Molecular Modelling Course

November 17. 2025 – 19. 2025 | Dr. Birkan Emrem, Dr. Ferdinand Jamitzky, Dr. Plamen Dobrev, Dr. Prasanth Ganta

Leibniz Supercomputing Centre

Outline

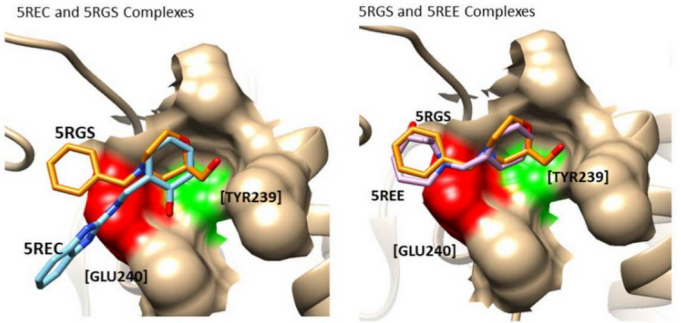
- Motivation
- Bio-molecular atomistic structures
- Molecular dynamics introduction
- Introduction to GROMACS
- Input file parameters and topologies for molecular dynamics simulations and in GROMACS in particular

Outline

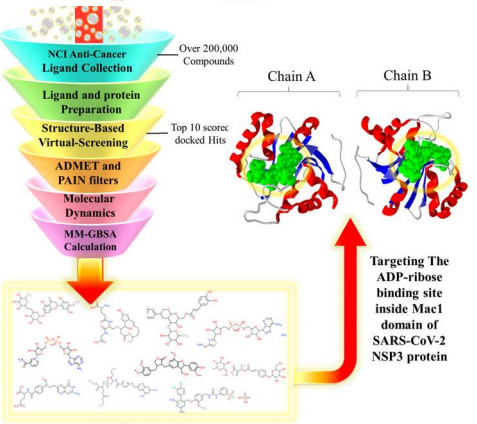
- Motivation
- Bio-molecular atomistic structures
- Molecular dynamics introduction
- Introduction to GROMACS
- Input file parameters and topologies for molecular dynamics simulations and in GROMACS in particular

Why do simulations?

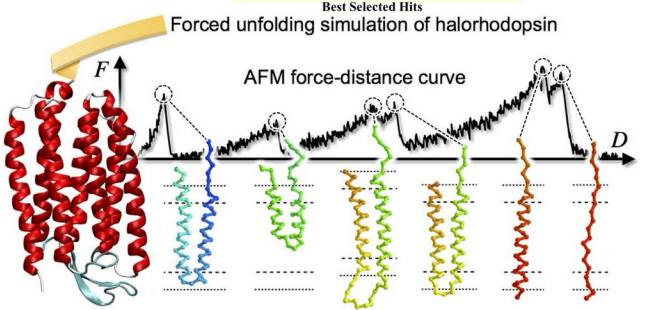
- To study observables not accessible by experimental approaches



- To carry out high throughput essays, studying system under multiple conditions or multiple systems simultaneously



- Reproduce experiment and validate models on microscopic level

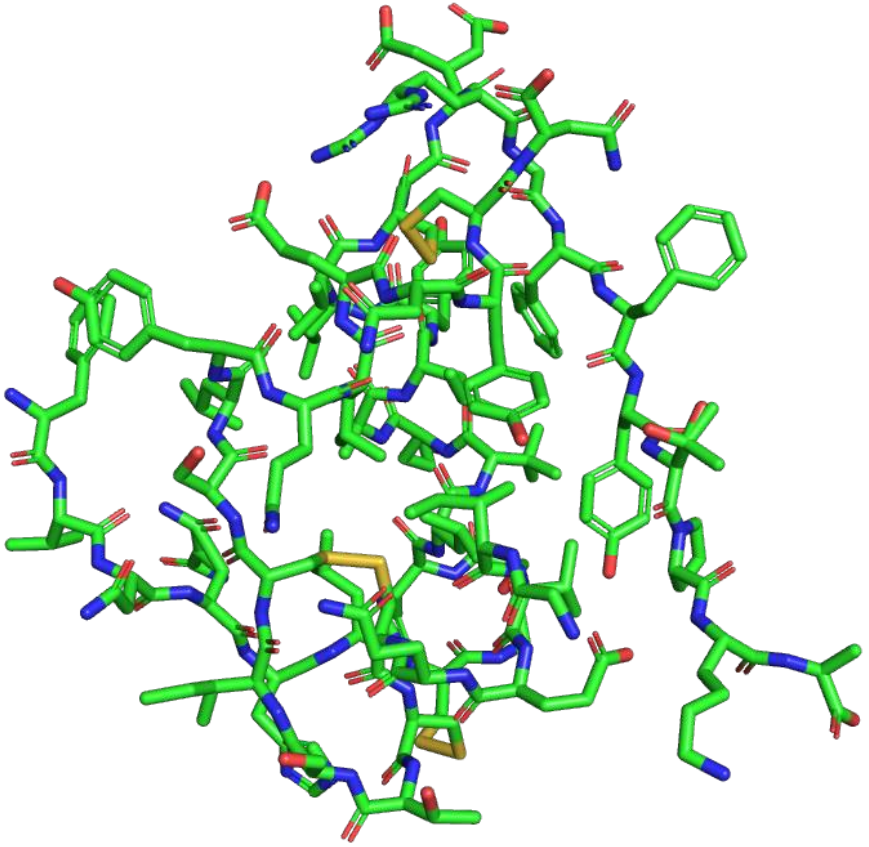


Viruses 2023, 15, 2291 https://doi.org/10.3390/v15122291
Scientific Reports 2021, 11, 7429 https://doi.org/10.1038/s41598-021-86471-0
Chemical Physics Letters 2018, 691, 276 https://doi.org/10.1016/j.cplett.2017.11.025

Outline

- Motivation
- Bio-molecular atomistic structures
- Molecular dynamics introduction
- Introduction to GROMACS
- Input file parameters and topologies for molecular dynamics simulations and in GROMACS in particular

Biomolecular Structures



- Structures of biological molecules are saved as a set of of atomic coordinates
- They are derived experimentally either by X-ray crystallography, atomic NMR or Cryo-EM, or *in silico* using AI methods (e.g. AlphaFold)
- Atomic resolution of a structure is crucial for obtaining physiologically relevant data during simulation

https://www.rcsb.org

RCSB PDB | Deposit | Search | Visualize | Analyze | Download | Learn | About | Careers | COVID-19 | Help | Contact us | MyPDB

RCSB PDB PROTEIN DATA BANK

236,060 Structures from the PDB

1,068,577 Computed Structure Models (CSM)

Enter search term(s), Entry ID(s), Ligand ID or sequence | Include CSM

Advanced Search | Browse Annotations | Help

PDB-101 | wwPDB | EMDataResource | NAKB | wwPDB Foundation | PDB-IHM

Redesigned PDB Statistics Support Enhanced Functionality | Explore Statistics

- Welcome
- Deposit
- Search
- Visualize
- Analyze
- Download
- Learn

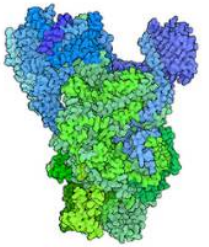
RCSB Protein Data Bank (RCSB PDB) enables breakthroughs in science and education by providing access and tools for exploration, visualization, and analysis of:

Experimentally-determined 3D structures from the **Protein Data Bank (PDB)** archive

Computed Structure Models (CSM) from AlphaFold DB and ModelArchive

These data can be explored in context of external annotations providing a structural view of biology.

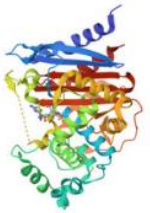
May **Molecule of the Month**



TOC-TIC Translocon

Latest Entries

As of Tue May 13 2025



9EHY

X-ray crystal structure of ADC-33 beta-lactamase in complex with ceftazidime in acyl and product forms

Features & Highlights

Announcement: Ligand Expo To Be Retired in 2025

Users should transition to RCSB PDB and wwPDB services as soon as possible

Watch the Webinar: Seeing Bird Flu in 3D

This virtual course will help you explore Influenza A H5N1 virus protein structures using RCSB.org tools

Watch the Webinar: Streamlining Access to RCSB PDB APIs with Python

This virtual course introduces and demonstrates usage of the rcsb-api Python package for easy interactions with RCSB PDB Search and Data API services

News | Publications

Access Computed Structure Model Annotations

Gene Ontology terms (for Molecular Function, Cellular Component, and Biological Process), InterPro protein family classifications, and Pharos disease associations are available for CSMs at RCSB.org

» *05/19/2025*

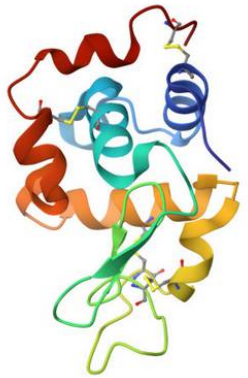
PDB-101 Focus: Peak Performance

PDB-101 materials explore the structural biology of nutrition.

» *05/16/2025*

Paper Published: A new chapter for Molecule of the Month

Biological Assembly 1



Explore in 3D: Structure | Sequence Annotations | Electron Density | Validation Report

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

Find Similar Assemblies

Biological assembly 1 assigned by authors.

Macromolecule Content

- Total Structure Weight: 14.33 kDa
- Atom Count: 1,082
- Modeled Residue Count: 129
- Deposited Residue Count: 129

Display Files | Download Files | Data API

2YVB | pdb_00002yvb

High resolution X-ray crystal structure of Tetragonal Hen egg white lysozyme

PDB DOI: https://doi.org/10.2210/pdb2YVB/pdb

Classification: HYDROLASE
Organism(s): Gallus gallus
Mutation(s): No

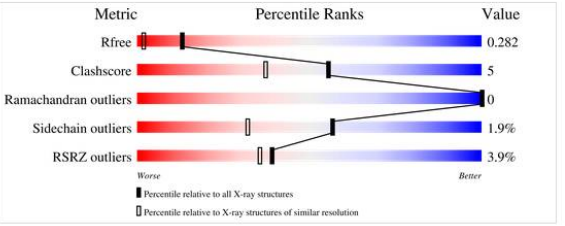
Deposited: 2007-04-10 **Released:** 2007-04-24
Deposition Author(s): Naresh, M.D., Jaimohan, S.M., Kumar, V.V., Mandal, A.B.

Experimental Data Snapshot

Method: X-RAY DIFFRACTION
Resolution: 1.62 Å
R-Value Free:
0.277 (Depositor), 0.280 (DCC)
R-Value Work:
0.229 (Depositor), 0.230 (DCC)
R-Value Observed:
0.231 (Depositor)

Starting Model: experimental
View more details

wwPDB Validation 3D Report | Full Report



This is version 1.6 of the entry. See complete history.

Literature

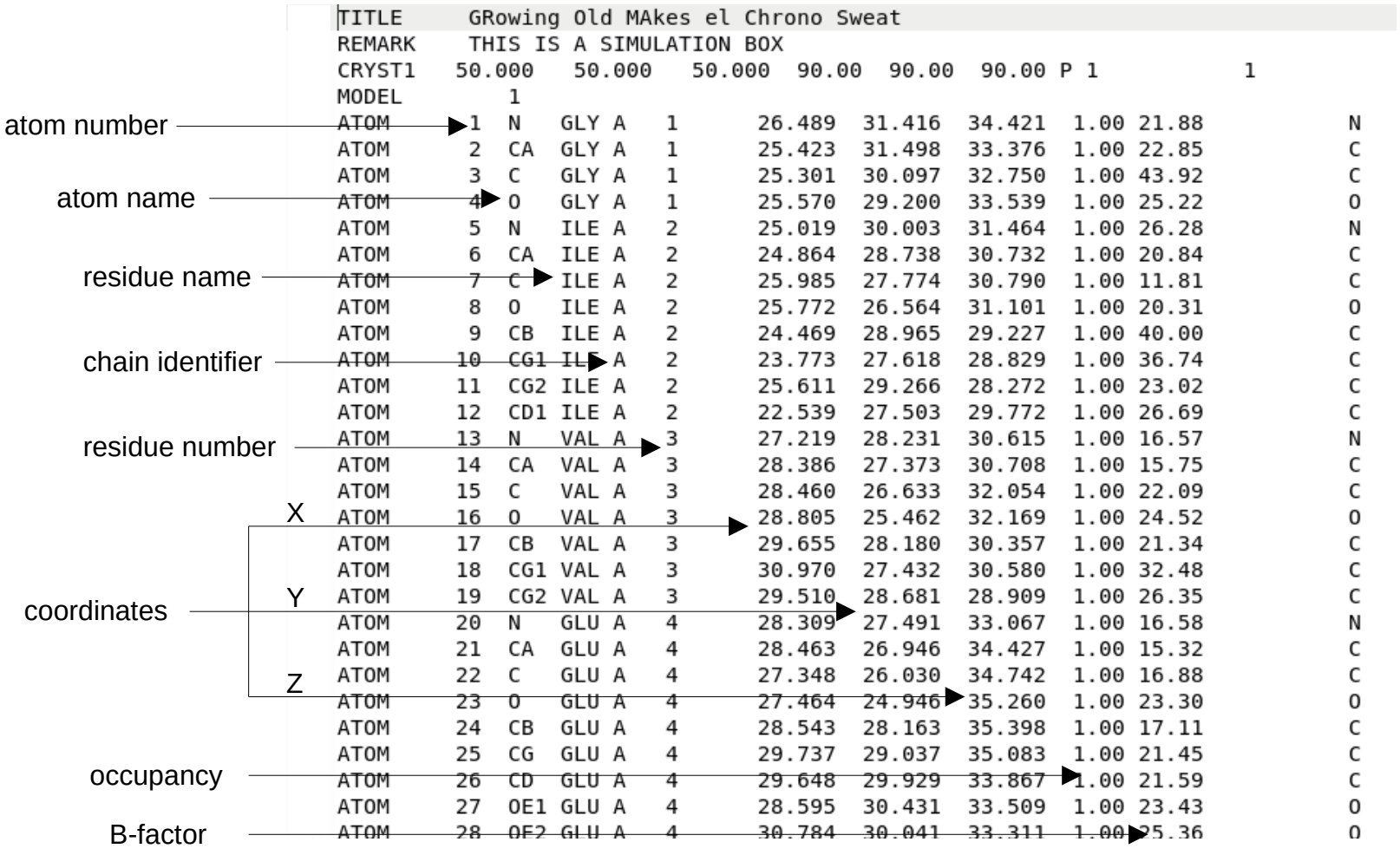
Download Primary Citation

High resolution X-ray crystal structure of Tetragonal Hen egg white lysozyme

Naresh, M.D., Jaimohan, S.M., Kumar, V.V.

To be published.

PDB file format



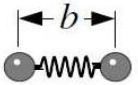

```
TITLE     GRowing Old MAkes el Chrono Sweat
REMARK    THIS IS A SIMULATION BOX
CRYST1   50.000   50.000   50.000  90.00  90.00  90.00 P 1           1
MODEL        1
ATOM      1  N   GLY A   1      26.489  31.416  34.421  1.00 21.88           N
ATOM      2  CA  GLY A   1      25.423  31.498  33.376  1.00 22.85           C
ATOM      3  C   GLY A   1      25.301  30.097  32.750  1.00 43.92           C
ATOM      4  O   GLY A   1      25.570  29.200  33.539  1.00 25.22           O
ATOM      5  N   ILE A   2      25.019  30.003  31.464  1.00 26.28           N
ATOM      6  CA  ILE A   2      24.864  28.738  30.732  1.00 20.84           C
ATOM      7  C   ILE A   2      25.985  27.774  30.790  1.00 11.81           C
ATOM      8  O   ILE A   2      25.772  26.564  31.101  1.00 20.31           O
ATOM      9  CB  ILE A   2      24.469  28.965  29.227  1.00 40.00           C
ATOM     10  CG1 ILE A   2      23.773  27.618  28.829  1.00 36.74           C
ATOM     11  CG2 ILE A   2      25.611  29.266  28.272  1.00 23.02           C
ATOM     12  CD1 ILE A   2      22.539  27.503  29.772  1.00 26.69           C
ATOM     13  N   VAL A   3      27.219  28.231  30.615  1.00 16.57           N
ATOM     14  CA  VAL A   3      28.386  27.373  30.708  1.00 15.75           C
ATOM     15  C   VAL A   3      28.460  26.633  32.054  1.00 22.09           C
ATOM     16  O   VAL A   3      28.805  25.462  32.169  1.00 24.52           O
ATOM     17  CB  VAL A   3      29.655  28.180  30.357  1.00 21.34           C
ATOM     18  CG1 VAL A   3      30.970  27.432  30.580  1.00 32.48           C
ATOM     19  CG2 VAL A   3      29.510  28.681  28.909  1.00 26.35           C
ATOM     20  N   GLU A   4      28.309  27.491  33.067  1.00 16.58           N
ATOM     21  CA  GLU A   4      28.463  26.946  34.427  1.00 15.32           C
ATOM     22  C   GLU A   4      27.348  26.030  34.742  1.00 16.88           C
ATOM     23  O   GLU A   4      27.464  24.946  35.260  1.00 23.30           O
ATOM     24  CB  GLU A   4      28.543  28.163  35.398  1.00 17.11           C
ATOM     25  CG  GLU A   4      29.737  29.037  35.083  1.00 21.45           C
ATOM     26  CD  GLU A   4      29.648  29.929  33.867  1.00 21.59           C
ATOM     27  OE1 GLU A   4      28.595  30.431  33.509  1.00 23.43           O
ATOM     28  OE2 GLU A   4      30.784  30.041  33.311  1.00 25.36           O
```

Outline

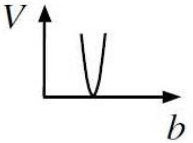
- Motivation
- Bio-molecular atomistic structures
- Molecular dynamics introduction
- Introduction to GROMACS
- Input file parameters and topologies for molecular dynamics simulations and in GROMACS in particular

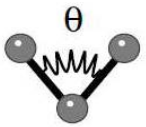
Molecular Dynamics Introduction

classic potentials

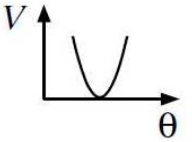


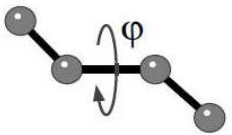
$$V^{\mathrm{B}} = \tfrac{1}{2} k_b (b - b_0)^2$$



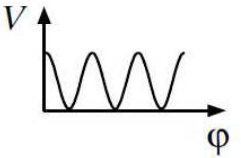


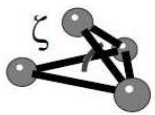
$$V^{\mathrm{W}} = \tfrac{1}{2} k_\theta (\theta - \theta_0)^2$$



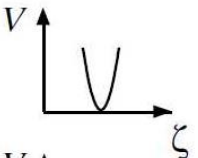


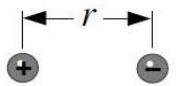
$$V^{\mathrm{D}} = k_\varphi [1 + \cos(n\varphi - \delta)]$$



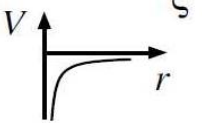


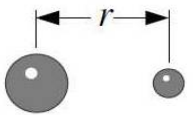
$$V^{\mathrm{E}} = \tfrac{1}{2} k_\zeta (\zeta - \zeta_0)^2$$



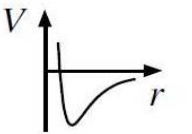


$$V^{\mathrm{C}} = q_i q_j / (4\pi\epsilon_0\epsilon_r r)$$



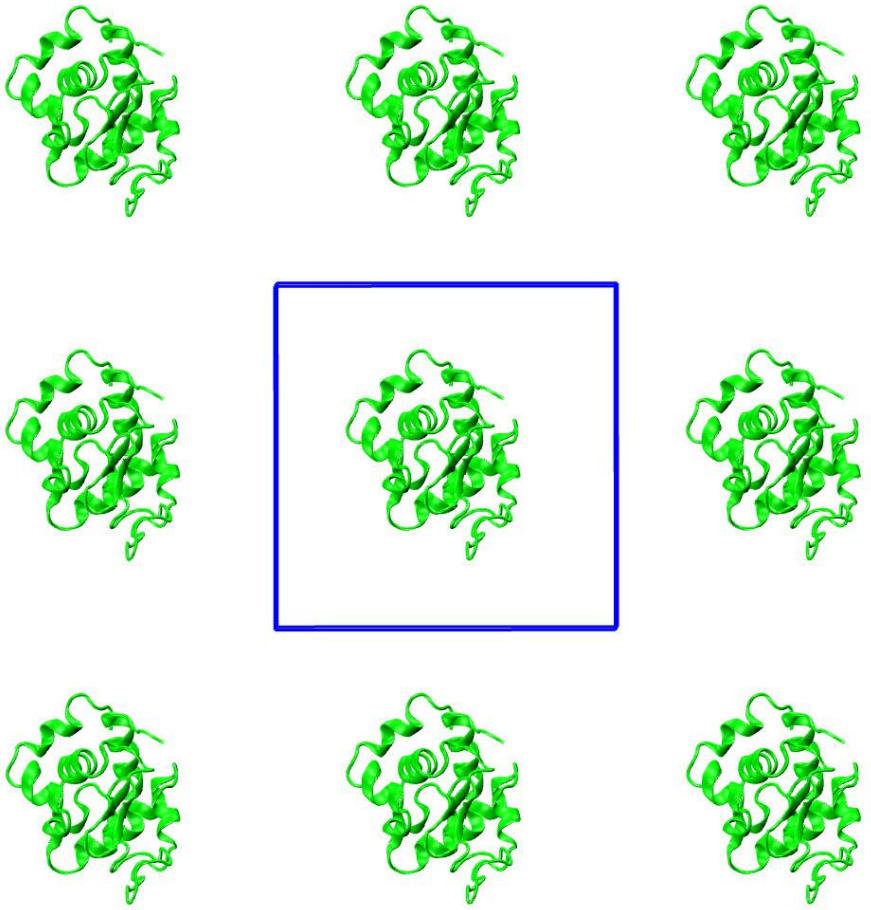


$$V^{\mathrm{LJ}} = C_{12}(i,j)/r^{12} - C_6(i,j)/r^6$$



Molecular Dynamics Introduction

periodic boundaries



Molecular Dynamics Introduction
thermostats and barostats

Berendsen thermostat

$$\frac{dT}{dt} = \frac{T_0 - T}{\tau}$$

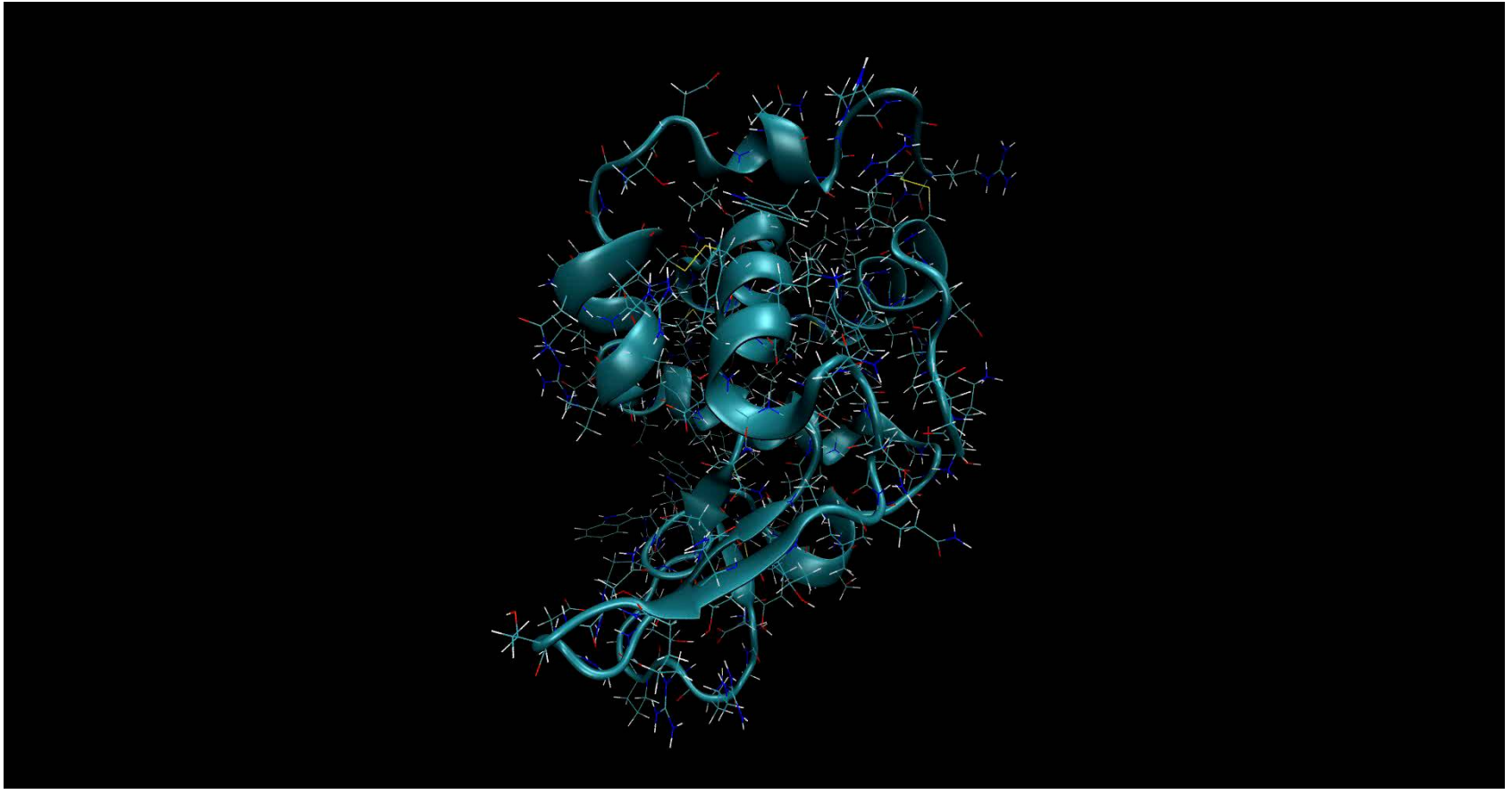
$$\lambda = \left[1 + \frac{n_{TC} \Delta t}{\tau_T} \left\{ \frac{T_0}{T\left(t - \frac{1}{2} \Delta t \right)} - 1 \right\} \right]^{1/2}$$

Berendsen barostat

$$\frac{dP}{dt} = \frac{P_0 - P}{\tau_p}$$

$$\mu_{ij} = \delta_{ij} - \frac{n_{PC} \Delta t}{3 \tau_p} \beta_{ij} \{ P_{0ij} - P_{ij}(t) \}$$

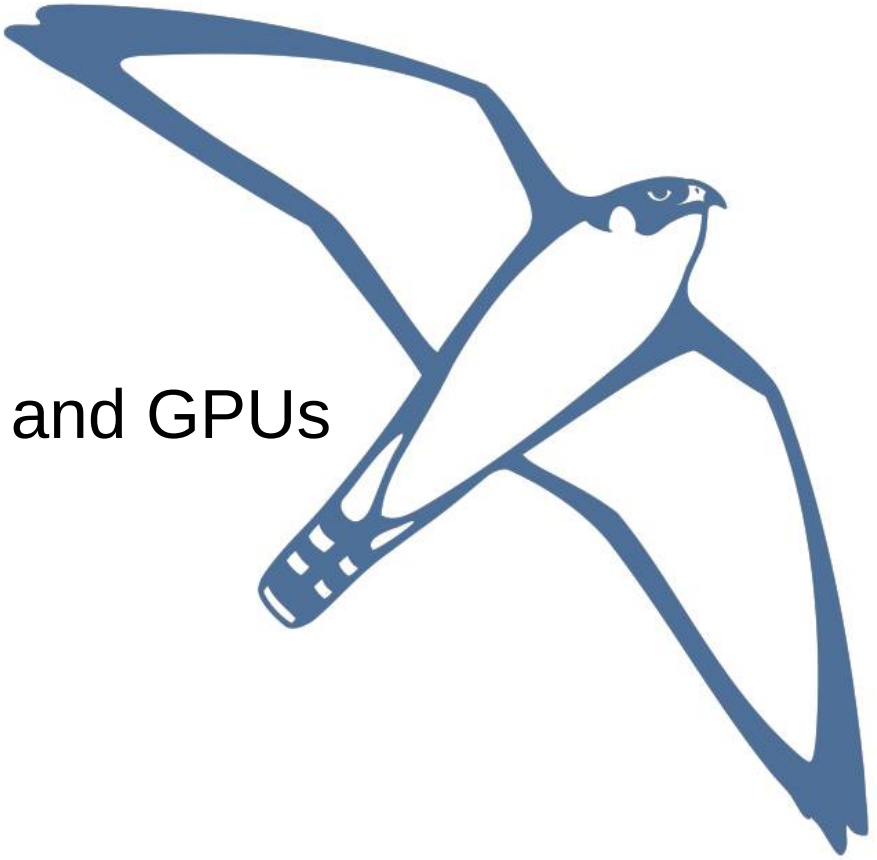
Molecular Dynamics Introduction
trajectories



Outline

- Motivation
- Bio-molecular atomistic structures
- Molecular dynamics introduction
- Introduction to GROMACS
- Input file parameters and topologies for molecular dynamics simulations and in GROMACS in particular

Introduction to GROMACS



- Open Source
- Highly parallel
- Supports almost all modern CPUs and GPUs
- Easy and straightforward to use

Outline

- Motivation
- Bio-molecular atomistic structures
- Molecular dynamics introduction
- Introduction to GROMACS
- Input file parameters and topologies for molecular dynamics simulations and in GROMACS in particular

Topologies

Molecular descriptions in MD software packages are stored in topologies

Introduction to GROMACS

topologies

```
; Include forcefield parameters
#include "amber99sb-ildn.ff/forcefield.itp"

[ moleculetype ]
; Name            nrexcl
Protein_chain_A     3

[ atoms ]
;   nr       type  resnr residue  atom   cgnr     charge       mass  typeB    chargeB      massB
; residue   1 LYS rtp NLYS q +2.0
     1         N3      1    LYS      N      1     0.0966      14.01
     2          H      1    LYS     H1      2     0.2165      1.008
     3          H      1    LYS     H2      3     0.2165      1.008
     4          H      1    LYS     H3      4     0.2165      1.008
     5         CT      1    LYS     CA      5    -0.0015      12.01
     6         HP      1    LYS     HA      6      0.118      1.008
     7         CT      1    LYS     CB      7     0.0212      12.01
     8         HC      1    LYS    HB1      8     0.0283      1.008
     9         HC      1    LYS    HB2      9     0.0283      1.008
    10         CT      1    LYS     CG     10    -0.0048      12.01
    11         HC      1    LYS    HG1     11     0.0121      1.008
    12         HC      1    LYS    HG2     12     0.0121      1.008
    13         CT      1    LYS     CD     13    -0.0608      12.01
    14         HC      1    LYS    HD1     14     0.0633      1.008
    15         HC      1    LYS    HD2     15     0.0633      1.008
    16         CT      1    LYS     CE     16    -0.0181      12.01
    17         HP      1    LYS    HE1     17     0.1171      1.008
```

```
[ bonds ]
;  ai    aj funct           c0           c1           c2           c3
    1     2     1
    1     3     1
    1     4     1
    1     5     1
    5     6     1
    5     7     1
    5    23     1
    7     8     1
    7     9     1
    7    10     1
   10    11     1
   10    12     1
   10    13     1
```

Introduction to GROMACS

topologies

```
[ angles ]
;  ai    aj    ak funct         c0         c1         c2         c3
    2     1     3     1
    2     1     4     1
    2     1     5     1
    3     1     4     1
    3     1     5     1
    4     1     5     1
    1     5     6     1
    1     5     7     1
    1     5    23     1

    6     5     7     1
```

```
[ dihedrals ]
;  ai    aj    ak    al funct         c0         c1         c2         c3         c4         c5
    2     1     5     6     9
    2     1     5     7     9
    2     1     5    23     9
    3     1     5     6     9
    3     1     5     7     9
    3     1     5    23     9

    4     1     5     6     9
```

Introduction to GROMACS

topologies

```
; Include Position restraint file
#ifdef POSRES
#include "posre.itp"
#endif

; Include water topology
#include "amber99sb-ildn.ff/spce.itp"

#ifdef POSRES_WATER
; Position restraint for each water oxygen
[ position_restraints ]
;  i funct       fcx        fcy        fcz
   1    1       1000       1000       1000
#endif

; Include topology for ions
#include "amber99sb-ildn.ff/ions.itp"

[ system ]
; Name
LYSOZYME C

[ molecules ]
; Compound        #mols
Protein_chain_A     1
SOL              6628
NA                 20
CL                 28
```

Introduction to GROMACS

topologies

```
; Include forcefield parameters
#include
"amber99sb-ildn.ff/forcefield.itp"

; Include chain topologies
#include "topol_Protein_chain_A.itp"
#include "topol_Protein_chain_B.itp"

; Include water topology
#include "amber99sb-ildn.ff/spce.itp"

#ifdef POSRES_WATER
; Position restraint for each water
oxygen
[ position_restraints ]
;  i funct       fcx        fcy        fcz
   1    1       1000       1000       1000
#endif

; Include topology for ions
#include "amber99sb-ildn.ff/ions.itp"

[ system ]
; Name
INSULIN (CHAIN A); INSULIN (CHAIN
B)

[ molecules ]
; Compound        #mols
Protein_chain_A     1
Protein_chain_B     1
SOL               31
```

Force field organization

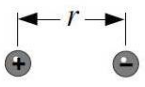
The potentials among all atom types inside a protein are described by set of parameters called force field

Introduction to GROMACS

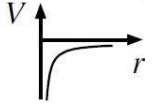
force fields

Molecule definition

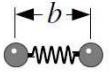
```
[ ALA ]
 [ atoms ]
    N    N                -0.41570     1
    H    H                 0.27190     2
   CA    CT                0.03370     3
   HA    H1                0.08230     4
   CB    CT               -0.18250     5
  HB1    HC                0.06030     6
  HB2    HC                0.06030     7
  HB3    HC                0.06030     8
    C    C                 0.59730     9
    O    O                -0.56790    10
[ bonds ]
    N    H
    N   CA
   CA   HA
   CA   CB
   CA    C
   CB  HB1
   CB  HB2
   CB  HB3
    C    O
   -C    N
[ impropers ]
   -C   CA    N    H
   CA   +N    C    O
```



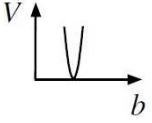
$$V^{\mathrm{C}} = q_i q_j / (4\pi\epsilon_0\epsilon_r r)$$



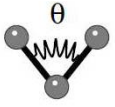
Bonded parameters



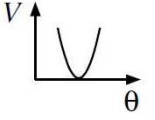
$$V^{\mathrm{B}} = \tfrac{1}{2} k_b (b - b_0)^2$$



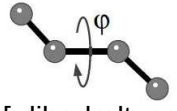

```
[ bondtypes ]
; i    j  func       b0          kb
  C    C      1    0.1525   259408.0 ; new99
  C   OS      1    0.1323   376560.0 ; new99
  C   H4      1    0.1080   307105.6 ; new99
```



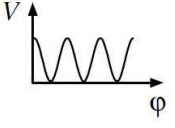
$$V^{\mathrm{W}} = \tfrac{1}{2} k_\theta (\theta - \theta_0)^2$$




```
[ angletypes ]
;  i    j    k  func       th0       cth
HW   OW   HW      1   104.520    836.800 ; TIP3P water
HW   HW   OW      1   127.740      0.000 ;
C    C    O       1   120.000    669.440 ; new99
C    C   OH       1   120.000    669.440 ; new99
CT   C    CT      1   117.000    527.184 ; new99
```



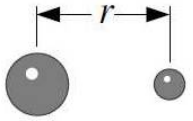
$$V^{\mathrm{D}} = k_\varphi [1 + \cos(n\varphi - \delta)]$$



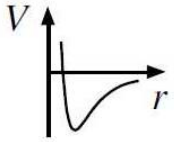

```
[ dihedraltypes ]
;i   j   k   l    func
 CT  CT  OS  CT    9      0.0      1.60247     3  ;
 CT  CT  OS  CT    9    180.0      0.41840     2  ;
 C   N   CT  C     9      0.0      1.12968     2  ; new for 99sb
 C   N   CT  C     9      0.0      1.75728     3  ; new for 99sb
```

Introduction to GROMACS

force fields



$$V^{\mathrm{LJ}} = C_{12}(i,j)/r^{12} - C_6(i,j)/r^6$$




```
[ atomtypes ]
; name      at.num  mass     charge ptype  sigma      epsilon
Br          35      79.90    0.0000  A   3.95559e-01  1.33888e+00 ; Converted from parm99.dat
C           6      12.01    0.0000  A   3.39967e-01  3.59824e-01
CA          6      12.01    0.0000  A   3.39967e-01  3.59824e-01
CB          6      12.01    0.0000  A   3.39967e-01  3.59824e-01
CC          6      12.01    0.0000  A   3.39967e-01  3.59824e-01
CK          6      12.01    0.0000  A   3.39967e-01  3.59824e-01
CM          6      12.01    0.0000  A   3.39967e-01  3.59824e-01
CN          6      12.01    0.0000  A   3.39967e-01  3.59824e-01
CQ          6      12.01    0.0000  A   3.39967e-01  3.59824e-01
CR          6      12.01    0.0000  A   3.39967e-01  3.59824e-01
CT          6      12.01    0.0000  A   3.39967e-01  4.57730e-01
CV          6      12.01    0.0000  A   3.39967e-01  3.59824e-01
```

Input parameters

Specifying the way the simulation would run and output data is descried by the input parameters

Introduction to GROMACS

simulation parameters

```
define              =  -DPOSRES
;disre              =  simple
;disre_fc           =  20
constraints         =  h-bonds
constraint-algorithm     = Lincs
lincs-order              = 6
integrator          =  md
dt                =  0.004     ; ps !
comm_mode           =  Linear
nsteps              =  2500000000;000000   ; total 10000  ps.
nstcomm             =  1

nstxout             =  100000
nstvout             =  100000
nstfout             =  0

nstxtcout = 5000

nstlog              =  1000
nstenergy           =  1000
nstlist             =  5
ns_type             =  grid
rlist               =  1
vdwtype             = Cut-off
cutoff-scheme       = Verlet
rvdw                = 1
coulombtype         = PME
fourierspacing      = 0.12
pme_order           = 4
rcoulomb            = 1
```

Introduction to GROMACS

simulation parameters

```
; Berendsen temperature coupling is on in
two groups
Tcoupl              =     v-rescale
tau_t               =     0.1           ;     0.5

tc-grps         =       System       ;
Protein    SOL_NA+_CL-
ref_t               =     300           ;    300

; Pressure coupling is not on
Pcoupl              =     c-rescale
pcoupltype = Isotropic
tau_p               =  1
compressibility     =   4.5e-5
ref_p               =  1.0
; Generate velocites is on at 300 K.
gen_vel             =  yes
gen_temp            =  300.0
gen_seed            =  -1
```

Introduction to GROMACS

simulation parameters, energy minimization

```
define          =  -DFLEXIBLE
constraints      =  none
integrator       =  steep
nsteps           =  100
;
;       Energy minimizing stuff
;
emtol            =  2000
emstep           =  0.01

nstcomm          =  1
ns_type          =  grid
rlist            =  1
rcoulomb         =  1.0
rvdw             =  1.0
Tcoupl           =  no
Pcoupl           =  no
gen_vel          =  no
```