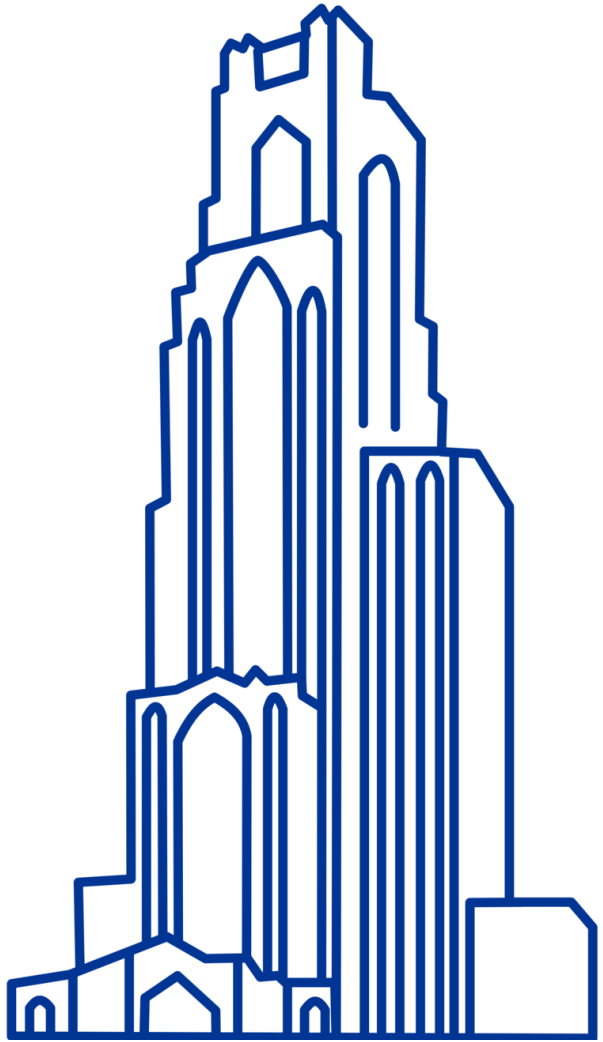


Computational Biology
(BIOSC 1540)

Lecture 10A

Atomistic insight

Foundations

Mar 18, 2025

Announcements

Assignments

- P02B is due Mar 28
- P02C is due Mar 28
- P03A is due Apr 4

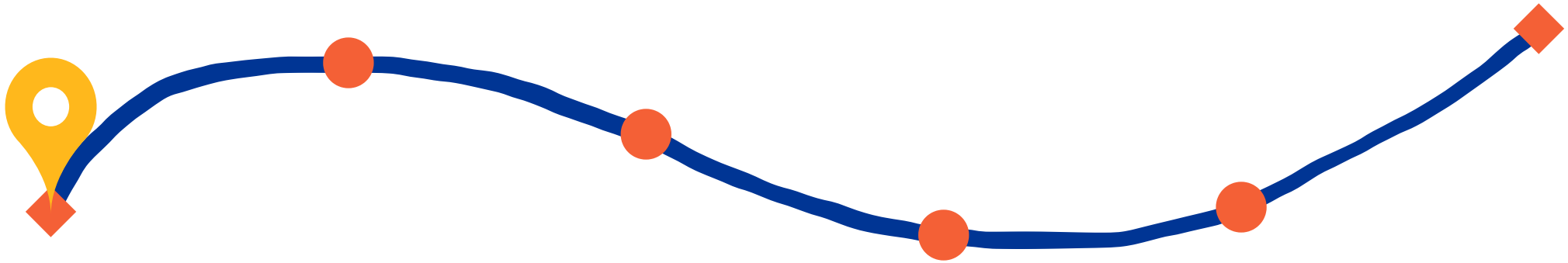
Quizzes

- Quiz 03 is today
- Quiz 04 (our last one) will be on Apr 8
- Remember that your lowest quiz is dropped

Final exam

- The final exam is on **Monday, Apr 28, at 4:00 pm in 244 Cathedral of Learning**
- The exam is cumulative and optional
- Will replace any quiz lower than your final exam grade

After today, you should have a better understanding of



Quiz 03

Please put away all materials as we distribute the quiz

Sit with an empty seat between you and your neighbors for the quiz

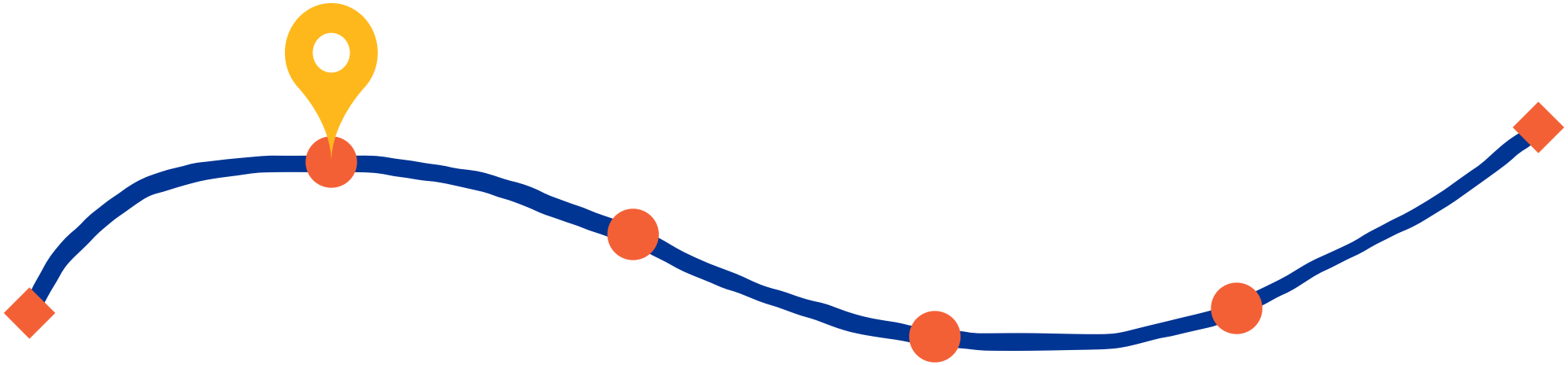
Fill out the cover page, and do not start yet

Quiz ends around **9:51 am**

https://www.clockfaceonline.co.uk/clocks/digital/

When you are finished, please hold on to your quiz and feel free to doodle or write anything on the last page

After today, you should have a better understanding of



Molecular ensembles and their relevance

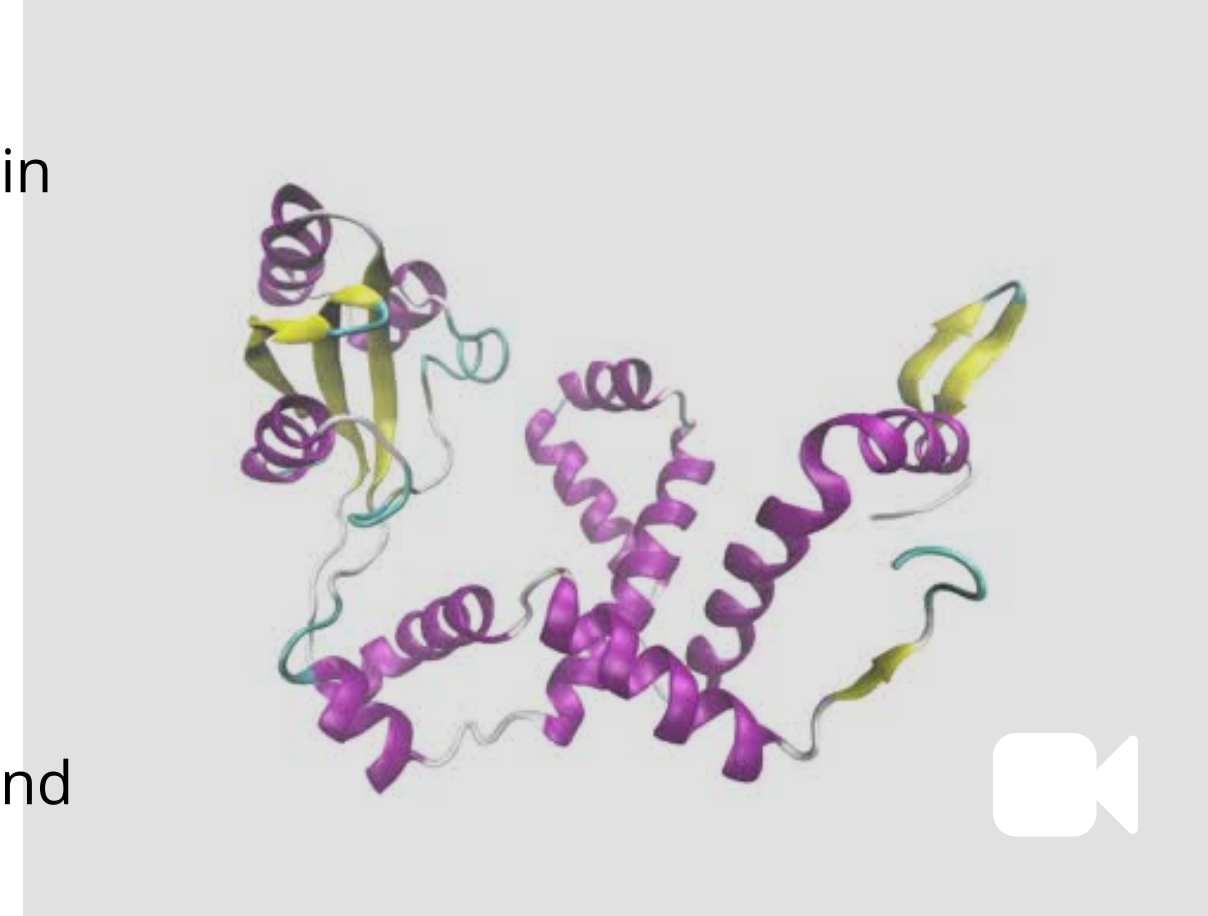
Macrostates

Physics at the molecular level is statistical

Number of Particles: Biological systems contain billions of atoms interacting simultaneously

Thermal Motion: Atoms and molecules are in constant motion due to thermal energy

Uncertainty and Variability: Exact positions and velocities of particles are inherently uncertain



Observable properties are averages of atomistic behaviors

Atomistic systems are **stochastic**, measurable properties are computed as averages

Microscopic level: Individual atoms and molecules

Macroscopic level: Bulk properties from collective behavior

Statistical mechanics uses statistical methods to relate microscopic properties to macroscopic observables

A Macrostate Describes the Overall Condition of a System

A macrostate is defined by macroscopic variables such as temperature, pressure, volume, and number of particles.

Example: Methanol and water

Temperature: 25 C

Pressure: 1.01325 bar

Volume: 100 mL

Composition: 70% methanol and 30% water by mass

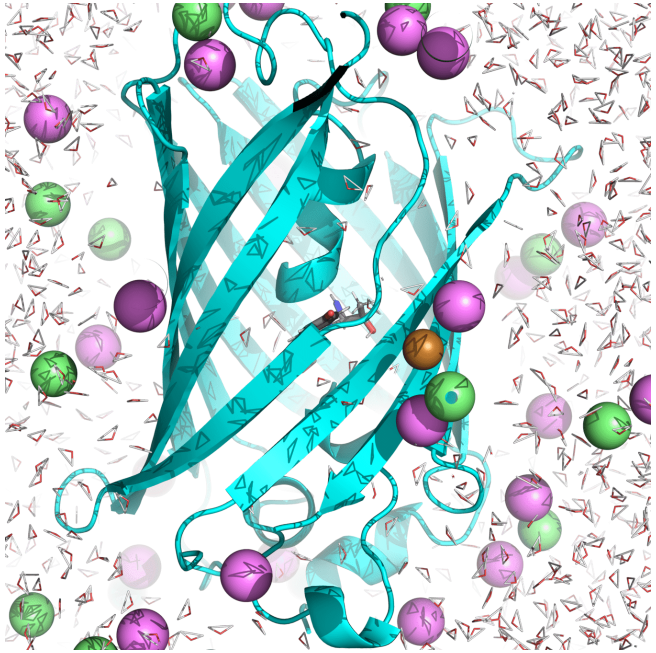
It provides a **coarse-grained system description**, ignoring the specific details of individual particles.

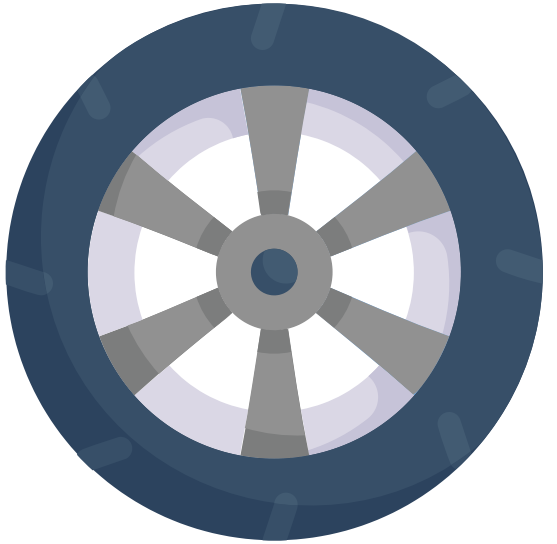
Changing any one of these values changes the macrostate

Macrostates Capture What We Can Measure

We cannot measure each molecule's exact position and velocity in a system.

Instead, we use macroscopic variables like density, energy, and composition, which summarize the system’s overall state.





Example: The pressure in a tire depends on the average behavior of gas molecules, not the exact motion of each one.

Changing a Macrostate Can Change a System's Properties

When a macrostate changes, the system may undergo phase transitions or shifts in observable properties.

Some macrostates are stable, while others are metastable (temporarily stable before changing).

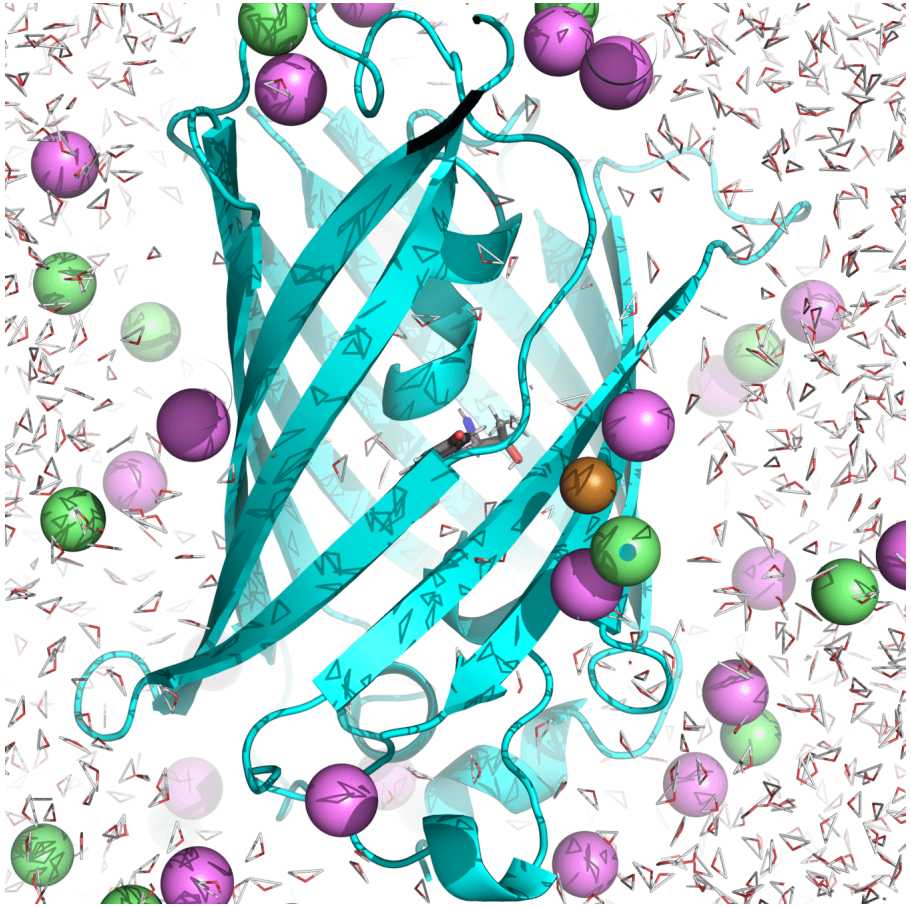
Example: Supercooled water can remain liquid below 0°C, but a small disturbance changes its macrostate to solid ice.

Molecular example: Protein simulations

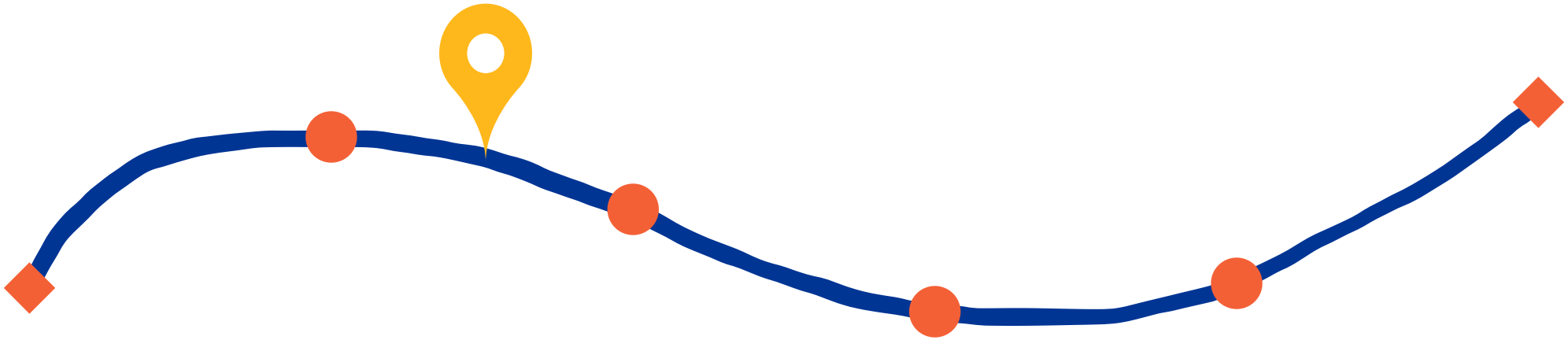
In structural biology, our molecular ensembles are normally defined with temperature, pressure, and chemical species

Chemical species are our proteins, solvent molecules, ions, etc.

Environmental factors such as pH will influence our chemical species



After today, you should have a better understanding of



Molecular ensembles and their relevance

Microstates

Molecular properties emerge from atomic interactions, but we cannot measure individual atoms directly

Biological and chemical properties arise from **atomic-scale interactions** like hydrogen bonding, electrostatic forces, and conformational changes.

Experimental techniques measure **averages** over many molecules, but they do not provide direct access to **individual atomic motions**.

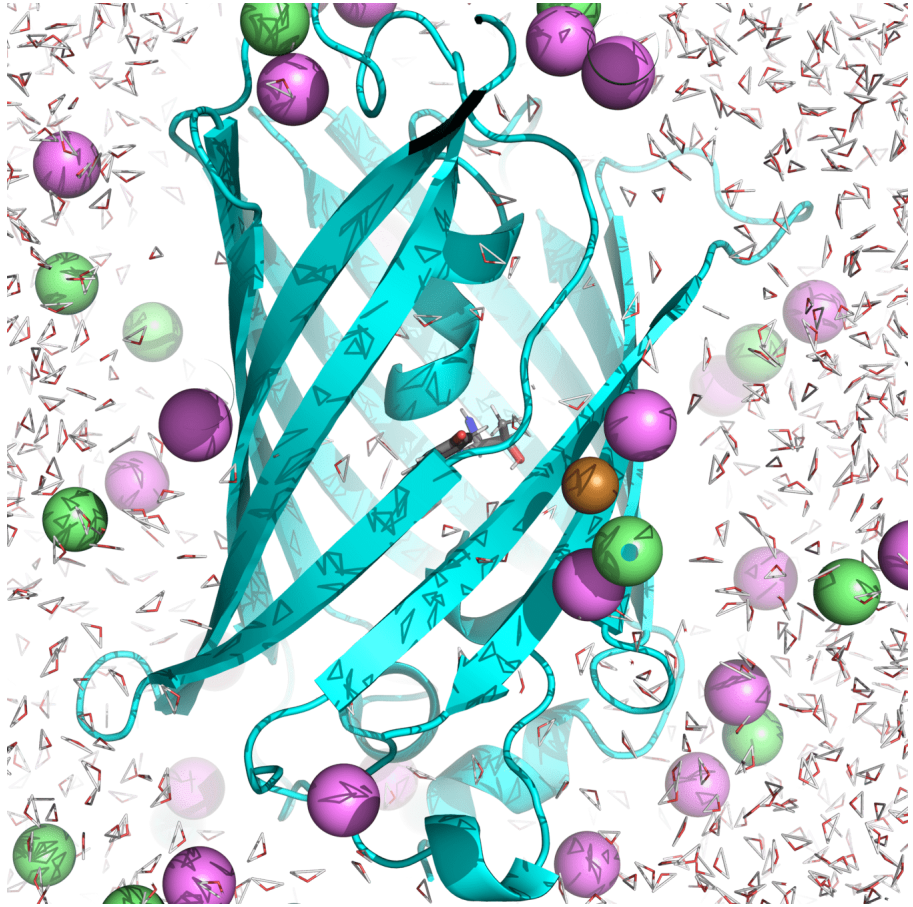
Computational methods, such as molecular simulations, allow us to track how **atoms move and interact over time**.

Ensembles Provide a Statistical View of Molecular Behavior

A system at a given **temperature, pressure, and volume** can exist in many possible microscopic configurations.

Each configuration (microstate) represents a unique arrangement of atomic positions and velocities.

By sampling an ensemble of microstates, we can determine **probability distributions** of molecular properties.

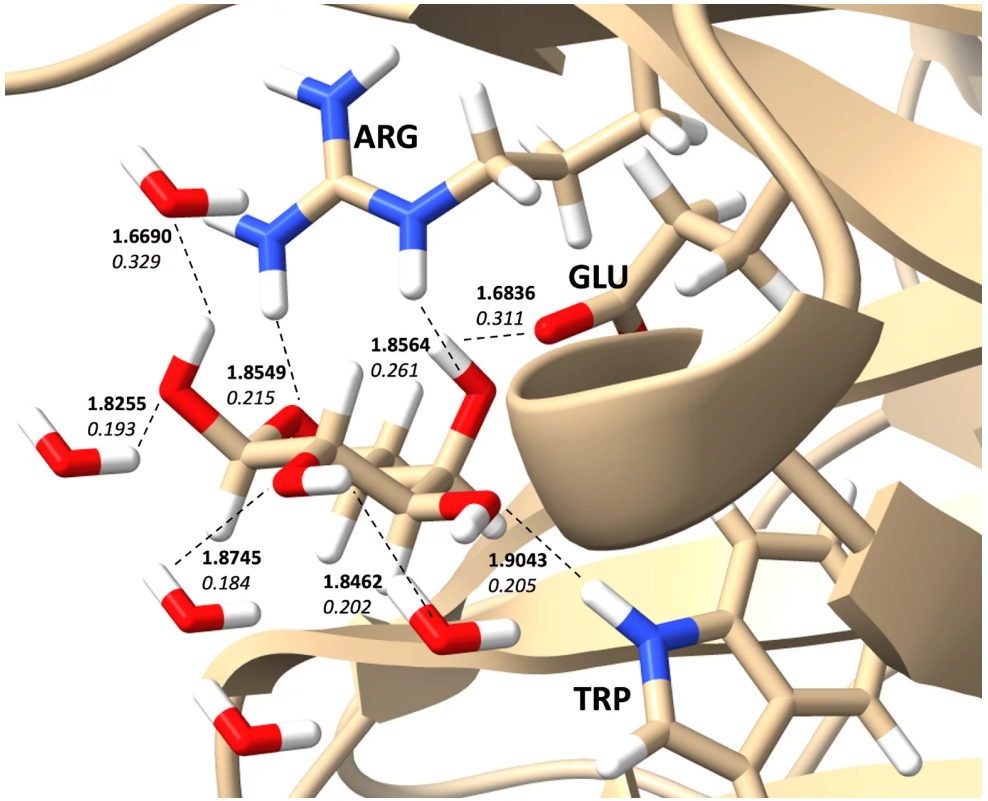


A microstate is a single arrangement of atoms and their velocities within a macrostate

Every microstate is one specific realization of atomic positions and momenta.

The system constantly moves between different microstates due to **thermal motion** and molecular interactions.

Example: A protein-ligand complex exists in many conformations—some tightly bound, others loosely interacting.

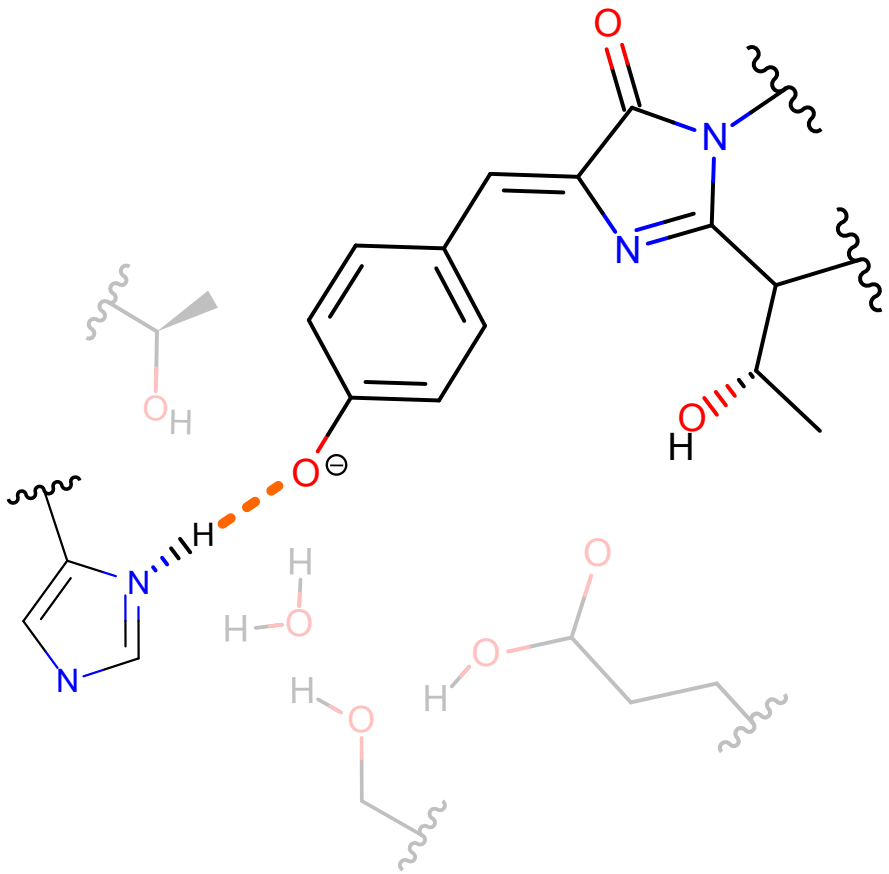


We can use ensembles to study the strength and length of a His148 hydrogen bond to the anionic chromophore

His148 stabilizes the anionic chromophore through **hydrogen bonding**, which influences fluorescence properties.

The **hydrogen bond length** fluctuates over time as atoms move between different microstates.

By sampling an **ensemble of molecular simulations**, we determine the **mean hydrogen bond length and energy**.

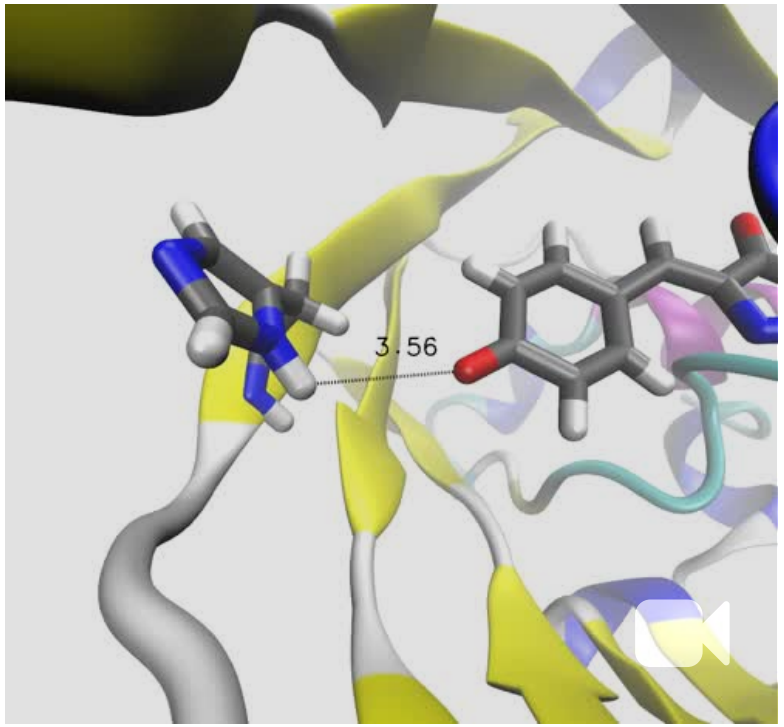


To measure the hydrogen bond, we must average over many microstates to get a meaningful result

Our macrostate: roGFP2 in water, with 150 mM NaCl at 300 K and 1 atm

A single microstate may show **a short or long bond length**, but this does not represent the overall behavior.

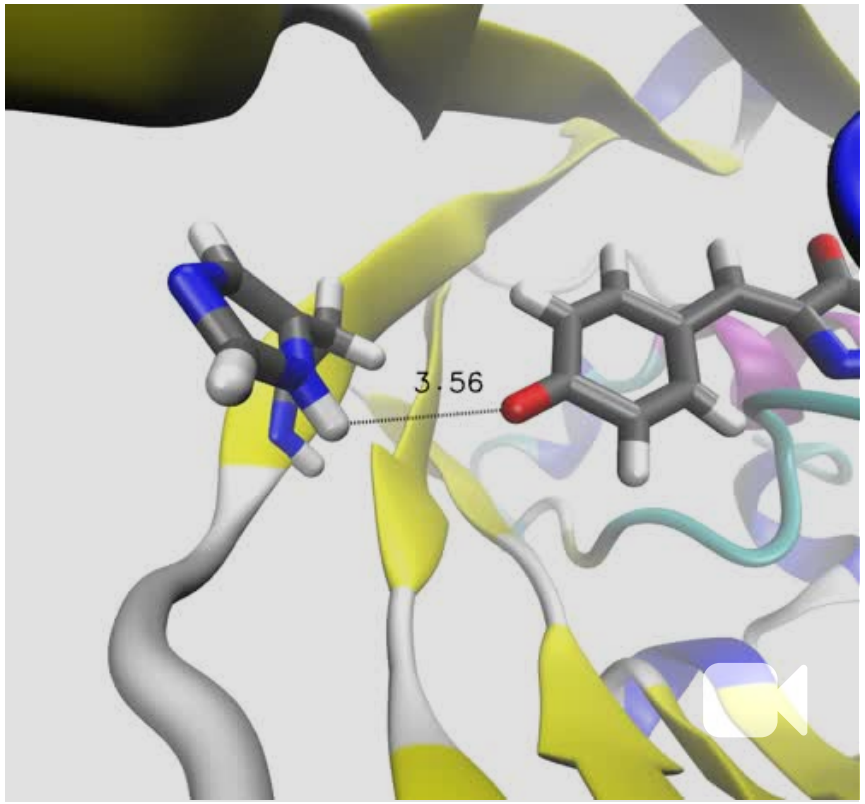
A properly sampled ensemble gives the **average bond length** and the **distribution of bond fluctuations**.

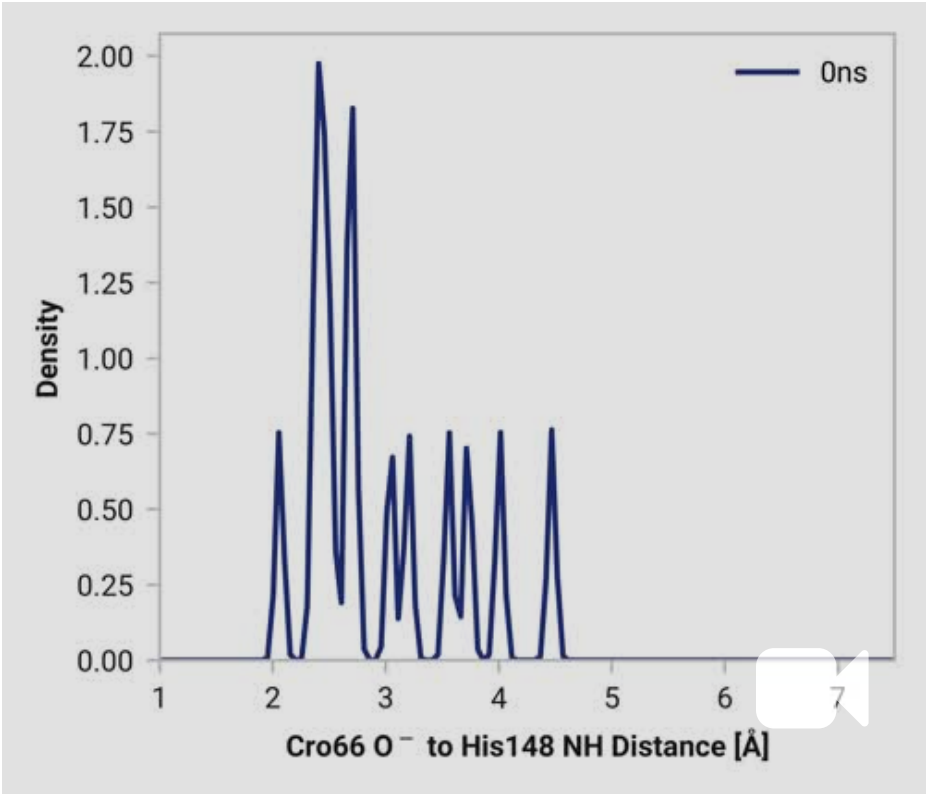


Here is the MD trajectory with a mean of 3.155 Å

Accurate results require statistical sampling across many microstates

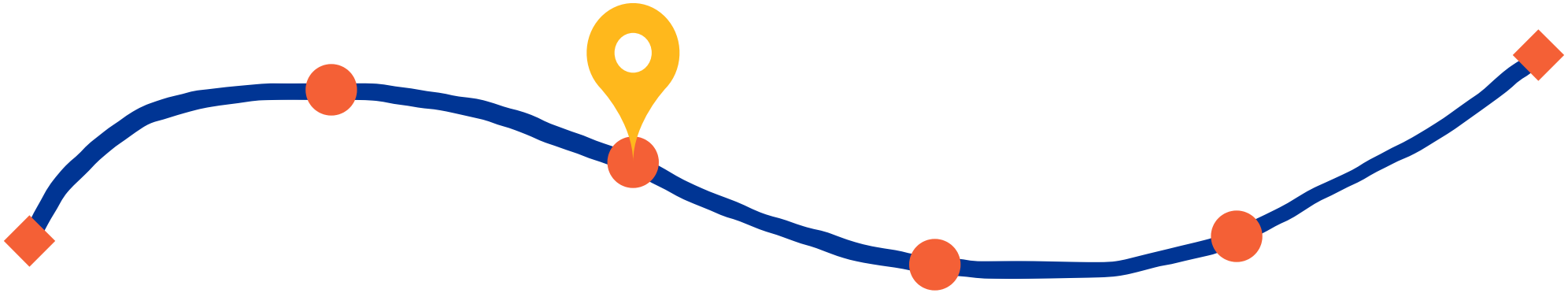
Observing one molecular snapshot is like looking at **one frame of a movie**—it does not capture the full dynamics.





By simulating thousands of microstates, we capture how the **hydrogen bond length varies over time**.

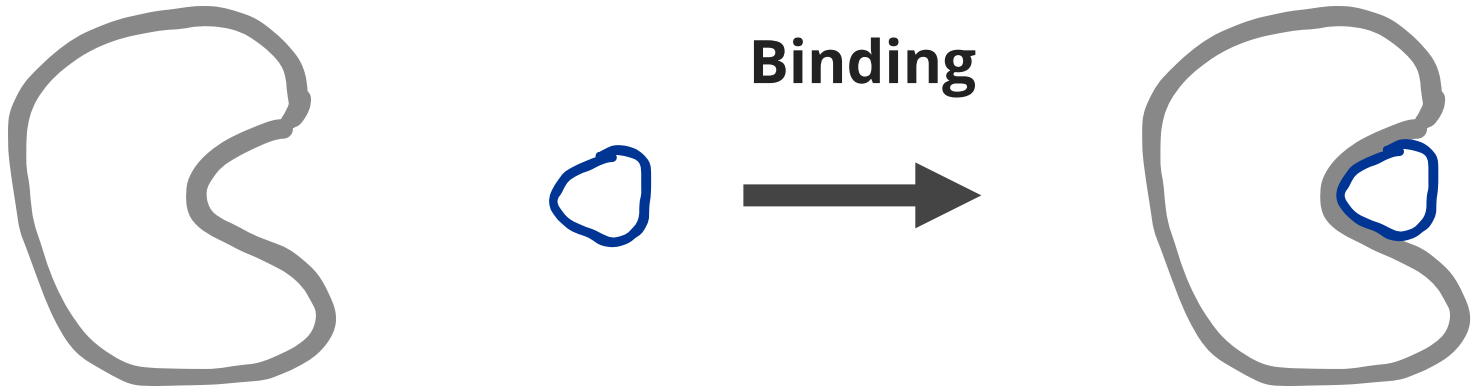
After today, you should have a better understanding of



Thermodynamics of binding

Selective binding to a protein is governed by thermodynamics (and kinetics)

Binding occurs when a compound/ligand interacts specifically with a protein



Protein **Ligand** **Protein-ligand**

We can model this as a reversible protein-ligand binding

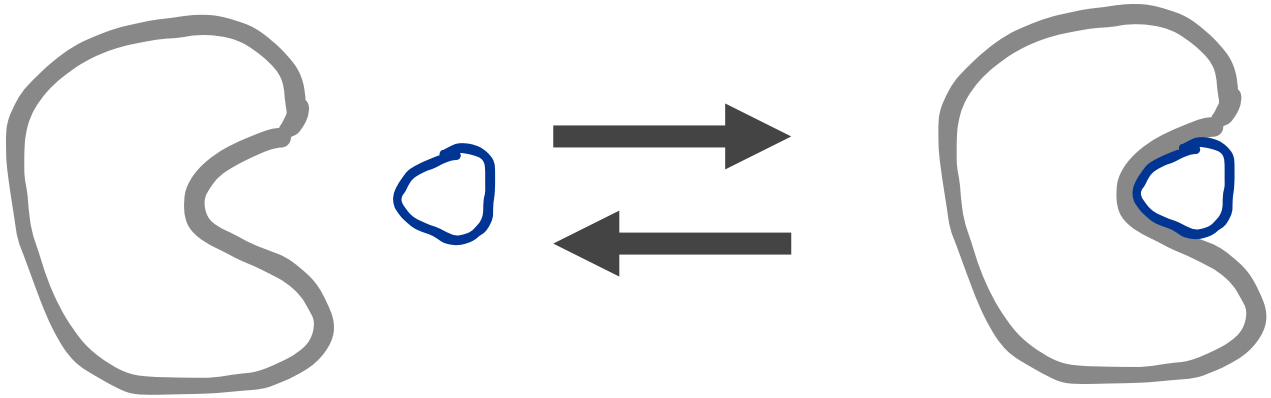
$$P + L \leftrightharpoons PL$$

Binding affinity is determined by the Gibbs free energy change

The change in free energy when a ligand binds to a protein

$$\Delta G_{bind} = G_{PL} - G_P - G_L$$

Determines binding process spontaneity



Gibbs free energy combines enthalpy and entropy

$$\Delta G_{bind} = \Delta H_{bind} - T\Delta S_{bind}$$

Enthalpy

ΔH_{bind}

Accounts for energetic interactions

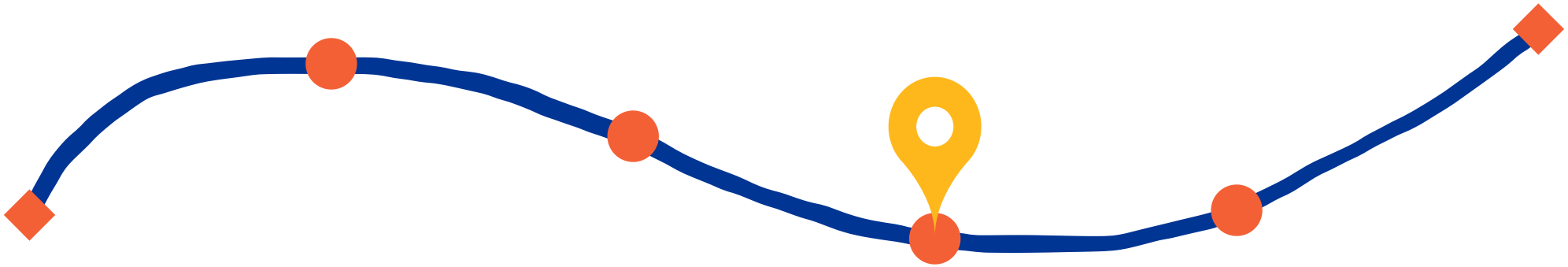
Entropy

ΔS_{bind}

How much conformational flexibility changes

Note: Simulations capture free energy directly instead of treating enthalpy and entropy separately

After today, you should have a better understanding of



Enthalpic contributions to binding

Enthalpy accounts for noncovalent interactions

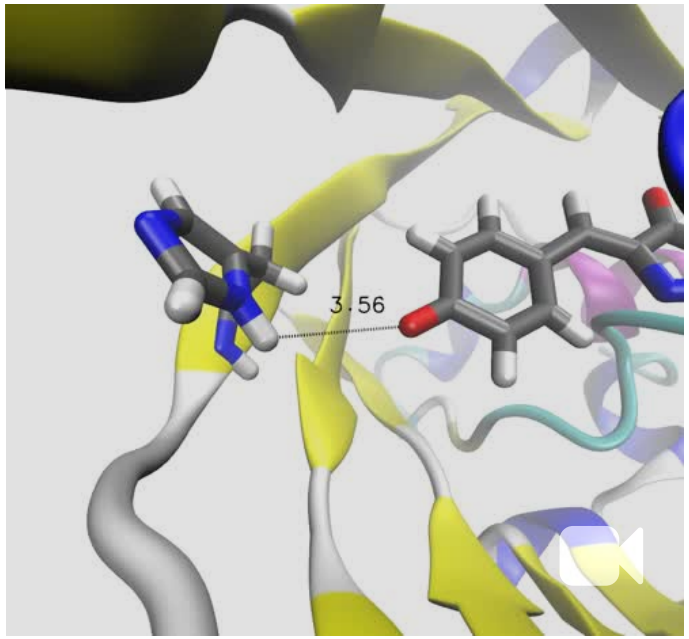
Ensemble differences in noncovalent interactions provide binding enthalpy

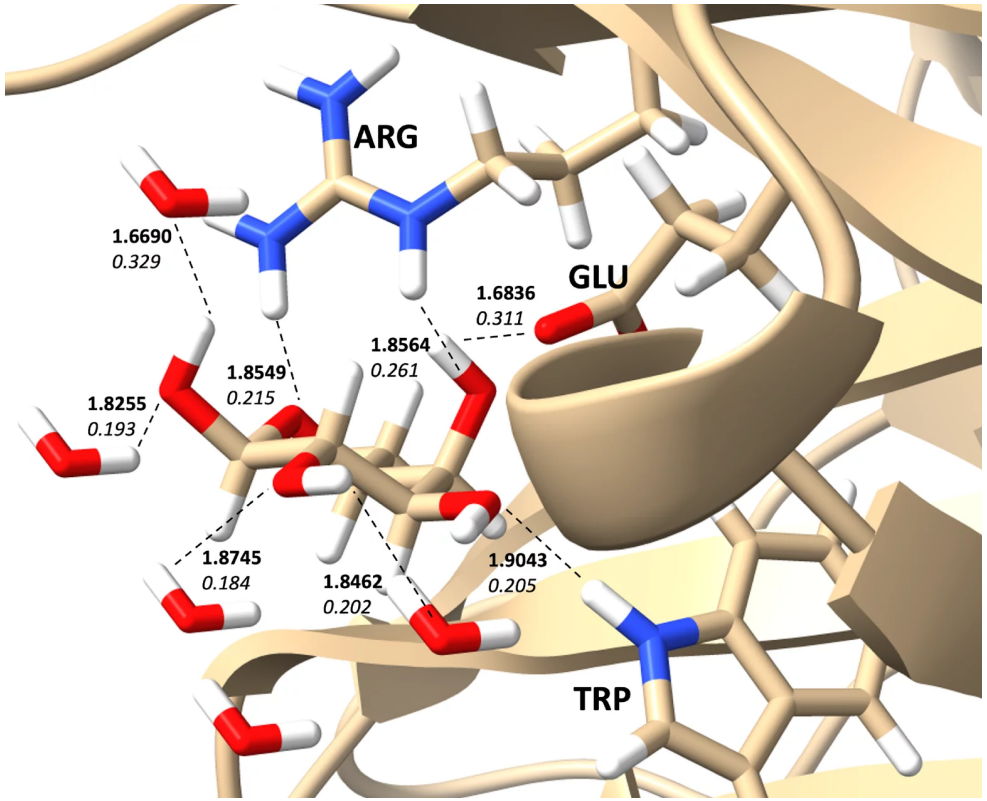
Noncovalent interactions: Electrostatics, hydrogen bonds, dipoles, π-π stacking, etc.

$$\Delta H_{bind} = \langle H_{PL} \rangle - \langle H_P \rangle - \langle H_L \rangle$$

$\langle \cdots \rangle$

Ensemble average



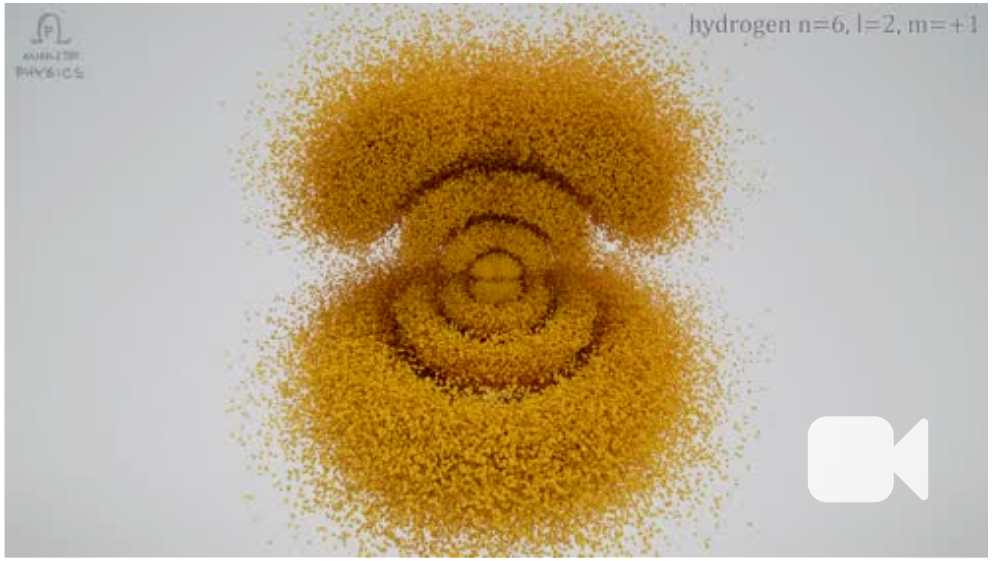


(We assume no covalent bond breaking)

Chemical interactions are determined by fluctuating electron densities

Molecular interactions are governed by their electron densities (Hohenberg-Kohn theorem)

This is rather difficult, so we often use conceptual frameworks to explain trends (e.g., hybridization and resonance)



Our noncovalent interactions conceptual framework:

1. Coulomb's law describes the interactions between charges

2. Molecular geometry uniquely specifies an electron density

3. Regions of increased electron density are associated with higher partial negative charges

4. Electrons are mobile and can be perturbed by external interactions

Electrostatic forces govern interactions between charged and polar regions

Charged molecules have a **net imbalance** between

- Positive charges in their nuclei
- Negative charges from their electrons

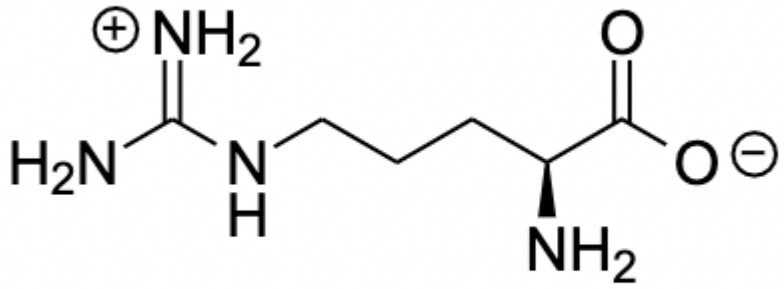
This leads to **net electrostatic attractions or repulsions** between different atoms or molecules

Role in binding

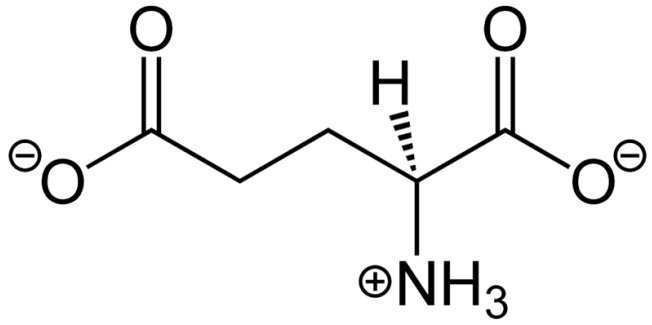
Long-Range Interaction: Can attract ligands to the binding site from a distance

Anchor Points: Often serves as key anchoring interactions in the binding site

Arginine



Glycine

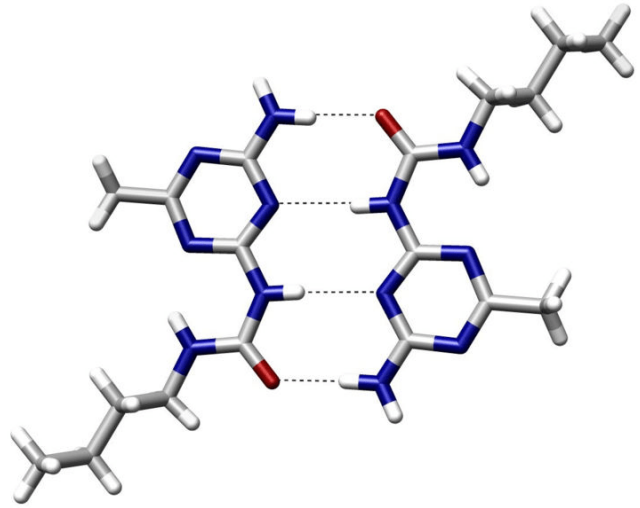


~5 to 20 kcal/mol per interaction

Hydrogen bonds are a type of electrostatic interactions

Attraction between a **(donor)** hydrogen atom covalently bonded to an electronegative atom and another **(acceptor)** electronegative atom with a lone pair

- **Common donors:** O-H, N-H groups
- **Common acceptors:** O and N atoms with lone pairs



Role in binding

Specificity: Precise orientation of the ligand

Stabilization: Moderately strong interactions

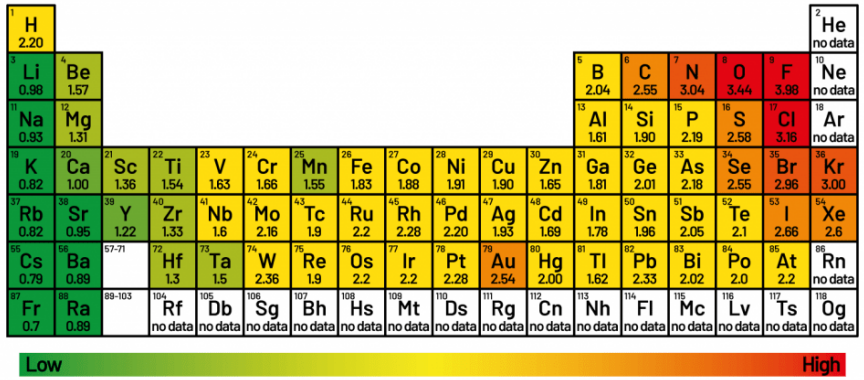
Dynamic: Allows for adaptability of ligands

Strongest when the hydrogen, donor, and acceptor atoms are colinear

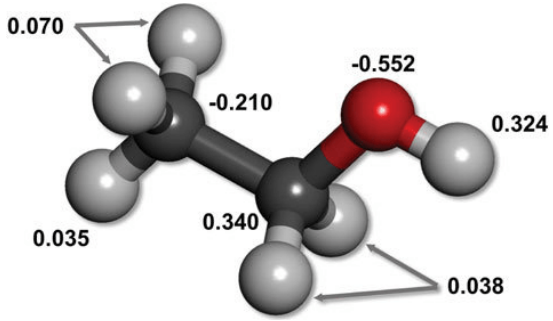
~2 to 7 kcal/mol per hydrogen bond

Uneven electron distribution creates partial charges and dipoles

Electronegativity differences lead to unequal distribution of electron density



Unequal distribution results in regions or partial positive or partial negative charges

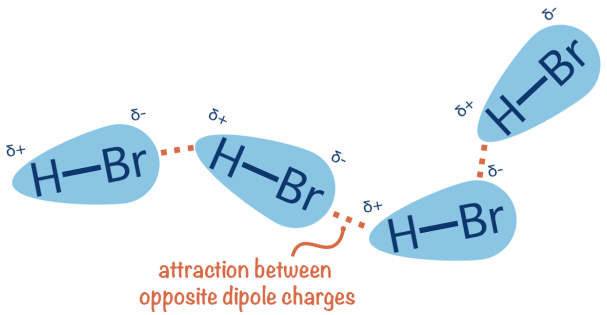


Role in binding

Directional binding: Highly directional, ensuring that the ligand aligns correctly

Flexibility: Can accommodate slight conformational changes

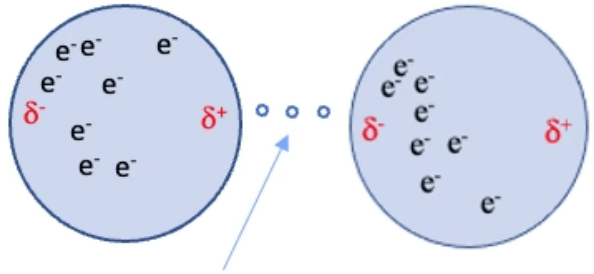
Consistent electron density spatial variation results in permanent dipoles



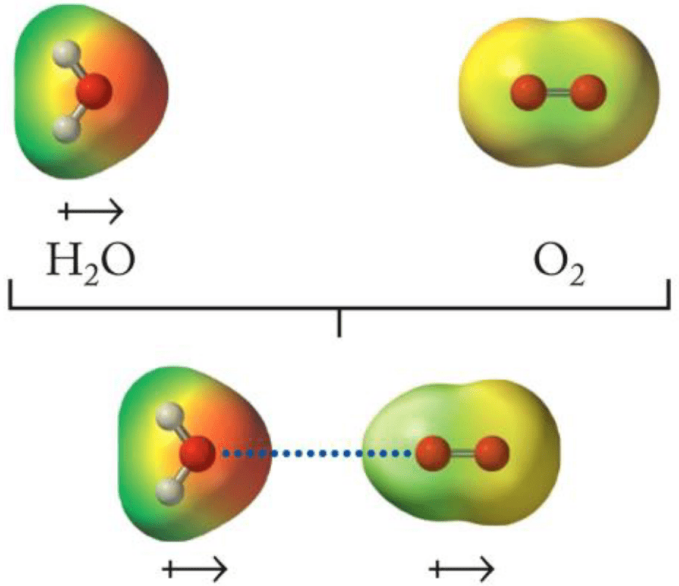
~0.01 to 1 kcal/mol per interaction

Van der Waals forces are weak, non-directional interactions

Dispersion: Electrons in molecules are constantly moving, leading to temporary uneven distributions that induce dipoles in neighboring molecules



Induction: The electric field of a polar molecule distorts the electron cloud of a nonpolar molecule, creating a temporary dipole



Role in binding

Complementary fit: Maximizes surface contact

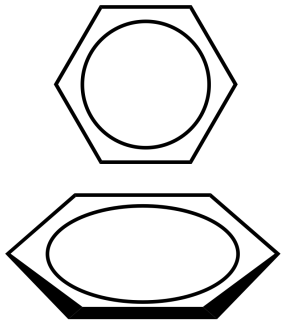
Flexibility: Allows small conformational changes

~0.4 to 4 kcal/mol per interaction

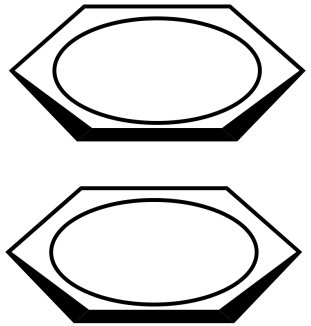
π-π interactions involve stacking of aromatic rings

Noncovalent interactions between aromatic rings due to overlap of π-electron clouds

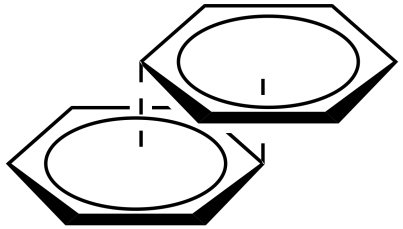
Edge-to-face



Face-to-face



Displaced



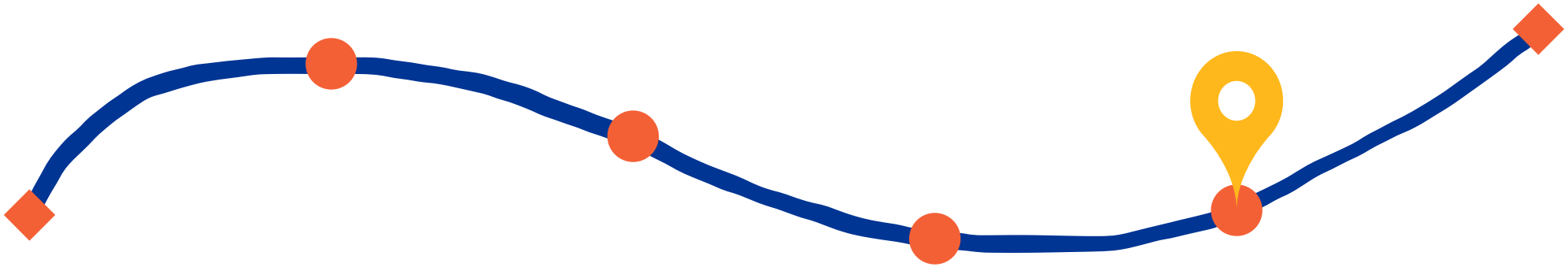
Role in binding

Orientation: Proper positioning of aromatics

Selectivity: Recognition of ligands

~1 to 15 kcal/mol per interaction

After today, you should have a better understanding of



Entropic contributions to binding

Entropy accounts for microstate diversity of a single system state

One of Alex's esoteric points: "Entropy is disorder," is a massive oversimplification that breaks down in actual practice

Entropy is formally defined as $S = k_B \ln \Omega$

Entropy is "energy dispersion"

Ω is the total number of microstates available to the system without changing the system state

Higher entropy implies greater microstate diversity

"System state" can be arbitrarily defined and compared as

- Unbound ligand vs. bound ligand
- Unfolded protein vs. folded protein
- Liquid water at 300 K vs. 500 K

Grid-based protein-ligand binding

Suppose I have a system with

- Protein receptor
- Ligands positioned on a grid

My macrostate (number and identity of particles, temperature, and pressure) remain constant

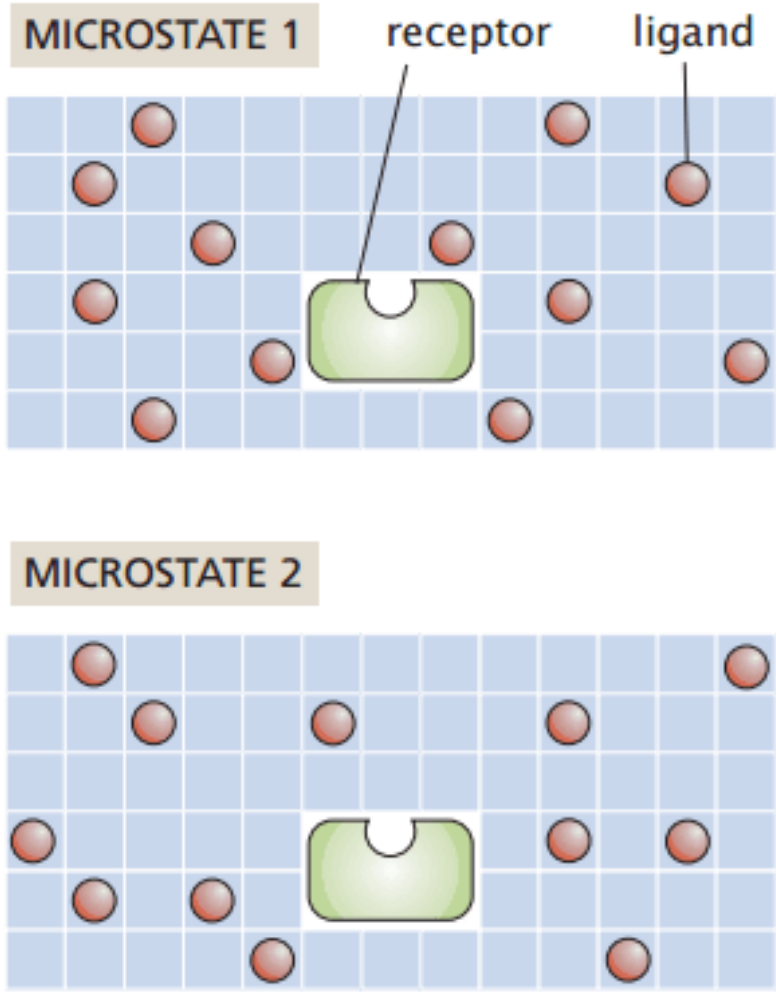
How many ways can I rearrange the ligands without binding to the receptor?

L Number of ligands

N Number of sites

$$\Omega = \frac{N!}{L!\,(N-L)!}$$

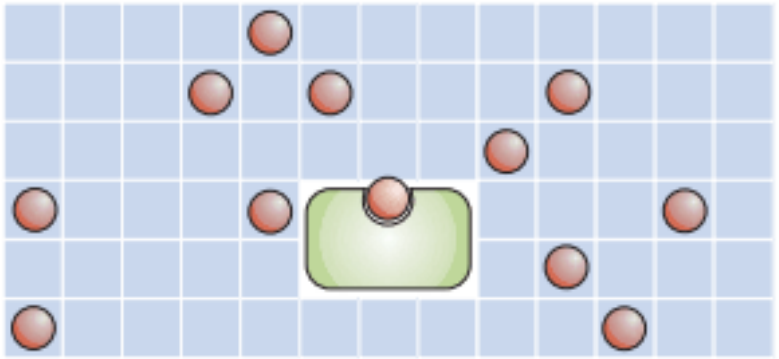
Number of ways to choose *L* grid sites out of *N* is the binomial coefficient



Grid-based protein-ligand binding

What if one ligand binds to the receptor?

$$\Omega = \frac{N!}{(L-1)!\,(N-L+1)!}$$

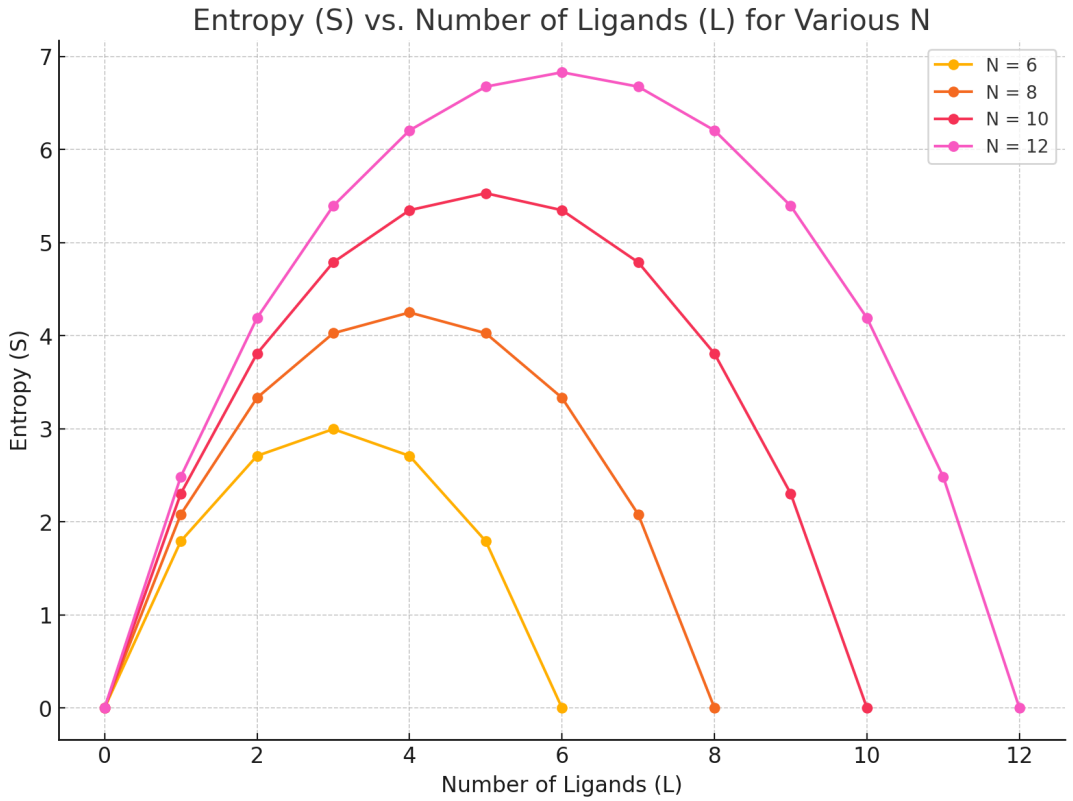


How does entropy change?

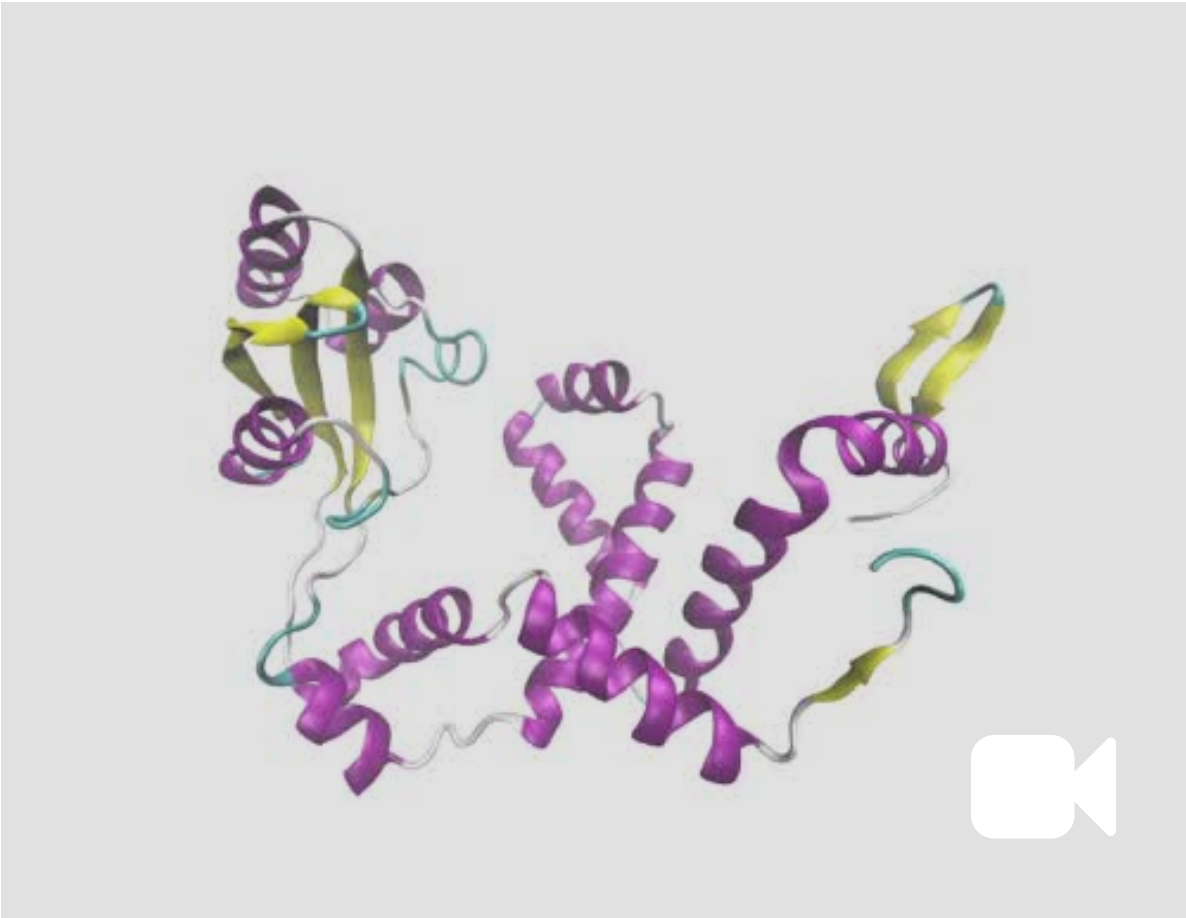
Increase	No change	Decrease

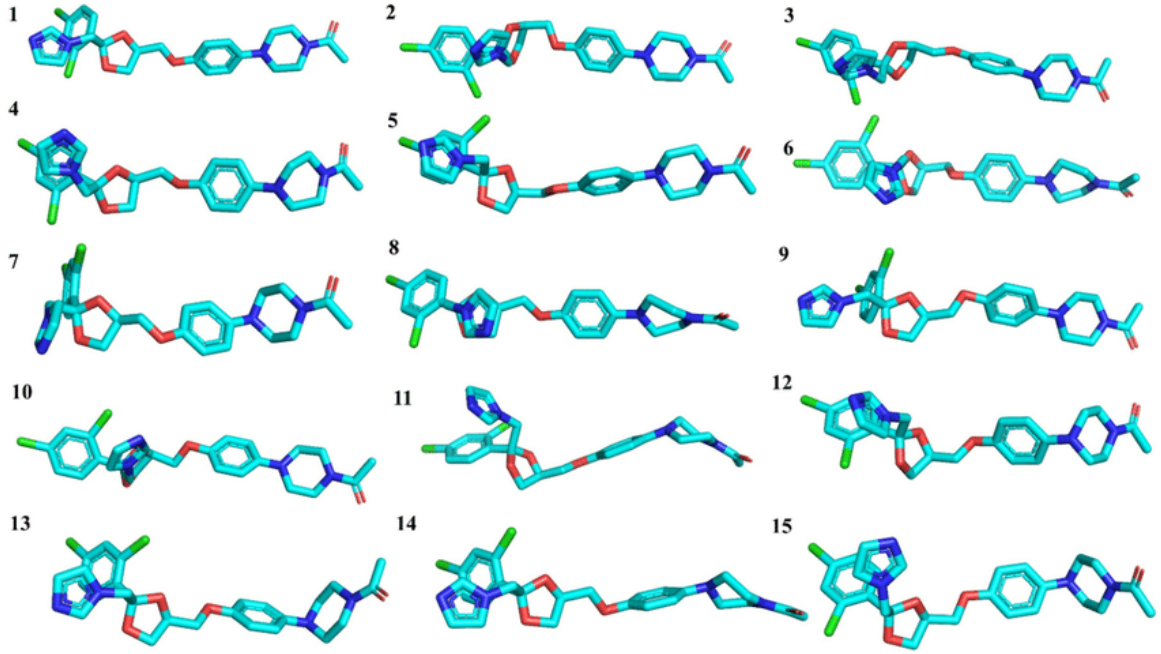
It depends on our ligand concentration!

How to interpret this: Pick a number of ligands and move to the right (L - 1), does entropy go up or down?



Entropic contributions involve accounting for the number of accessible microstates/configurations for protein and ligand





Before the next class, you should

Lecture 10A:

Atomistic insights - Foundations

Lecture 10B:

Atomistic insights - Methodology

Today

Thursday

- Work on P02B and P02C