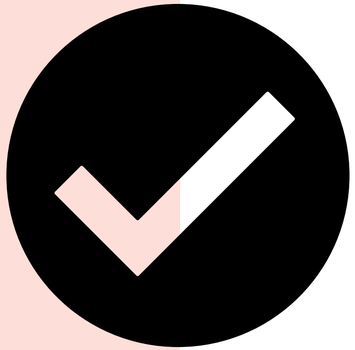
Cellular and Molecular Biology I

BIO-205-9

Camille Goemans - 2024

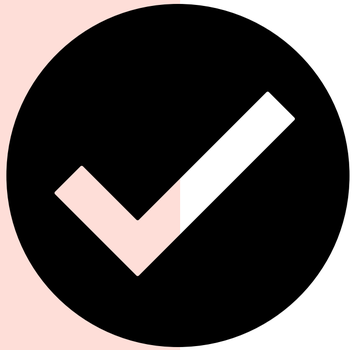
Quick recap

- Lecture 1 - Structure and packing of DNA
- Lecture 2 - Structure of chromosomes and how genomes evolve
- Lecture 3 - Mobile genetic elements and DNA replication
- Lecture 4 - DNA replication and DNA repair



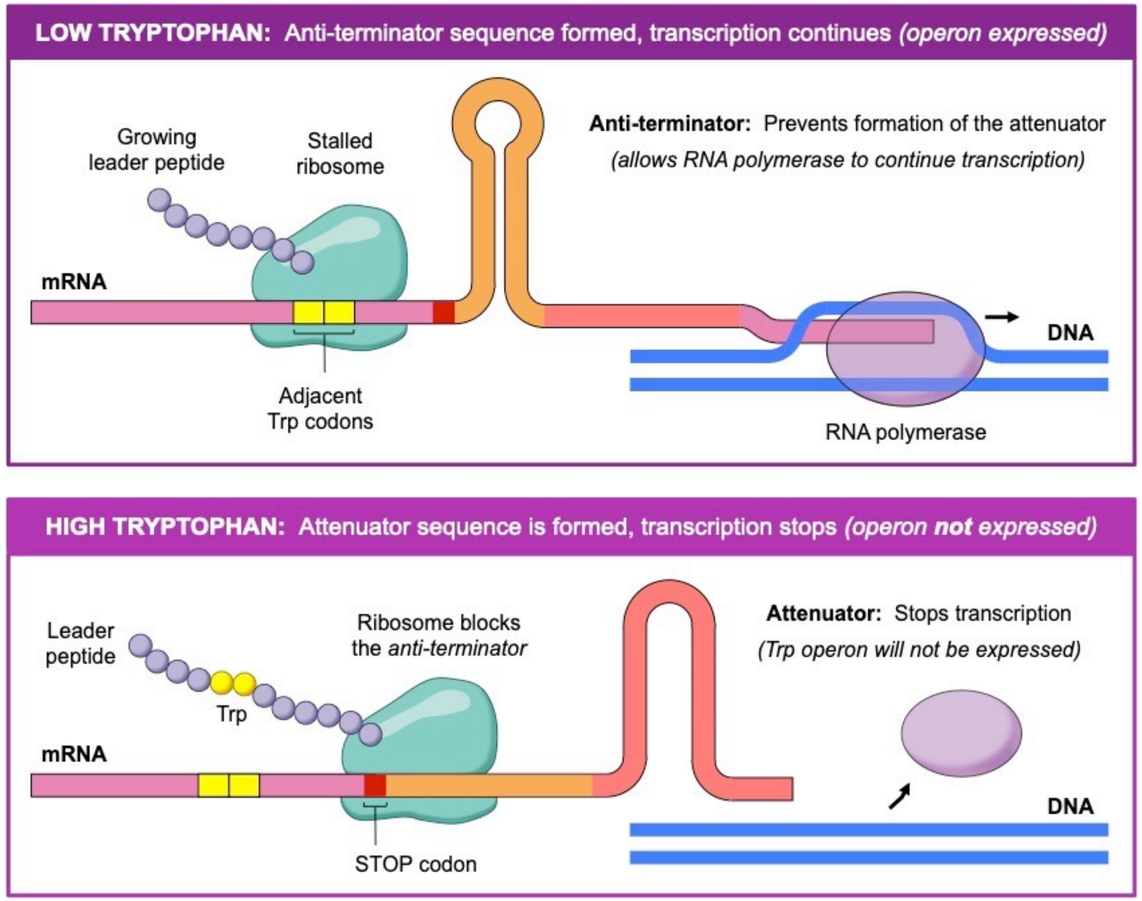
Quick recap

- Lecture 5 - Transcription and RNA processing
- Lecture 6 - Translation and protein folding
- Lecture 7 - Transcriptional control
- Lecture 8 - Post-transcriptional control



Recap: Transcription attenuation

The expression of some genes is inhibited by **premature termination of transcription**



Recap: Transcription attenuation

Attenuation is an alternative method of reducing the expression of the trp operon in prokaryotic cells

- It relies on the capacity for prokaryotes to undertake transcription and translation at the same time (as the DNA is cytosolic)
- Whereas the trp repressor system blocks the initiation of transcription, attenuation blocks the completion of transcription

Located between the operator and trp genes is a leader sequence that encodes a short polypeptide containing two Trp residues

- When transcribed, this leader sequence contains several self-complementary sections that are capable of forming hairpin loops
- One such loop is called the attenuator and functions as a site of transcriptional termination when it is formed (transcription halts)
- Alternatively, an earlier loop called the anti-terminator may form, which prevents attenuator formation (so transcription continues)

The presence of tryptophan determines whether the attenuator or anti-terminator is formed within the leader sequence

- If tryptophan is in low supply, the ribosome will pause at the two Trp codons within the leader sequence
- This pause allows time for the anti-terminator to form on the mRNA transcript, so transcription will continue (operon is expressed)
- If tryptophan is in high supply, the ribosome will not need to pause at the two Trp codons within the leader sequence
- This means the anti-terminator is unable to form and the attenuator will form instead (transcription halted ; operon not expressed)

Recap: miRNA, siRNA and piRNA

MicroRNA (miRNA), small interfering RNA (siRNA), and piwi-interacting RNA (piRNA) are all small RNA molecules involved in gene regulation, primarily by silencing or down-regulating target genes. However, they differ in their origins, structures, mechanisms, and biological roles.

Recap: miRNA, siRNA and piRNA

1. miRNA (MicroRNA)

- Origin: miRNAs are endogenously encoded in the genome. They are transcribed by RNA polymerase II, forming primary miRNAs (pri-miRNAs) that are processed in the nucleus and cytoplasm to mature miRNAs.

- Structure: They are usually 21-23 nucleotides long and have imperfect base-pairing, leading to stem-loop structures in precursor molecules.

- Function: miRNAs primarily regulate gene expression by binding to complementary sequences in the 3' untranslated region (3' UTR) of target mRNAs. They usually result in translational repression or mRNA degradation, although the exact mechanism depends on the level of sequence complementarity.

- Role: miRNAs are involved in various cellular processes, including development, differentiation, and disease (e.g., cancer).

Recap: miRNA, siRNA and piRNA

2. siRNA (Small Interfering RNA)

- Origin: siRNAs can be endogenous (e.g., from transposons or viruses) but are often exogenous, such as those introduced experimentally or derived from viral infection.

- Structure: siRNAs are typically 20-25 nucleotides long and have perfect or near-perfect complementarity to their target RNA. They are often derived from longer double-stranded RNAs (dsRNAs) that are processed by the enzyme Dicer into siRNA duplexes.

- Function: siRNAs guide the RNA-induced silencing complex (RISC) to degrade specific mRNAs through a perfect base-pairing mechanism. This direct pairing leads to the target mRNA's cleavage and degradation, reducing gene expression.

- Role: siRNAs are part of the cellular defense against viral infection and transposons. They are also widely used as a tool in research to "knock down" genes in experiments by selectively silencing specific mRNAs.

Recap: miRNA, siRNA and piRNA

3. piRNA (Piwi-interacting RNA)

- Origin: piRNAs are derived from specific genomic regions called piRNA clusters, which are often transposon-rich. They do not require Dicer for processing but are instead processed through a unique biogenesis pathway.

- Structure: piRNAs are generally longer than miRNAs and siRNAs, about 24-31 nucleotides, and show a preference for starting with a uridine at the 5’ end. They do not form duplexes like siRNA and miRNA and do not have a defined secondary structure.

- Function: piRNAs interact with Piwi proteins (a subclass of the Argonaute family) to silence transposable elements, particularly in germ cells, through both transcriptional and post-transcriptional silencing.

- Role: piRNAs play a crucial role in protecting the integrity of the germline genome by silencing transposons and other repetitive elements. They are essential for fertility and maintaining genomic stability in reproductive cells.

Recap: miRNA, siRNA and piRNA

Summary of Differences

Feature	miRNA	siRNA	piRNA
Source	Endogenous, genome-encoded	Endogenous or exogenous (e.g., viruses)	Endogenous, piRNA clusters
Length	~21-23 nucleotides	~20-25 nucleotides	~24-31 nucleotides
Structure	Imperfectly paired, stem-loop	Perfectly paired, double-stranded	Single-stranded, no defined structure
Processing	Dicer-dependent	Dicer-dependent	Dicer-independent
Protein Complex	Argonaute (AGO) family	Argonaute (AGO) family	Piwi subfamily of Argonaute
Function	Translational repression or degradation	mRNA cleavage	Transposon silencing
Primary Role	Gene regulation in development, differentiation	Antiviral defense, gene knockdown	Genome integrity in germ cells

Recap: p-bodies and stress granules

P-bodies, or processing bodies, are dynamic, membrane-less granules within the cytoplasm of eukaryotic cells. They serve as hubs for mRNA metabolism, particularly for mRNA decay and storage. P-bodies are involved in post-transcriptional regulation, influencing gene expression by controlling the stability, translation, and degradation of mRNA.

Stress granules are dynamic, membrane-less cytoplasmic aggregates of RNA and protein that form in response to cellular stress, such as oxidative stress, heat shock, or nutrient deprivation. They play a crucial role in helping cells survive by temporarily halting translation and preserving select mRNAs under adverse conditions. This helps conserve energy and resources, allowing the cell to prioritize immediate stress response needs while protecting non-essential mRNAs for future use.

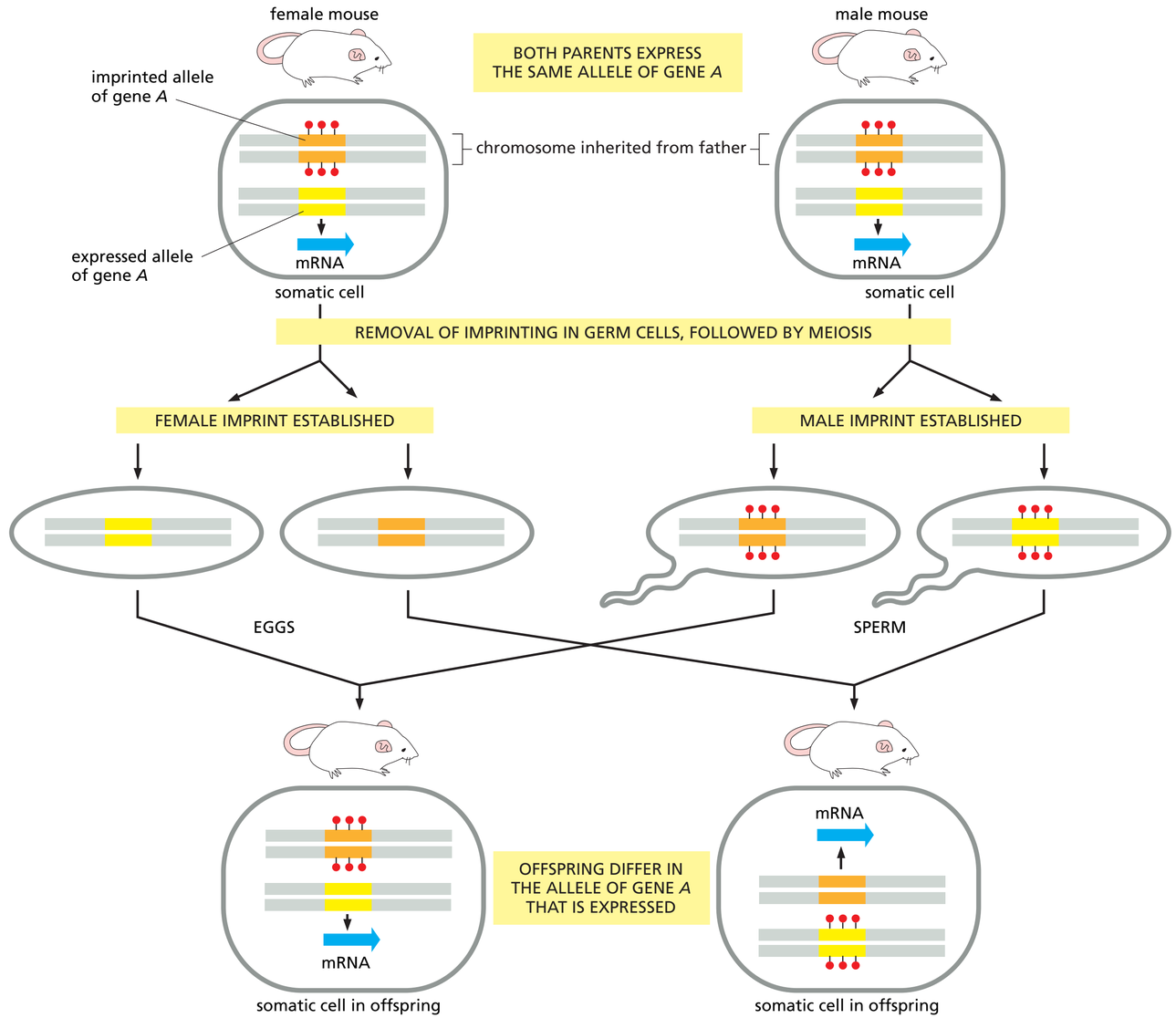
Recap: X-inactivation

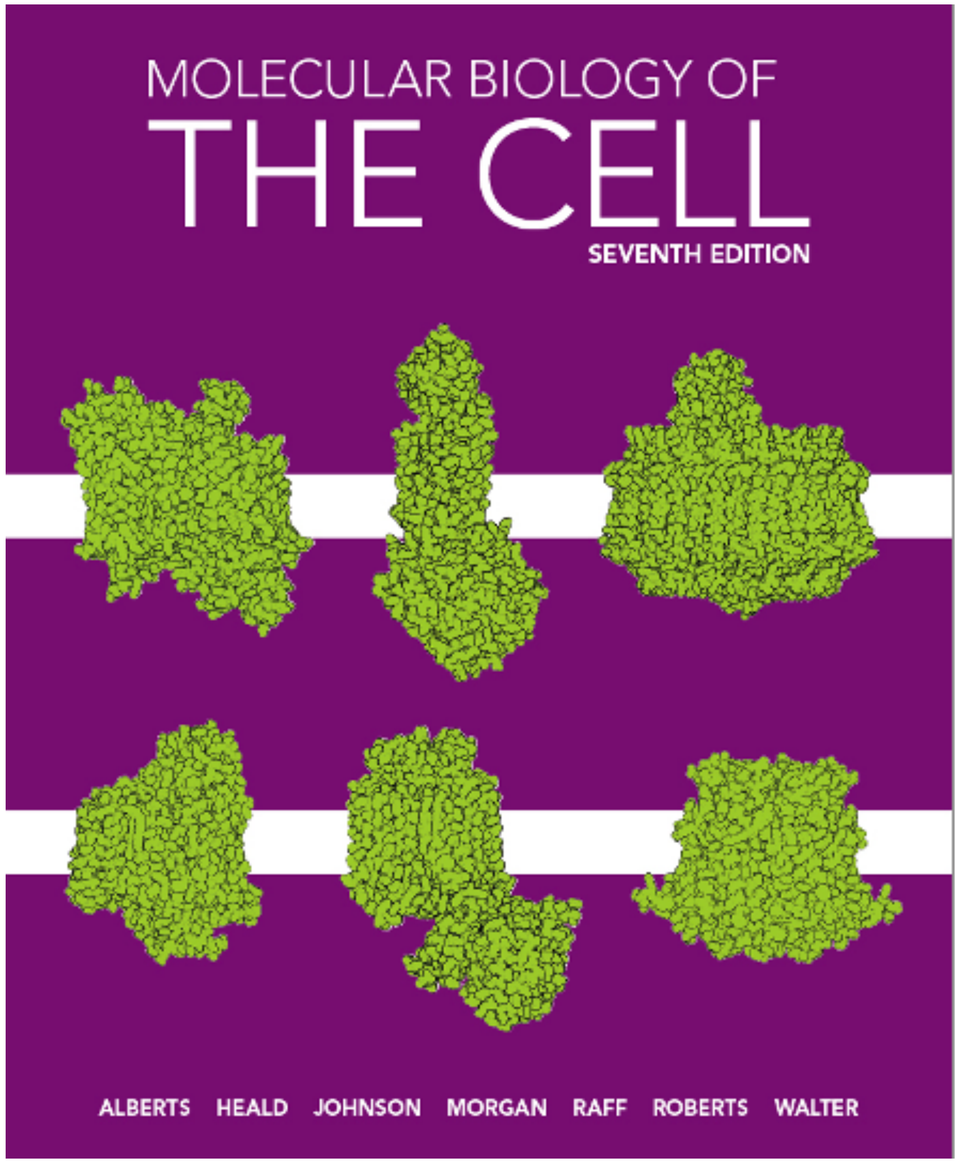
X-inactivation, a process occurring in female mammals to ensure dosage compensation, randomly silences one of the two X chromosomes in each somatic cell. This ensures that females (XX) have the same effective dose of X-linked genes as males (XY), who have only one X chromosome.

The exact **mechanism by which the X chromosome to be inactivated is chosen** is still not fully understood, but several factors contribute:

- **Counting Mechanism**: The cell has a counting mechanism that ensures only one X chromosome remains active. If more than two X chromosomes are present (as in some cases of aneuploidy, like XXX individuals), all but one X will be inactivated.
- **Mutual Repression of Xist and Tsix**: Xist has an antisense partner called Tsix, which is also expressed from the XIC and acts to suppress Xist expression on the chromosome that remains active. Tsix and Xist are in a mutually exclusive relationship—only one X chromosome will end up expressing Xist (and become inactivated), while the other expresses Tsix (and remains active).
- **Stochastic Expression and Epigenetic Factors:** The random choice may involve stochastic (random) fluctuations in the levels of Xist and Tsix. Once one chromosome "wins" by expressing higher levels of Xist or Tsix, this choice is stabilized by epigenetic modifications (such as DNA methylation and histone modification) to ensure only one X remains active.

Genomic imprinting: somatic vs. germinal cells





Chapter 8

Analyzing Cells, Molecules, and Systems

Copyright © 2022 W. W. Norton & Company, Inc.

How do know all of this?

Progress in **Science** are linked to **advances in technology**

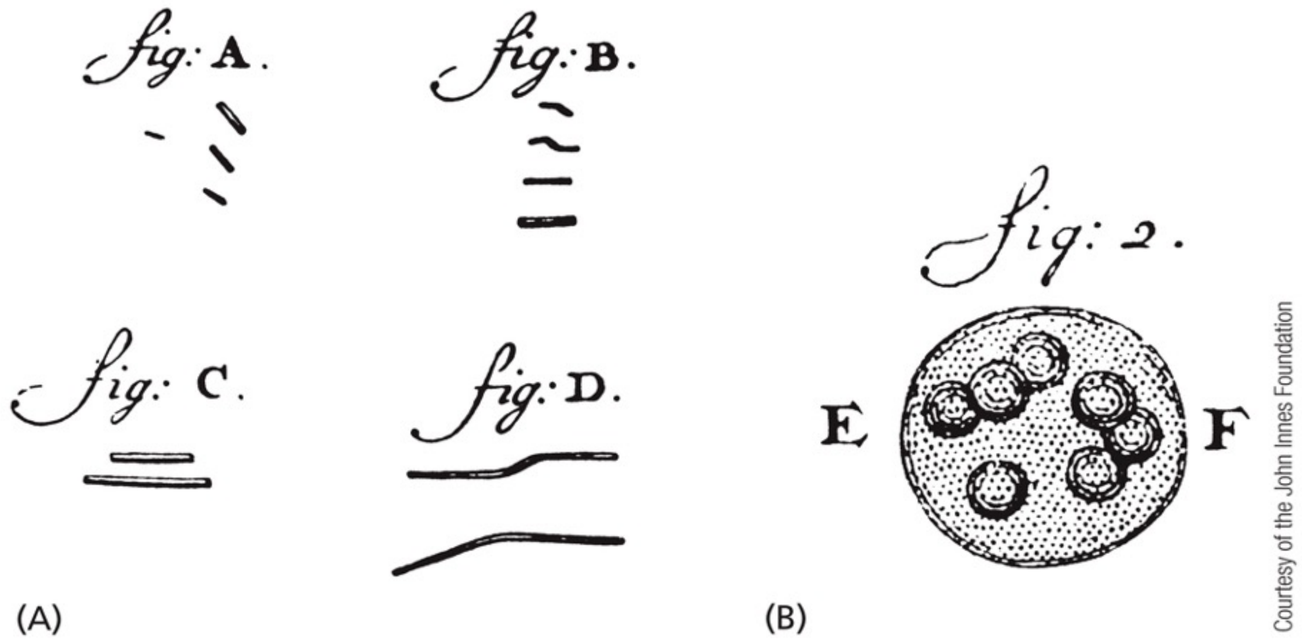


Figure 8–1 Microscopic life. A sample of “diverse animalcules” seen by van Leeuwenhoek using his simple microscope. (A) Bacteria seen in material he excavated from between his teeth. Those in fig. B he described as “swimming first forward and then backwards” (1692). (B) The eukaryotic green alga Volvox (1700). (Courtesy of the John Innes Foundation.)

How do know all of this?

Biochemistry

Genetics

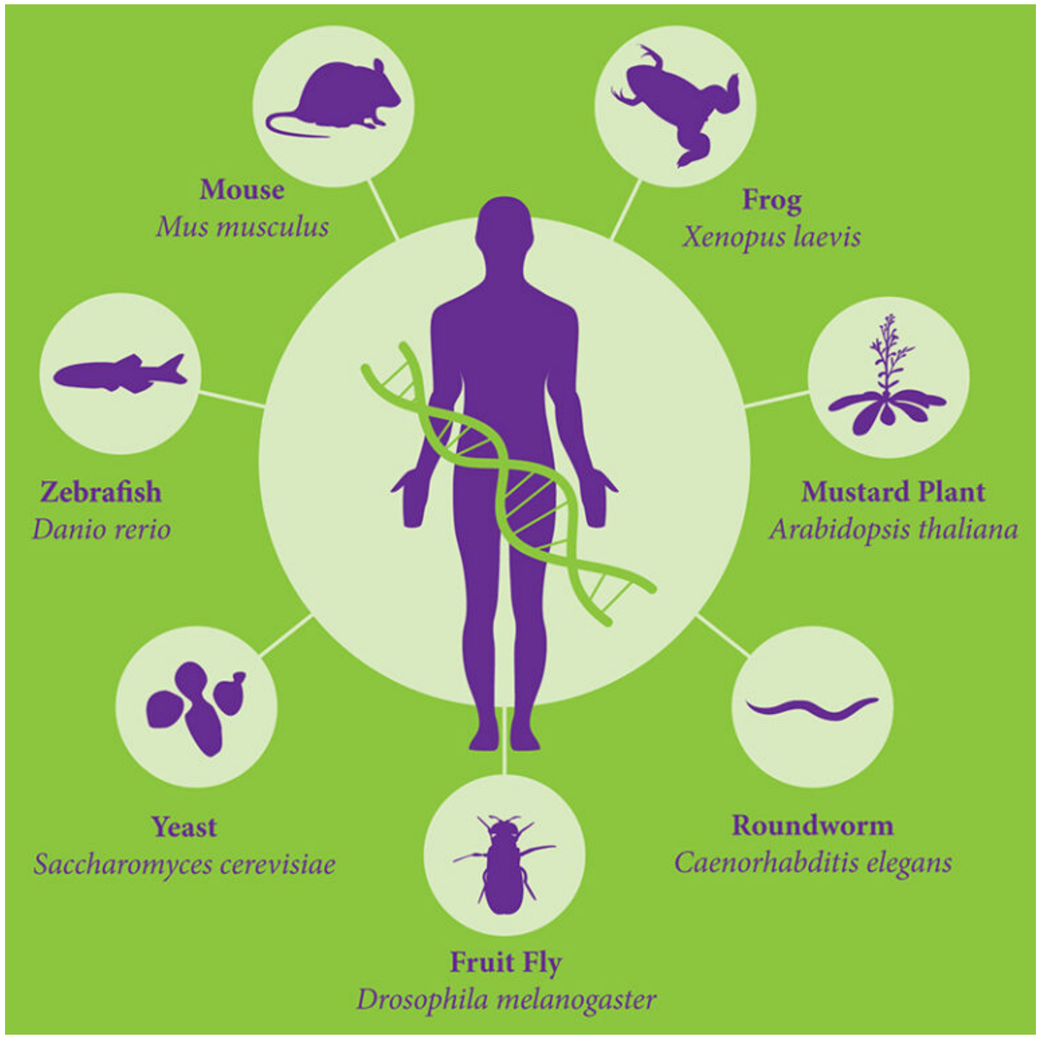
Molecular Biology - need to simplify the systems

Plan

- **Model organisms**
- Isolating cells and growing them in culture
- Studying proteins
    - Protein sequence
    - Protein purification
    - Protein structure
    - Protein visualization
    - Mass spectrometry

Model organisms

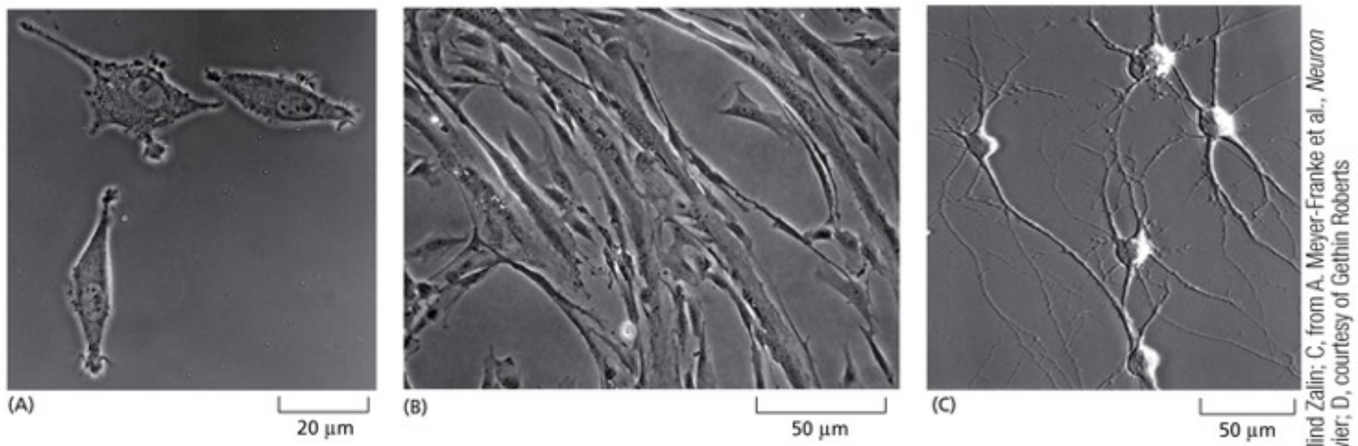
Bacteria: *Escherichia coli, Bacillus subtilis*

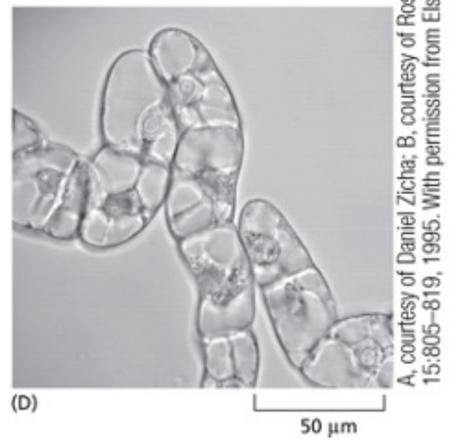


Plan

- Model organisms
- **Isolating cells and growing them in culture**
- Studying proteins
    - Protein sequence
    - Protein purification
    - Protein structure
    - Protein visualization
    - Mass spectrometry

Why do we need to isolate cells?





A, courtesy of Daniel Zicha; B, courtesy of Rosalind Zalin; C, from A. Meyer-Franke et al., *Neuron* 15:805–819, 1995. With permission from Elsevier; D, courtesy of Gethin Roberts

- To produce **proteins** in large amounts (unicellular organisms)
- To get information about **specific cell types** in a tissue (human or mammals cells)
- Cultured cells provide a **homogeneous population** to work with in the lab or from which to extract material
- Experiments carried on cells are typically referred to as ***"in vitro"*** as opposed to experiments with intact organisms (***in vivo***)

How to isolate cells?

- Disrupt the **extracellular matrix and cell-cell junctions**
    - with **proteolytic enzymes** (trypsin and collagenase) to digest proteins in the extracellular matrix
    - with agents like **EDTA** (ethylenediaminetetraacetic-acid) that chelate (bind) Ca^{2+} necessary for cell-cell adhesion
- Many cells require a **solid surface** to grow on and divide (plastic culture dish)
- Many cells require **specific materials to coat the culture dish** (polylysine or extracellular matrix components) to grow and proliferate

Primary cultures are prepared directly from the organism. They can be grown and re-cultured (**passaged**) repeatedly into **secondary cultures** for weeks or months. This way, most cells keep their original properties.

Embryonic stem cells are **pluripotent**. They can differentiate into **any cell type** in the body. *In vitro*, they need the **appropriate signalling factors and nutrients** to differentiate into specific cell types.

Eukaryotic cell lines

- Cell cultures have one problem: cells **die;** they stop dividing after a number of division (replicative cell senescence)
- **Fibroblasts** divide **25 to 40 times in culture**
- This reflects the **shortening and uncapping of telomeres,** following the **decreased production** of the telomerase
- By providing a **telomerase,** fibroblasts can be propagated as an **immortalized cell line**
- For some other cell types, the process is **more complicated**
- Cell lines are often derived from **cancer cells** (transformed cell lines)
- Can be stored in **liquid nitrogen** at -196 degrees

TABLE 8–1 Some Commonly Used Cell Lines

Cell line*	Cell type and origin
3T3	Fibroblast (mouse)
BHK21	Fibroblast (Syrian hamster)
MDCK	Epithelial cell (dog)
HeLa	Epithelial cell (human)
PtK1	Epithelial cell (rat kangaroo)
L6	Myoblast (rat)
PC12	Chromaffin cell (rat)
SP2	Plasma cell (mouse)
COS	Kidney (monkey)
293	Kidney (human); transformed with adenovirus
CHO	Ovary (Chinese hamster)
DT40	Lymphoma cell for efficient targeted recombination (chick)
R1	Embryonic stem cell (mouse)
E14.1	Embryonic stem cell (mouse)
H1, H9	Embryonic stem cell (human)
S2	Macrophage-like cell (*Drosophila*)
BY2	Undifferentiated meristematic cell (tobacco)

*Many of these cell lines were derived from tumors. All of them are capable of indefinite replication in culture and express at least some of the special characteristics of their cells of origin.

Hybridoma cell lines and monoclonal antibodies

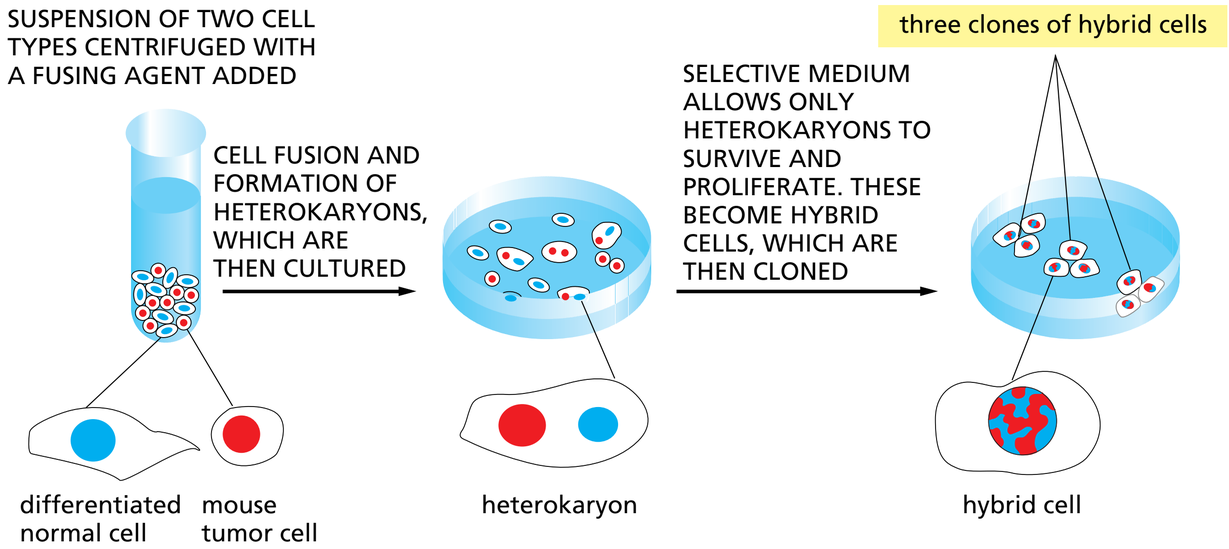
- **Antibodies** are extremely useful tools in the lab and are now also used to treat cancers
- Their **great specificity** allows to **detect selected proteins** among thousands
- They are produced by **inoculating an animal** with the **purified protein** of interest and further isolation of the antibodies from the **serum**

Polyclonal antibodies are a mixture of antibodies that recognise **different antigenic sites** (epitopes) on the protein of interest

Hybridoma cell lines and monoclonal antibodies

Monoclonal antibodies are identical antibodies that recognise **one antigenic site** on the protein of interest

- Needs the propagation of a **clone of cells** from a **single antibody-secreting B lymphocyte**
- B lymphocytes have a **limited life-span** so they are isolated from animals and **fused with a transformed (cancer-derived) lymphocyte cell line**
- Each hybrid cell produces a single type of **monoclonal antibody**

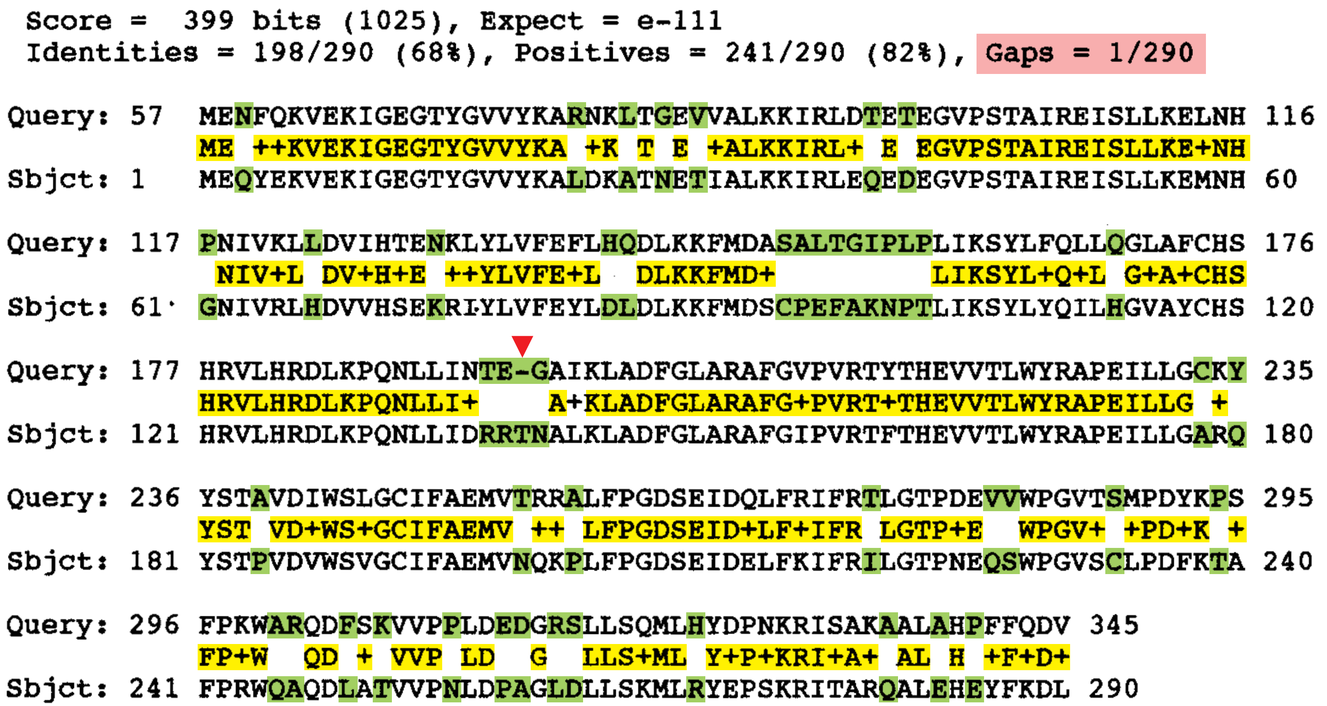


Plan

- Model organisms
- Isolating cells and growing them in culture
- **Studying proteins**
    - **Protein sequence**
    - Protein purification
    - Protein structure
    - Protein visualization
    - Mass spectrometry

Protein sequence

- Identification of a **protein** but no information on what it does
- Comparison of proteins in **databases (Basic Local Alignment Tool, BLAST)**



Plan

- Model organisms
- Isolating cells and growing them in culture
- **Studying proteins**
    - Protein sequence
    - **Protein purification**
    - Protein structure
    - Protein visualization
    - Mass spectrometry

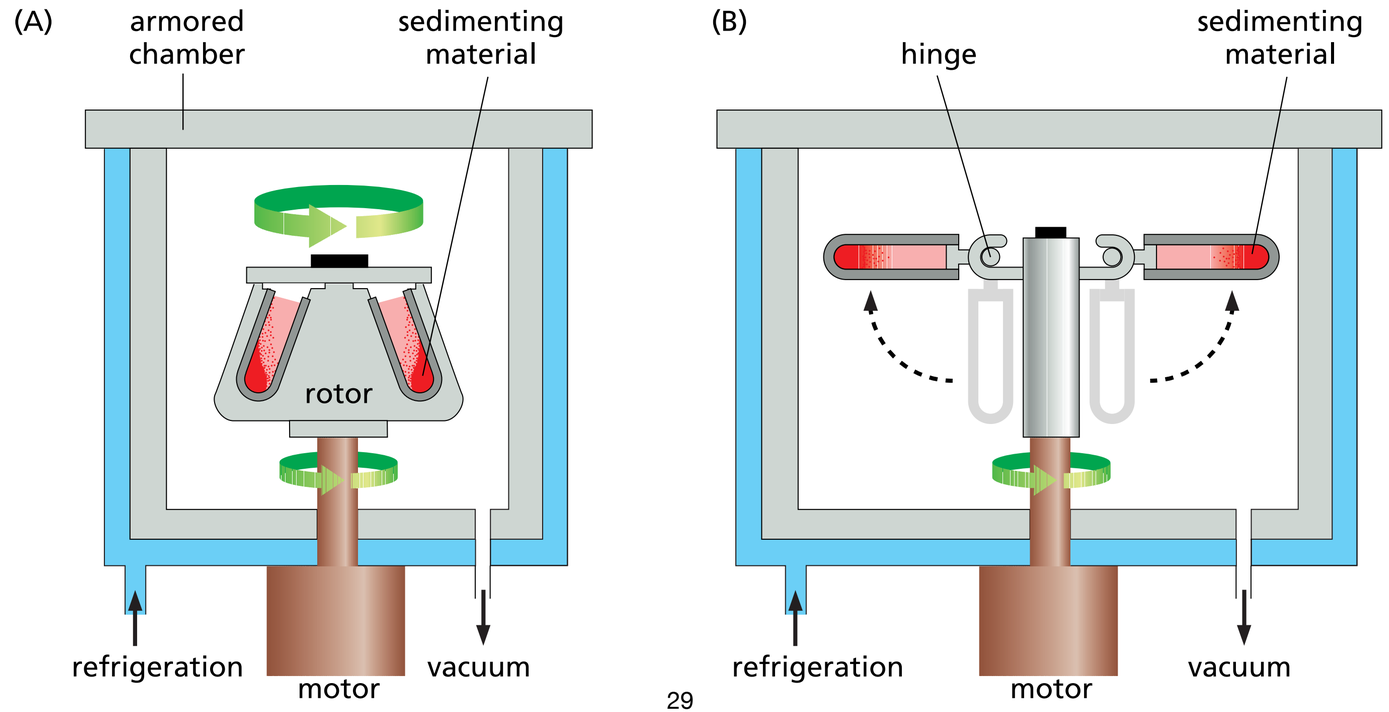
Protein purification

Why do we need to purify proteins?

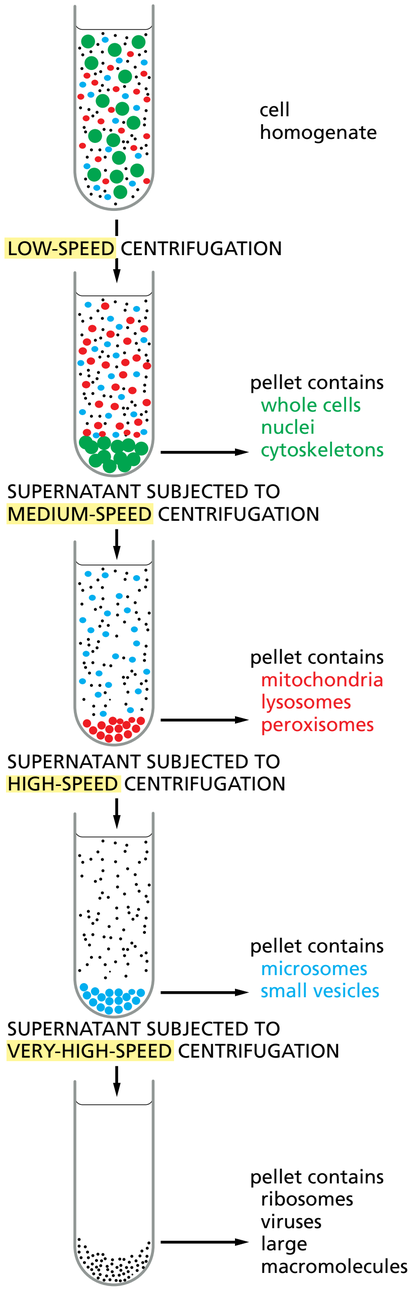
- To **separate** them from the other thousands of proteins in each cell
- To **study** them *in vitro* (perform biochemical assays)
- To get their **3D structure**

Step 1: opening the cells

- Osmotic shock, ultrasonic vibrations, ground up in a blender, …
- These processes **break the cell membranes** and release the content inside small membrane vesicles
- These processes **leave most of the organelles intact**
- The components of this **“homogenate”** can be separated by **ultracentrifugation,** that separates the elements by **size and density**

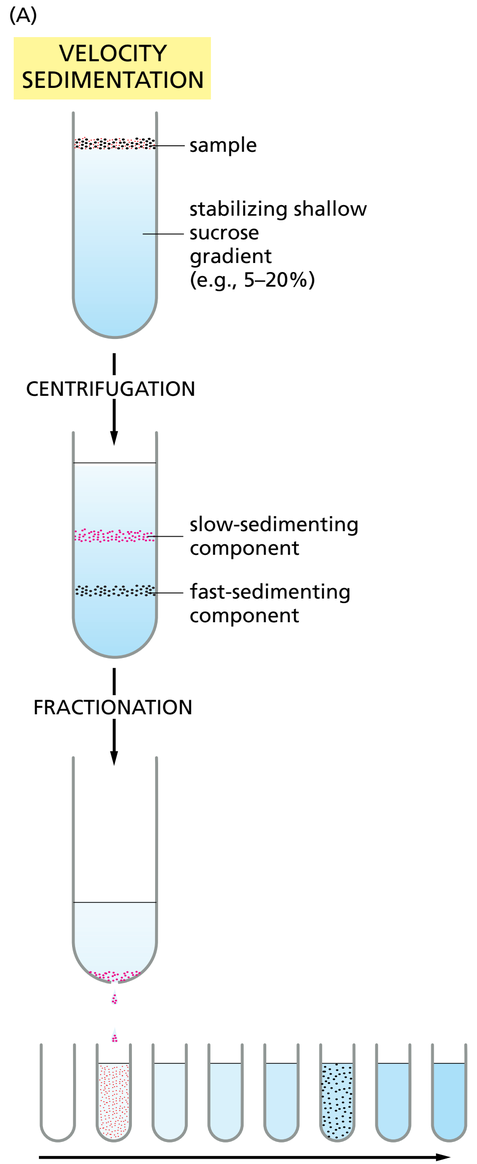


Step 1: opening the cells



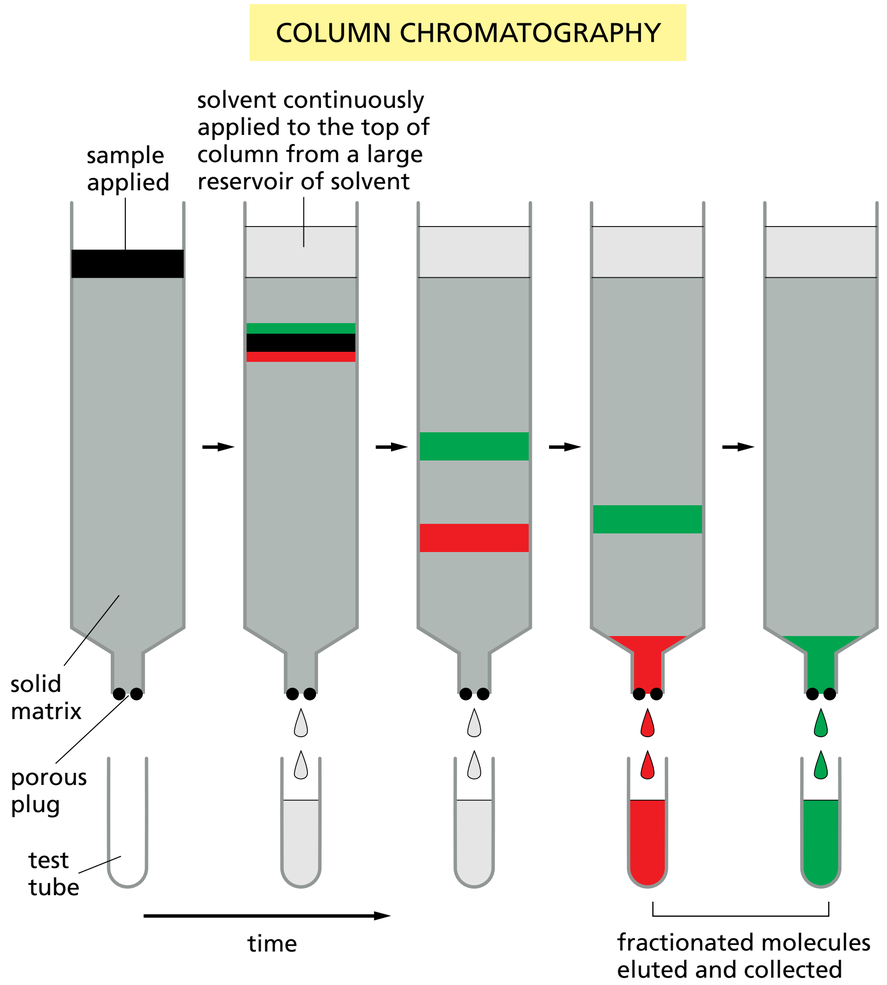
Step 1: opening the cells

- Improving separation using **gradients**



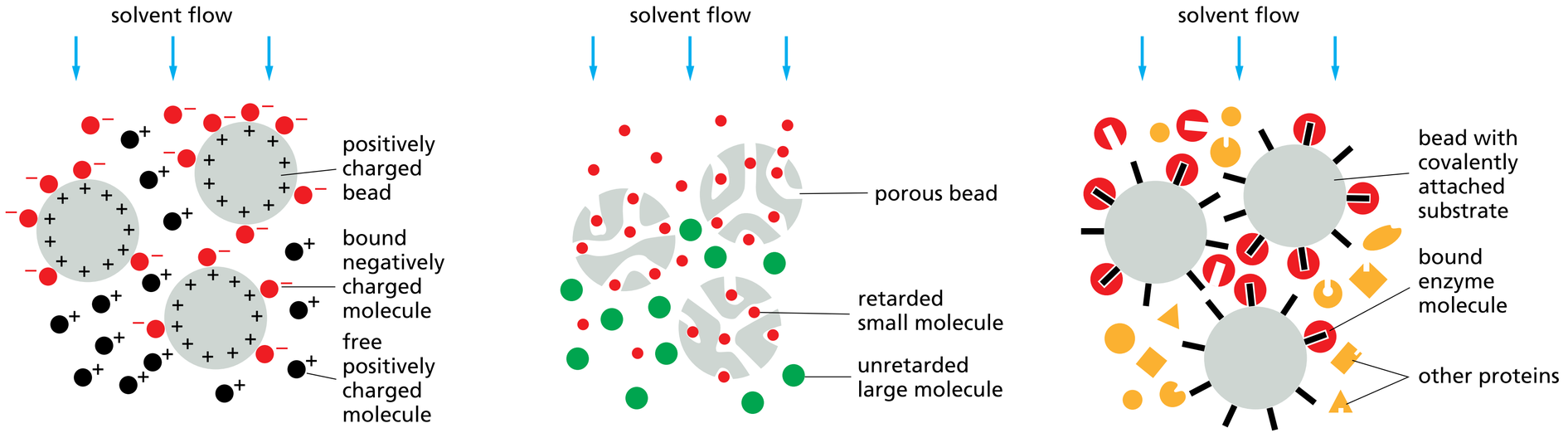
Step 2: separating the proteins

- Using **column chromatography**
- A mixture of proteins is passed through a **porous gel matrix**
- Different proteins are **retarded to different extents** depending on their **interaction with the matrix**
- They are **collected separately as they come out** of the column

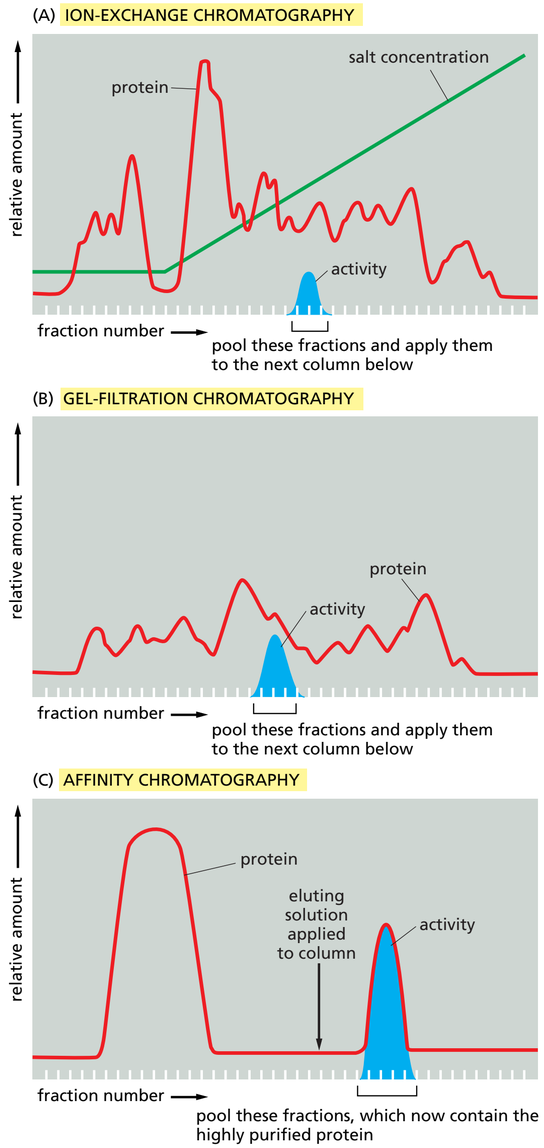


Step 2: separating the proteins

- Depending on the **matrix**, proteins can be separated according to
  - their **charge** (ion-exchange chromatography),
  - their **hydrophobicity** (hydrophobic chromatography)
  - their **molecular weight** (gel filtration = size-exclusion chromatography)
  - their **ability to bind small molecules** (affinity chromatography)



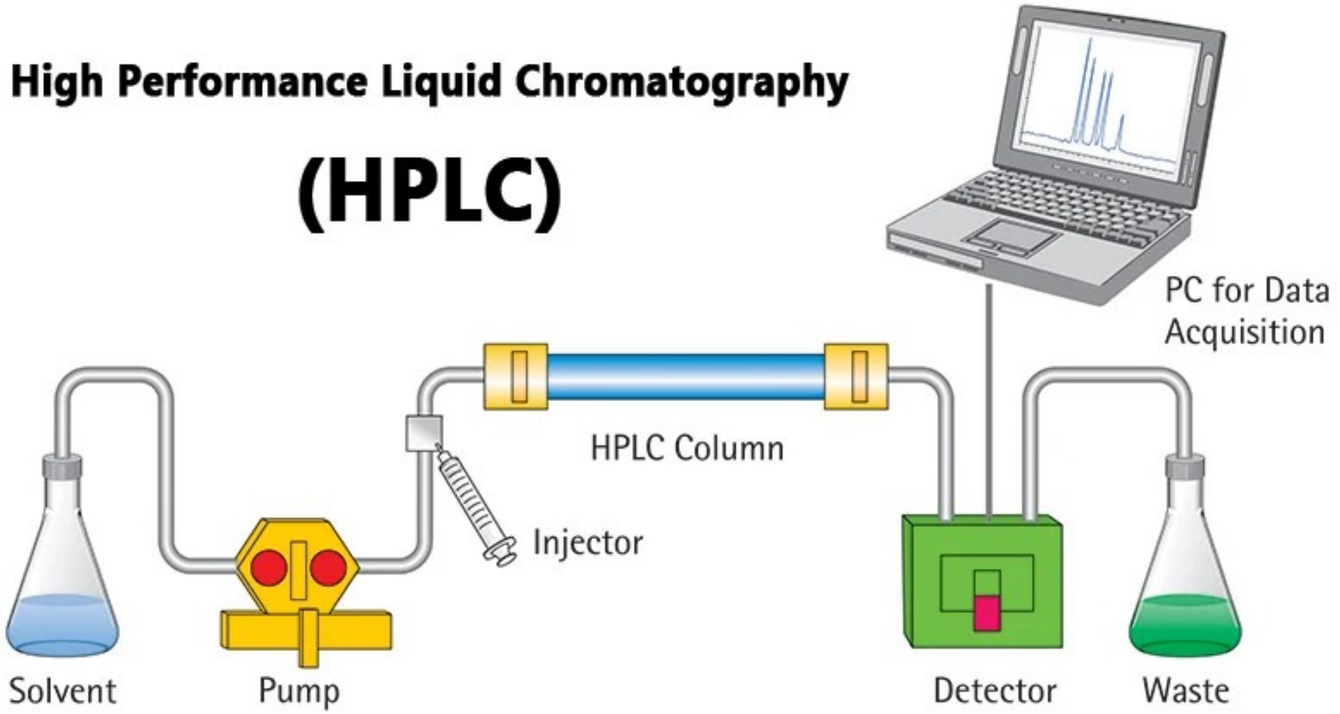
Step 2: separating the proteins



Step 2: separating the proteins

- The **resolution** of classical chromatography columns is **limited**

- Special resins (silica-based) composed of tiny spheres (3-10 um) can be packed to form a uniform column bed to attain a high degree of resolution = **high-performance liquid chromatography (HPLC)**

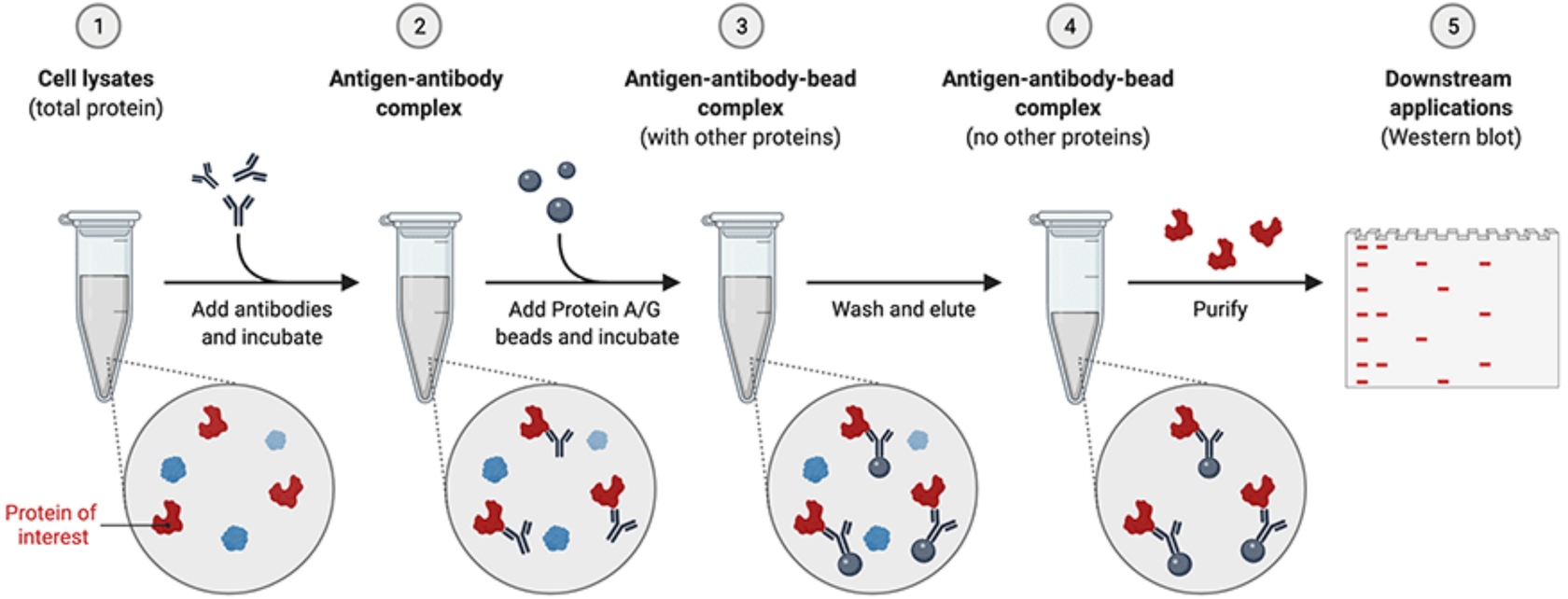


https://noguchi.ug.edu.gh/services/hplc-analysis/

https://microbenotes.com/high-performance-liquid-chromatography-hplc/

Step 2: separating the proteins

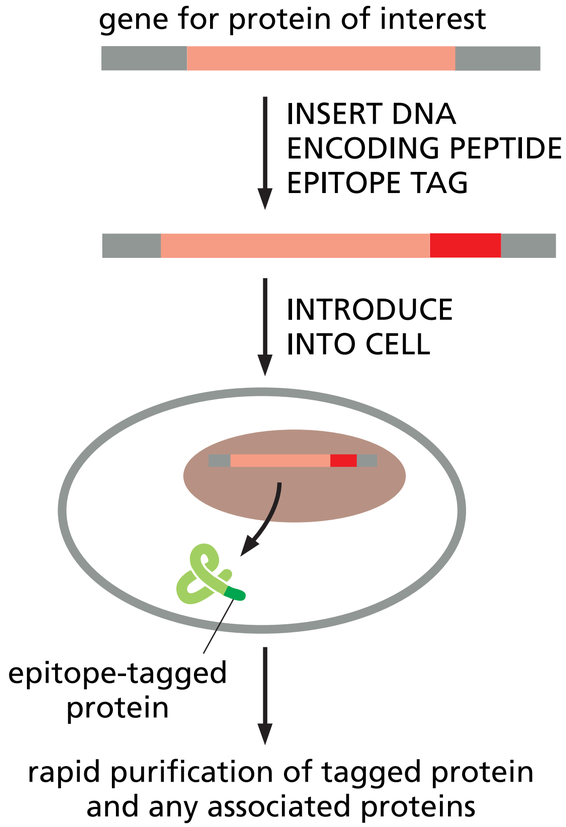
- **Immunoprecipitation (IP)** as a rapid affinity purification



https://www.rockland.com/resources/immunoprecipitation-technique/

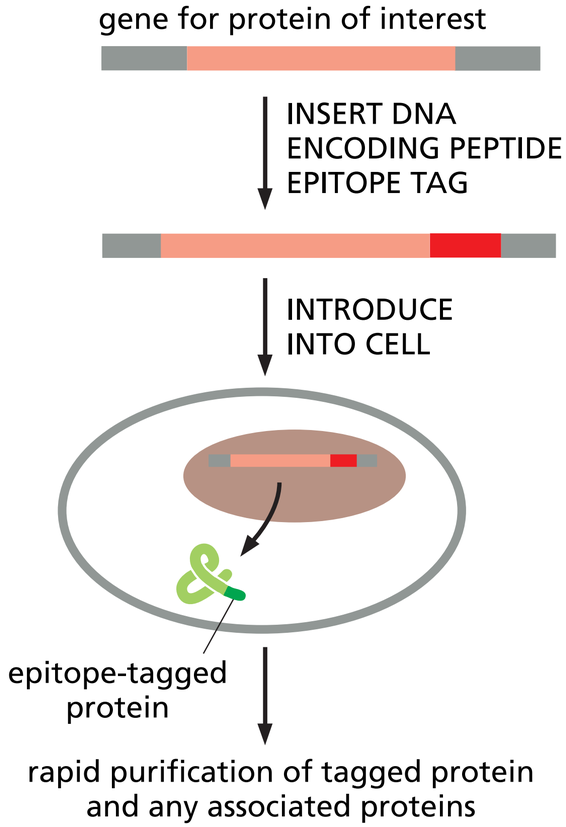
Step 2: separating the proteins

- **Protein tags** provide an easy way to purify proteins
  - His-Tag (addition of 8 His residues) strongly binds to Nickel
  - GST (glutathione-S-transferase, entire protein)
  - **Cleavage site** between the protein and the tag



Step 2: separating the proteins

- **Protein tags** provide an easy way to purify proteins
  - His-Tag (addition of 8 His residues) strongly binds to Nickel
  - GST (glutathione-S-transferase, entire protein)
  - **Cleavage site** between the protein and the tag



Plan

- Model organisms
- Isolating cells and growing them in culture
- **Studying proteins**
    - Protein sequence
    - Protein purification
    - **Protein structure**
    - Protein visualization
    - Mass spectrometry

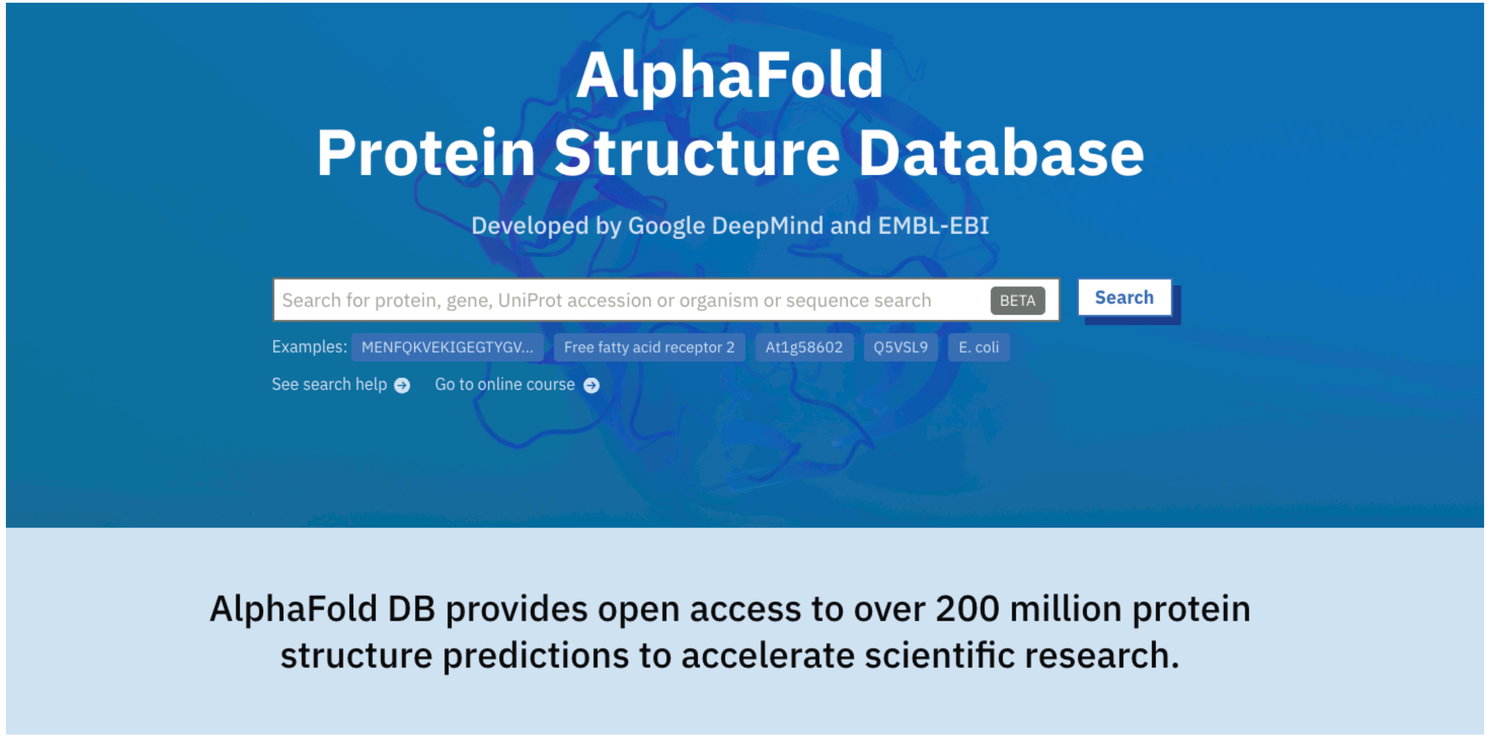
Solving protein structures

➡ Mapping the **precise position** of each amino-acid within the **3D structure**

- **AlphaFold**
- **X-ray cristallography**
- **Nuclear Magnetic Resonance Spectroscopy**
- **CryoEM**

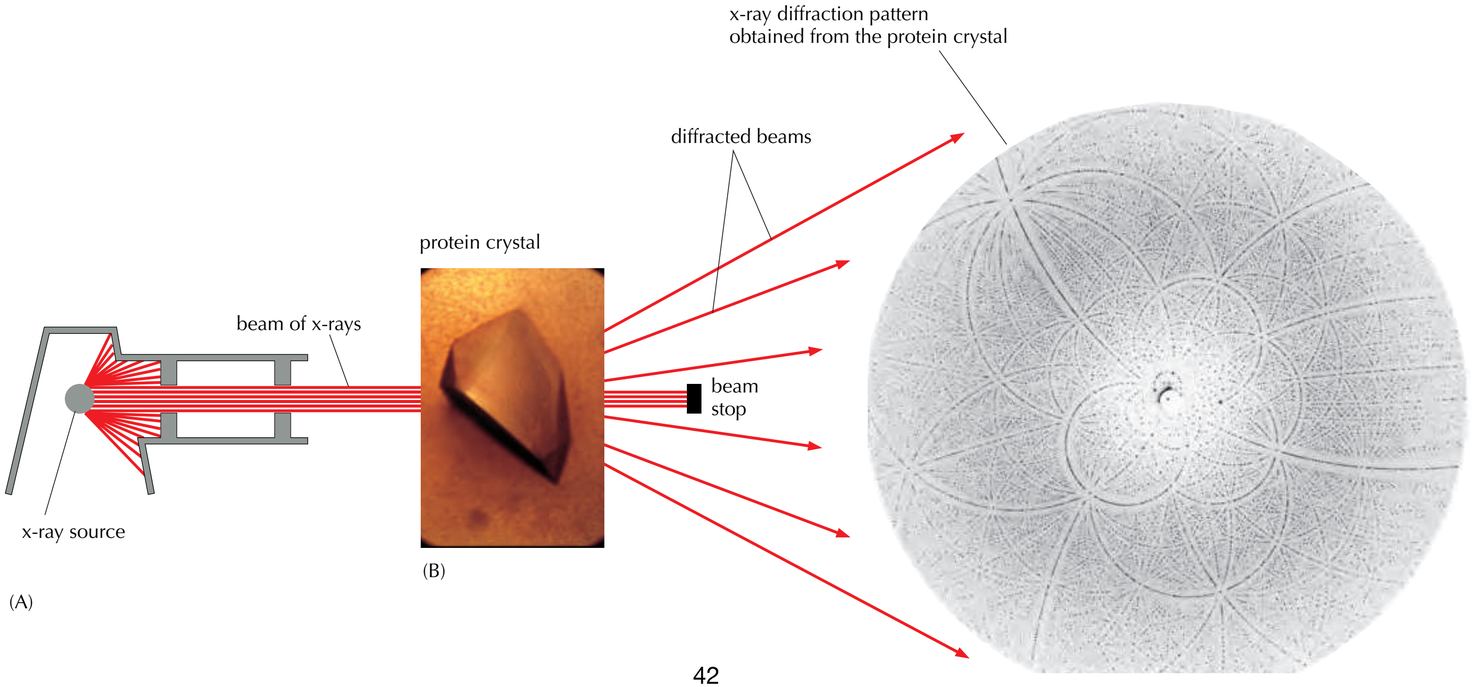
Solving protein structures

- AlphaFold



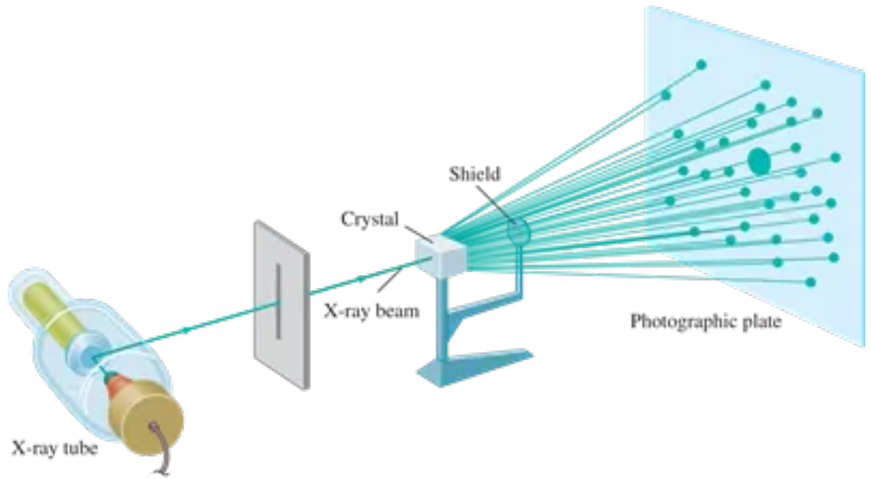
Solving protein structures

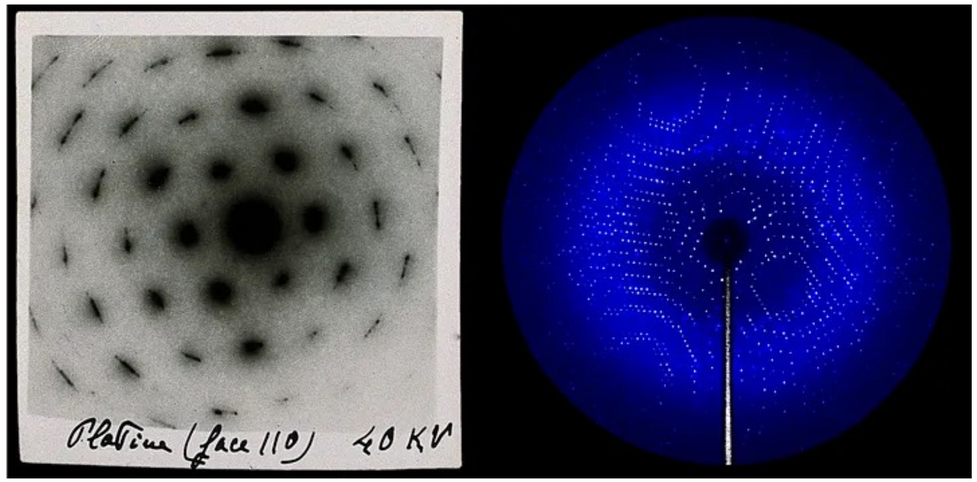
- **X-ray cristallography**
  - X-rays are a form of **electromagnetic radiation** with short wavelength (~0,1 nm)
  - If a narrow beam of X-rays is directed at a sample of **pure proteins**, most pass through it but a small fraction is **scattered by the atoms of the protein**
  - If the sample is a **well-ordered crystal,** the scattered waves reinforce each other and **diffraction spots** appear on the detector



Solving protein structures

- **X-ray cristallography**
  - Slowest step: getting a **protein crystal** (large amounts of very pure protein)
  - The position and intensity of each spot of the **diffraction pattern** contain information on the **location of the atoms in the crystal**
  - Computational methods to generate a **3D electron-density map**
  - By trial and error, the map and sequence are correlated by a computer to provide the **best possible fit**





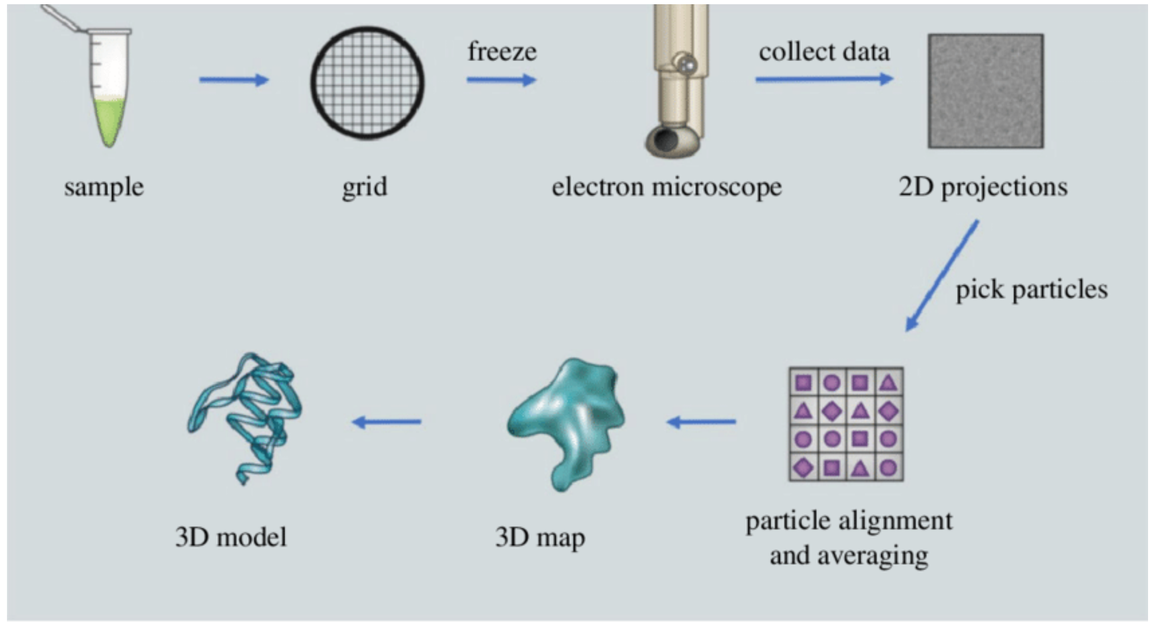
https://macromoltek.medium.com/what-is-x-ray-crystallography-1e186bc3d180

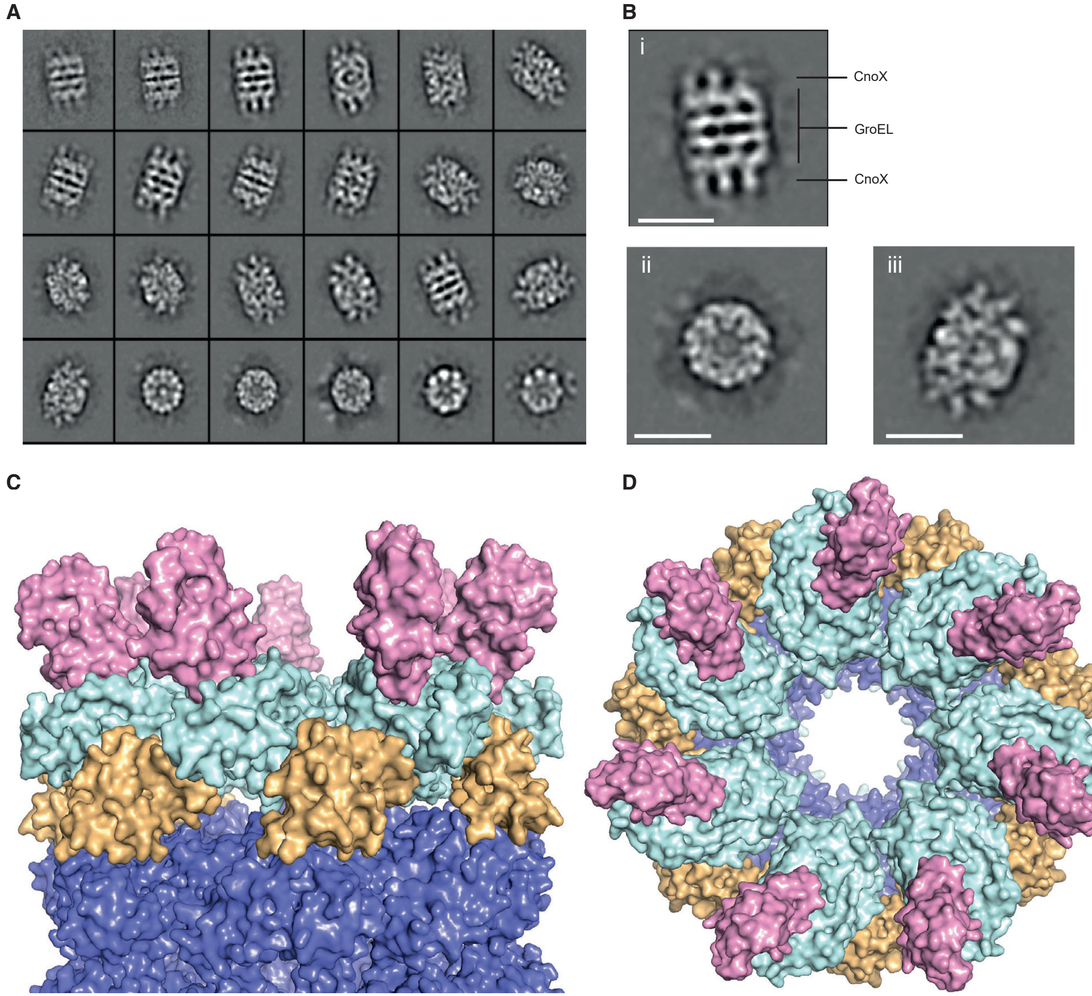
Solving protein structures

- **Nuclear magnetic resonance (NMR) spectroscopy**
  - Does not depend on a **protein crystal**
  - Only suited for **small proteins**
  - Requires a small volume of a **concentrated protein** placed in a **strong magnetic field**
  - Hydrogen nuclei behave as **magnets** and spin when exposed to magnetic field. The way they spin depends **on their environment**
  - Interested to know more? https://www.youtube.com/watch?v=Sn3dNMv-67k

Solving protein structures

- CryoEM





Plan

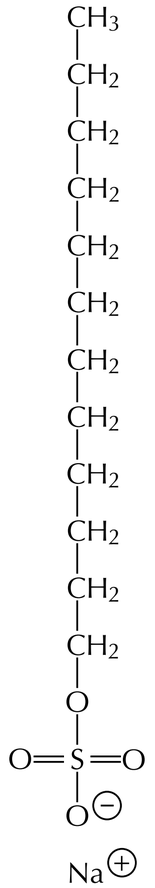
- Model organisms
- Isolating cells and growing them in culture
- **Studying proteins**
    - Protein sequence
    - Protein purification
    - Protein structure
    - **Protein visualization**
    - Mass spectrometry

Visualizing/Separating proteins

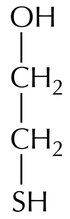
- SDS-PAGE = **sodium dodecyl sulfate polyacrylamide gel electrophoresis**
- Proteins have a **net** positive or negative **charge** depending on the amino-acid they contain
- When an **electric field** is applied to a solution containing proteins, they will **migrate** depending on this charge, size and shape
- SDS-PAGE uses a **cross-linked gel of polyacrylamide** through which the proteins migrate, the pore size of the gel can be adjusted

Proteins can be separated on SDS-PAGE

- The proteins are dissolved in a powerful negatively charged detergent, **SDS**
- SDS binds to **hydrophobic regions** of the protein, causing their **unfolding and release** from other proteins or membranes
- **B-mercaptoethanol** removes **disulfide bonds**



SDS



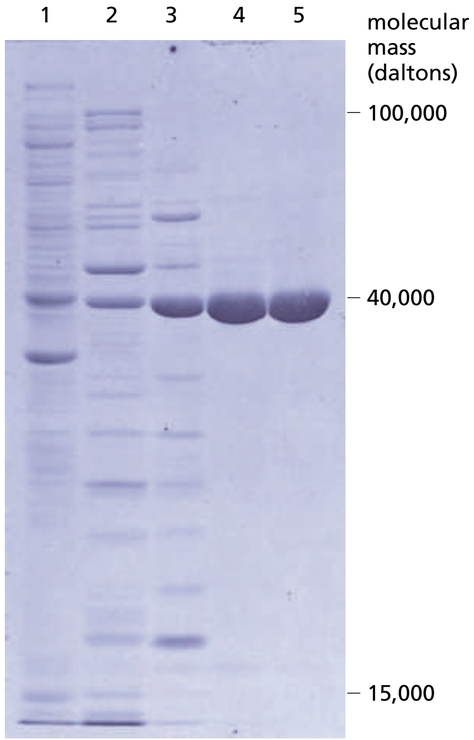
β-mercaptoethanol

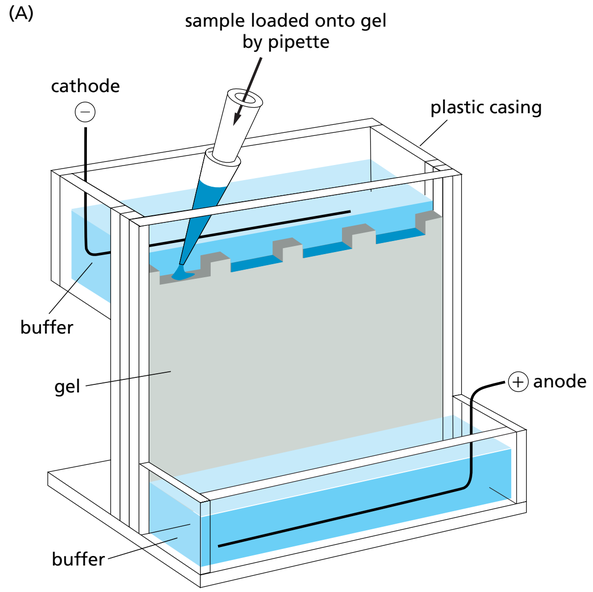
Proteins can be separated on SDS-PAGE

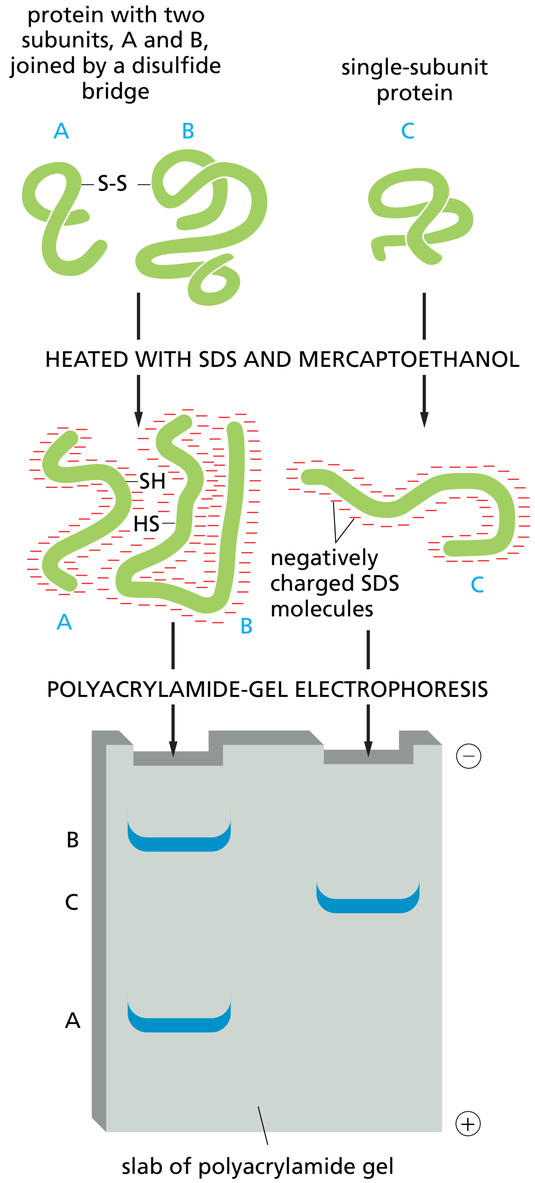
- Binding of SDS masks the **intrinsic charge of the protein** and causes it to migrate towards the **positive electrode** during SDS-PAGE
- Proteins with the **same size** move at the **same speed**
    - they are **unfolded** so their shape is the same
    - they bind the **same amount of SDS** so their charge is the same

Proteins can be separated on SDS-PAGE

- **Larger proteins are retarded** more than small ones in the acrylamide mesh
- Proteins can be stained (for example using **Coomassie blue)**

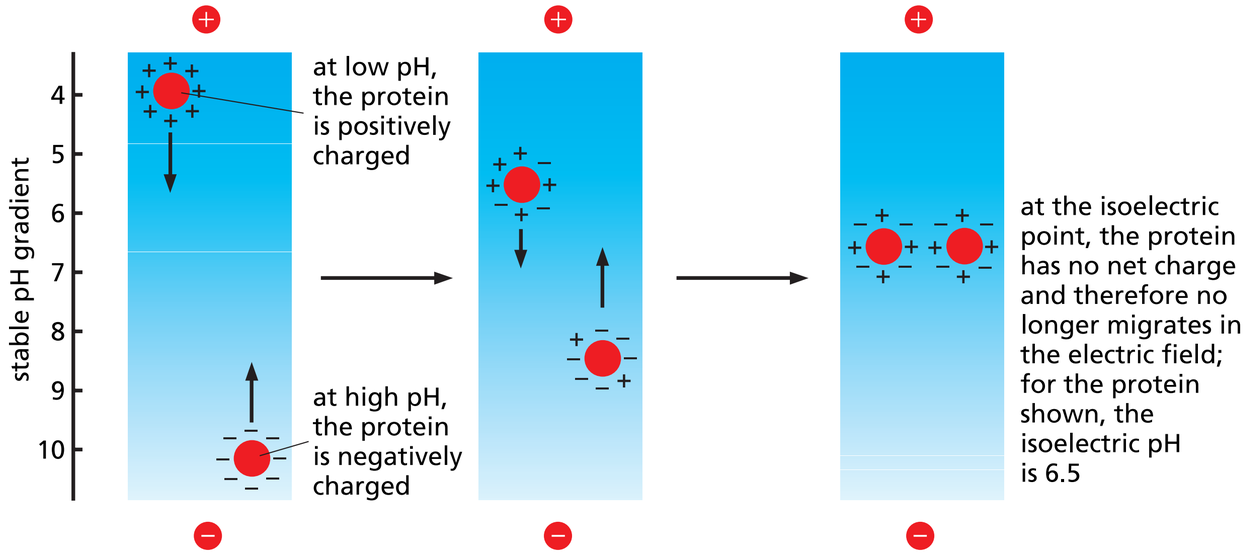






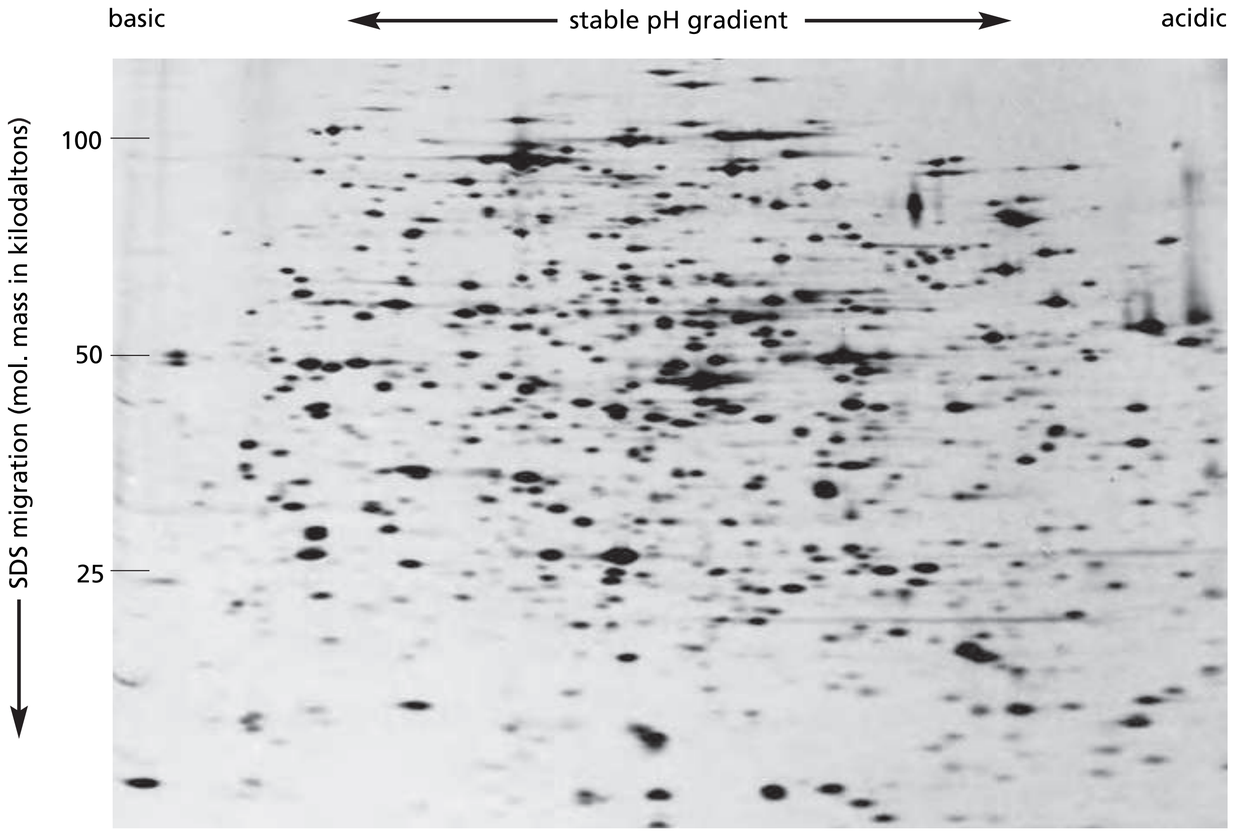
2D-gel electrophoresis

- Two proteins can have the **same size/shape/mass/charge**
- **2D-gel electrophoresis** combines **two separation methods**
  - **D1:** separation by **intrinsic charge**
  - **D2:** classical **SDS-PAGE**



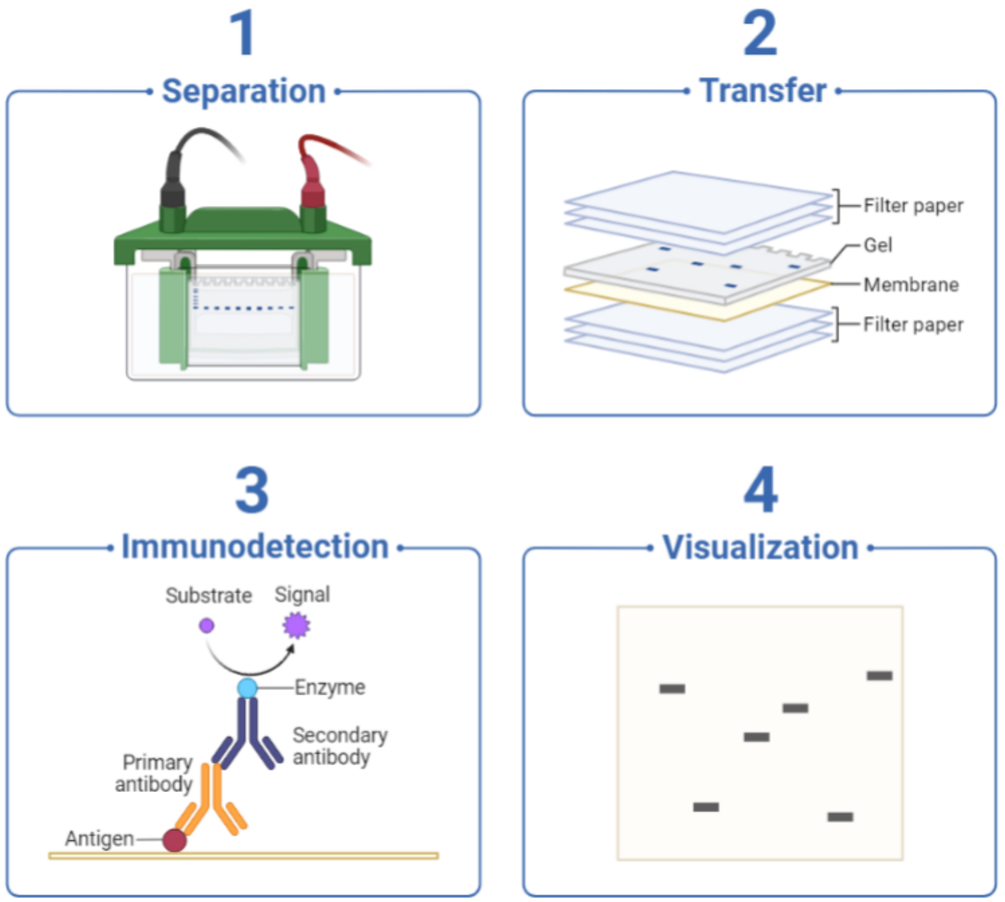
2D-gel electrophoresis

- Two proteins can have the **same size/shape/mass/charge**
- **2D-gel electrophoresis** combines **two separation methods**
  - **D1:** separation by **intrinsic charge**
  - **D2:** classical **SDS-PAGE**

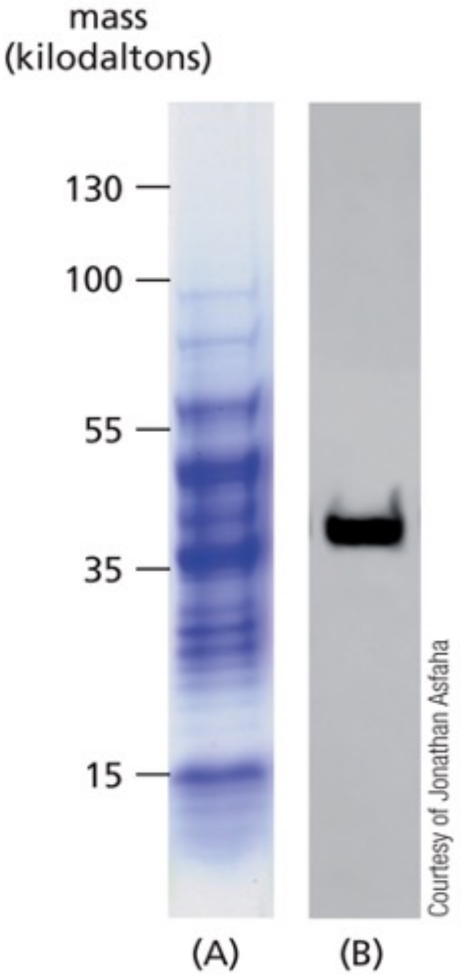


Detecting proteins by Western Blot

- A specific protein can be **detected** on an **SDS-PAGE** using a specific **labelled antibody**
- **Transfer** of the proteins from the gel onto a nitrocellulose/nylon **membrane** using an **electric current**
- The membrane is soaked in **antibody solution**



https://www.biomol.com/resources/applications/western-blot/



Detecting proteins by Western Blot

	Southern Blot	Northern Blot	Western Blot
Target molecule	DNA	RNA	Protein
Sample preparation	DNA extraction enzymatic digestion	RNA isolation	Protein extraction
Separation	Electrophoresis	Electrophoresis	Electrophoresis
Membrane material	Nylon	Nylon	Nitrocellulose or PVDF
Probe	Nucleic acid probe with sequence homologous to target	RNA, DNA, or oligodeoxynucleotide	Primary antibody
Probe label	Radiolabel, enzyme	Radiolabel, enzyme	Enzyme
Detection methods	X-ray film, chemiluminescence	X-ray film, chemiluminescence	Film, cooled CCD, camera, LED, or infrared imaging system

Table 1: Comparing Southern, Northern, and Western Blots.

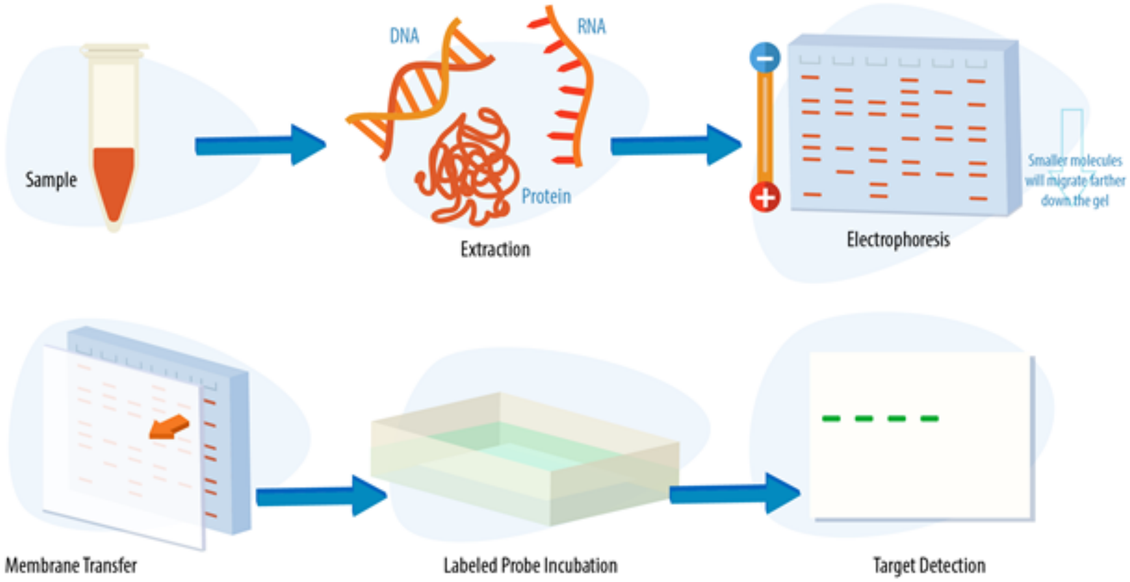


Figure 1: Gel electrophoresis, transfer, incubation, and detection.

https://www.labmanager.com/southern-vs-northern-vs-western-blotting-techniques-854

Plan

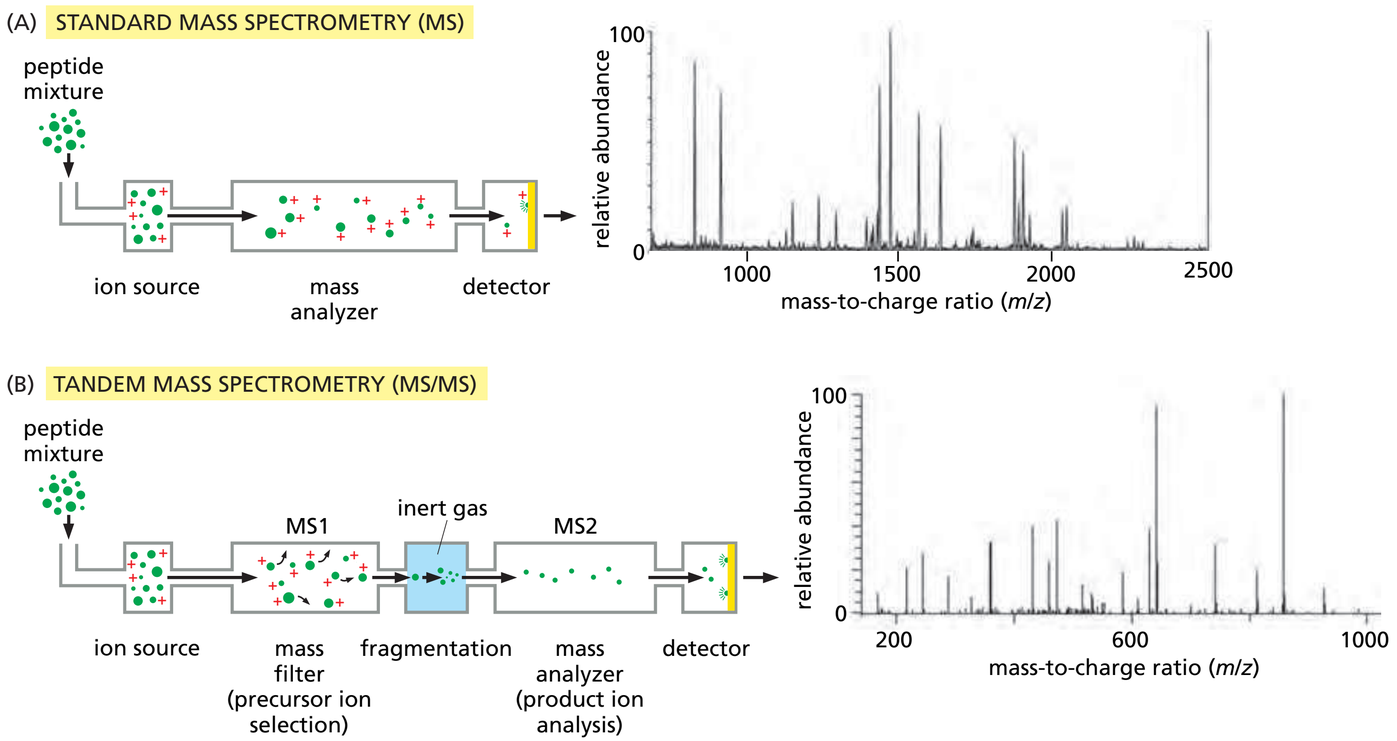
- Model organisms
- Isolating cells and growing them in culture
- **Studying proteins**
    - Protein sequence
    - Protein purification
    - Protein structure
    - Protein visualization
    - **Mass spectrometry**

Mass spectrometry to identify unknown proteins

- How to **identify** a protein obtained from one of the purification methods?
- **Genomes** of most experimentally used organisms are known - **catalogue of possible proteins** is known
- Use of **mass spectrometry** combined with **searches of databases**
- Charged particles have **very precise dynamics** when subjected to **electric and magnetic fields in a vacuum**
- **Mass spectrometry** separates ions according to their mass-to-charge ratio (m/z)
- Extremely **sensitive** method

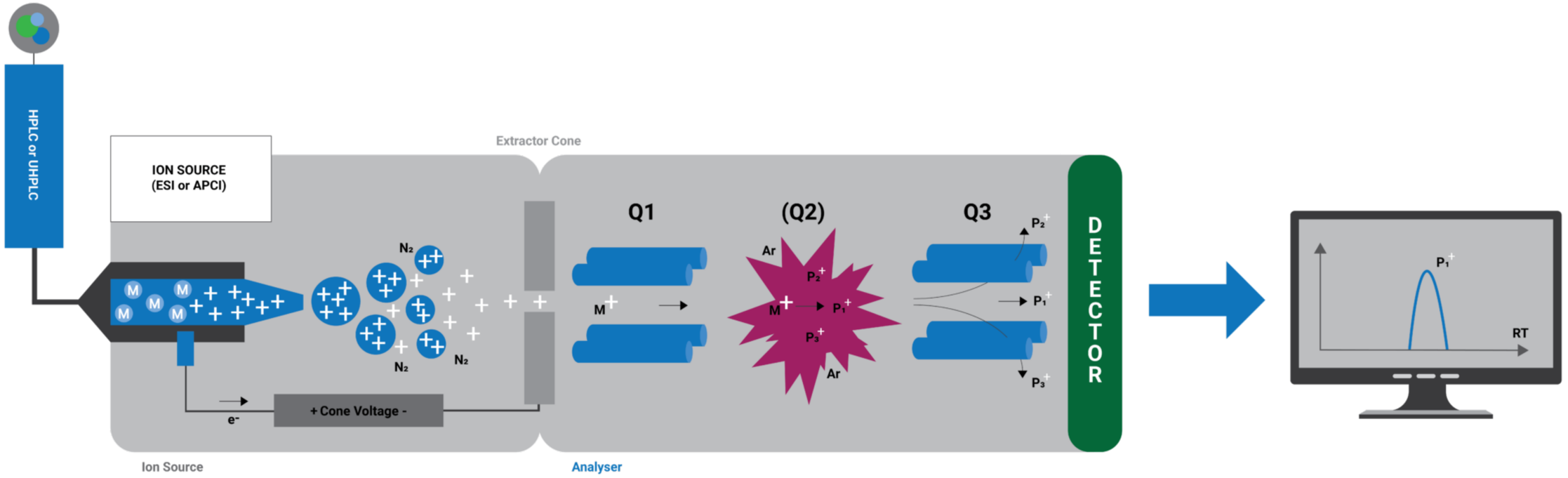
Mass spectrometry to identify unknown proteins

- **Ion source** transforms tiny amounts of peptide into a gas containing individual charged peptides
- **Mass analyzer** where ions are accelerated by electric or magnetic fields in which ions are separated according to their mass-to-charge ratio
- **Detector** on which the ions collide which generate a mass spectrum
- **Different types** of mass spectrometer, with different ion sources and mass analyzers (e.g. MALDI-TOF)



Mass spectrometry to identify unknown proteins

- Use of **LC-MS/MS** for **complex protein mixtures**



https://blog.perkinelmer.com/posts/mass-spectrometry-explained-how-lc-ms-ms-works/

Summary

- Model organisms
- Isolating cells and growing them in culture
- Studying proteins
    - Protein sequence
    - Protein purification
    - Protein structure
    - Protein visualization
    - Mass spectrometry

Have a nice day!