Futuro in AREA

>>sharing experience

promoting network

Understanding molecular mechanisms using molecular dynamics (MD) simulations

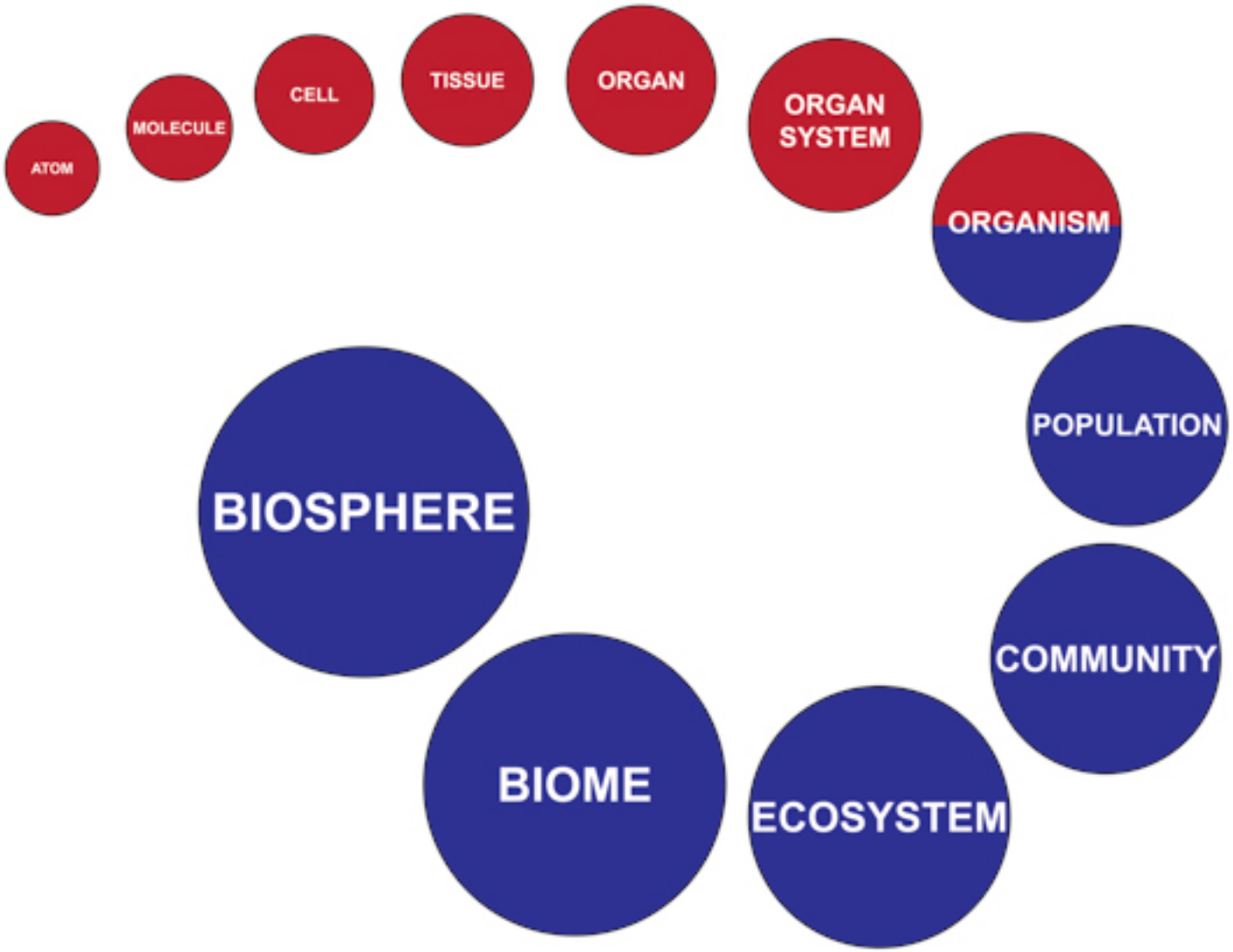
Delre Pietro, Institute of crystallography (IC)

23/10/2019

Consiglio Nazionale delle Ricerche

Molecular Dynamics (MD)

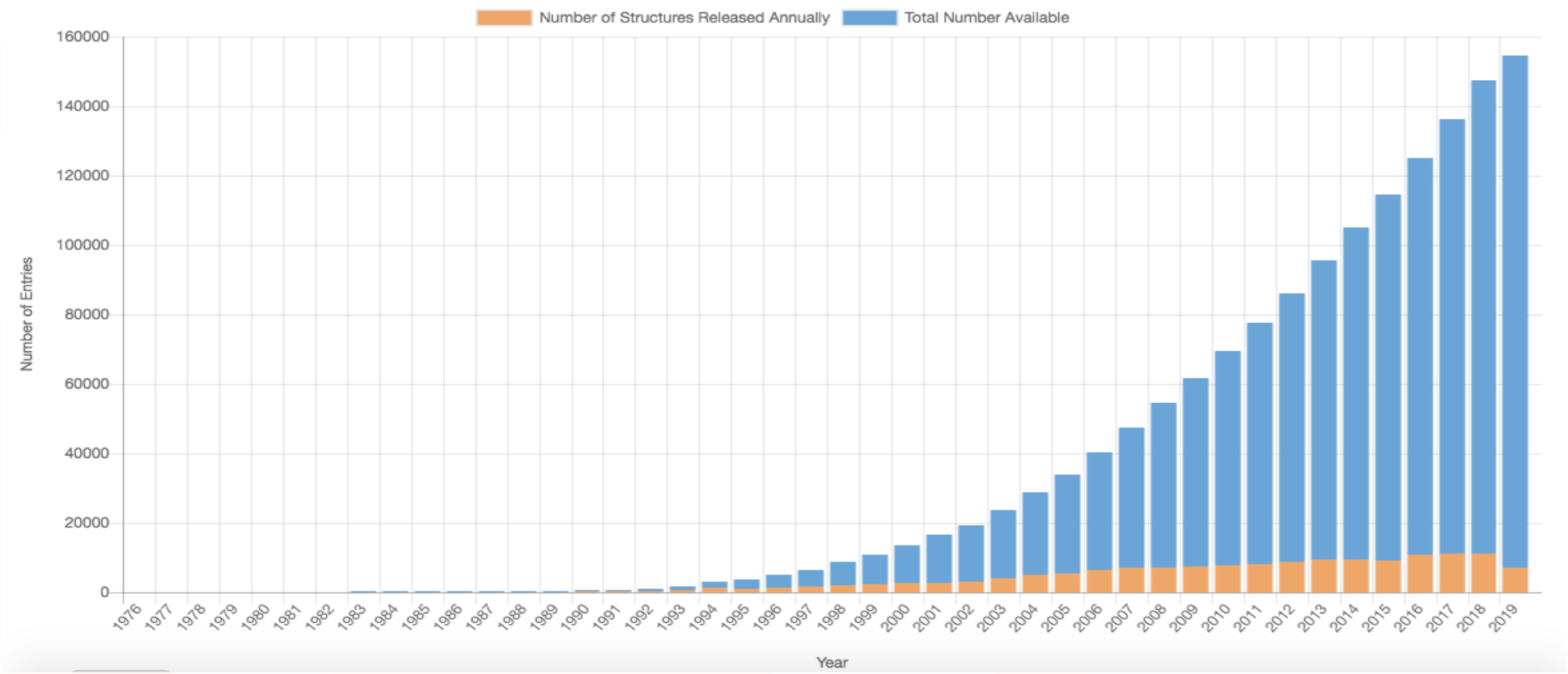
1. Basic Principles
2. Examples
3. Applications
4. Limitations



https://www.livingoceansfoundation.org/education/portal/course/ecology/#biological-hierarchy-of-life

RCSB PDB
PROTEIN DATA BANK

154735 Biological Macromolecular Structures Enabling Breakthroughs in Research and Education



https://www.rcsb.org/stats/growth/overall

ATOM	63	CB	LYS	A	3	66.750	29.725	6.408	1.00	20.42	C
ATOM	64	CG	LYS	A	3	67.656	29.230	7.510	1.00	21.60	C
ATOM	65	CD	LYS	A	3	68.223	27.854	7.207	1.00	22.32	C
ATOM	66	CE	LYS	A	3	69.119	27.347	8.319	0.01	22.13	C
ATOM	67	NZ	LYS	A	3	69.699	26.028	7.962	0.01	22.19	N
ATOM	68	N	ILE	A	4	65.984	32.679	5.545	1.00	20.69	N
ATOM	69	CA	ILE	A	4	65.858	33.533	4.385	1.00	21.50	C
ATOM	70	C	ILE	A	4	67.244	33.620	3.722	1.00	22.43	C
ATOM	71	O	ILE	A	4	68.209	33.831	4.452	1.00	22.94	O
ATOM	72	CB	ILE	A	4	65.330	34.969	4.663	1.00	20.49	C
ATOM	73	CG1	ILE	A	4	63.969	34.898	5.367	1.00	20.36	C
ATOM	74	CG2	ILE	A	4	65.223	35.784	3.360	1.00	19.99	C
ATOM	75	CD1	ILE	A	4	63.362	36.305	5.694	1.00	20.97	C
ATOM	76	N	ASP	A	5	67.251	33.383	2.445	1.00	23.57	N

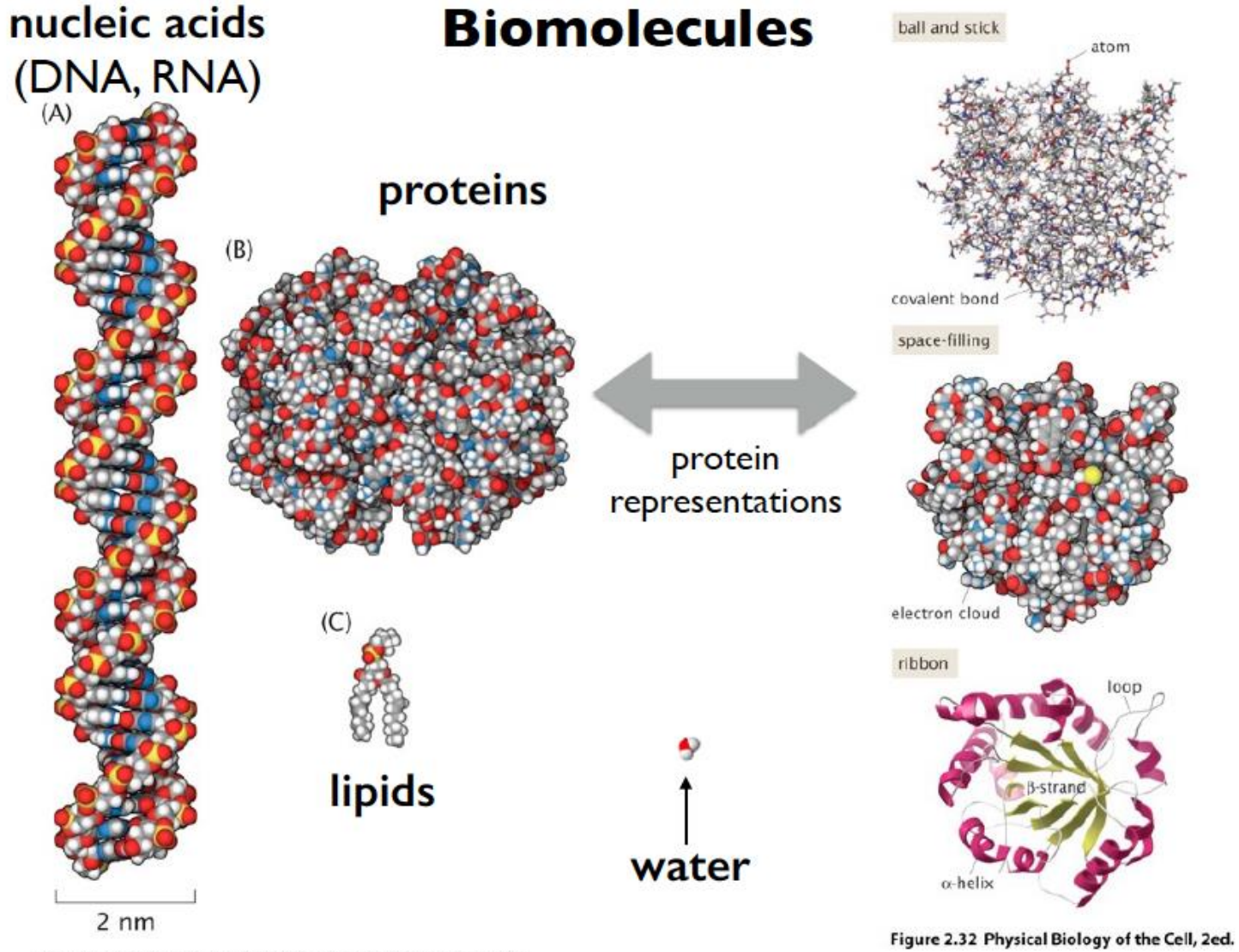
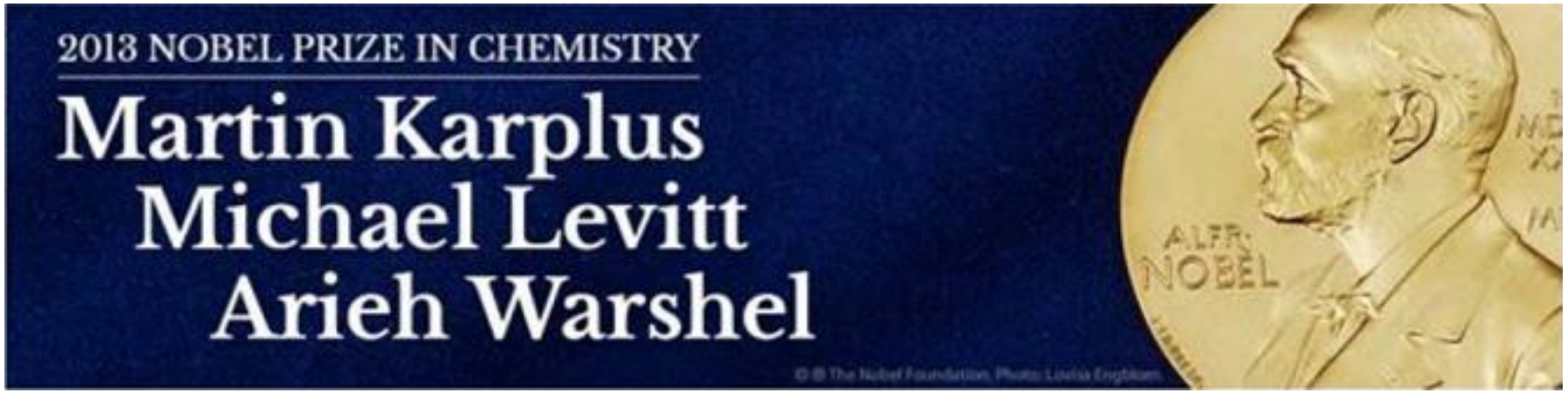


Figure 1.1 Physical Biology of the Cell, 2ed. (© Garland Science 2013)

Figure 2.32 Physical Biology of the Cell, 2ed.

"Everything that living things do can be understood in terms of the jiggling and wiggling of atoms." — Richard Feynman*

Molecular dynamics simulation (MD)



https://www.cup.uni-freiburg.de/de/aktuelles/nachrichten/nobelpreise/2013

Some mathematics…

- Newton's equation of motion

$$F_i = m_i a_i$$

- The force can be written as the gradient of the potential energy

$$F = -\nabla_i V$$

- Combine the two equations to get

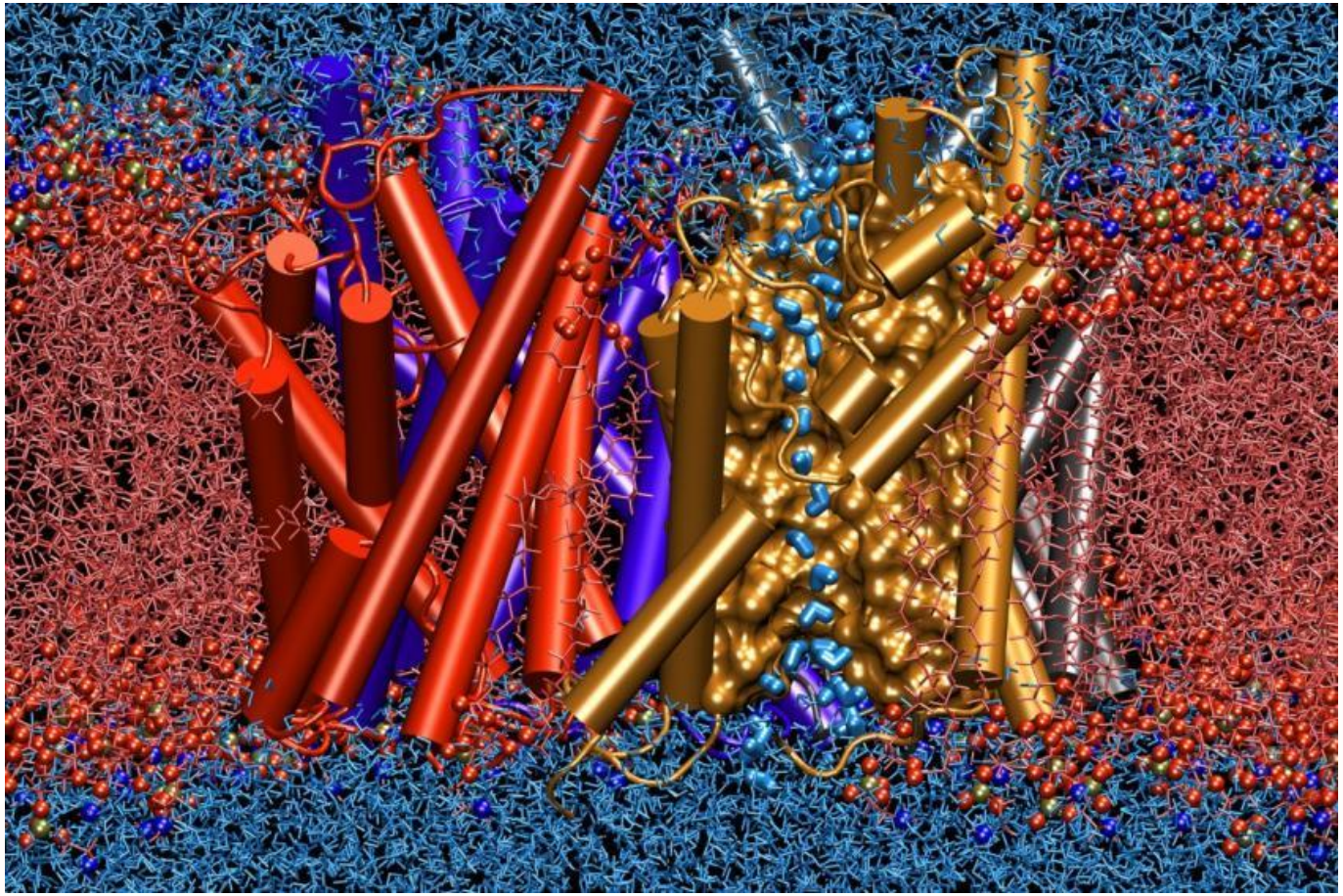
$$\frac{dV}{dr} = -m_i \frac{d^2 r}{dt^{2i}}$$

$$\mathbf{r}_i(t + \Delta t) = 2\mathbf{r}_i(t) - \mathbf{r}_i(t - \Delta t) + \frac{\mathbf{F}_i}{m_i}\Delta t^2$$

- A trajectory is obtained by solving this differential equation

$$\begin{matrix} \mathbf{r}_1(0), \ldots, \mathbf{r}_N(0) \\ \mathbf{r}_1(\Delta t), \ldots, \mathbf{r}_N(\Delta t) \\ \mathbf{r}_1(2\Delta t), \ldots, \mathbf{r}_N(2\Delta t) \\ \mathbf{r}_1(3\Delta t), \ldots, \mathbf{r}_N(3\Delta t) \end{matrix} \quad \text{trajectory} \quad \begin{matrix} (\mathbf{r}_1(t), \ldots, \mathbf{r}_N(t)) \\ 0 \le t \le \tau \end{matrix}$$

Case of study: Water Channels in Cell Membranes



-starting point: aquaporin tetramer (PDB)

- careful setup of the system (lipid bilayer membrane and hydrated by slabs of water on both sides)
- system of 106,000 atoms.
- Full atomic simulation: use of supercomputer

- Multinanosecond trajectory

trajectory

$$(\mathbf{r}_1(t), \ldots, \mathbf{r}_N(t))$$

$$0 \leq t \leq \tau$$

https://www.ks.uiuc.edu/Highlights/supplement/2003-10a.html

MD ANALYSIS

NumPy

python™

simulation trajectory

dcd, xtc, trr, ncdf, trj, pdb, pqr, gro, crd, dms, trz, mol2, xyz, config, history, gms, ...

psf, tpr, prmtop, dms, mol2, hoomd xml, ...

"accessible" structured data

analysis algorithm

processed data

tables

images

graphs

MDAnalysis: A Python Package for the Rapid Analysis of Molecular Dynamics Simulations | SciPy 2016

Analysis module: MDAnalysis.analysis

- standard analysis functionality (RMSD, RMSF, distances, density, hydrogen bonds, native contacts…)

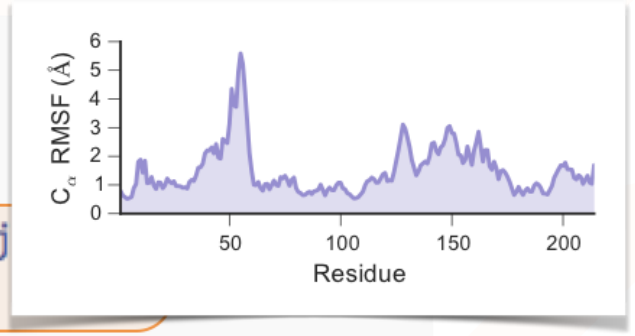
C_α RMSF

$$\rho_i = \sqrt{\left\langle (\mathbf{x}_i(t) - \langle \mathbf{x}_i \rangle)^2 \right\rangle}$$

MD ANALYSIS

NumPy

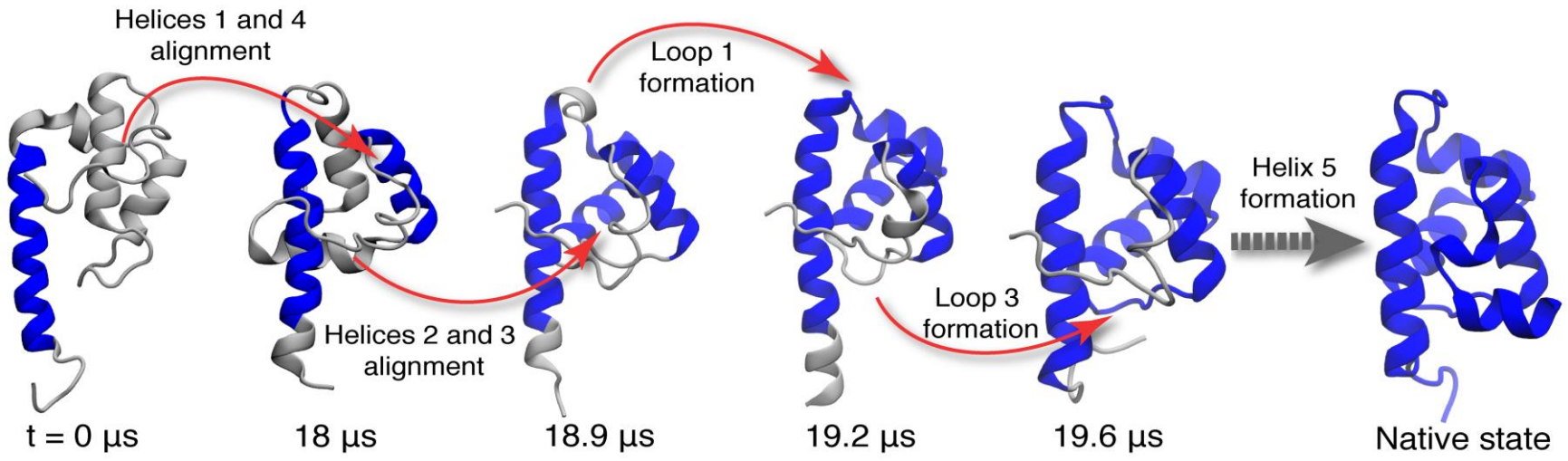
```
import numpy as np
import MDAnalysis as mda

u = mda.Universe("topol.tpr", "trj
ca = u.select_atoms("name CA")
means = np.zeros((len(ca), 3))
sumsq = np.zeros_like(means)
for k, ts in enumerate(u.trajectory):
    sumsq += (k/(k+1.0)) * (ca.positions - means)**2
    means[:] = (k*means + ca.positions)/(k+1.0)
rmsf = np.sqrt(sumsq.sum(axis=1)/(k+1.0))

matplotlib.pyplot.plot(ca.residues.resids, rmsf)
```



MDAnalysis: A Python Package for the Rapid Analysis of Molecular Dynamics Simulations | SciPy 2016

MD simulations allow studying:

- Ion transport in biological systems (case study)
- Protein stability
- Conformational changes
- Protein folding

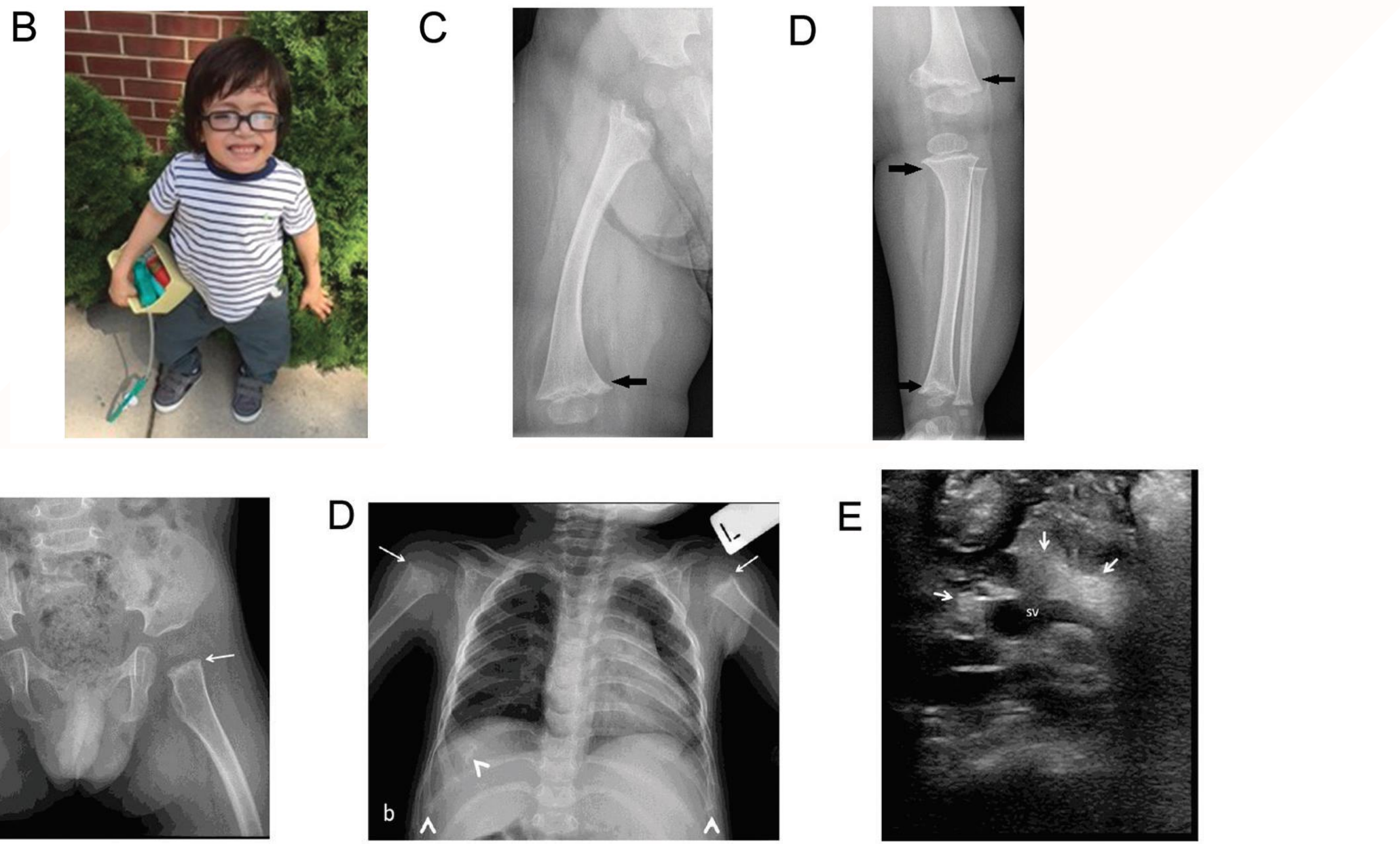


http://www.ks.uiuc.edu/Research/folding/

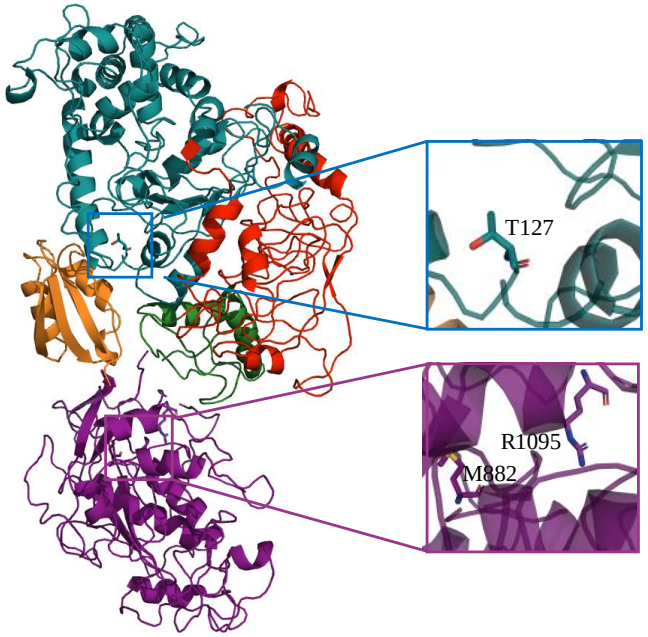
- The difference beetwen native and mutant proteins

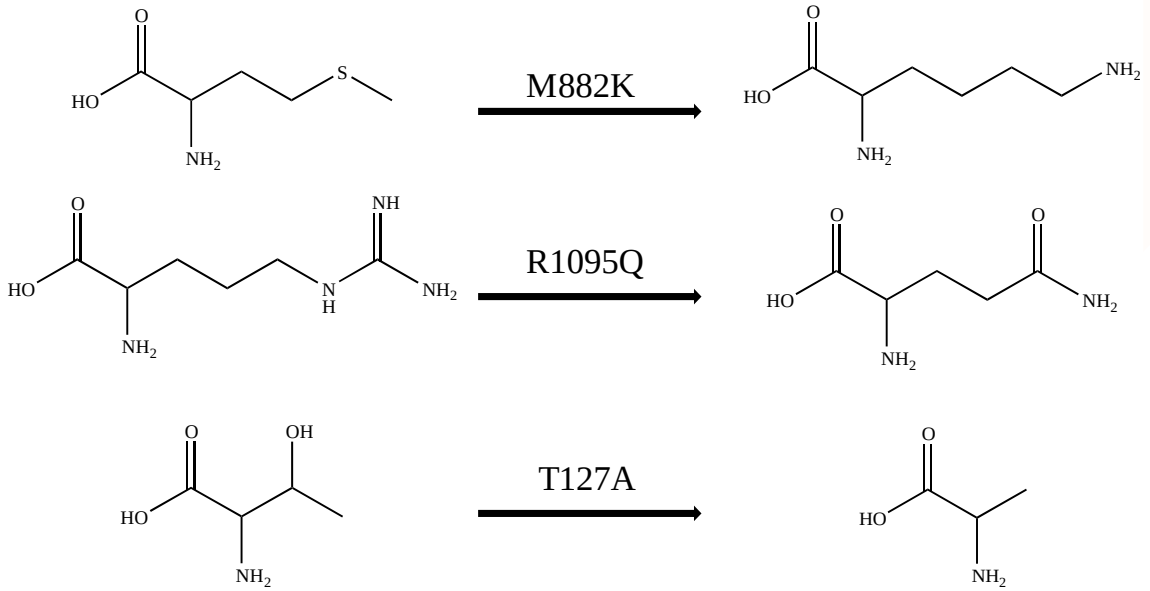
Role of Elongation Factor-like 1 (EFL1) in Shwachmann-Diamond Syndrome pathogenesis: insights from MD simulations

Pietro Delre, Domenico Alberga, Abril S, Nuria Sanchez, Orazio Nicolotti, Michele Saviano, Dritan Siliqui, Giuseppe Felice Mangiatordi*

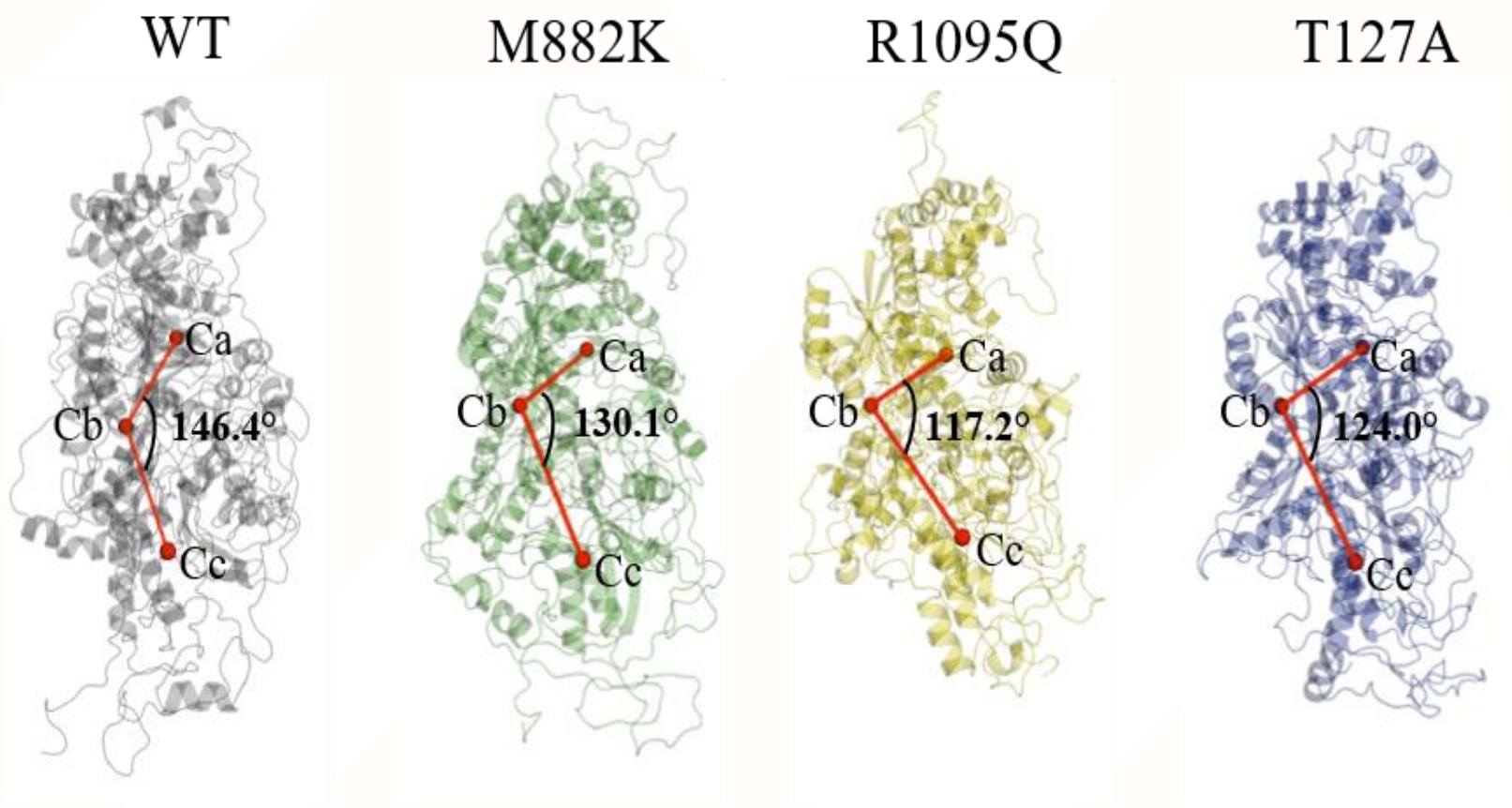


Stepensky P, Chacón-Flores M, Kim KH, et al Mutations in EFL1, an SBDS partner, are associated with infantile pancytopenia, exocrine pancreatic insufficiency and skeletal anomalies in aShwachman-Diamond like syndrome Journal of Medical Genetics 2017;54:558-566.



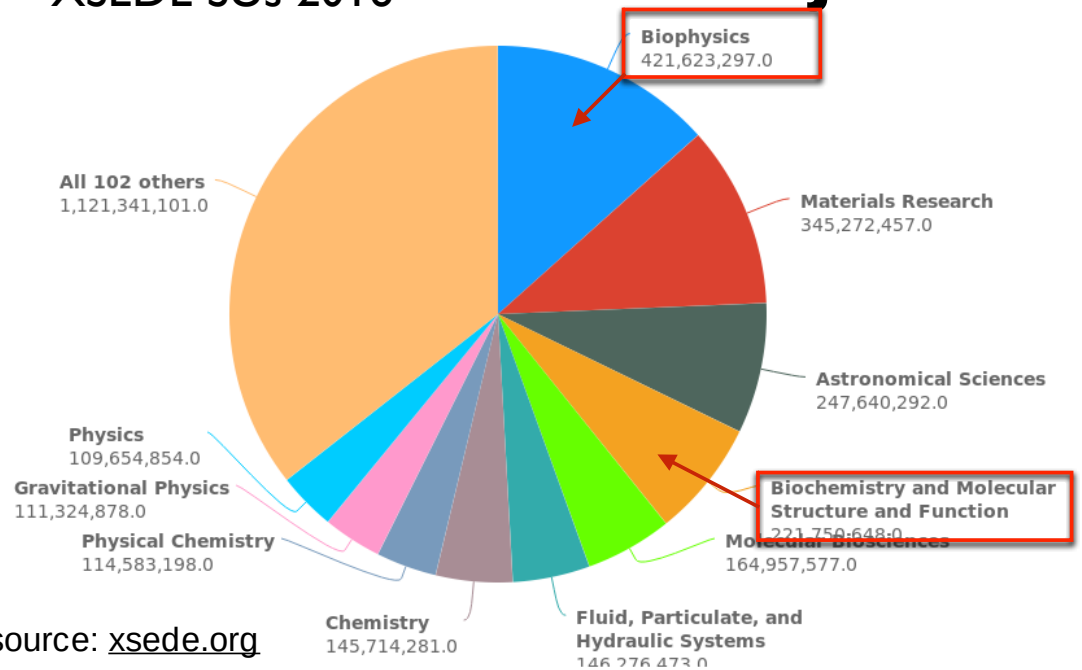


- Starting point: EFL1 WT (pdb code: 5ANC)
- mutant creation starting from the EFL1 WT
- Simulation of all systems (**about 105,000 atoms per system**)
- obtaining the trajectory for each system (time = **200ns**, frame = **10000**, Δt= **2fs**)
- Trajectories analysis


WT
Ca
Cb
146.4°
Cc
M882K
Ca
Cb
130.1°
Cc
R1095Q
Ca
Cb
117.2°
Cc
T127A
Ca
Cb
124.0°
Cc

MD trajectories can be big-ish

XSEDE SUs 2016



source: xsede.org

trajectory

$$(\mathbf{r}_1(t), \ldots, \mathbf{r}_N(t))$$

$$0 \leq t \leq t$$

	max (2016)	typical (2016)
atoms N	~10^7	~10^5
simulated time τ	~10 μs	0.1–1 μs
trajectory frames	~10^9	~10^5
trajectory size	< 10 TiB	150 GiB

see also: T. Cheatham and D. Roe. Computing in Science Engineering, 17:30–39, 2015.

MD limitations

- Limited size of the system under investigation due to the high computational time required

- MD simulations are computationally costly, they require hundreds or even millions of CPU years…..

- Supercomputers needed

Thank You

For Your Attention

Istituto Di Cristallografia CNR
pietro.delre@ic.cnr.it