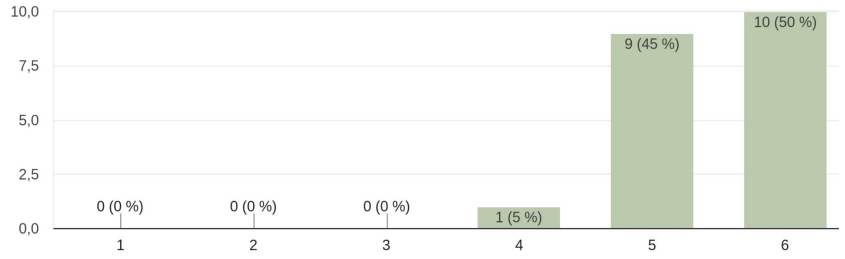
Feedback of lecture 4

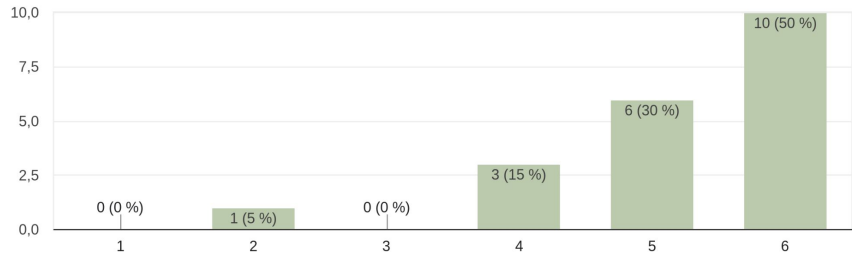
How was your overall impression of the fourth lecture?

20 Antworten



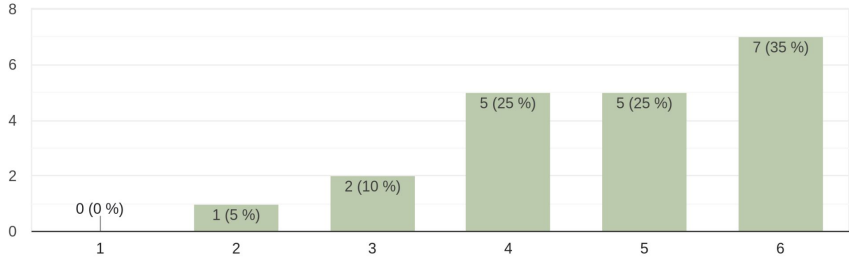
How well could you understand and follow David (the lecturer)?

20 Antworten



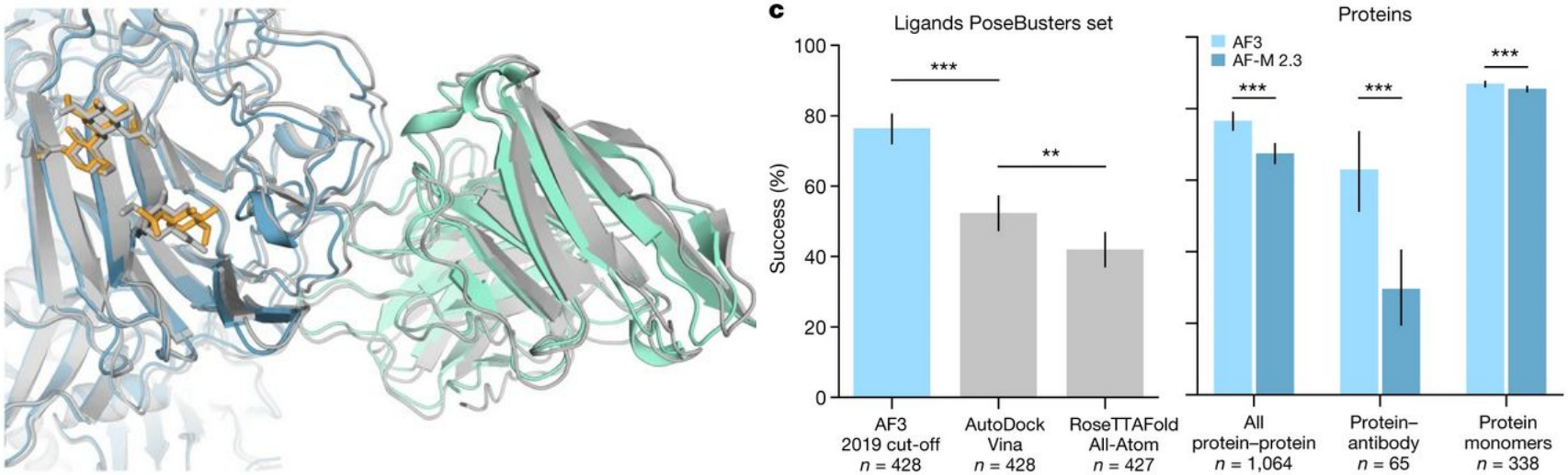
Do you think a 30- to 45-minute session of Ask Me Anything, some time in the semester, would be of value for you?

20 Antworten



\+ Topic difficult but delivery enjoyable
\+ Slides helped understanding
\+ *Ask Me Anything* would be a good idea
? AMA session outside of the lecture?
? More stories & reading & less math expected
\- Give time to read slides and think
\- Add videos for visuals
\- Include more group work
\- Make session more interactive

AMIDD 2025 Lecture 5: Protein as Drug Targets



Left: human coronavirus spike protein (left) bound to neutralization antibody (right), predicted by *AlphaFold3* (predictions in color, ground truth in gray). Right: Performance of *AlphaFold3* for protein-ligand interaction and protein-protein interaction prediction. AF-M: AlphaFold Multimer. Adapted from Abramson, …, Hassabis, Jumper, *Accurate structure prediction of biomolecular interactions with AlphaFold 3*, Nature (2024)

Dr. Jitao David Zhang, Computational Biologist

[1] Pharmaceutical Sciences, Pharma Research and Early Development, Roche Innovation Center Basel, F. Hoffmann-La Roche

[2] Department of Mathematics and Informatics, University of Basel

Ill. Niklas Elmehed © Nobel Prize Outreach

John J. Hopfield

Ill. Niklas Elmehed © Nobel Prize Outreach

Geoffrey E. Hinton

The Nobel Prize in Physics 2024: "for foundational discoveries and inventions that enable machine learning with artificial neural networks"

Ill. Niklas Elmehed © Nobel Prize Outreach

David Baker

Ill. Niklas Elmehed © Nobel Prize Outreach

Demis Hassabis

Ill. Niklas Elmehed © Nobel Prize Outreach

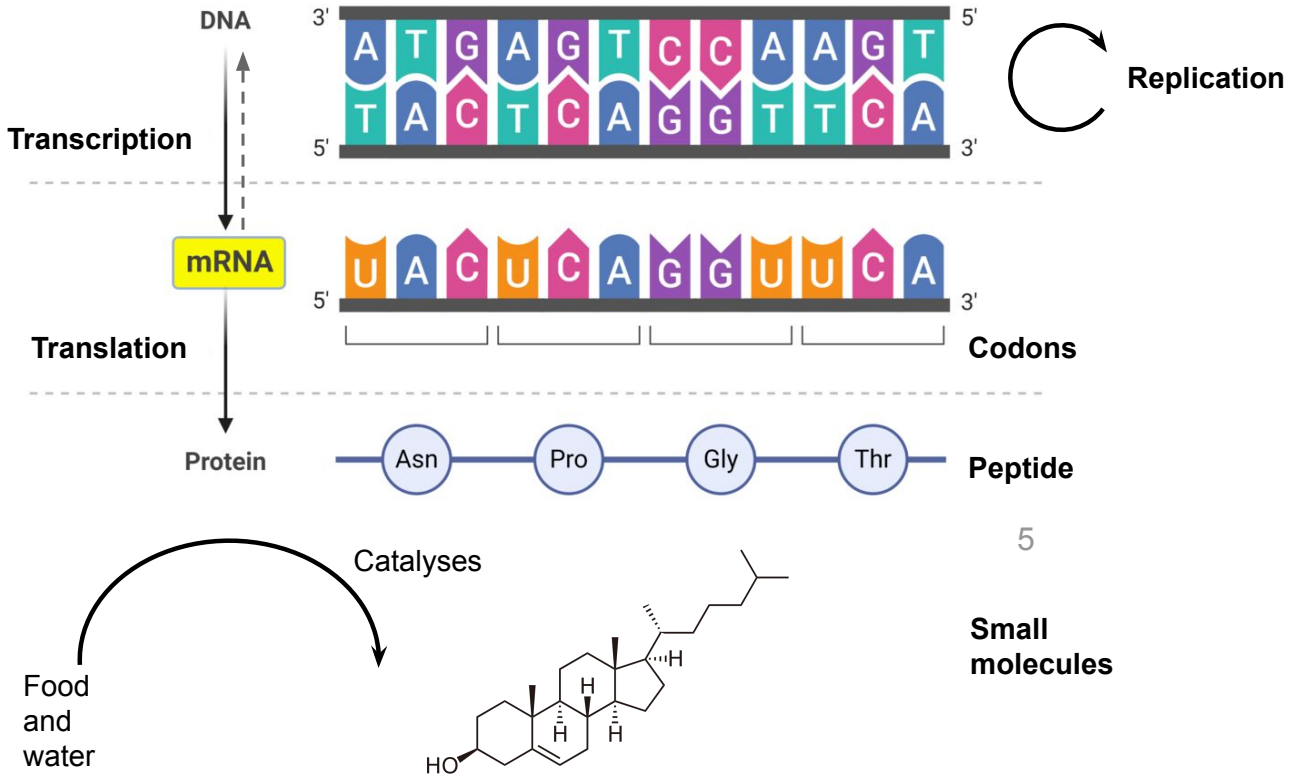
John M. Jumper

The Nobel Prize in Chemistry 2024 was divided, one half awarded to David Baker "for computational protein design", the other half jointly to Demis Hassabis and John M. Jumper "for protein structure prediction"

Topics of lecture 5

- **Protein, ligand, and protein-ligand interaction**
- **ODE-based mechanistic models**

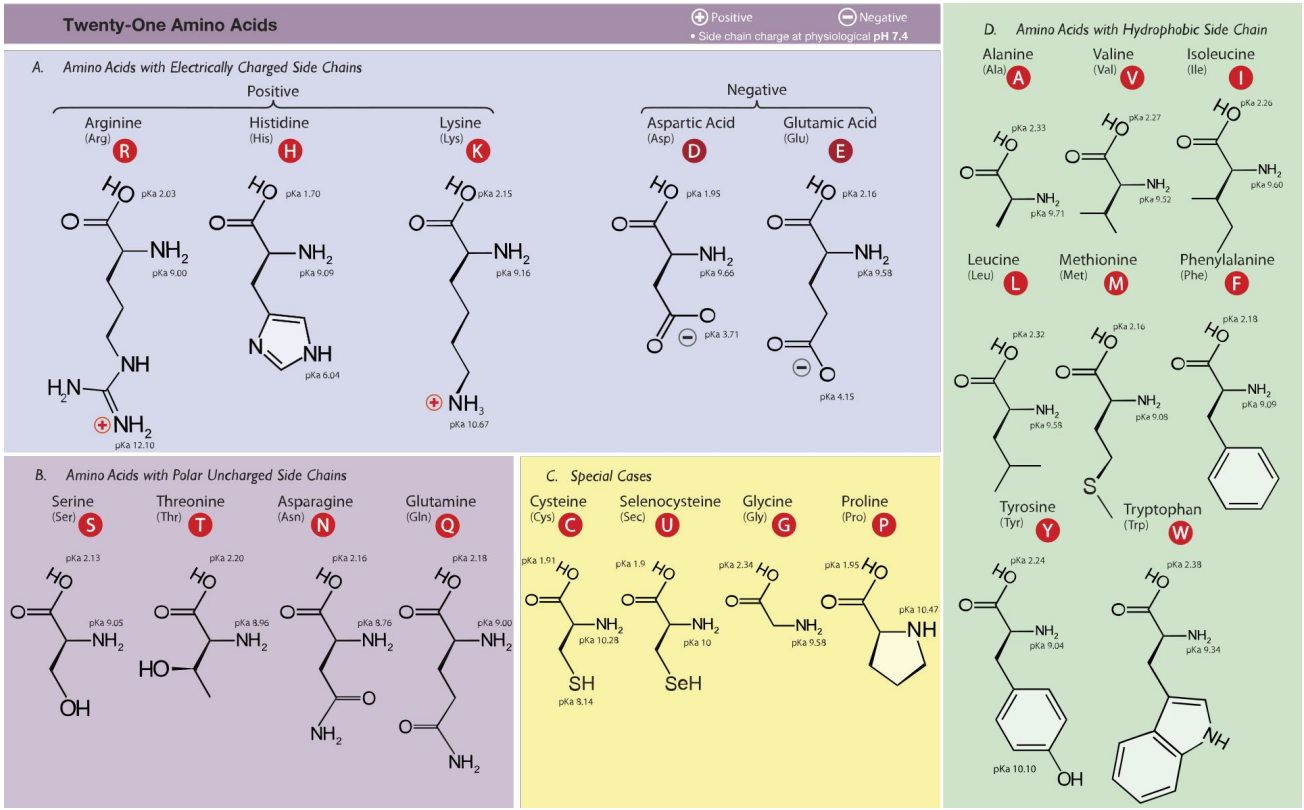
Central Dogma revisited



Central dogma as an information channel: nodes and edges can all be targeted by drugs

Target	Example drugs
Small molecules	Dietary supplements
Catalysis	Enzyme inhibitors
Protein	Receptor agonists/antagonists, ion channel blockers, antibodies
Translation	Antimicrobial protein synthesis inhibitors
RNA	Antisense oligonucleotides (ASO), vaccines
Transcription	Antimicrobials (e.g. actinomycin D and α-Amanitin), splicing modifiers (e.g. Risdiplam/Evrysdi)
Reverse transcription	Antivirals (e.g. reverse transcriptase inhibitors AZT/Zidovudine)
DNA	Gene therapies (e.g. chimeric activated receptors in T-cells, CAR-T)
DNA replication	Topoisomerase inhibitors (e.g. quinolones) and chemotherapy agents

Amino acids are building blocks of proteins and form peptide bonds



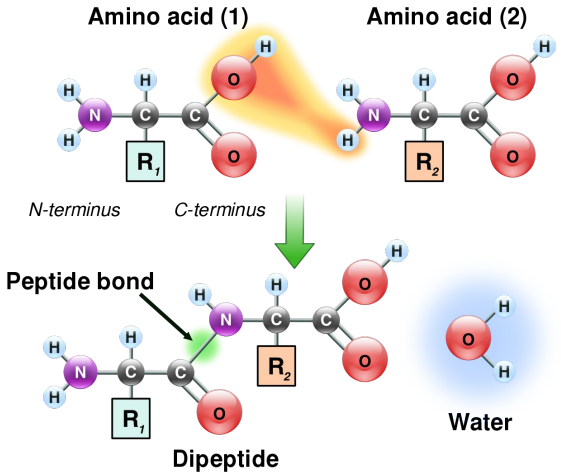
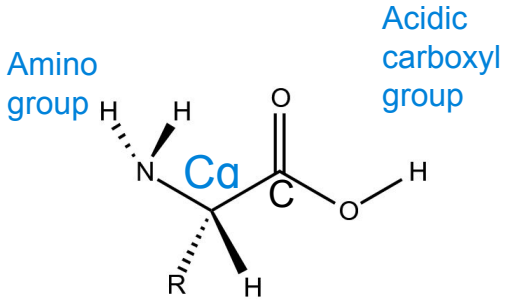
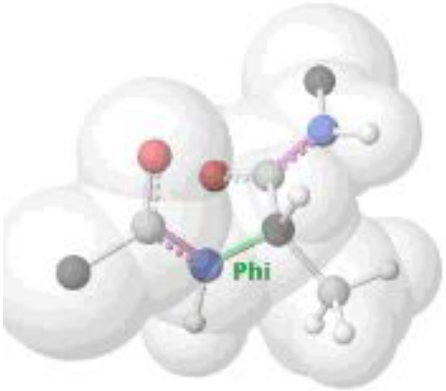


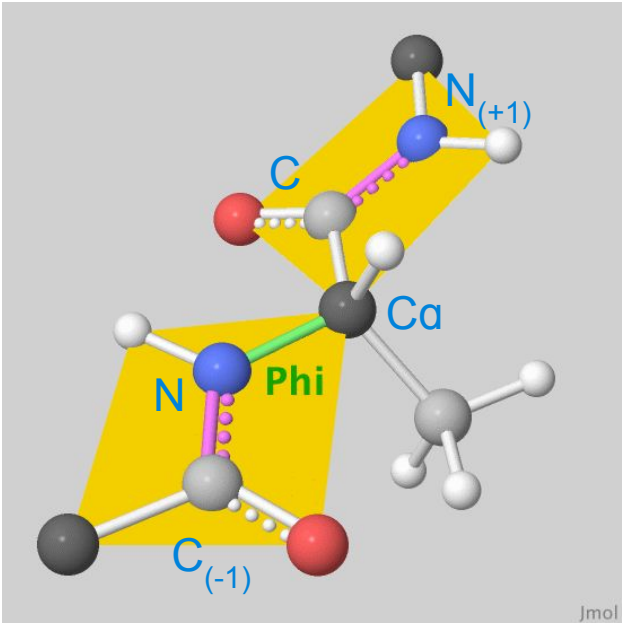
Figure by Dan Cojocari. Reused with CC license from wikimedia

Primary structure of proteins

- (Top left) Human proteins are chains of amino acids (AAs). The backbone remains the same while the side chain varies among AAs.
- (Right) The amino group and the carboxyl group of adjacent amino acids form peptide bonds. Proteins are therefore called *polypeptides*.
- C-Cα bonds and Cα-N bonds can rotate at two *dihedral angles,* Ψ (psi) and ϕ (phi), respectively.
- (Bottom left) Due to steric collisions, only a subset of combinations of Ψ and ϕ is possible



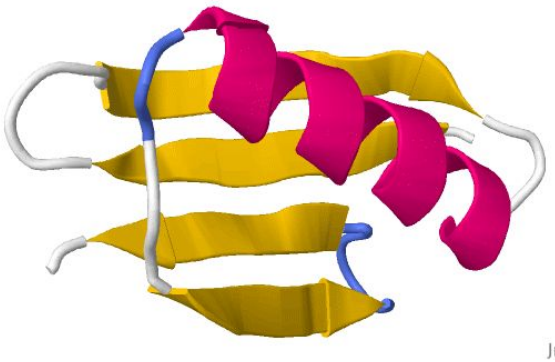


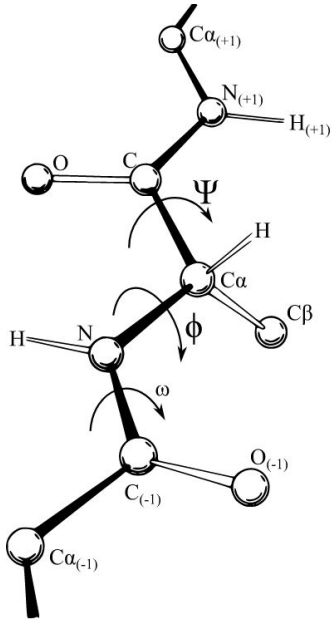


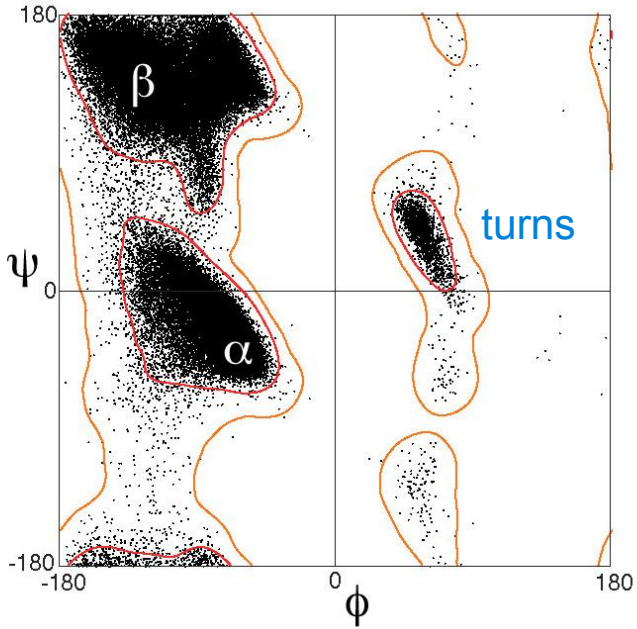
The Ramachandran Principle: alpha helices, beta strands, and turns are the most likely confirmations for a polypeptide

Most other conformations are impossible to due to clashes, known as *steric collisions*, between atoms.

To learn more about the topic, check out the YouTube video tutorial or the Slides by Eric Martz, and finish the Quiz.



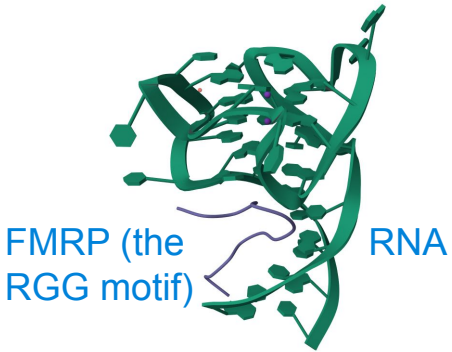




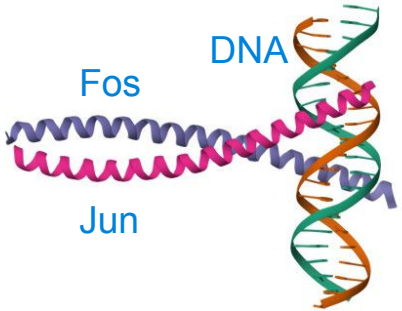
100,000 dots taken from high-resolution crystallographic structures. Wikimedia Commons, courtesy Jane and David Richardson (Proteins 50:437, 2003). This plot excludes glycines, prolines, and amino acids preceding prolines.

Proteins specifically and tightly bind to other molecules

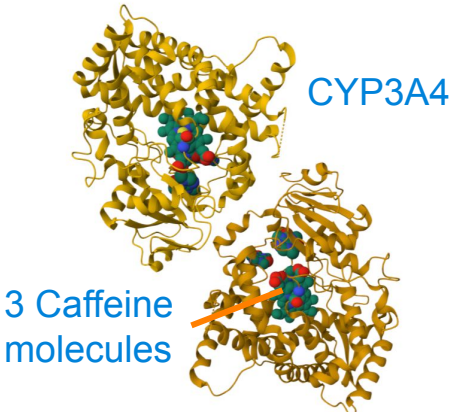
One protein binds to another protein PDB 3iol



Protein binds to RNA. Protein FMRP is encoded by gene *FMR1*. Mutations associated with *FMR1* induce the fragile X syndrome.PDB 5DE5



Protein complex binds to DNA. The complex Fos:Jun is known as AP-1, a transcription factor. PDB 1FOS.



Protein binds to small molecule. Cytochrome P450 3A4 (CYP3A4) is a major drug metabolizing enzyme, which also metabolizes caffeine. PDB 8so1

Major protein classes by functions

Top: an antigen presenting cell; **Bottom**: a T cell; **Red dots**: viral peptides

Enzymes: catalysis of chemical reactions.

- To learn the basics of enzymes, watch the video *How Enzymes Work*.

Transporters: moving ions, small molecules, and proteins across membranes.

- Probably missing in the drawing. To learn the basics of transporters and other ways cell transport material across membranes, Watch the video *Biology: Cell Transport*.

Receptors and kinases: signalling allows cells adapt to the environment.

- To learn the basics of cellular signaling, watch the video *Common cell signaling pathway*.

Structural proteins: stiffness, rigidity, and mechanistical forces.

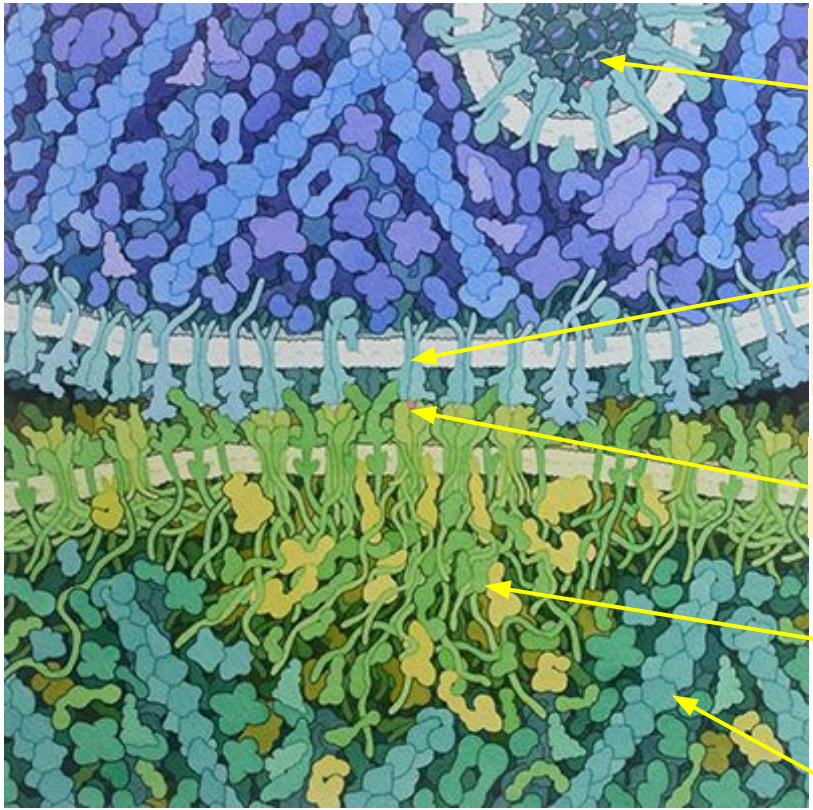
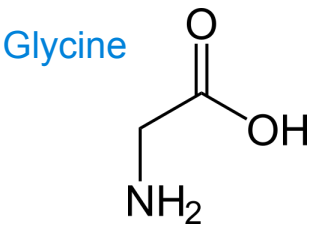


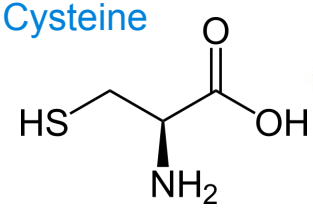
Figure: Immunological Synapse, David S. Goodsell, 2020

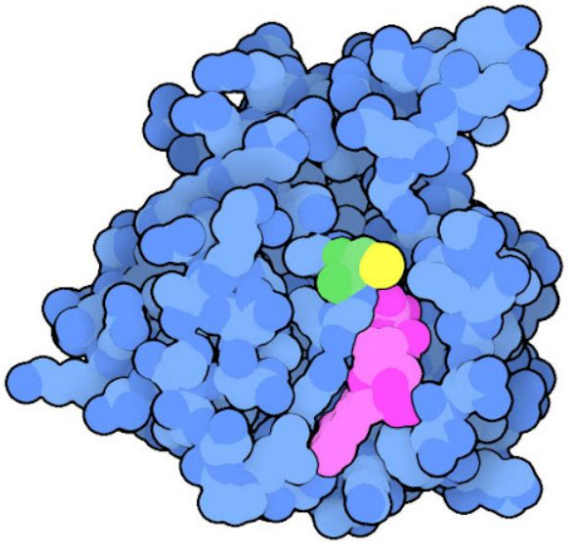
Some diseases are caused by changes in single protein

Glycine

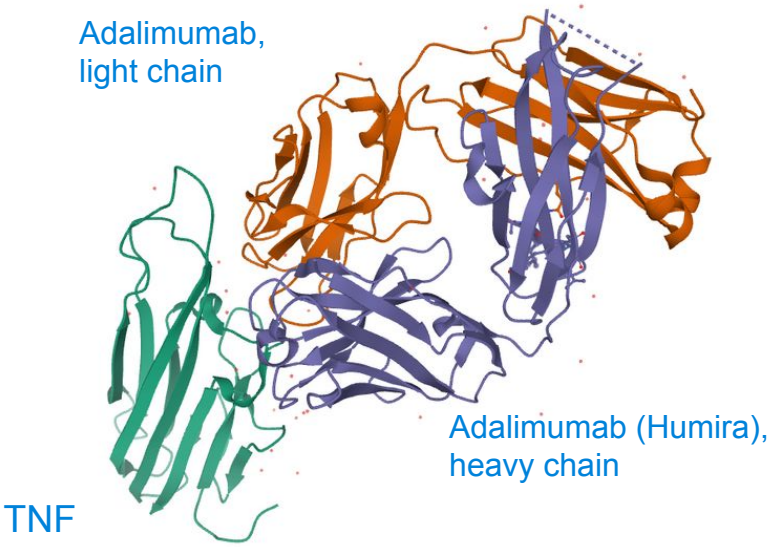


Cysteine





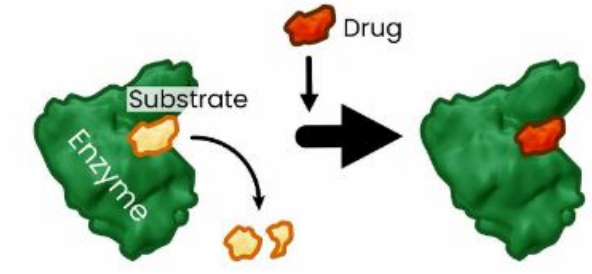
Mutation of glycine (G) to cysteine (M) at position 12 (green, with sulfur in yellow) in Ras protein leads to a protein that is continually activated. The structure of the oncogenic mutant (PDB ID 4ldj) reveals that the mutation modifies the interaction with GDP (magenta) and GTP, which acts as a switch that turns the protein on and off.



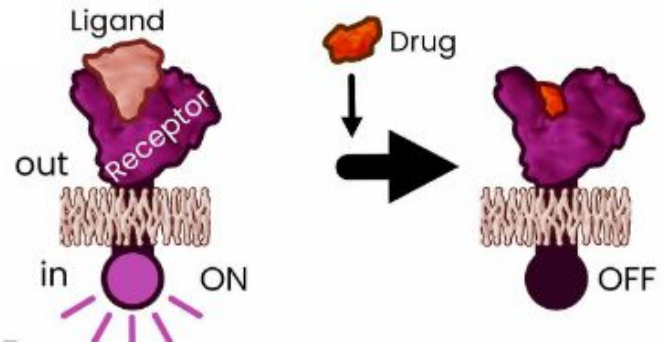
Tumor necrosis factor (TNFα) promotes the inflammatory response in autoimmune diseases. Its level is elevated in diseases including inflammatory bowel disease and rheumatoid arthritis. Monoclonal antibodies against TNFα, for instance adalimumab (Humira), are used for such indications. PDB 3WD5

About 80-90% small-molecule and biological drugs are supposed to work by competitive inhibition

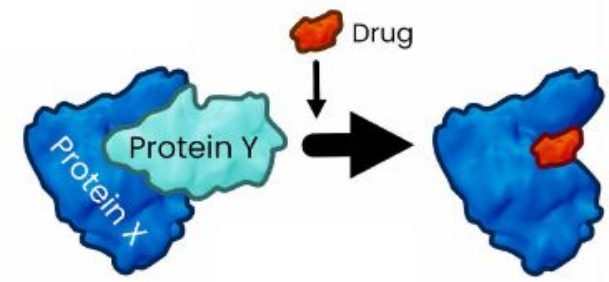
(1) Drug inhibits enzyme binding to substrate



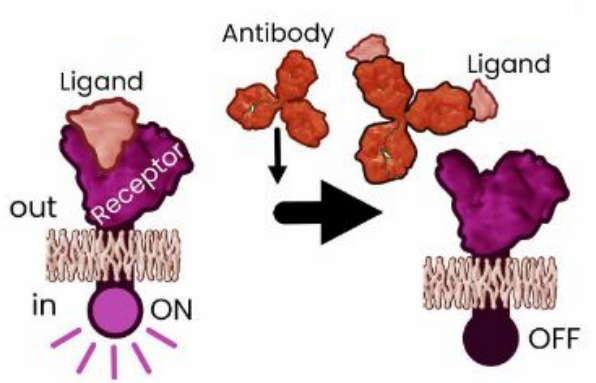
(2) Drug inhibits receptor binding to ligand



(3) Drug inhibits protein X binding to protein Y

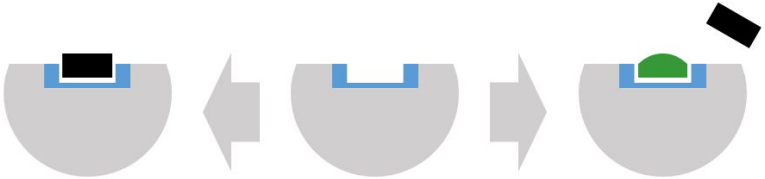


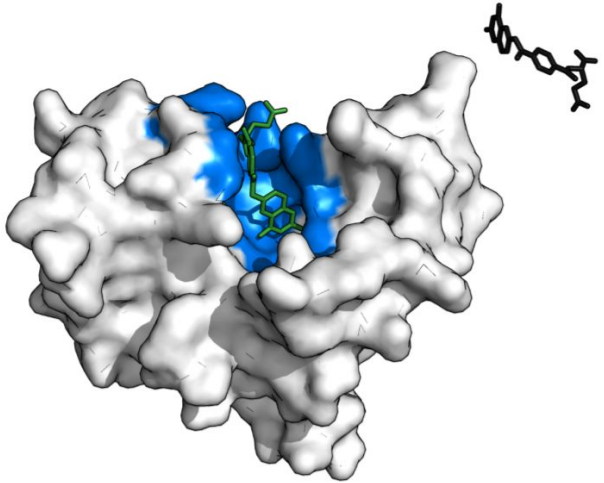
(4) Drug inhibits ligand binding to receptor



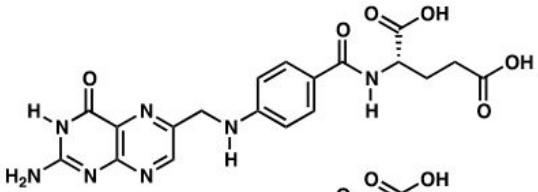
Deshaies, Raymond J., and Patrick Ryan Potts. "Load and Lock: An Emerging Class of Therapeutics That Influence Macromolecular Dissociation." Science 389, no. 6762 (2025)

Methotrexate is a competitive inhibitor of DHFR

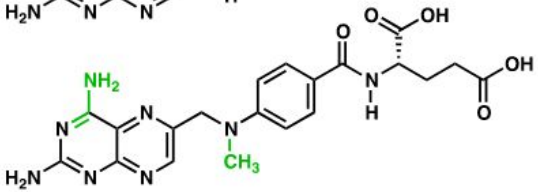




Work by Thomas Shafee, Shared under CC-AS-4.0, and work by Boghog. Based on PDB record 4QI9.



Dihydrofolic acid



MTX

The protein: Dihydrofolate reductase (DHFR) converts dihydrofolic acid into tetrahydrofolate. The process is important for cell proliferation and cell growth. DHFR is a drug target for oncology (cancer) and autoimmune diseases.

The natural substrate: Dihydrofolic acid (vitamin B9), in black. Dihydrofolic acid is the *natural ligand* of DHFR.

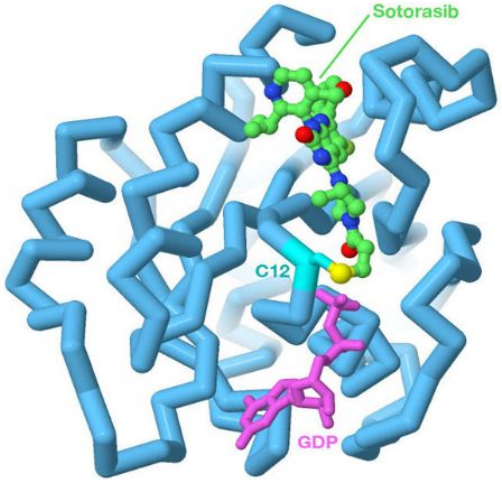
The drug: Methotrexate (MTX), in green, is a *synthesized ligand* of DHFR, and it is a *competitive inhibitor* of DHFR.

The binding site: where the enzyme binds its substrate and catalyses the chemical reaction, in blue.

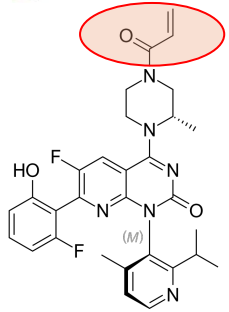
Sotorasib, Pertuzumab, Trastuzumab: examples of small and large molecules inhibiting signaling

Left: The drug *sotorasib* binds covalently to the sulfur atom in cysteine 12 of the *Ras* protein, blocking its action. The drug is shown with carbon atoms in green, the cysteine sulfur is in yellow, and GDP is in magenta. Image created in Jmol using PDB ID 6oim.

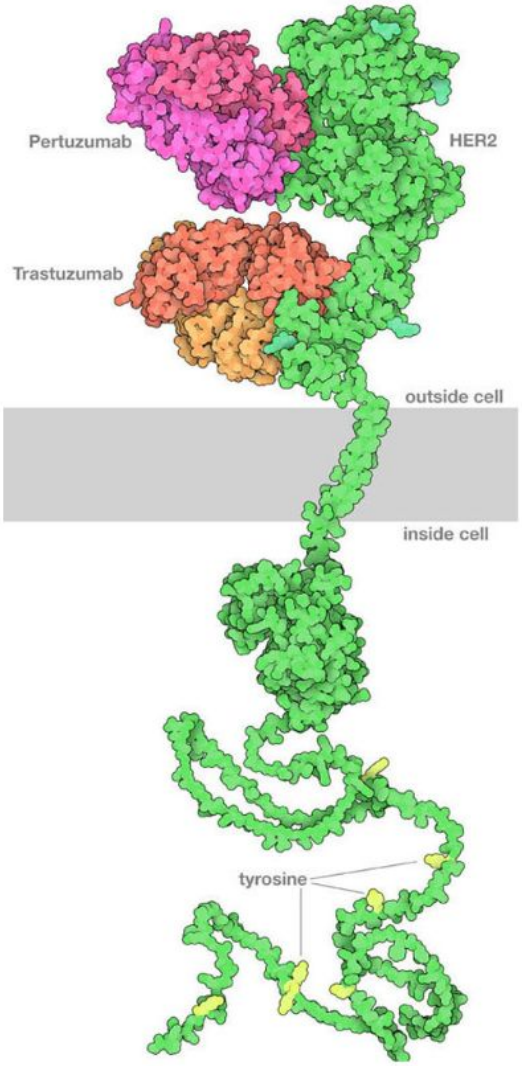
Right: The extracellular domain of HER2 bound to two therapeutic antibodies: *pertuzumab* and *trastuzumab*. The antibodies block the formation of active dimers of the receptor, thus blocking the growth signal (PDB 6ogi). The transmembrane domain is from PDB 2ksi. The kinase domain inside the cell is from PDB ID 3pp0, and the unstructured tail at bottom is predicted by *AlphaFold2*.



Sotorasib



Highlight: covalent warhead (the acrylamide group)



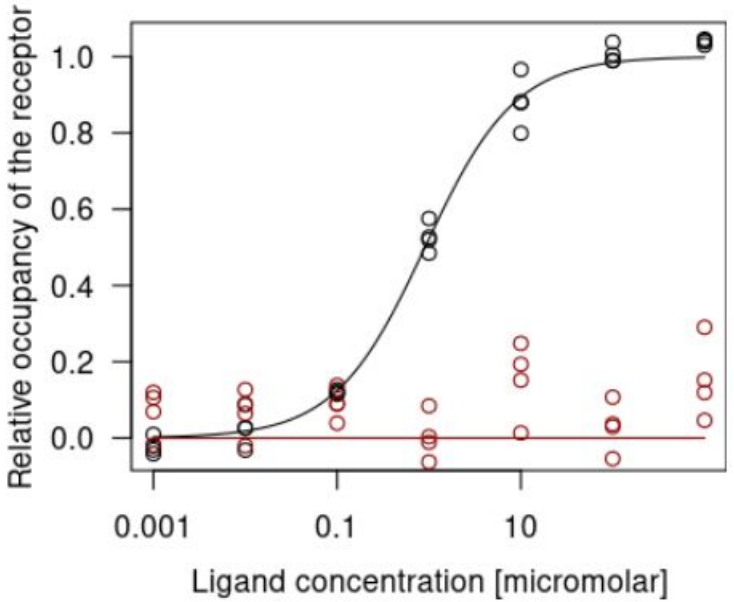
Concentration-occupancy curves characterize protein-ligand binding

X-axis: ligand concentration. Common units: molar (M), micromolar (mM, 10^{-6} M), nanomolar (nM, 10^{-9} M), picomolar (pM, 10^{-12} M).

Y-axis: relative occupancy of the receptor. Alternative values are possible, for instance response (more about that later).

Points: individual measurements. In this plot: mean value of replicates with error bars indicating variability.

Lines: fitted sigmoidal curves using the Hill function or its variants.

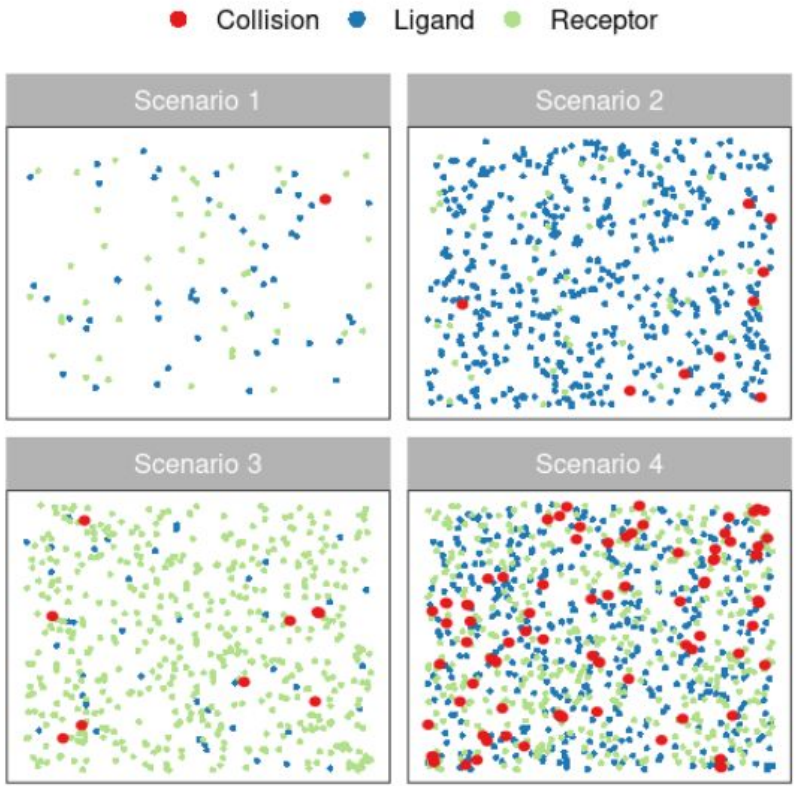


A simple mathematical model addresses a key question: how is a receptor occupied by varying concentrations of drugs?

$$\mathrm{L} + \mathrm{R} \underset{k_2}{\overset{k_1}{\rightleftharpoons}} \mathrm{LR}$$

$$\frac{d[LR]}{dt} = k_1[L][R] - k_2[LR]$$

- Ligand binding to receptor is a reversible reaction.
- **The law of mass action**: the rate of the chemical reaction is directly proportional to the product of the activities or concentrations of the reactants. The proposition can be derived from the *collision theory*. See the right graph for an illustration.



An ordinary differential equation (ODE) model quantifies receptor occupancy by varying concentrations of ligands

$$\mathrm{L} + \mathrm{R} \underset{k_2}{\overset{k_1}{\rightleftharpoons}} \mathrm{LR}$$

The law of mass action

$$\frac{d[LR]}{dt} = k_1[L][R] - k_2[LR]$$

At equilibrium, no net change of [LR]

$$k_1[L][R] = k_2[LR]$$

$R_{total} = [R] + [LR]$

$$k_1[L]([R_{total}] - [LR]) = k_2[LR],$$

$$[LR] = \frac{k_1[L][R_{total}]}{k_1[L] + k_2}$$

$K_D \equiv k_2/k_1$

$$[LR] = [R_{total}]\frac{[L]}{[L] + K_D}$$

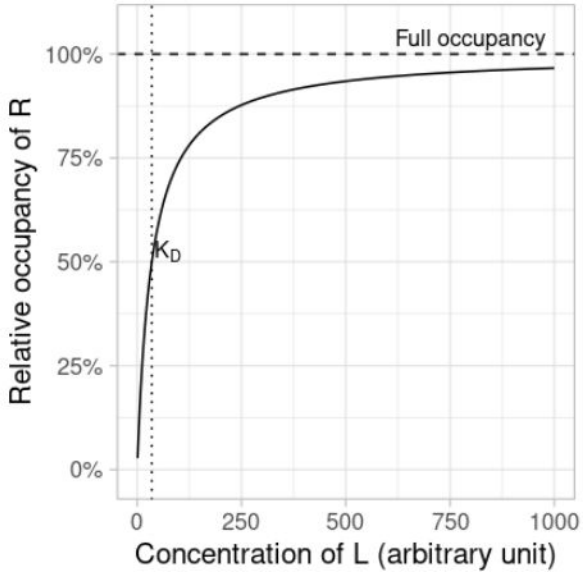
The ODE model induces the simplest form of the *Hill-Langmuir Equation*

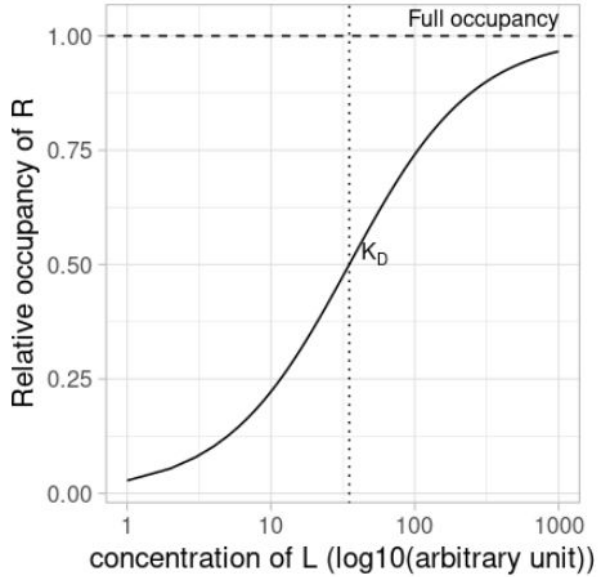
$$\mathrm{L} + \mathrm{R} \underset{k_2}{\overset{k_1}{\rightleftharpoons}} \mathrm{LR} \qquad K_D \equiv k_2/k_1$$

The Hill-Langmuir function describes the occupancy of receptors by natural ligands of drugs. We will meet it again.

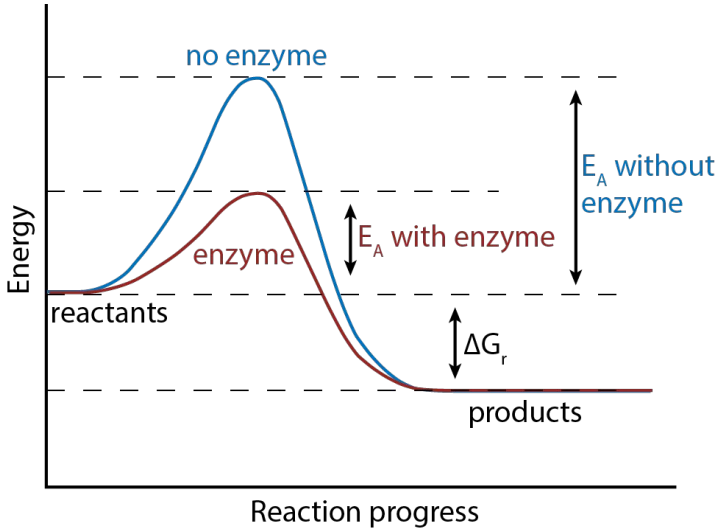
We can interpret K_D (the dissociation constant) both mathematically and physically & chemically. Mathematically, K_D represents (1) the ratio of reaction speeds, and (2) the concentration required to occupy half of the receptors.

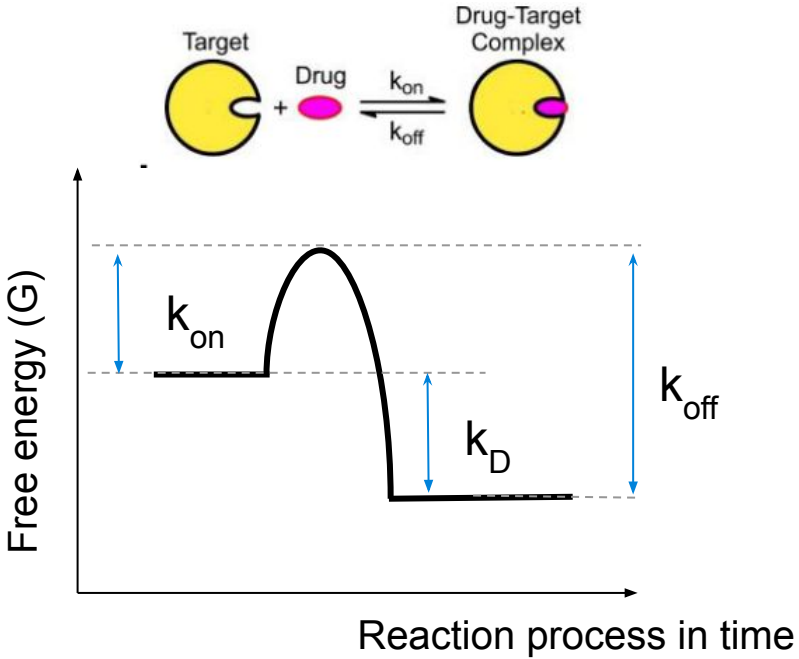
Hill-Langmuir function: $[LR] = [R_{total}]\frac{[L]}{[L] + K_D}$





The thermodynamic interpretation: K_D is directly associated with the free energy of the reaction



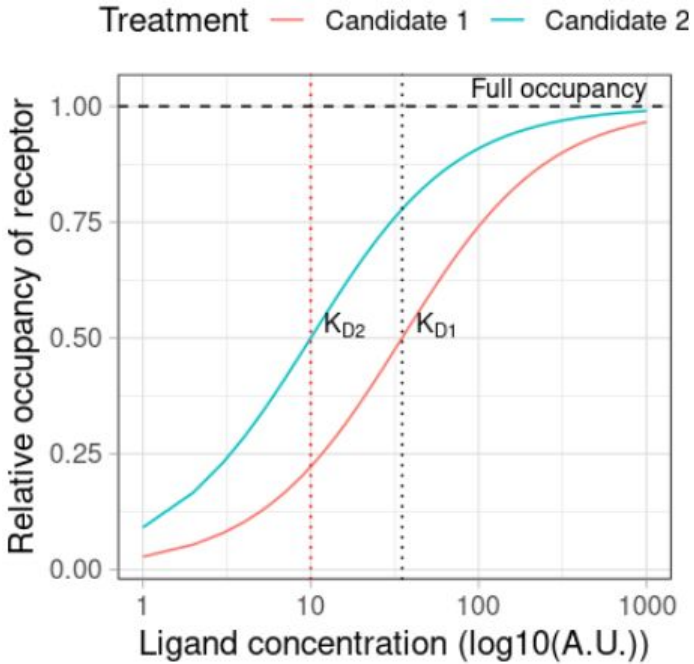


Left: LibreTexts, Role of Enzymes,
Matthew F Kirk, Kansas State University

$$\Delta G = RT \ln K_D$$

(R: gas constant, T: absolute temperature)

Question: all other conditions the same, which drug candidate is more favorable? Why?



The Lotka-Volterra model of predator-prey relationships

- The Lotka-Volterra equations modelling predator-prey relationships.

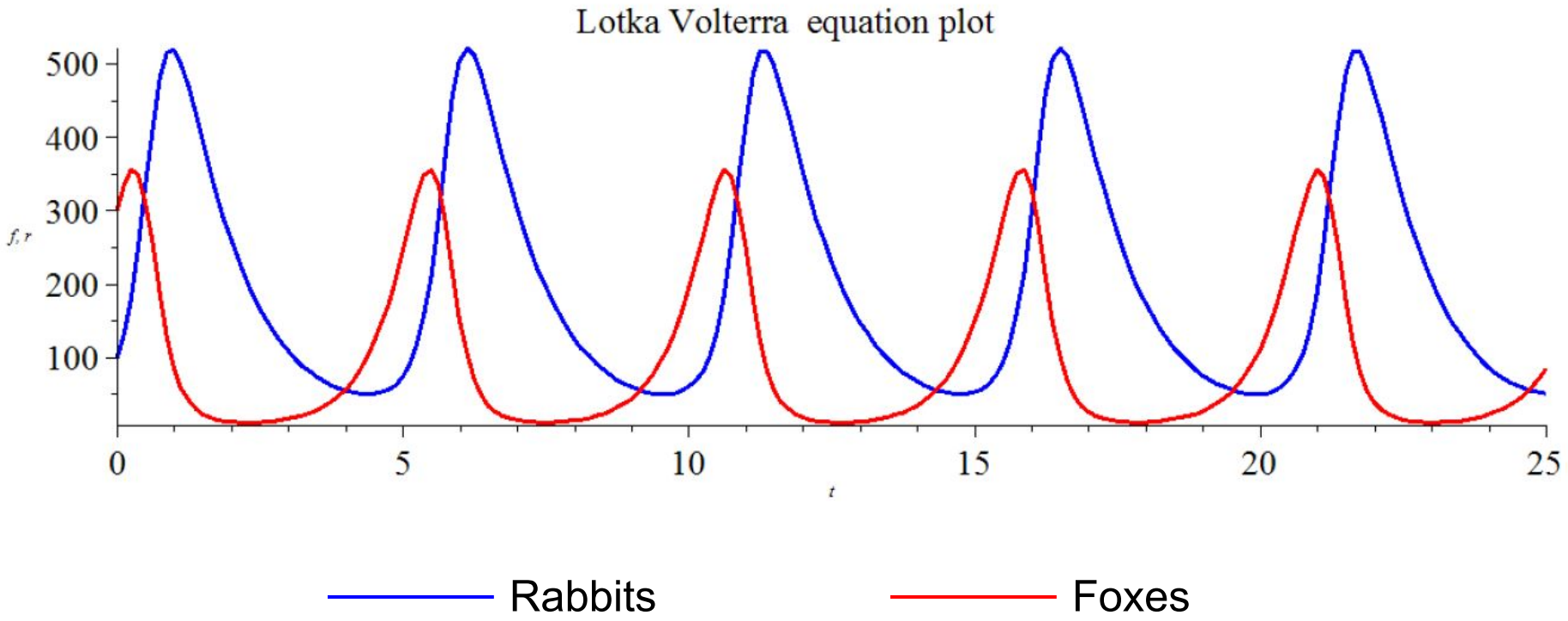
$$\frac{dx}{dt} = \alpha x - \beta xy, \tag{1}$$

$$\frac{dy}{dt} = -\gamma y + \delta xy, \tag{2}$$

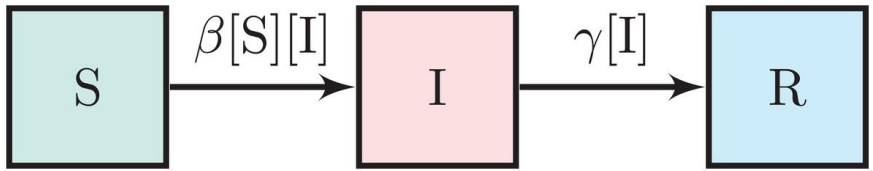
where

- x is the number of prey (*e.g.* rabbits),
- y is the number of predator (*e.g.* foxes),
- $\frac{dx}{dt}$ and $\frac{dy}{dt}$ represent growth rates of the two populations,
- t represents time,
- α, β, γ, and δ are real parameters specifying the interaction of the two species.

The Lotka-Volterra equations, visualized



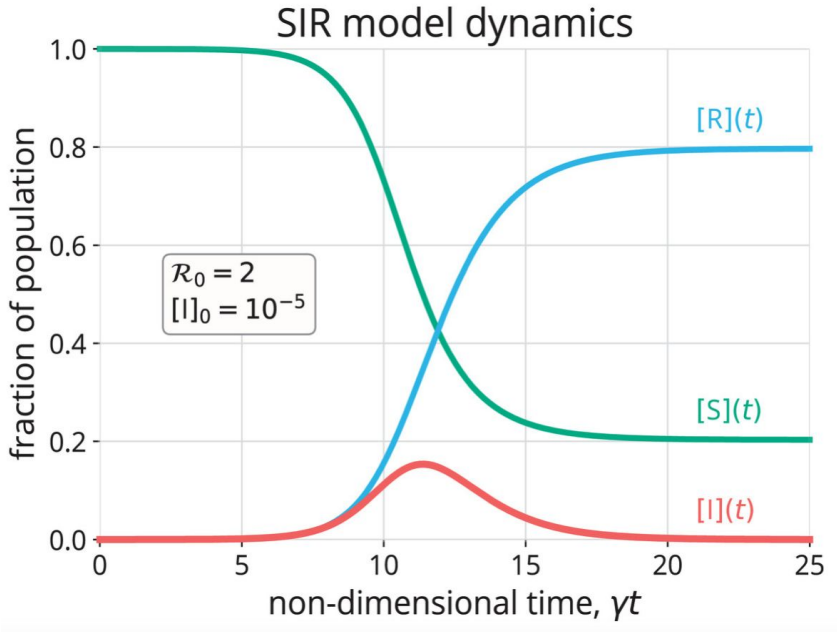
The SIR model of epidemiology models population behavior of viral infection and recovery



The SIR model of epidemiology

- S: Susceptible
- I: Infectious
- R: Removed

$$\frac{dS}{dt} = -\frac{\beta IS}{N},$$
$$\frac{dI}{dt} = \frac{\beta IS}{N} - \gamma I,$$
$$\frac{dR}{dt} = \gamma I$$

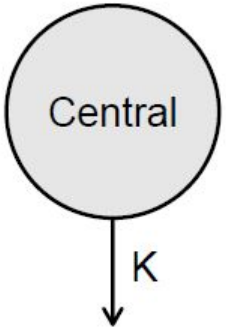


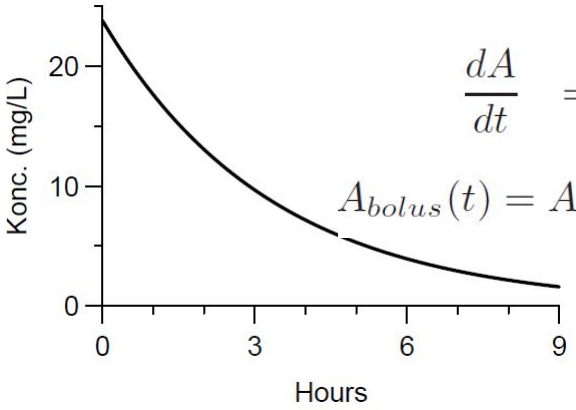
R_0, the basic reproduction number, is the number of people infected by the initial infectious individual. It is defined as β/γ.

Simon CM. 2020. The SIR dynamic model of infectious disease transmission and its analogy with chemical kinetics. PeerJ Physical Chemistry 2:e14https://doi.org/10.7717/peerj-pchem.14

ODE-based mechanistic models are often used in pharmacokinetic modelling

- Pharmacokinetics (PK) describes how the drug is absorbed, distributed, metabolised, and excreted by the body.

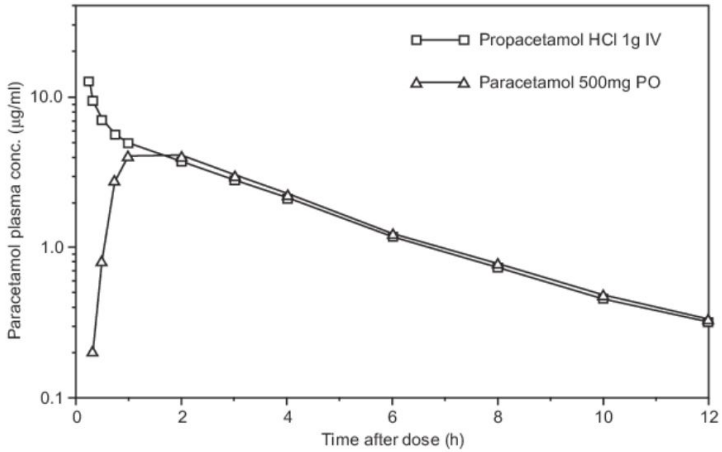




$$\frac{dA}{dt} = -K \cdot A$$

$$A_{bolus}(t) = A_0 \exp(-K \cdot t)$$

- A basic mathematical model of PK is a compartment model, i.e. one or more ordinary differential equations that describe the relationship between drug concentration and time. The simplest model is the decay model of bolus (injection).
  - A_0: initial concentration
  - A(t): drug concentration at time *t*
  - K: rate of clearance
- A real-world example: PK of propacetamol, a pro-drug of paracetamol, delivered via IV.

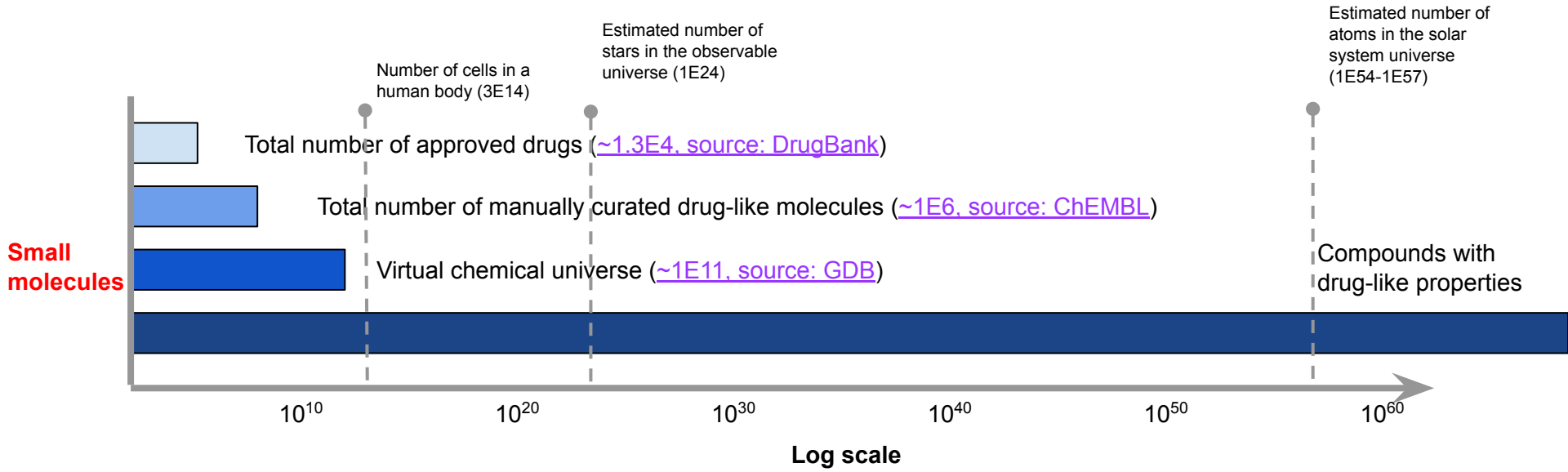


Conclusions and outlook

- **We reviewed the central dogma from the drug discovery's perspective.**
- **We learned examples of ODE-based mechanistic models.**
- **Next time, we shall continue learning statistic and causal models.**

Backup slides

Why drug *discovery*? 1. The chemical space is huge

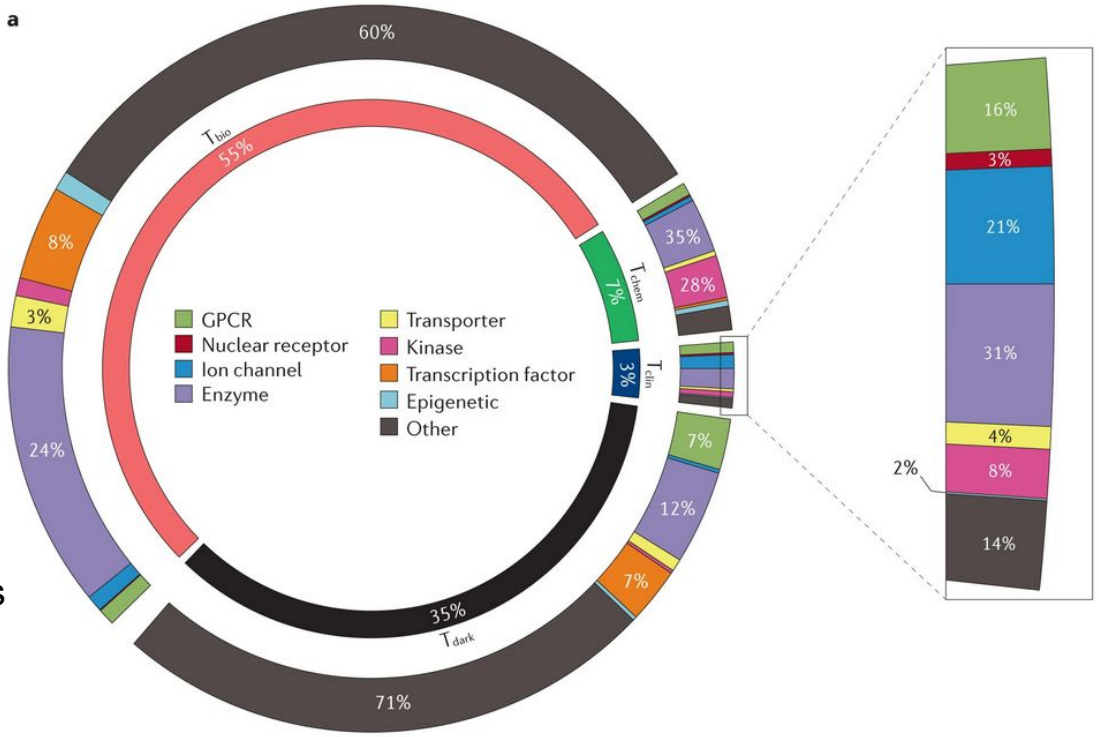


Why drug *discovery*? 2. The druggable proteome is huge - even excluding mutations, transcriptome, genome, …

There are about 20,000 proteins encoded by the human genome. We can classify them by (1) our knowledge of them, and (2) whether we have reliable chemical tools, biological tools, or even drugs to manipulate them.

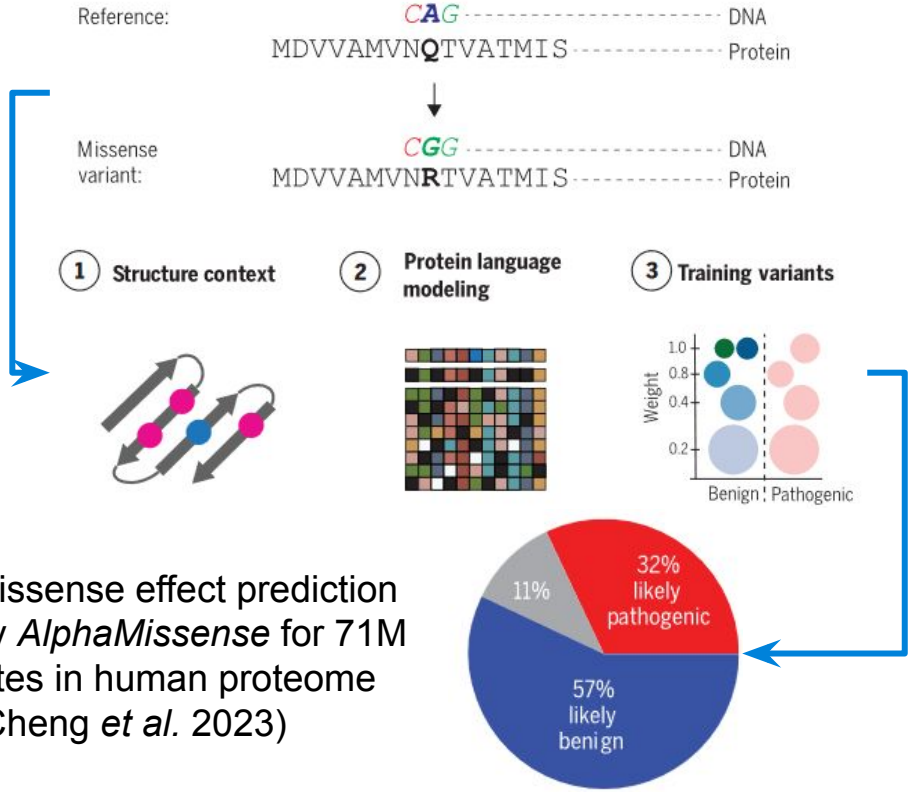
Inner ring: percentages of the whole proteome, classified by whether we have drugs (T_{clin}), whether we have chemical tool compounds (T_{chem}), whether we have biological compounds (T_{Bio}), or we are in the dark (T_{Dark}). Currently, we have only drugs for a few hundred proteins.

Outer ring: protein families.

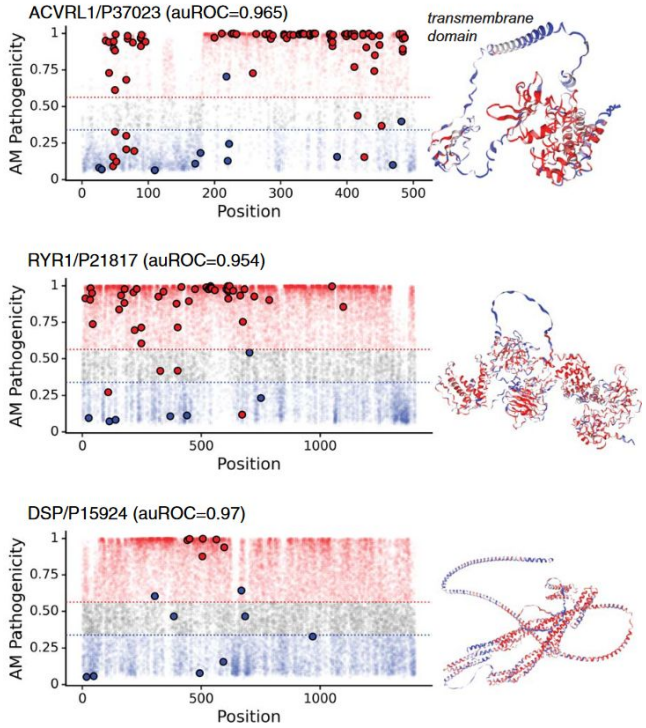


Oprea, et al. "Unexplored Therapeutic Opportunities in the Human Genome." Nature Reviews Drug Discovery 17 (February 23, 2018): 317–32.

Why drug *discovery*? 2. The druggable proteome is huge - now consider the mutations with predicted pathogenicity

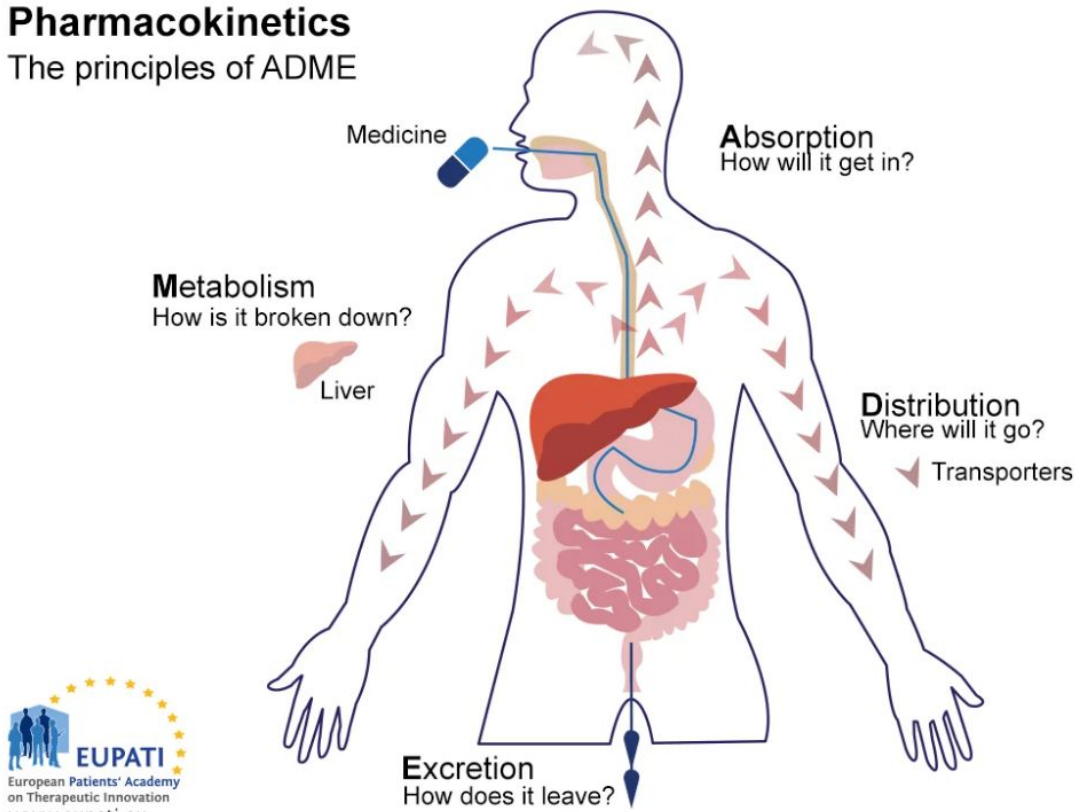


Missense effect prediction by *AlphaMissense* for 71M sites in human proteome (Cheng *et al.* 2023)



Example proteins chosen from ACMG clinically actionable genes

Why drug *discovery*? 4. The drug have to be absorbed and distributed in order to have systemic and organ-specific effects

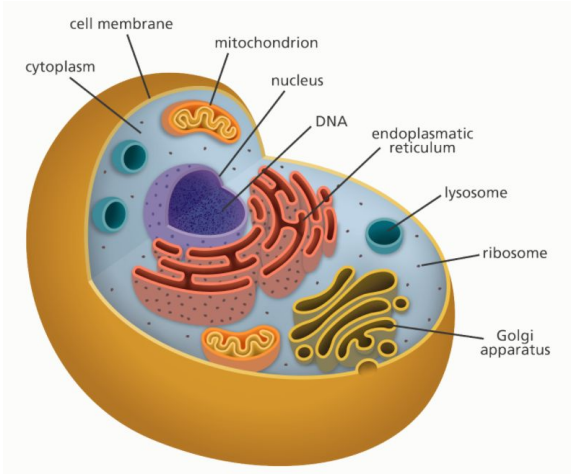


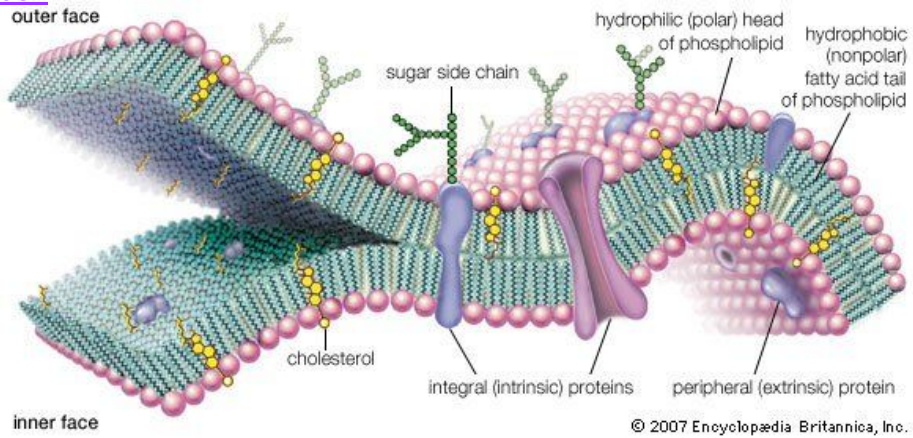
Why drug *discovery*? 5. Drugs have to reach the targets - despite physical barriers

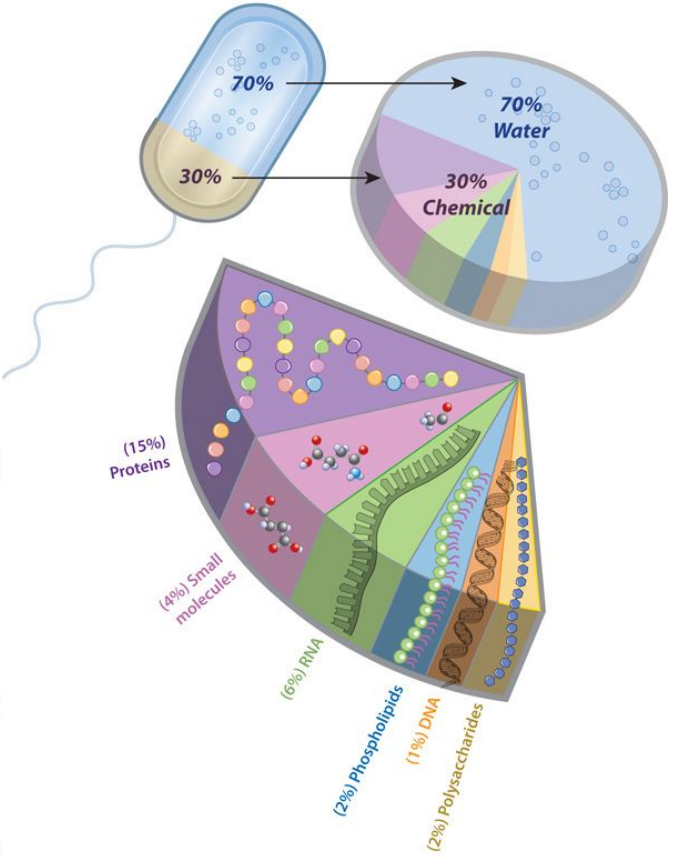
Bottom: Cell membrane, copyright of Encylopedia Britannica, Inc.

Top: Figure from The Human Protein Atlas

Right: Chemical composition of a human cell, by Scitable Nature Education.



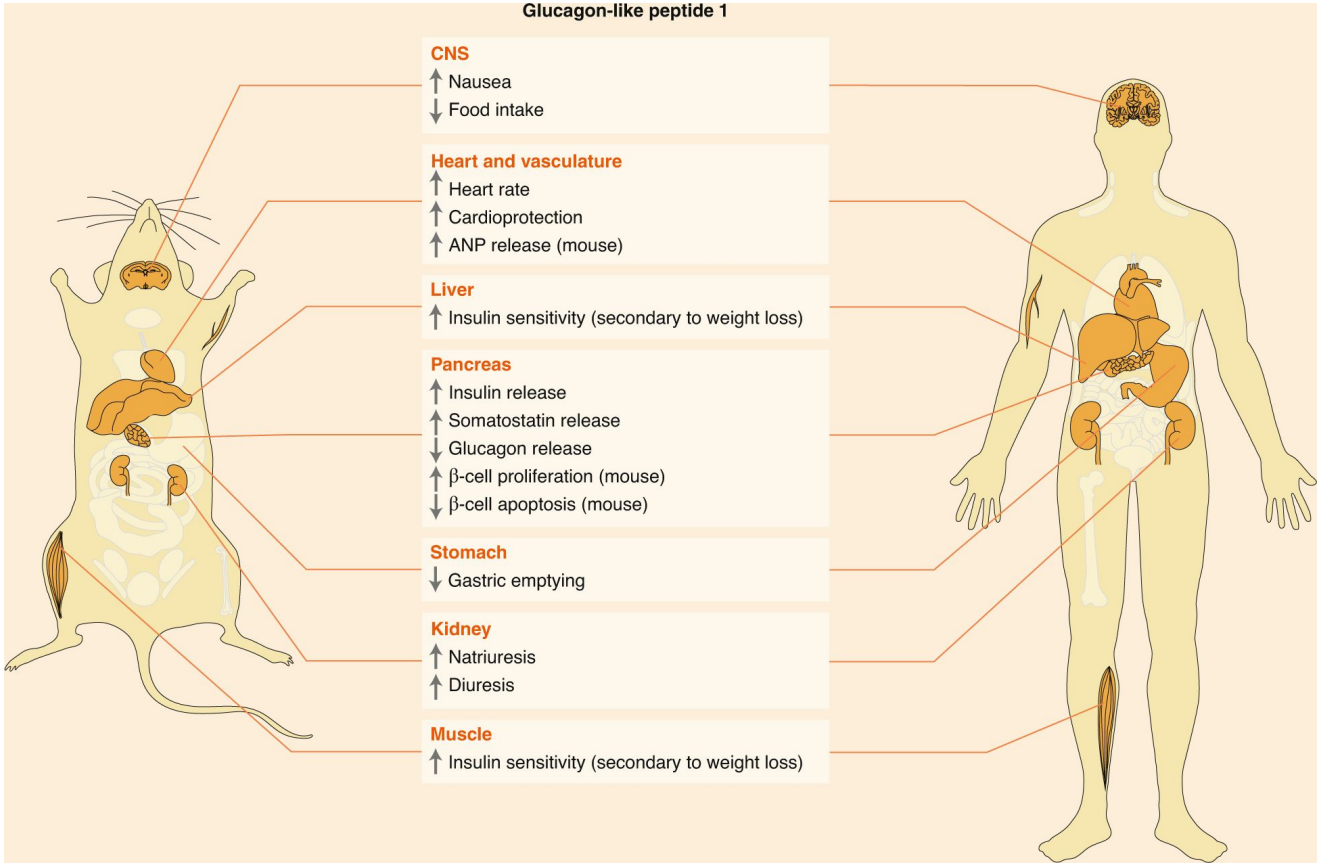




Why drug *discovery*? 6. The drug can have organ-specific and systemic effects, causing either benefits or risks

Direct effects of Glucagon-like peptide (GLP-1) and GLP1 receptor agonists (GLP1-RA) like semaglutide.

Gribble, Fiona M., and Frank Reimann. "Metabolic Messengers: Glucagon-like Peptide 1." Nature Metabolism 3, no. 2 (February 2021): 142–48.



Why drug *discovery*? 7. Do all patients benefit from the drug, or only some of them? Learn from the story of Herceptin

Link to the video

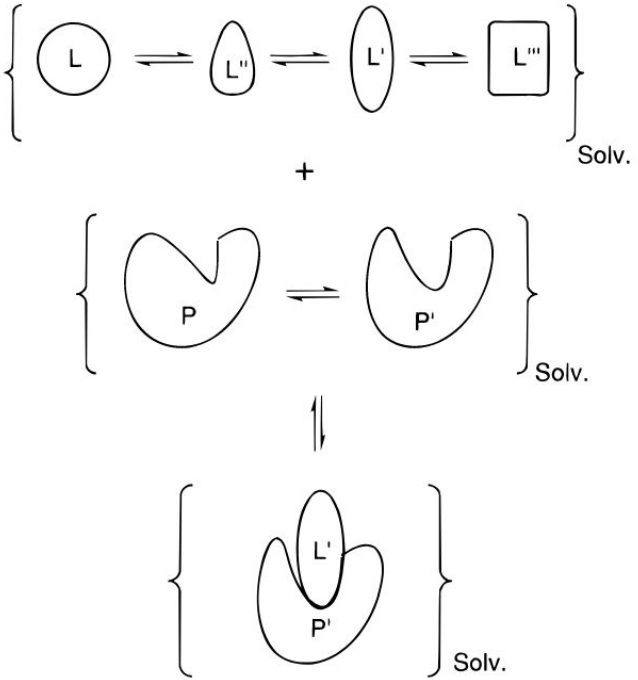
Questions for the video

1. What is the **indication** of *Herceptin*? What is its generic (USAN, or United States Adopted Name) name?
2. What is the **gene target** of Herceptin?
3. Which class best describes the target: Enzyme, Ion channel, Receptor and Kinase, or Structural protein?
4. In which year was the **target** of Herceptin described? When was Herceptin **approved**?
5. What was the **improvement** of Herceptin compared with earlier antibodies?
6. Why does a **biomarker** matter besides developing drugs?
7. In the clinical trial of *Herceptin* for **metastatic breast cancer**, how much improvement in the **median survival** did Herceptin achieve? And how much improvement is in the **adjuvant setting** (Herceptin applied directly after operation)?

Questions for further thinking

- Susan Desmond-Hellmann summarizes successful drug development in four aspects: (1) having a deep understanding of the basic science and the characteristics of the drug, (2) targeting the right patients, (3) setting a high bar in the clinic, and (4) working effectively with key regulatory decision makers. Where do you think mathematics and computer science play a crucial role?
- She emphasized the importance of collaboration. What skill sets do we need for that?
- How do you like her presentation? Anything that you can learn from her about presentation and storytelling?

Why drug *discovery*? 8. Conformational changes, water, and precision make modelling of protein-ligand interactions challenging



rel K_d	ΔG (kcal mol^{-1})
5	0.96
10	1.37
29	2.00
100	2.73
156	3.00
840	4.00
4526	5.00

Free energy change equals the sum of entropic and enthalpic changes. Forming a complex reduces entropy: highly favorable enthalpic contacts between the protein and the ligand are therefore necessary. Small ΔG differences translate to huge K_D differences (see table above), therefore a computational model must have very high accuracy (ideally ±1-2 kcal/mol) modelling a complex system to predict K_D well.

Babine, Robert E., and Steven L. Bender. "Molecular Recognition of Protein−Ligand Complexes: Applications to Drug Design." Chemical Reviews 97, no. 5 (1997): 1359–472. Recent review: Cournia, Zoe, Bryce Allen, and Woody Sherman. "Relative Binding Free Energy Calculations in Drug Discovery: Recent Advances and Practical Considerations." Journal of Chemical Information and Modeling 57, no. 12 (2017).