Partial charges, bonds, valence and torsional angles, nonbonding interactions



R.E.D. server

The latest stable interface of PyRED

RED Development

RESP ESP charge Derive Server

Home | Register | Mini How-To | Documentation | Demo | Assistance | Submit | Download

- Home
- Overview
- License
- Gallery
- Golden book
- Citation NEW
- Help
- FAQ
- Bibliography
- Tutorials
- Bugs & new features
- PyRED news
- Qstat

q4md links

- REDDB
- R.E.D.
- AmberFFC
- FFParmDev
- People

Useful links

PyRED RED DB

RESP ESP charge Derive Server
Development
Home Page

R.E.D. Server *Development* is a new service designed to generate force field parameters and build force field libraries for new molecules/molecular fragments. This Internet site provides the software and hardware resources required for the generation of AMBER and GLYCAM type empirical force fields, that can be used by computational biologists involved in force field based structural studies. CHARMM and OPLS force fields can also be targeted by applying user-defined options.

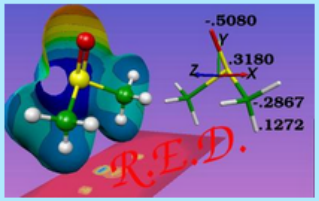
R.E.D. Server *Development*

Interface to PyRED

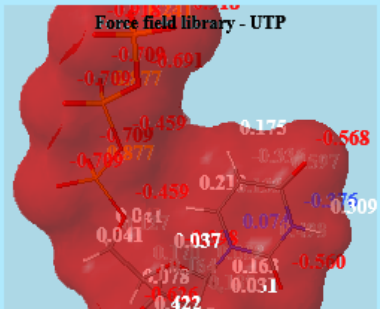
Project initiated: November 2011

Release: May 2014

Web server open to all. NEW



Done using Molekel...



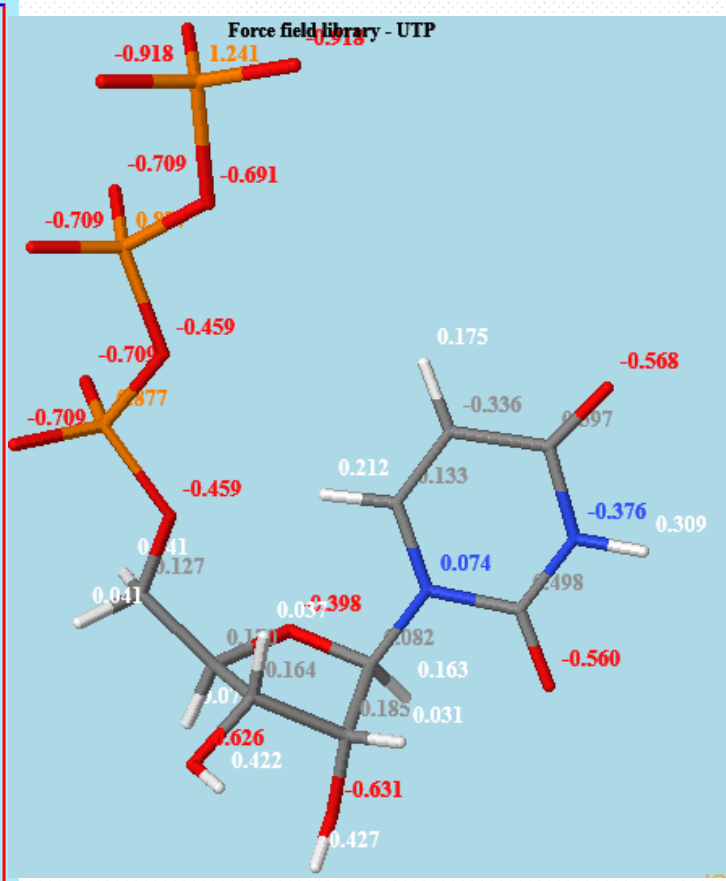
Wireframe | Stick | Ball & Stick | Space-filling

Reset to original orientation | Zoom

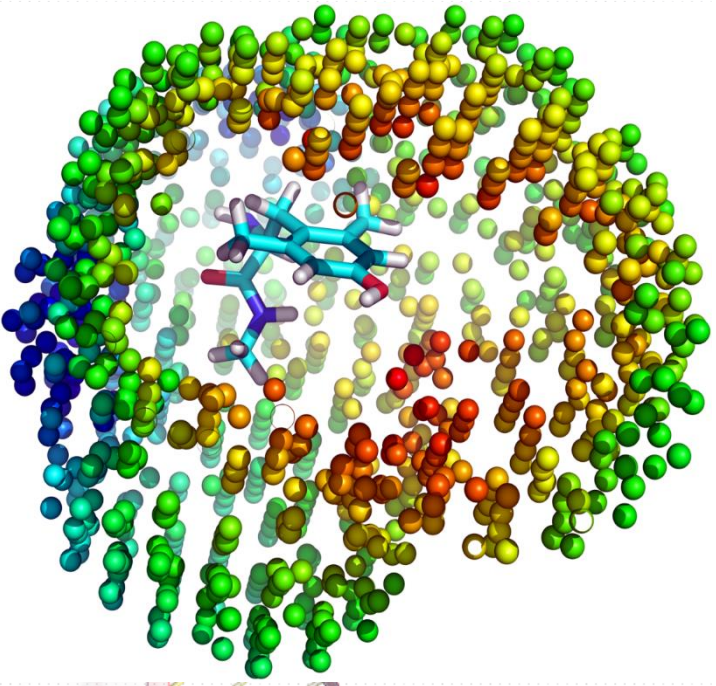
Spin | Axes | Boundbox

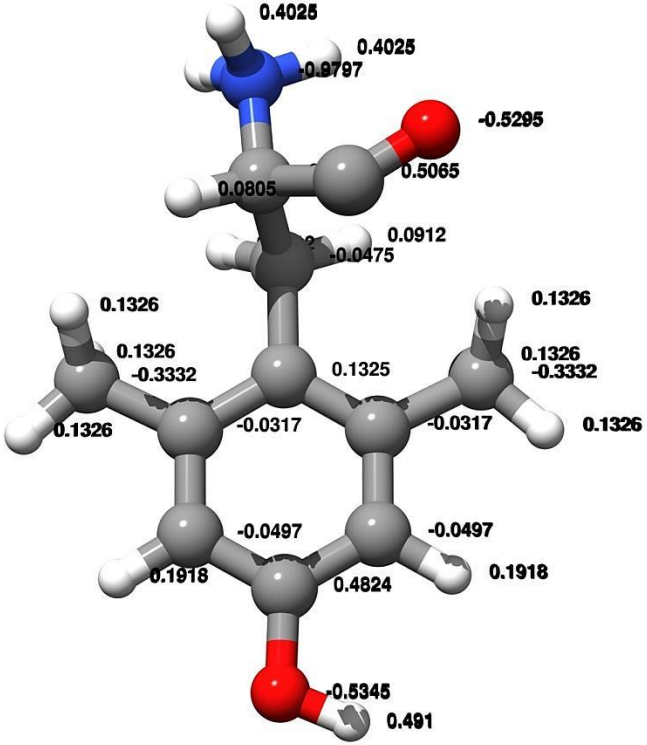
background lightblue

Atomic charges | Atom types | Atom names | Atom indexes | All off

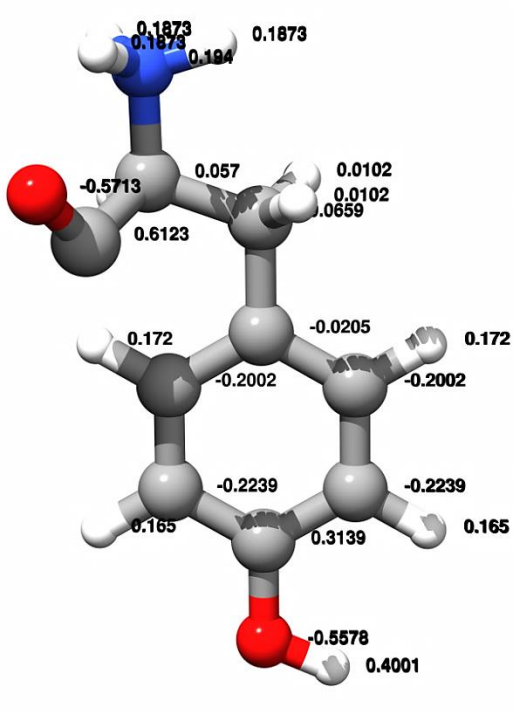


RESP method





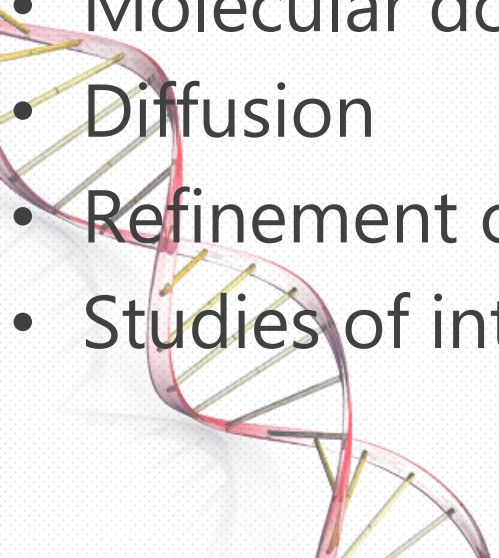
Dmt



N-Tyr

Applications of MD

- Protein folding
- Stability checks
- Studies of the dynamics
- Calculation of the binding free energies
- Conformational changes
- Molecular docking
- Diffusion
- Refinement of structures
- Studies of interactions and affinity

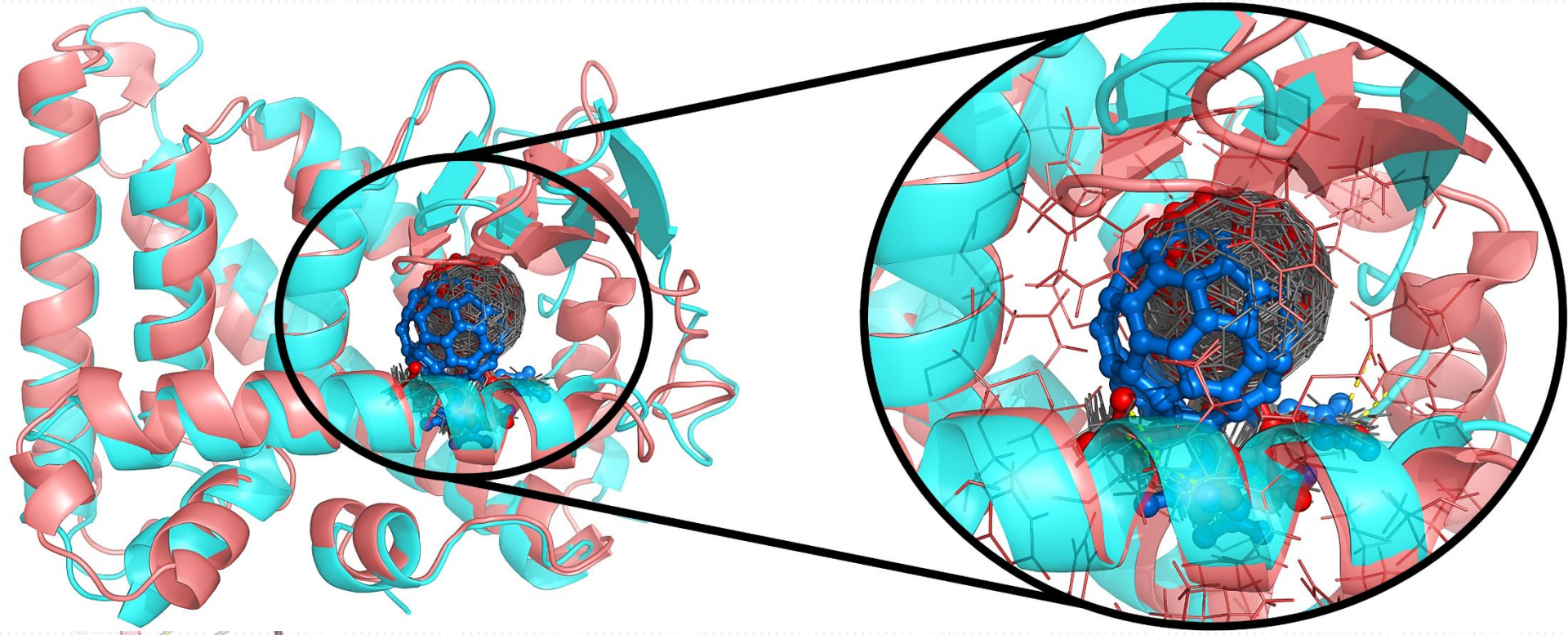


Applications of MD

- Equilibrium properties
  - Thermodynamic properties (average energy, heat capacity),
  - Mechanical properties (pressure, density, volume, radius of gyration),
  - Structural properties (correlation function),
- Non-equilibrium properties (transport)
  - Diffusion coefficients,
  - Shear viscosity,
  - Heat conductivity.

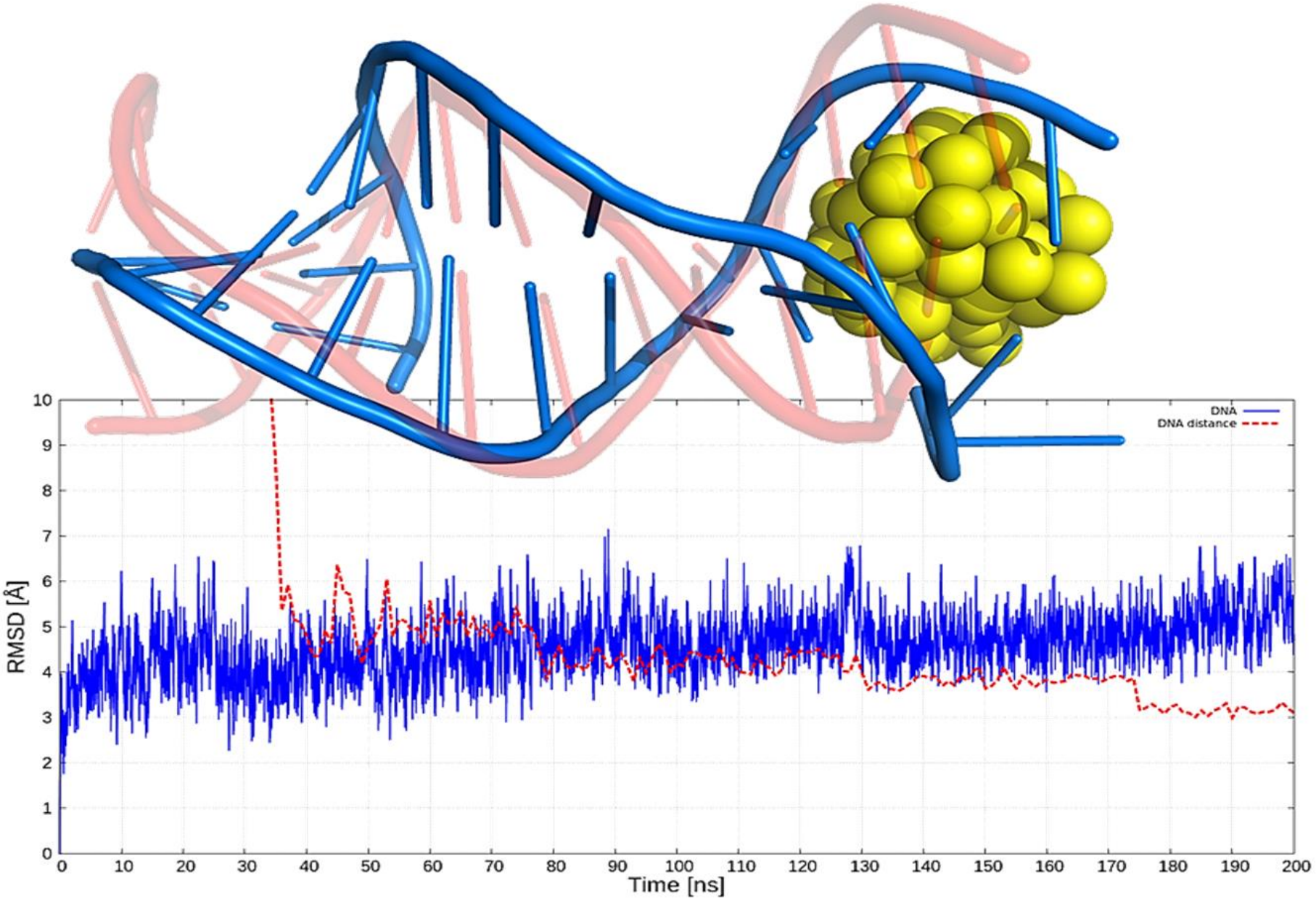
Practical applications

- Studies of the stability of complexes

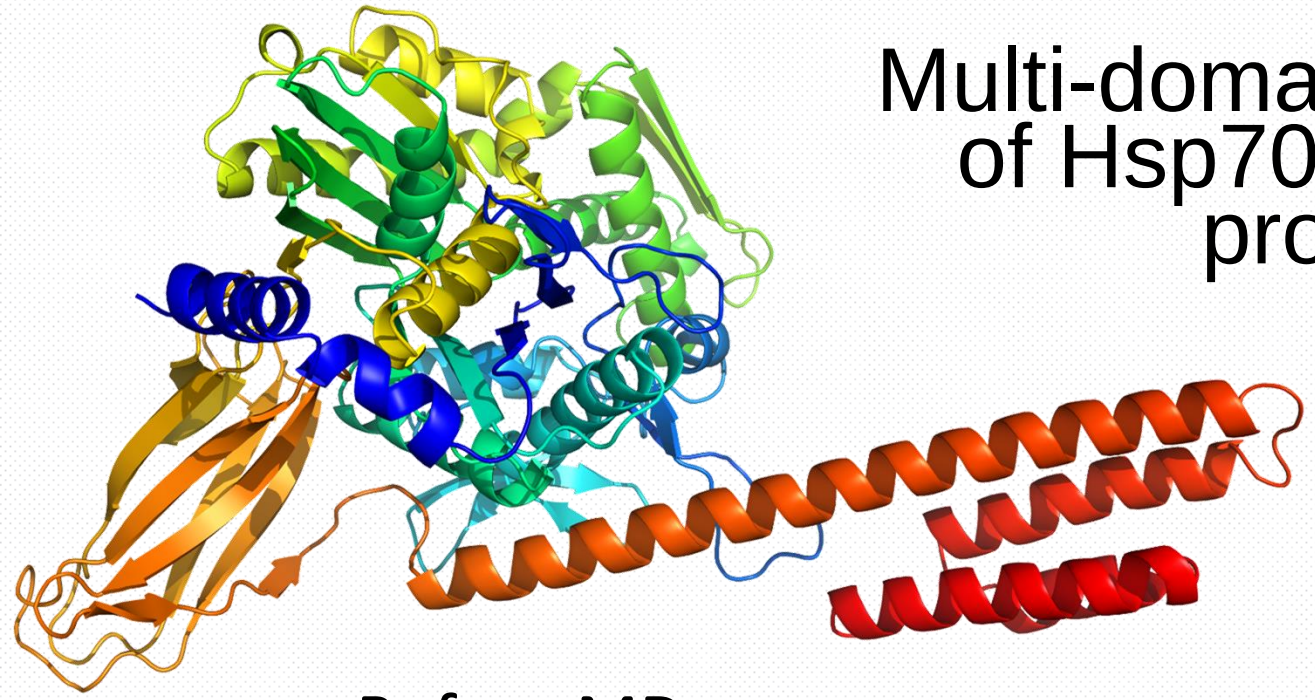


Inhibitors or toxins? Large library target-specific screening of fullerene-based nanoparticles for drug design purpose, L Ahmed, B Rasulev, S Kar, P Krupa, MA Mozolewska, J Leszczynski, Nanoscale 9 (29), 10263-10276, 2017.

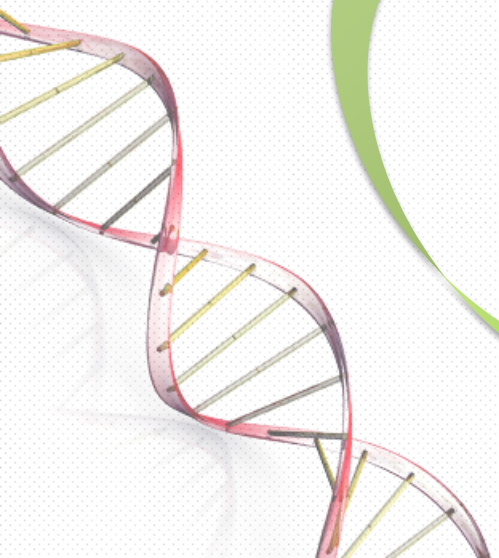
Practical applications

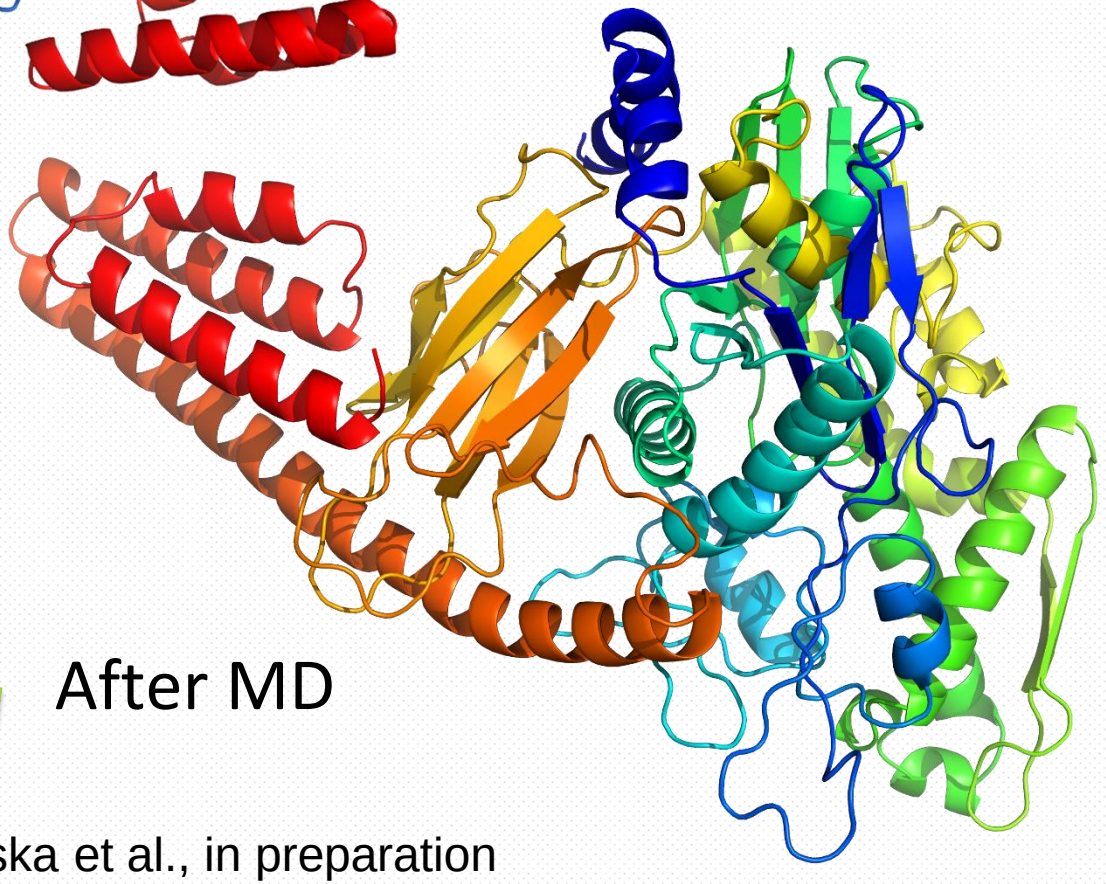


Multi-domain dynamics of Hsp70 chaperon protein



Before MD

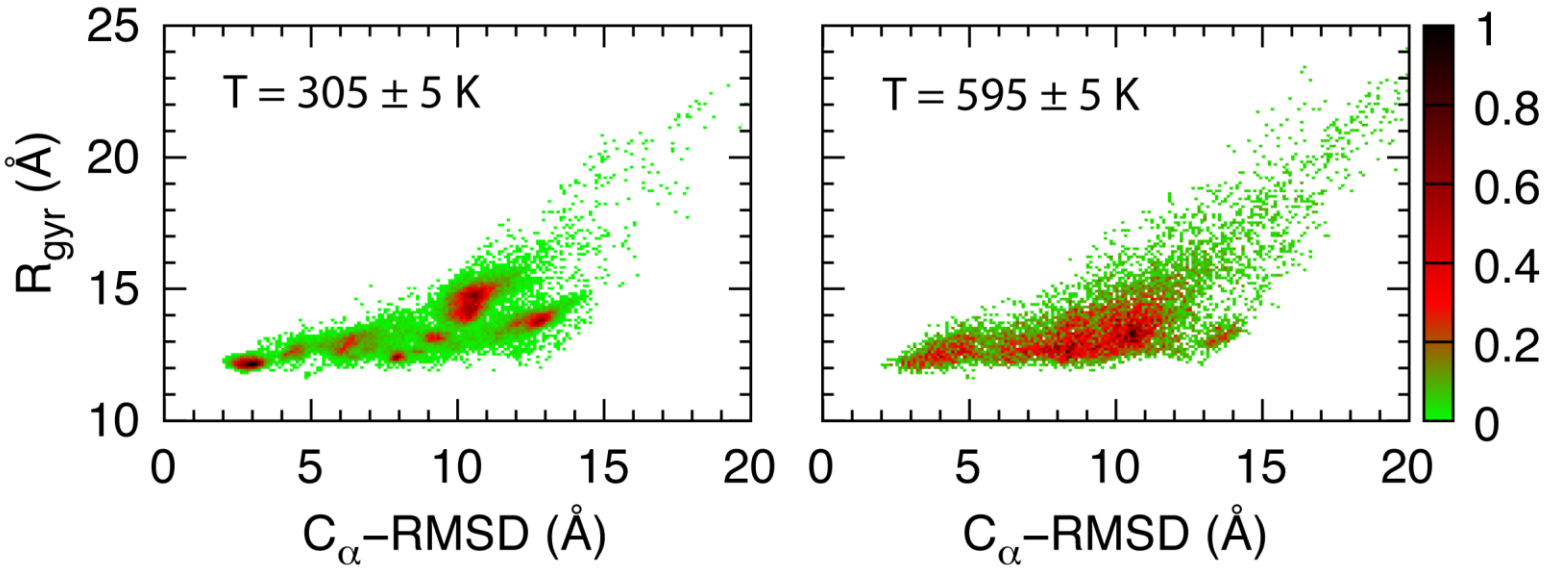


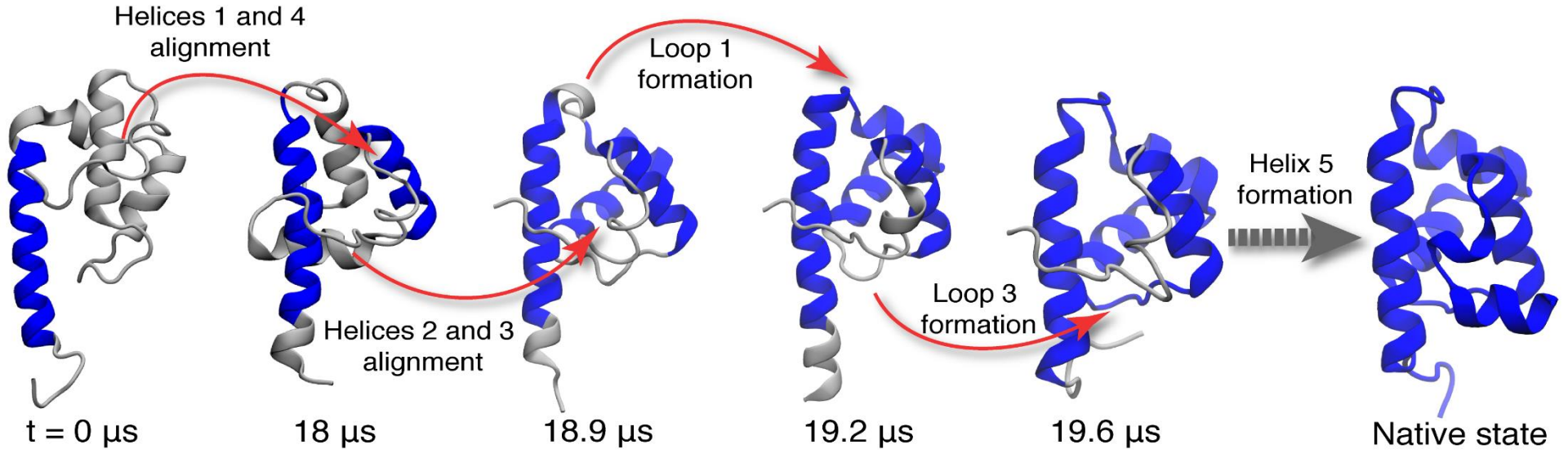


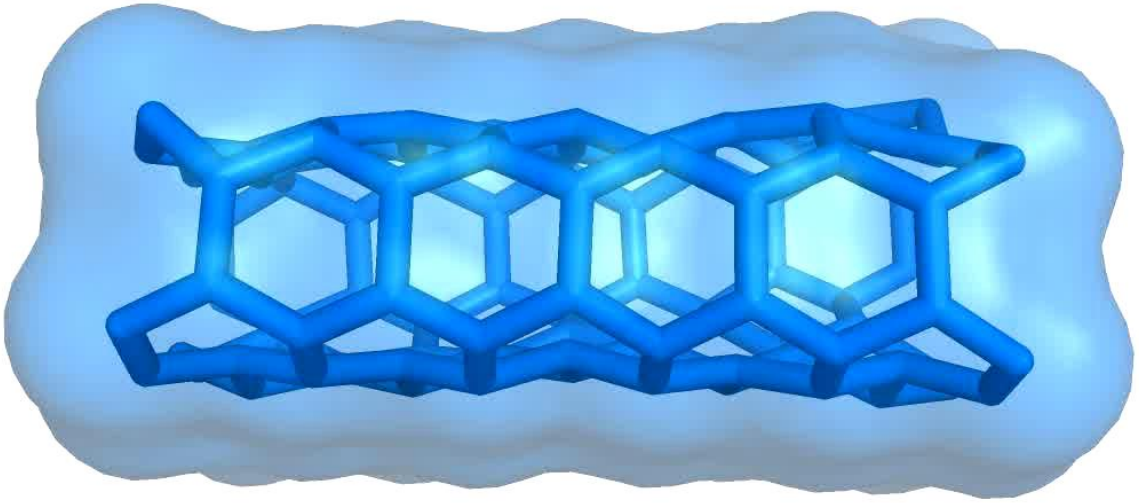
After MD

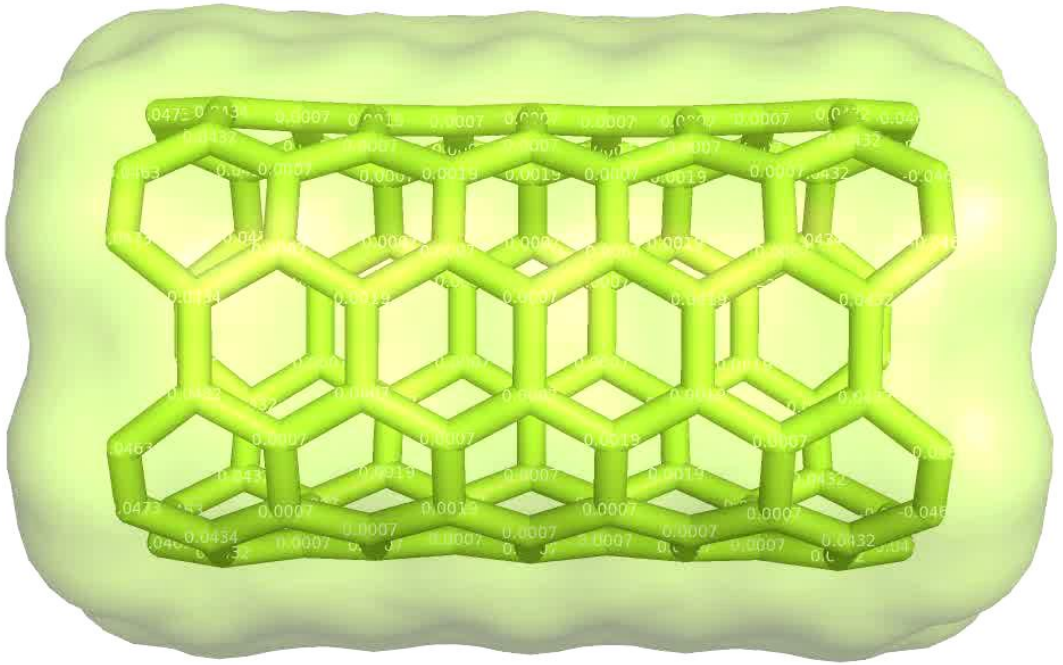
Mozolewska et al., in preparation

Protein folding

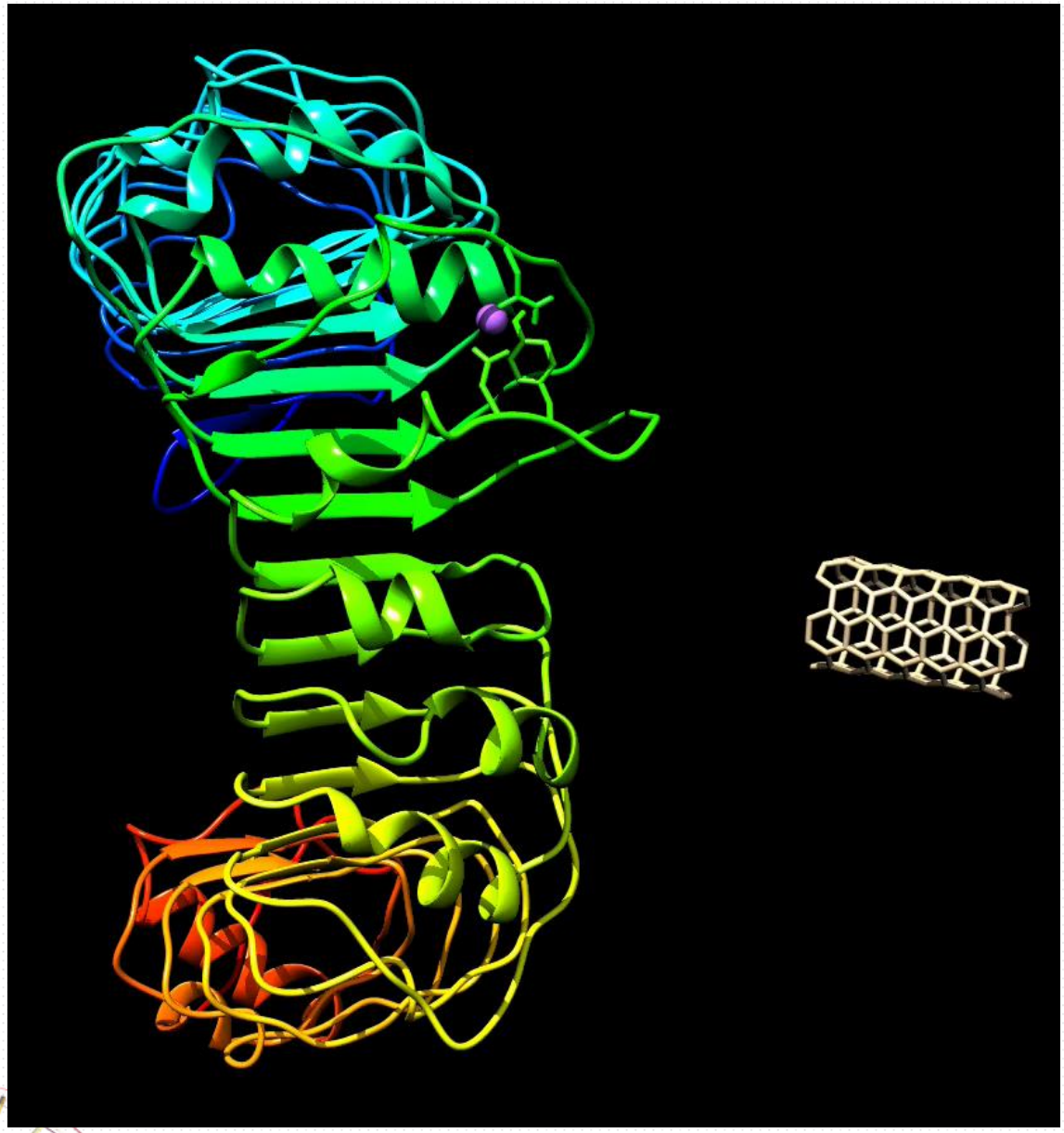


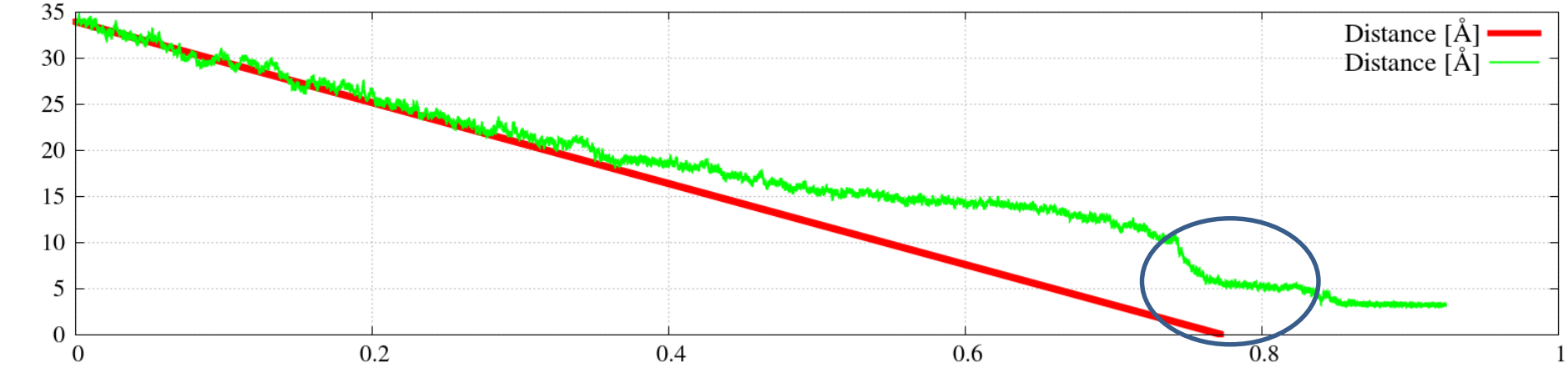


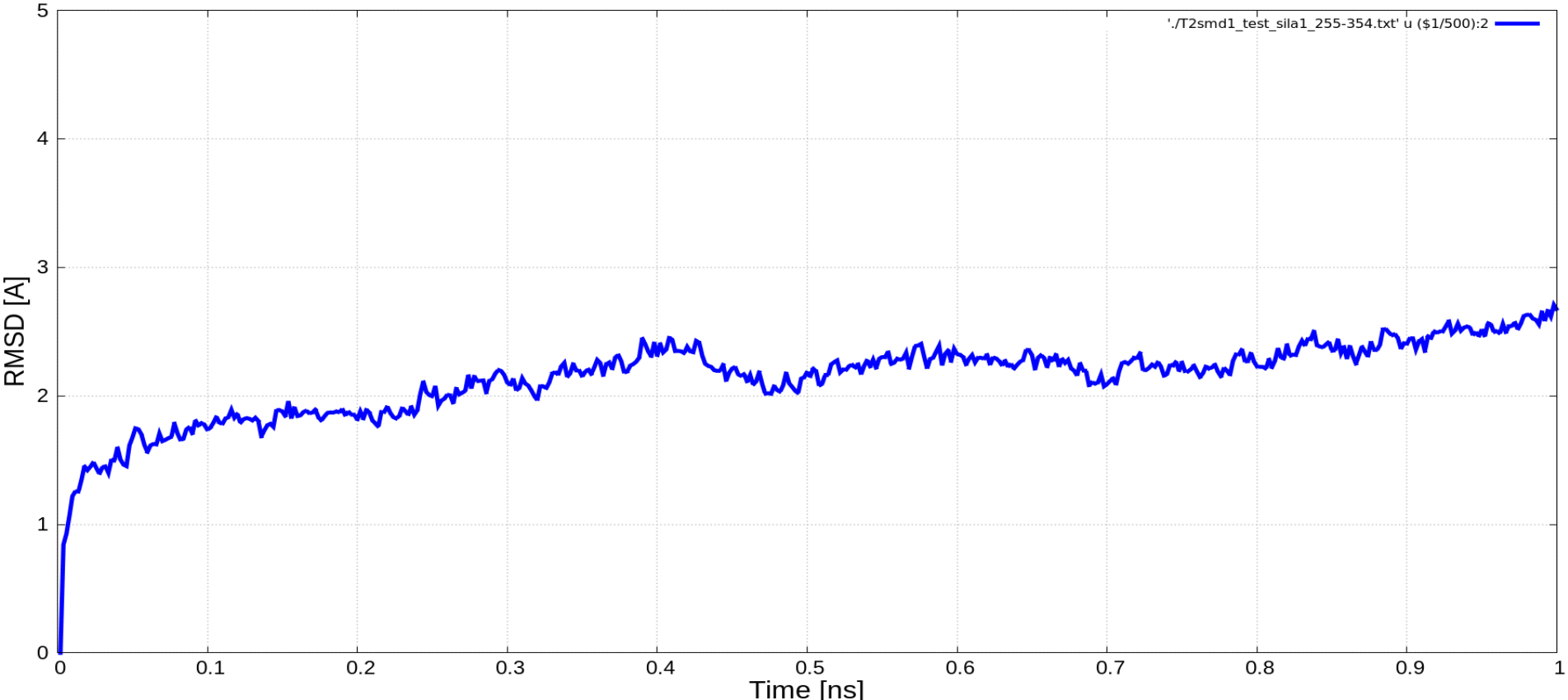




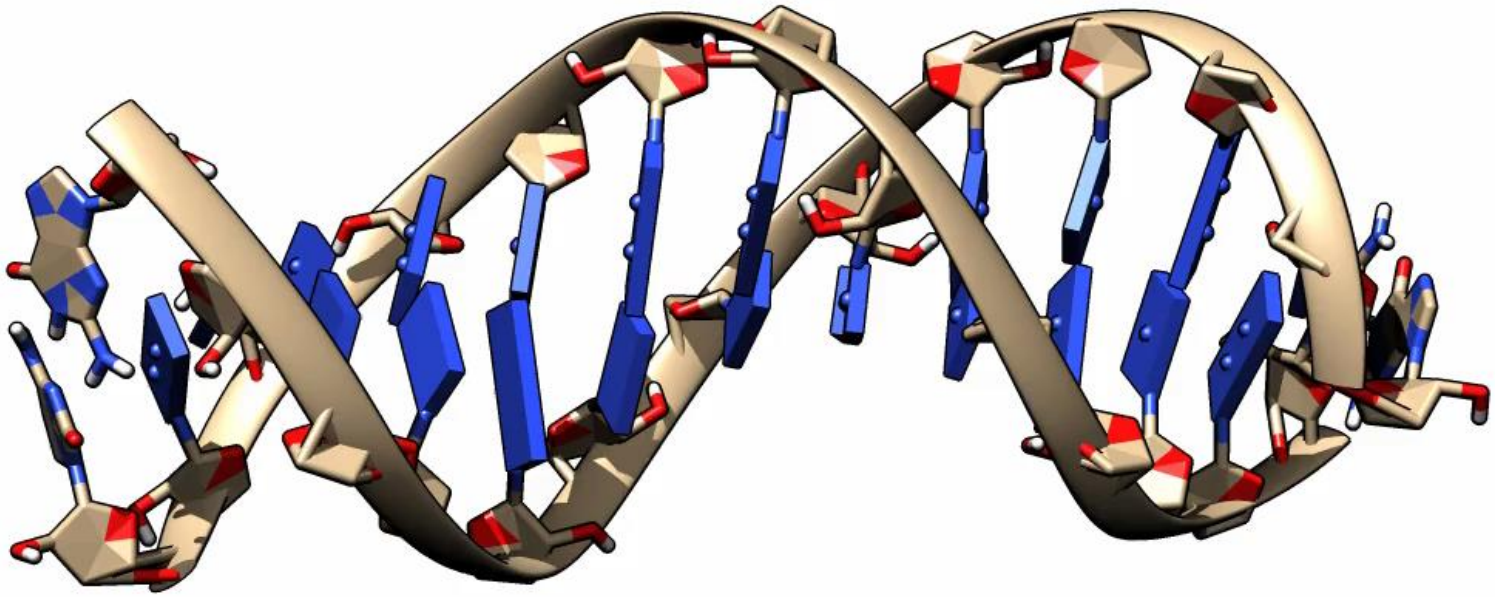
SMD




Distance [Å]
Distance [Å]
35
30
25
20
15
10
5
0
0
0.2
0.4
0.6
0.8
1

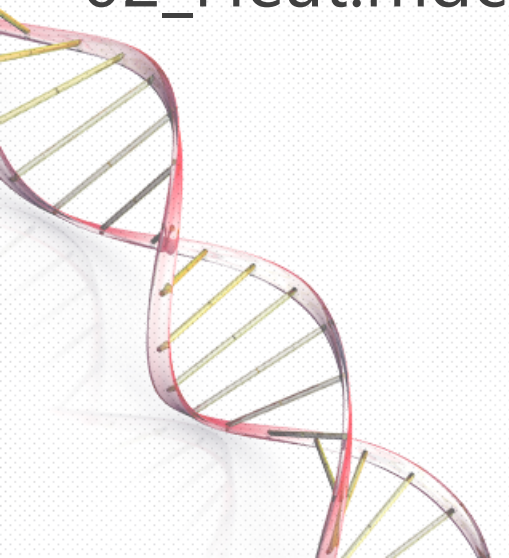

'./T2smd1_test_sila1_255-354.txt' u ($1/500):2
RMSD [A]
Time [ns]
5
4
3
2
1
0
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1

DNA



Pliki wynikowe

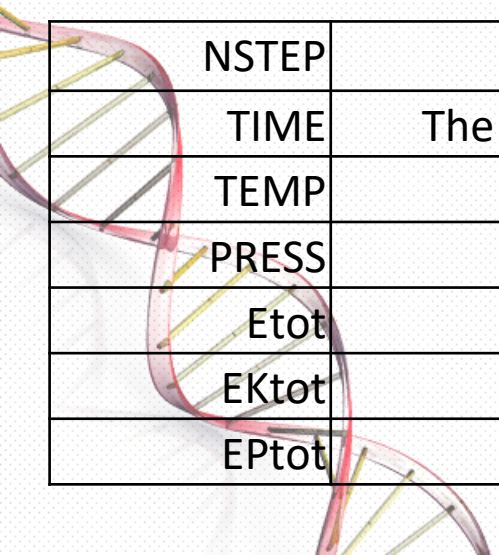
- 02_Heat.out
- 02_Heat.rst- output restart file with coordinates and velocities
- 02_Heat.mdcrd - trajektoria



02_Heat.out

In the **02_Heat.out** file you will find the output from the heating MD. You should be able to see system information including timestep energies, and temperature. For example on the 1000 time step:

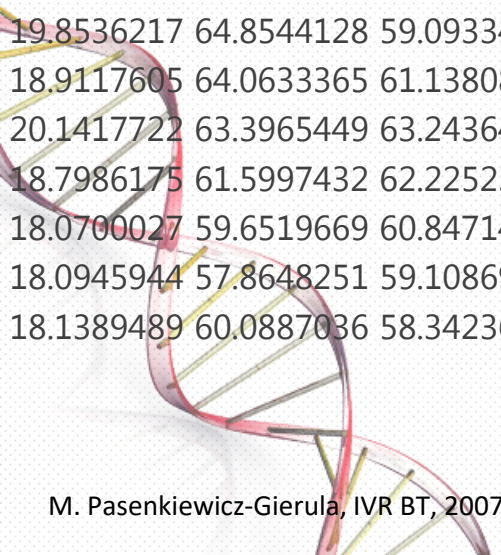
```
 NSTEP =     1000   TIME(PS) =       2.000  TEMP(K) =    29.48  PRESS =     0.0
 Etot   =     -6944.9552  EKtot   =       112.3015  EPtot      =     -7057.2567
 BOND   =         1.0442  ANGLE   =         1.7653  DIHED      =         9.4906
 1-4 NB =         2.6284  1-4 EEL =        46.3073  VDWAALS    =      1448.7074
 EELEC  =     -8567.1999  EHBOND  =         0.0000  RESTRAINT  =         0.0000
 Ewald error estimate:   0.4641E-03
 ------------------------------------------------------------------------------

 NMR restraints: Bond =    0.000   Angle =     0.000   Torsion =     0.000
================================================================================
```

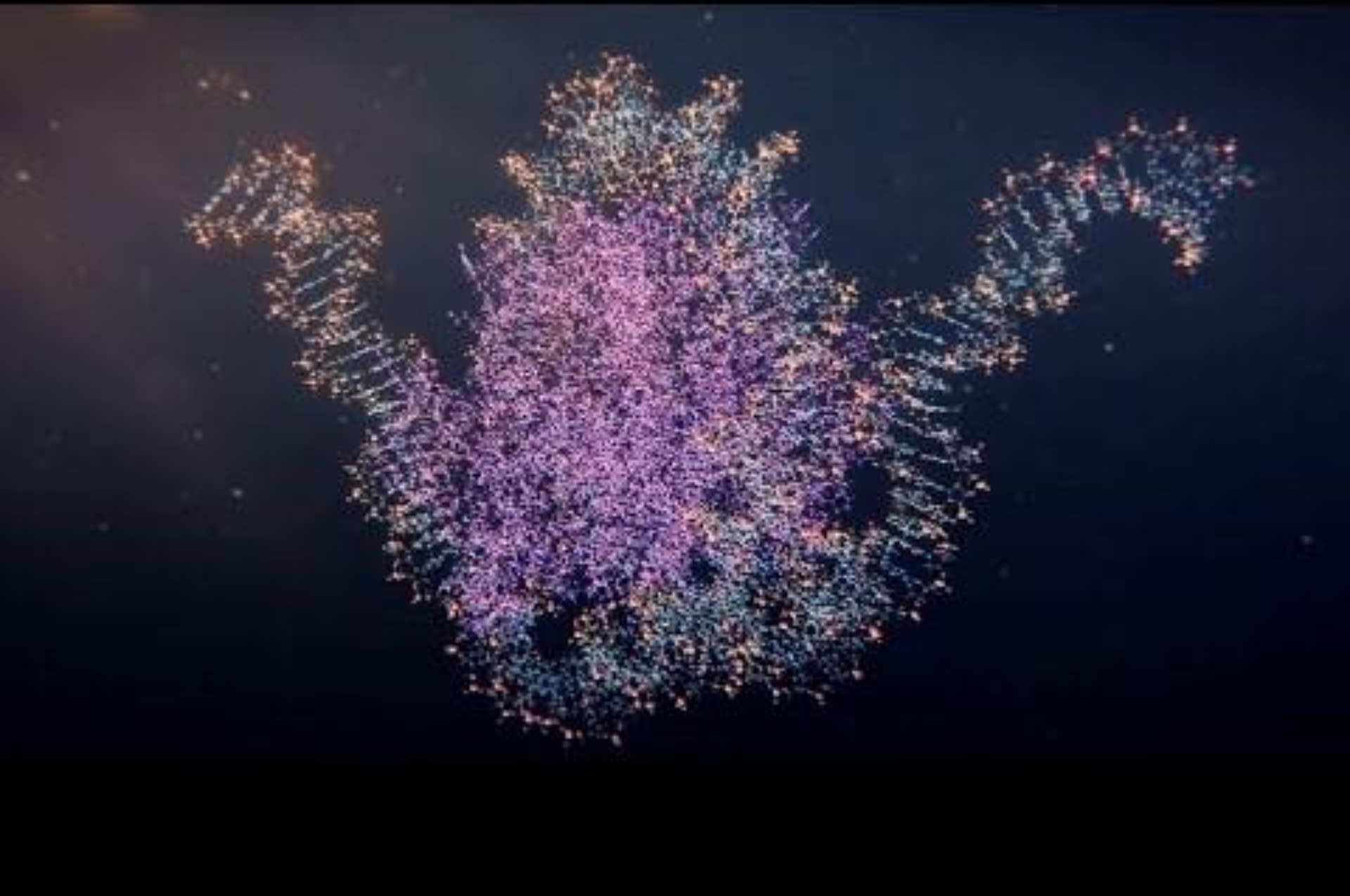


NSTEP	The time step that the MD simulation is at
TIME	The total time of the simulation (including restarts)
TEMP	System temperature
PRESS	System pressure
Etot	Total energy of the system
EKtot	Total kinetic energy of the system
EPtot	Total potential energy of the system

Trajektoria układu (zbiór położeń atomów w funkcji czasu)

6*6*2 LAYER made of POPC A, built of 6 residues; united atom
31949 0.9466000E+03
23.6689567 64.7803866 41.0944437 24.2084846 64.1276521 42.3503141
23.3005781 64.3412346 43.4894241 21.9890438 63.5834207 43.3622669
21.0339638 63.8854831 44.5671746 19.7936690 63.2203522 44.1329786
18.5972271 63.2837801 45.1084069 18.9041160 62.4152601 46.3808788
17.8836478 62.6568132 47.4873951 18.1484763 62.4949013 48.7911322
19.4947570 61.9605219 49.3301369 19.8400878 62.4858792 50.7388486
18.7550477 62.4269931 51.7337683 19.1719719 63.0012332 53.0614203
18.1688185 62.8697911 54.0943868 18.5051257 63.7985092 55.2943021
17.6161882 63.5432961 56.5497843 18.0276000 64.4250195 57.7689137
17.8397159 65.6610175 57.6811261 19.0840960 63.8578614 58.4151114
19.8536217 64.8544128 59.0933408 20.1222681 64.3236762 60.4909903
18.9117605 64.0633365 61.1380845 18.9059599 63.1863742 62.4971722
20.1417722 63.3965449 63.2436460 17.6672145 63.7205446 63.1361581
18.7986175 61.5997432 62.2252364 17.9526703 61.1540734 61.1771877
18.0700027 59.6519669 60.8471474 17.6959929 59.2458115 59.4893770
18.0945944 57.8648251 59.1086975 16.2484431 59.2799871 59.4928062
18.1389489 60.0887036 58.3423607 21.1512760 65.2581816 58.4208670





Pytania i odpowiedzi na zadania

m.mozolewska@ipipan.waw.pl

Zdjęcia, schematy i rysunki zostały zaczerpnięte z internetu.

Zadania (maksymalnie 10 punktów)

1. Na podstawie pliku top, mdcrd i netcdf zrób następujące operacje:
   a) Wczytaj za pomocą pakietu
   b) Ile jest atomów i klatek w trajektorii?
   c) Stwórz plik pdb
   d) Wygeneruj na podstawie pliku pdb mapę kontaktów
   e) Na podstawie trajektorii zrób PCA
   f) Wyrysuj RMSD trajektorii
   g) Oblicz promień żyracji