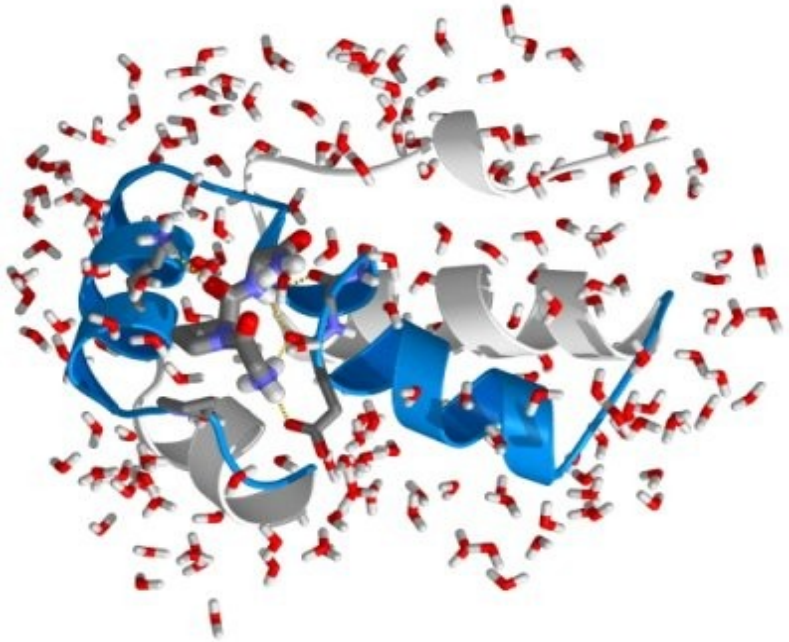
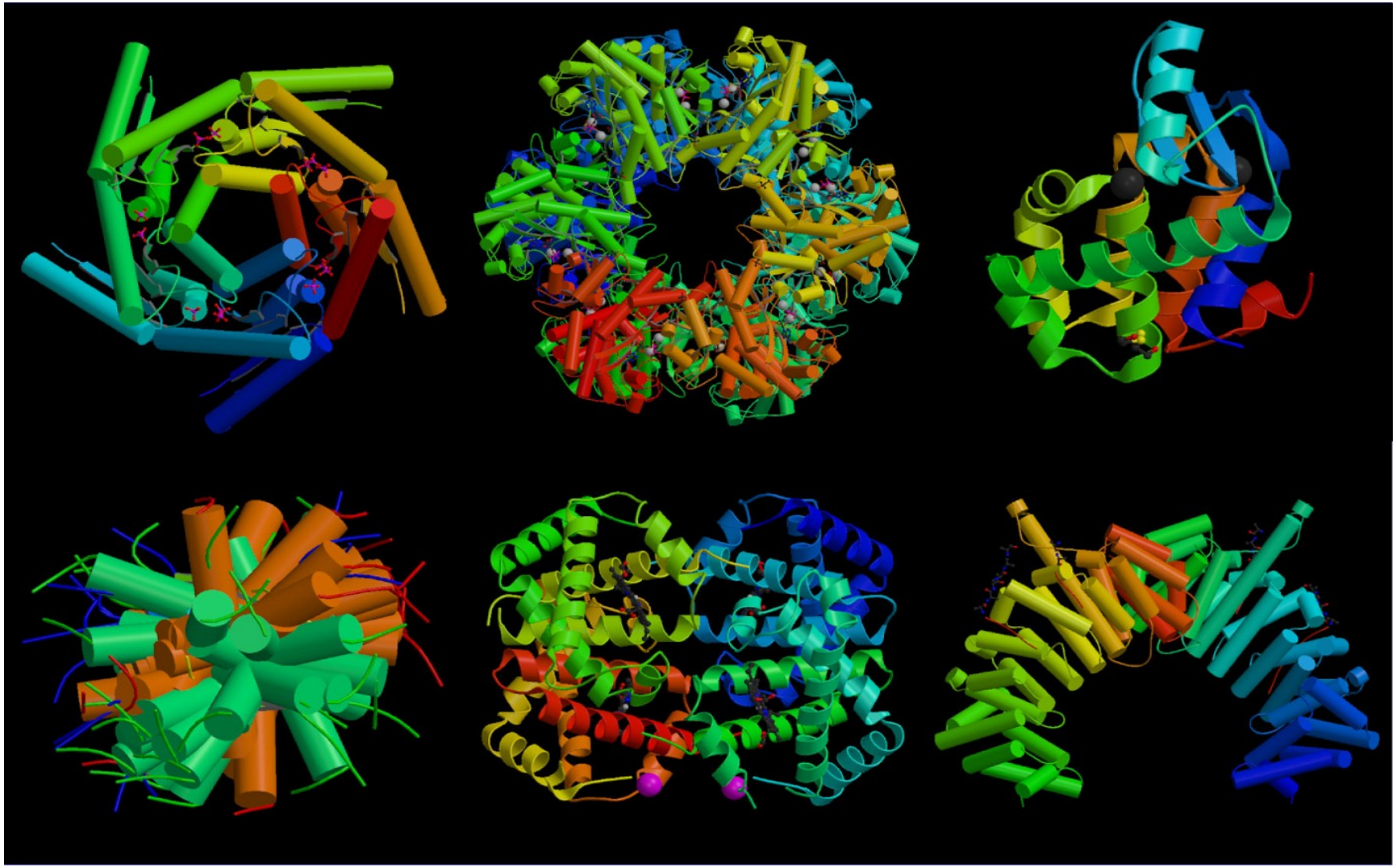
Molecular Dynamics - GROMACS

GROningen **MA**chine for **C**hemical **S**imulations

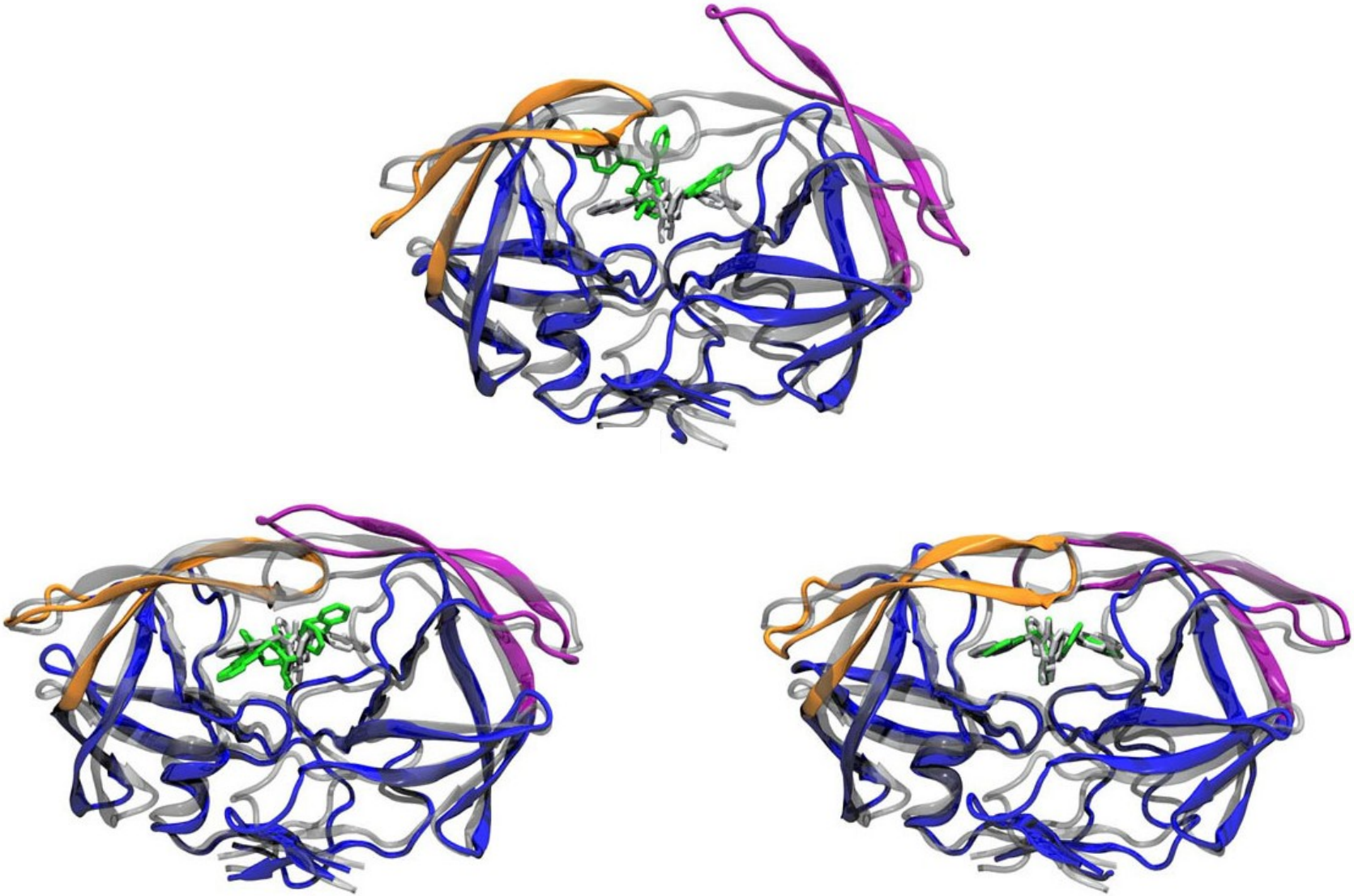


CS334
Seminar in Computational Biology
16 October 2008
Filip Jagodzinski

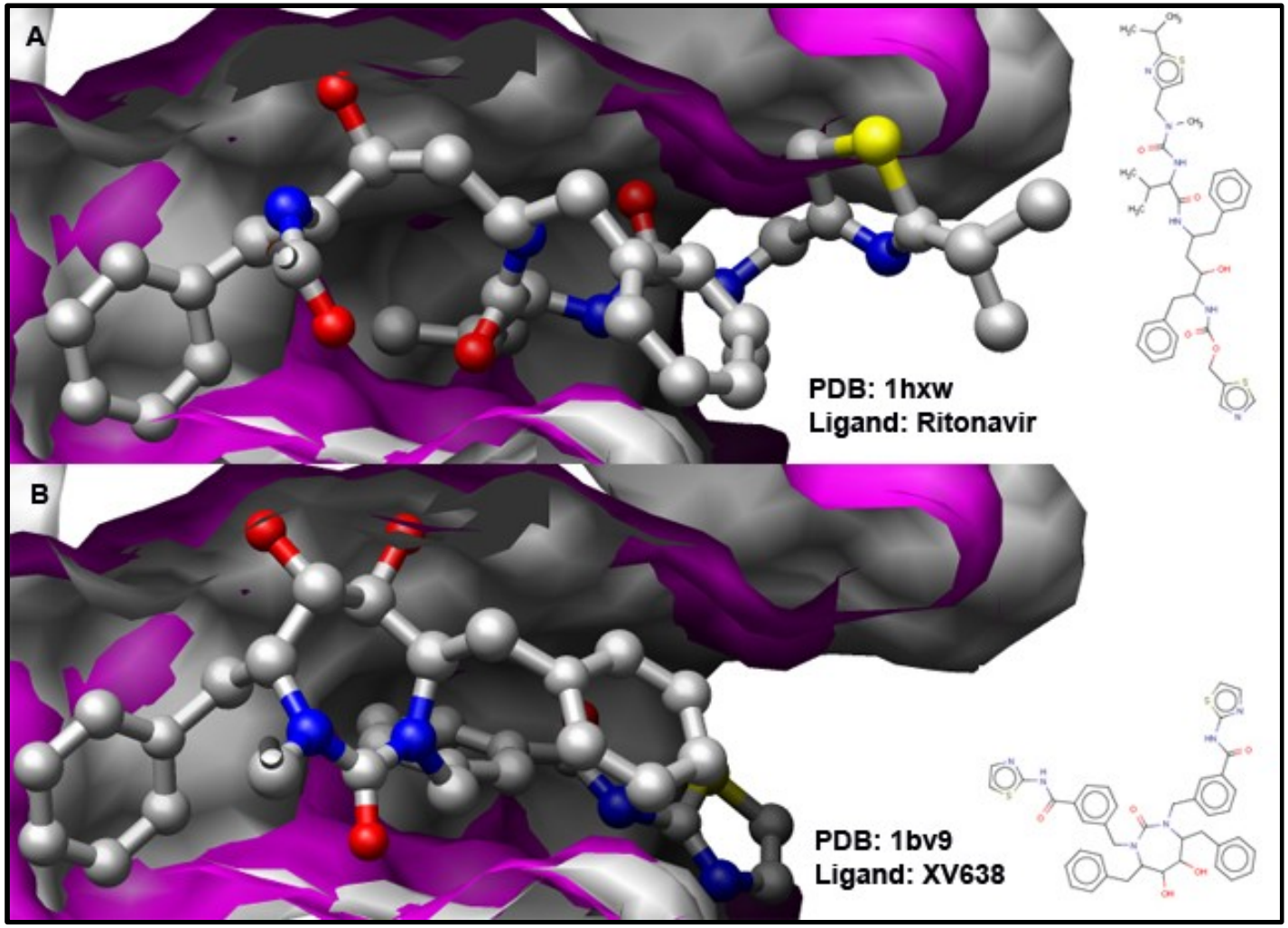
Introduction – Proteins



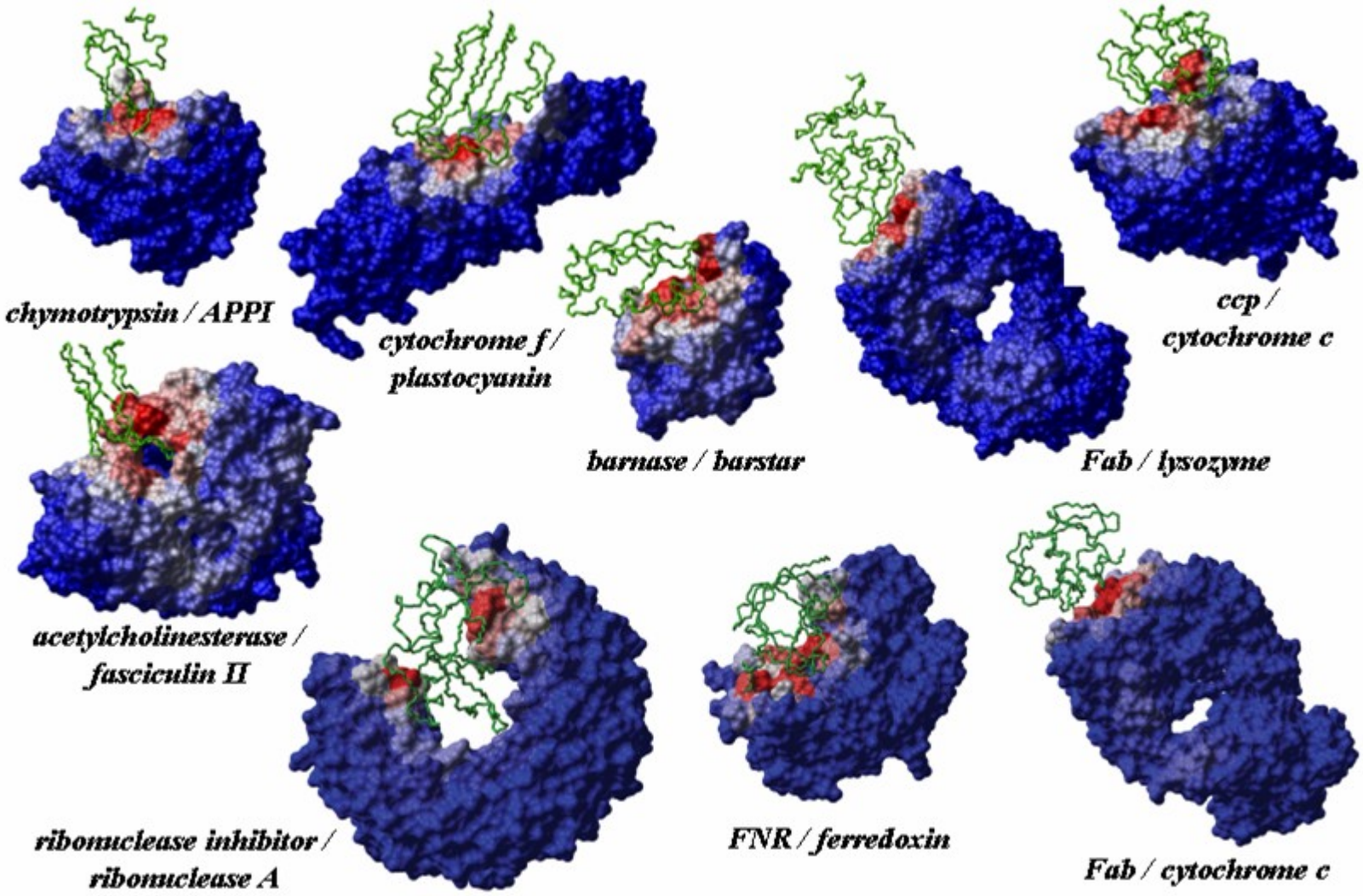
Introduction – Protein Function



Introduction – Medicine



Introduction – Predicting protein ligand interactions



Simulating Protein Motion – Timescales & Atom/Protein Motion

femtoseconds	10^{-15} seconds	Atomic Fluctuations
picoseconds	10^{-12} seconds	Side chain Motion
nanoseconds	10^{-9} seconds	Helix Motion
microseconds	10^{-6} seconds	Domain Motion
milliseconds	10^{-3} seconds	Protein Folding/Unfolding

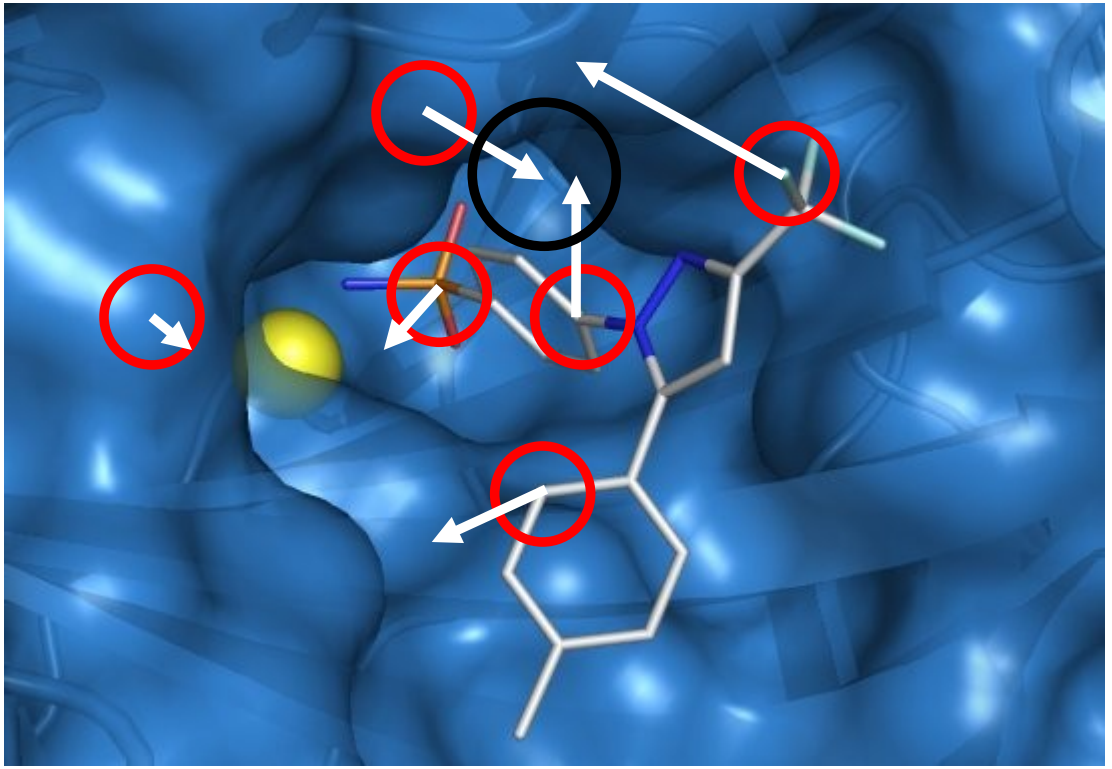
Simulating Protein Motion – A system of particles

What is needed to simulate the motion / interactions of proteins?

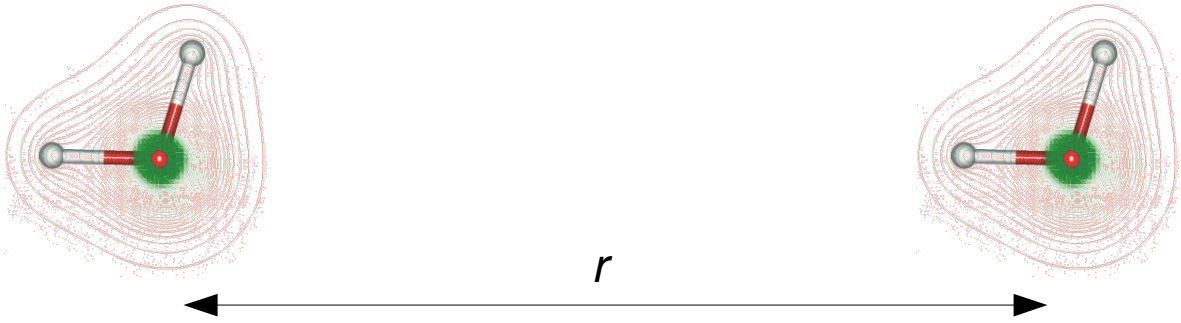
Particle **Positions**
(x, y, z)
(-32.3, 12.3, 6.4)

Particle **Velocities**
(x, y, z)
(3.3, 4.3, -0.4)

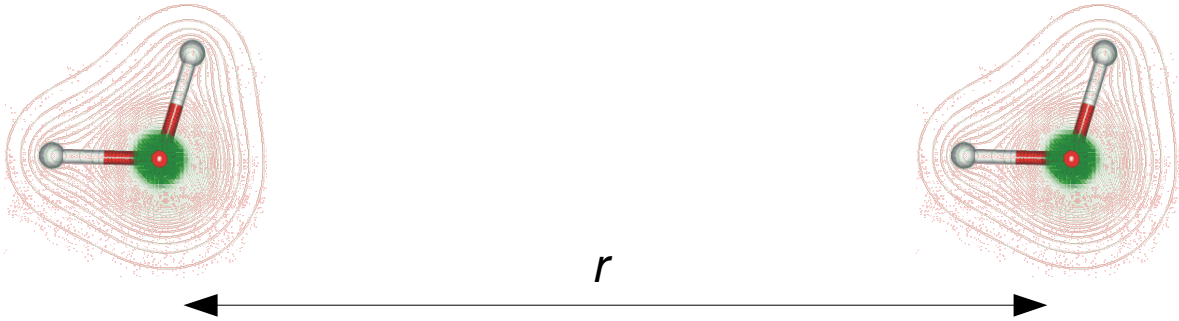
Particle **Interactions**
$\mathbf{F}_{ij}$

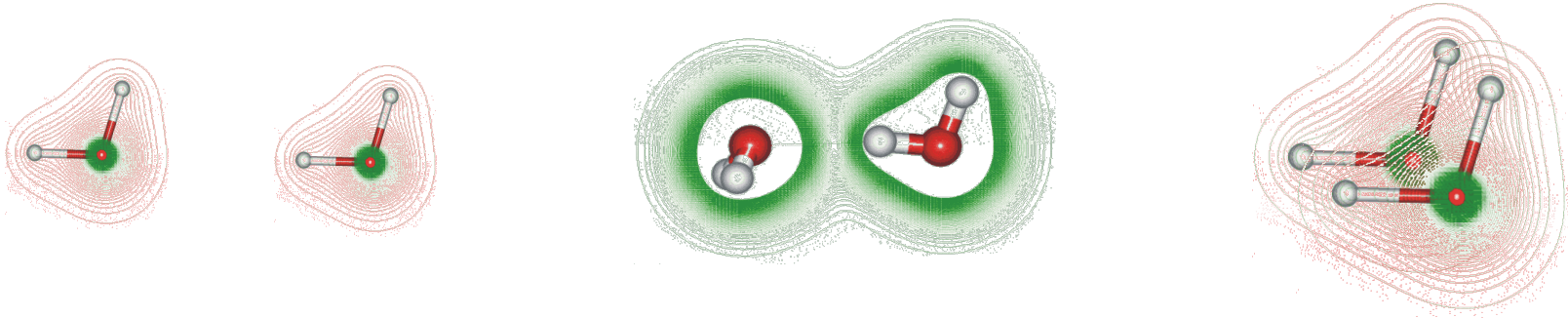


Simulating Protein Motion – Atom interactions – proximity



Simulating Protein Motion – Atom interactions – proximity



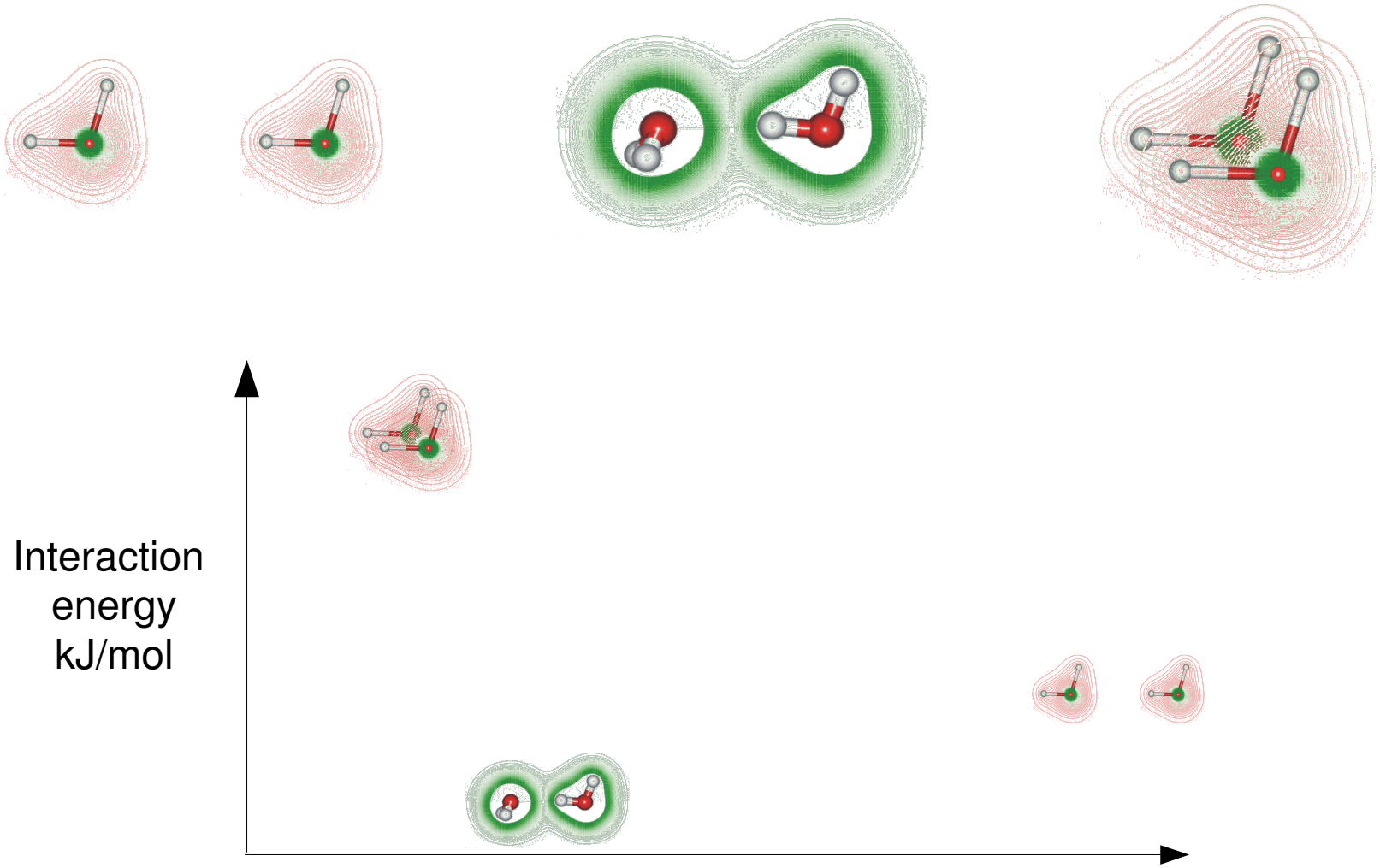


Separation = far
No Interaction

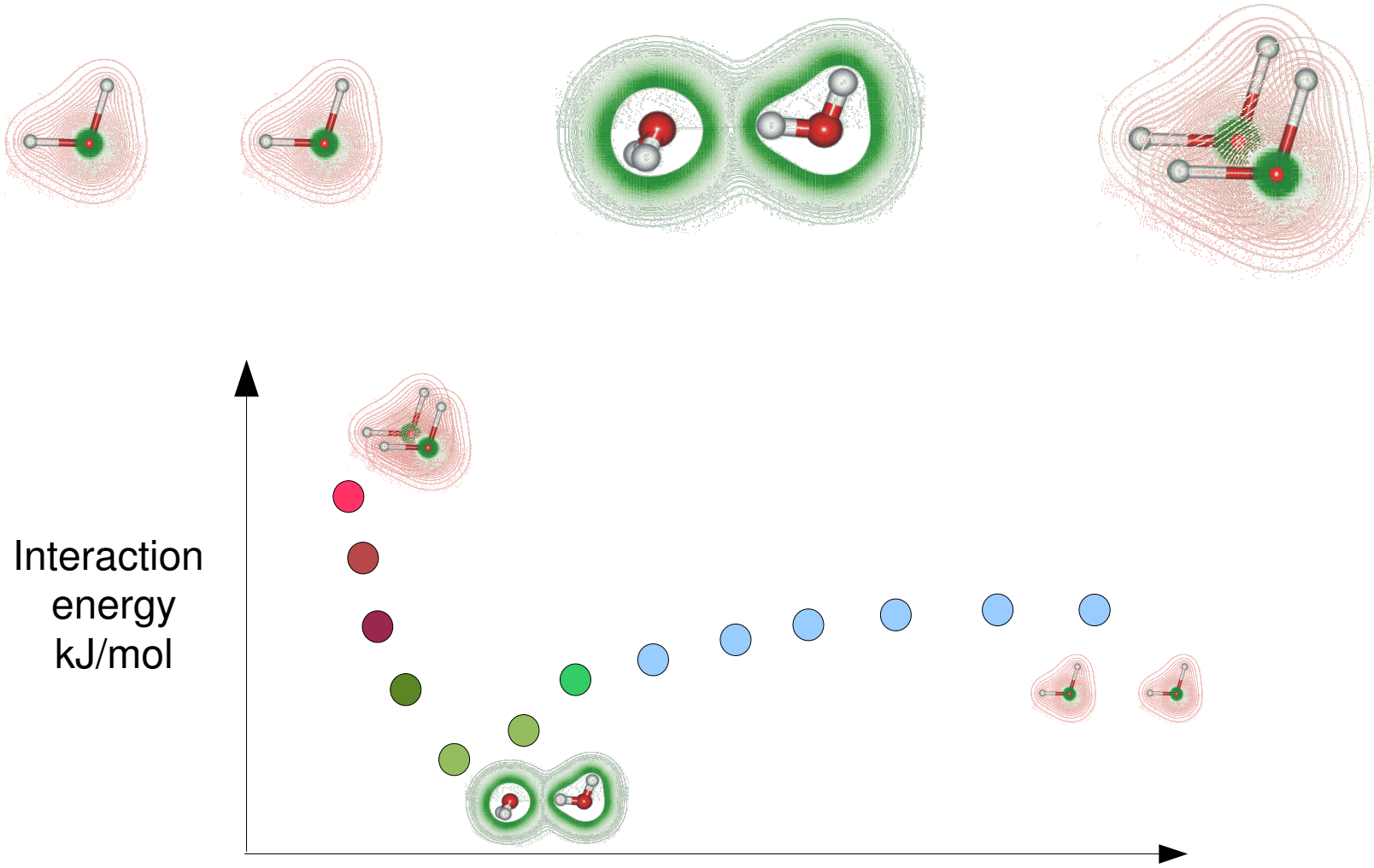
Separation = “just right”
“ideal” Interaction

Separation = “very small”
“repulsive” Interaction

Simulating Protein Motion – Atom interactions – proximity

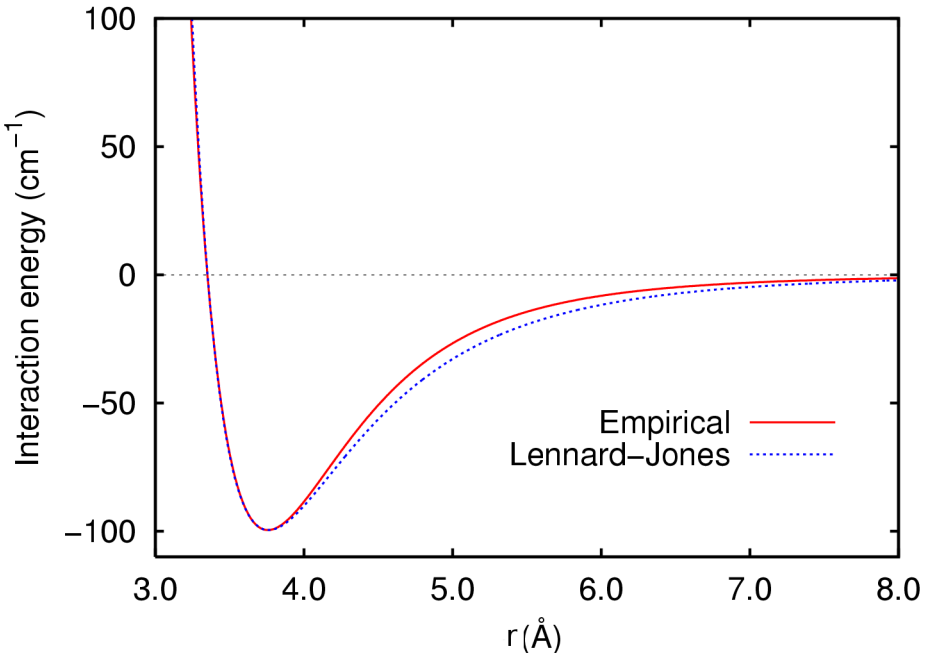


Simulating Protein Motion – Atom interactions – proximity



Simulating Protein Motion – Atom Interactions – Lennard-Jones Approximation

A pair of neutral atoms are subject to two distinct forces between them: an **attractive force** at long ranges (van der Waals force), and a **repulsive force** at short ranges (Pauli repulsion force). The **Lennard-Jones potential** is an approximation of these two forces:



$$\mathbf{V}(r) = 4\epsilon\left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6}\right]$$

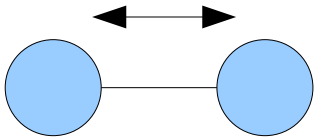
σ = finite distance at which interparticle potential is zero
ε = depth of potential well
r = separation distance

the **12** term describes repulsion
the **6** term describes attraction

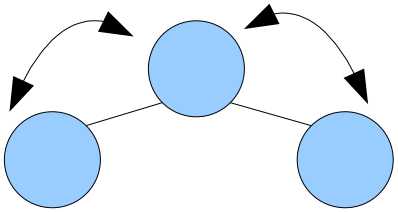
For water, **σ** = 0.316 nm, **ε** = 0.6501 kJ/mol
(1nm = 10Angstroms)

Simulating Protein Motion – Atom Interactions – Bonded and non-bonded

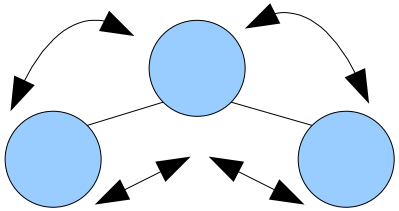
The Lennard-Jones approximation models only one type of interaction, the **non-bonded interaction**, but **bonded interactions** exist, too, for which there are also mathematical models:



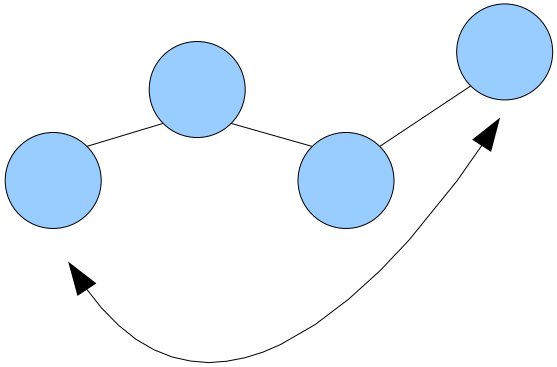
Bond stretching



Bond Bending



Bond Bending
and Stretching



Torsion Strain

Simulating Protein Motion – Atom Interactions – Potential Energy

The total potential energy of two molecules is the sum of the energy, **E**, of the **bonded** interactions and the **non-bonded** interactions between them. The Potential Energy, **V**, of a system is a function of the positions, **r**, of the individual atoms and the energies between them.

$$\mathbf{V}(\mathbf{r}) = E_{\text{bonded intractions}} + E_{\text{non-bonded interactions}}$$

$$E_{\text{bonded interactions}} = E_{\text{bond stretching}} + E_{\text{bond bending}} + E_{\text{torsion}} \ldots$$

$$E_{\text{non-bonded interactions}} = E_{\text{van der Waals}} + E_{\text{Pauli Repulsion}}$$

Lennard-Jones Potential

Simulating Protein Motion – Physics Review

How do we relate the interaction energies of the atoms and molecules to the motion of individual atoms?

Newton's **Second Law:**

$$\mathbf{F} = ma$$

Velocity: Instantaneous change in position

$$\frac{d\mathbf{r}_i(t)}{dt} = \mathbf{v}_i(t)$$

Acceleration: Instantaneous change in velocity

$$\frac{d\mathbf{v}_i(t)}{dt} = \frac{\mathbf{F}_i(t)}{m_i}$$

Simulating Protein Motion – Molecular Dynamics – Leap Frog

The force $\mathbf{F}_i$ exerted on atom i by the other atoms in the system is given by the negative gradient of the potential energy function **V** which in turn depends on the coordinates of all N atoms in the system

$$\mathbf{F}_i(t) = \frac{-\partial \mathbf{V}(\mathbf{r}_1(t), \mathbf{r}_2(t), ..., \mathbf{r}_N(t))}{\partial \mathbf{r}_i(t)}$$

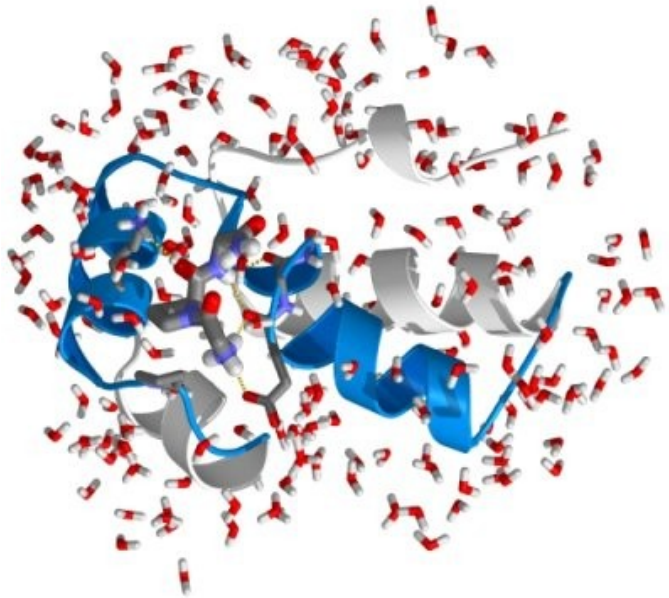
The Velocity and Positions of the particles in a system can be determined once the potential Energy of the system is known:

$$\mathbf{v}_i(t + \delta t/2) = \mathbf{v}_i(t - \delta t/2) + \frac{\mathbf{F}_i(t)}{m_i}\delta t$$

$$\mathbf{r}_i(t + \delta t) = \mathbf{r}_i(t) + \mathbf{v}_i(t + \delta t/2)\delta t$$

The time step, *t* , is generally set to 2 femtoseconds so that any affect due to the change in the position of a single atom can be detected.

Using GROMACS

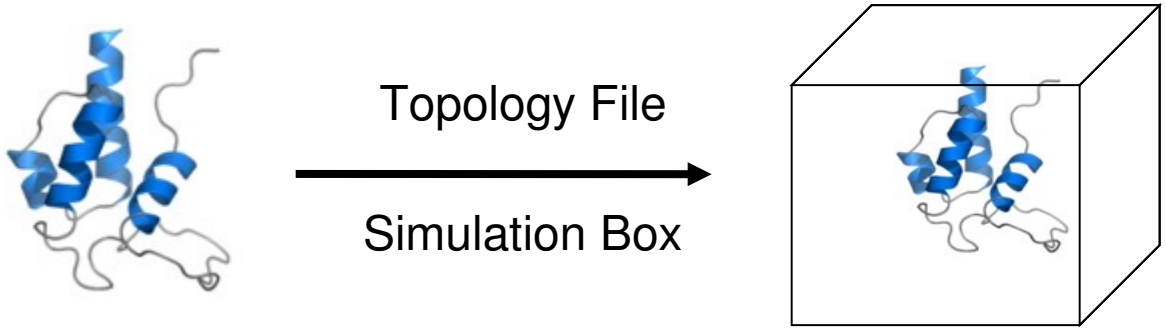


gromacs.org

The goal of molecular modeling is to perform computer simulations, or to do chemical “experiments”, with a computer rather than a laboratory bench.

GROMACS –Generate Simulation Box

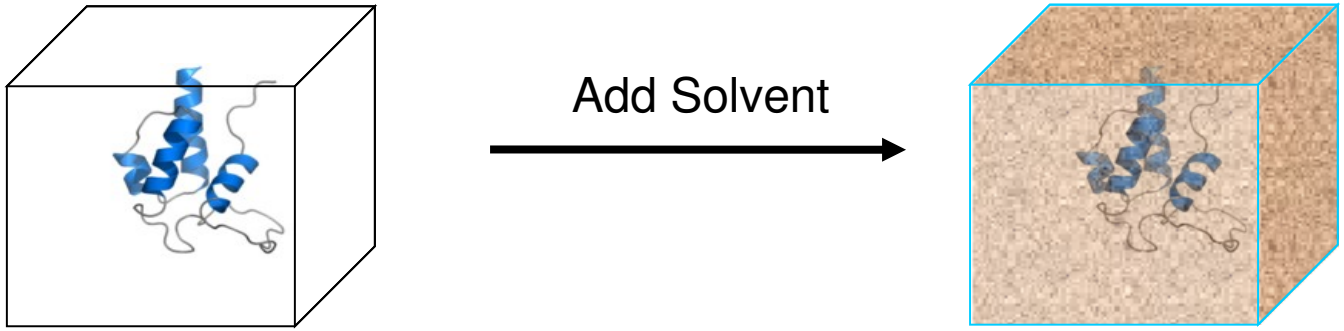
We need a simulation environment where the energy calculations will be performed ... let us choose a simulation box.



0 | pdb2gmx -f protein.pdb -o protein.gro -p protein.top
editconf -f protein.gro -o protein.gro -d 0.5

GROMACS – Solvate the simulation box

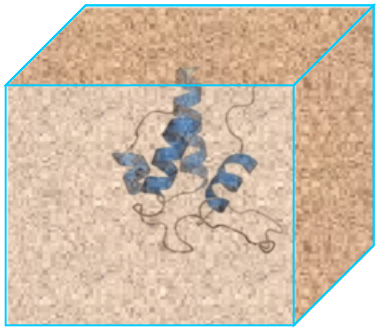
We need to solvate the simulation box to mimic real cellular conditions.



genbox -cp protein.gro -cs -o protein_b4em.gro -p protein.top

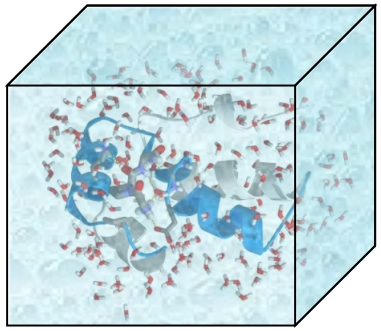
GROMACS – Energy Minimize simulation Box

We have placed water molecules indiscriminately into the simulation box, and unfavorable clashes may exist. We need to remove these clashes to mimic the “real” world.



Energy Minimize Water

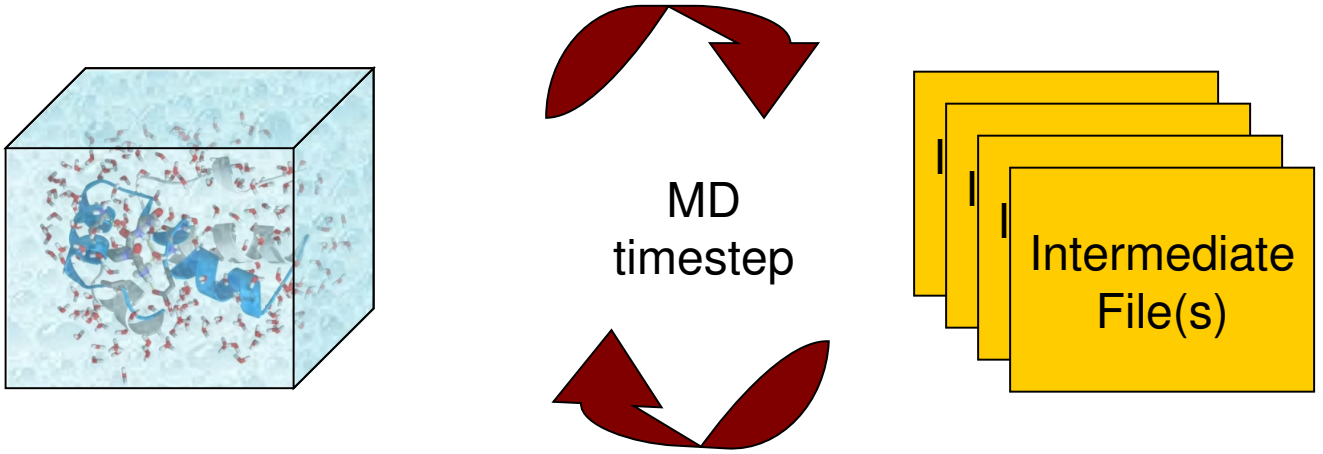
Position Restrain MD System



```
grompp -f em -c protein_b4em -p protein -o protein_em
mdrun -nice 4 -s protein_em -o protein_em -c protein_b4pr -v
grompp -f pr -c protein_b4pr -r protein_b4pr -p protein -o protein_pr
mdrun -nice 4 -s protein_pr -o protein_pr -c protein_b4md -v
```

GROMACS – Perform the MD Integration Steps

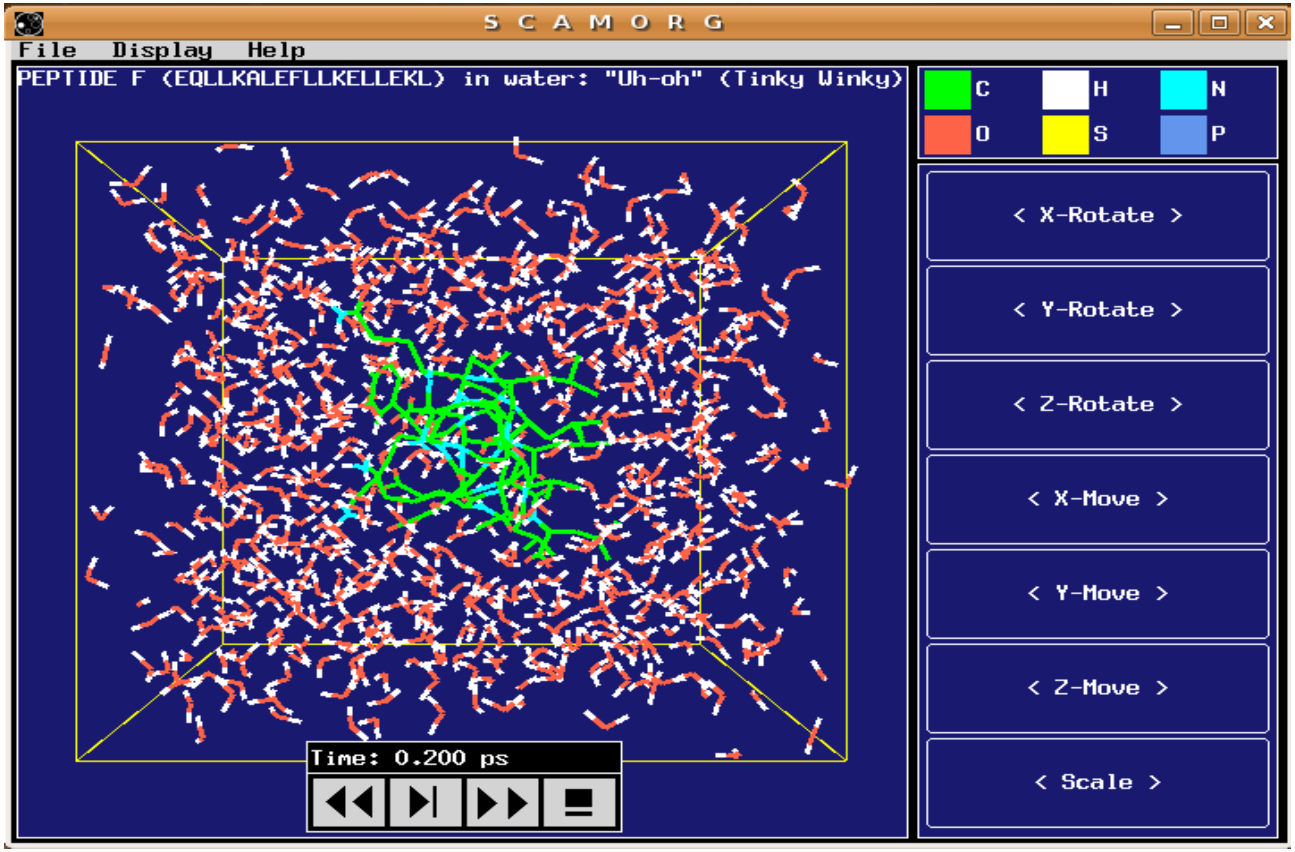
Now we are ready to perform the MD simulation




```
grompp -f md -c protein_b4md -p protein -o protein_md
mdrun -nice 4 -s protein_md -o protein_md -c protein_after_md -v
```

GROMACS – Visualize the results

ngmx -f protein_md -s protein_md



GROMACS – Discussion – MD Drawbacks, advantages, etc.

Possible questions/Comments:

- What are the advantages/disadvantages of MD?
- Are there “other” methods that can be used to help find cures for important diseases?
- What if the crystal structure of a protein is not available. What then?
- How can we “simply” the system – rigidity?
- Does simulating the motion of a single protein in isolation have any relevance to how that protein actually functions within the actual cellular and extracellular environment, where there are many interactions with other cells and other proteins.
- Lead discussion into using rigidity.