Cellular and Molecular Biology I

BIO-205-9

Camille Goemans

CpG and CpG islands*

A **CpG** is a specific DNA sequence where a Cytosine (C) nucleotide is immediately followed by a Guanine (G) nucleotide on the same DNA strand, in the 5' → 3' direction.

- The p in "CpG" stands for the phosphate that links the C and G.
- CpGs are the primary sites of DNA methylation in vertebrates — cytosine can be chemically modified to 5-methylcytosine.

Why CpGs are important:

- Gene regulation: Methylation at CpG sites can silence gene expression.
- Mutation hotspot: Methylated CpGs mutate to TpG at a high rate, making CpGs relatively rare across the genome.

CpG and CpG islands*

A **CpG island** is a DNA region enriched in CpG sites.

- CpG islands tend to be unmethylated in normal cells.
- ~70% of human gene promoters have CpG islands.

Why CpG islands matter:

- Gene expression: Unmethylated CpG islands typically allow active transcription.
- Cell identity: Their methylation status helps define cell differentiation.
- Cancer biology: Abnormally hypermethylated CpG islands can silence tumor-suppressor genes.

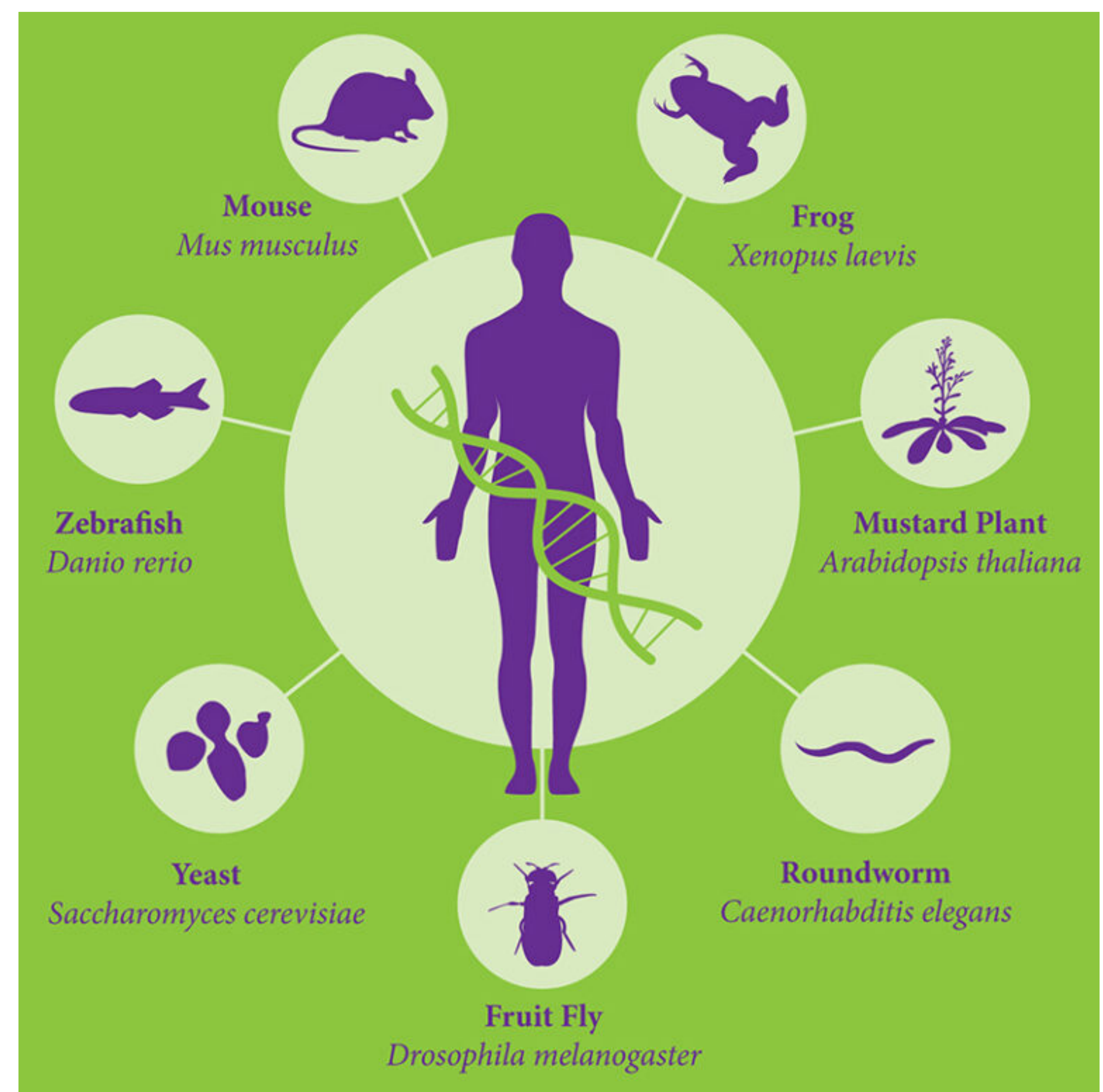
V. Molecular and Cellular Biology in the lab

1. **Model organisms**
2. Cell cultures
3. Studying proteins
    - Protein sequence
    - Protein purification
    - Protein visualization
    - Protein structure
    - Mass spectrometry

Model organisms

- Model organisms are **species that are extensively studied** to understand particular biological processes, with the assumption that discoveries made in these organisms can be **generalized** to other species, including humans.

- They are chosen for their **short generation times**, **ease of care**, **well-understood genetics**, and the availability of **research tools**



Model organisms

Model Organism	Scientific Name	Type	Key Uses in Research
Bacterium – Escherichia coli	E. coli	Prokaryote (bacterium)	- Gene expression and regulation - DNA replication and repair - Basic metabolism - Biotechnology and genetic engineering (e.g., plasmid vectors)
Yeast – Saccharomyces cerevisiae	S. cerevisiae	Eukaryote (unicellular fungus)	- Cell cycle and division - Aging and metabolism - Eukaryotic gene expression - Protein folding and secretion
Nematode – Caenorhabditis elegans	C. elegans	Multicellular animal (worm)	- Developmental biology - Nervous system and behavior - Apoptosis (programmed cell death) - Genetic regulation of lifespan
Fruit Fly – Drosophila melanogaster	D. melanogaster	Insect	- Genetics and heredity - Embryonic development - Neurobiology and behavior - Evolutionary biology
Zebrafish – Danio rerio	D. rerio	Vertebrate (fish)	- Developmental biology and organogenesis - Genetics of vertebrates - Disease modeling (heart, brain, cancer) - Regeneration studies
Mouse – Mus musculus	M. musculus	Mammal	- Human disease models (cancer, diabetes, neurodegeneration) - Immunology - Genetics and genomics - Drug testing and pharmacology
Plant – Arabidopsis thaliana	A. thaliana	Plant	- Plant genetics and development - Photosynthesis and metabolism - Stress responses - Evolutionary biology of plants
Frog – Xenopus laevis / Xenopus tropicalis	Xenopus species	Amphibian (vertebrate)	- Embryonic development - Cell cycle regulation - Signal transduction pathways

Model organisms

Model organisms are super important because:

- **Conservation** of genes and pathways. Many fundamental biological processes are conserved across species.
- Ethical and practical **advantages**: Easier and cheaper to study than humans or other large animals.
- **Genetic tools:** Many have fully sequenced genomes, genetic mutants, and tools like CRISPR or RNAi.

V. Molecular and Cellular Biology in the lab

1. Model organisms
2. **Cell cultures**
3. Studying proteins
    - Protein sequence
    - Protein purification
    - Protein visualization
    - Protein structure
    - Mass spectrometry

Cell culture

- Cell culture is the process of **growing cells under controlled laboratory conditions** — typically outside their natural environment (in vitro).
- Cells are maintained in a **nutrient-rich liquid medium** within sterile containers, such as flasks or Petri dishes, under specific temperature, gas (CO_2), and humidity conditions.
- This technique allows us to study the **behavior, physiology, and biochemistry of cells** in a controlled setting.

Cell culture is used in many areas of biology, biotechnology, and medicine:

- Research: Studying cell growth, signalling, differentiation, and genetic regulation.
- Medicine: Testing drugs, vaccines, and studying disease mechanisms.
- Biotechnology: Producing proteins, antibodies, and vaccines.
- Tissue engineering: Growing cells for artificial tissues or regenerative medicine.
- Toxicology: Testing effects of chemicals or cosmetics on cell health.

Cell culture - types of cell cultures

Type	Description	Example Uses
Primary Cell Culture	Cells taken directly from tissues and cultured for the first time. They closely resemble the original tissue but have a limited lifespan.	Studying normal cell physiology, metabolism, and cell–cell interactions.
Secondary/ Subcultured Cells	Cells that have been transferred (subcultured) from a primary culture to new growth media.	Used to expand the culture or continue experiments over time.
Cell Lines	Cells that have adapted to continuous growth in vitro. Derived from primary cultures that have undergone transformation (spontaneously or induced).	Used in long-term studies, drug testing, and biotechnology production.

Cell culture - types of cell lines

Type	Description	Examples
Finite Cell Line	Limited number of divisions before senescence (cell aging).	Human fibroblasts, kidney cells.
Continuous (Immortal) Cell Line	Can divide indefinitely due to transformation (natural or artificial).	HeLa (human cervical cancer), HEK293 (human embryonic kidney), CHO (Chinese hamster ovary).

Cell culture - cell isolation

- 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

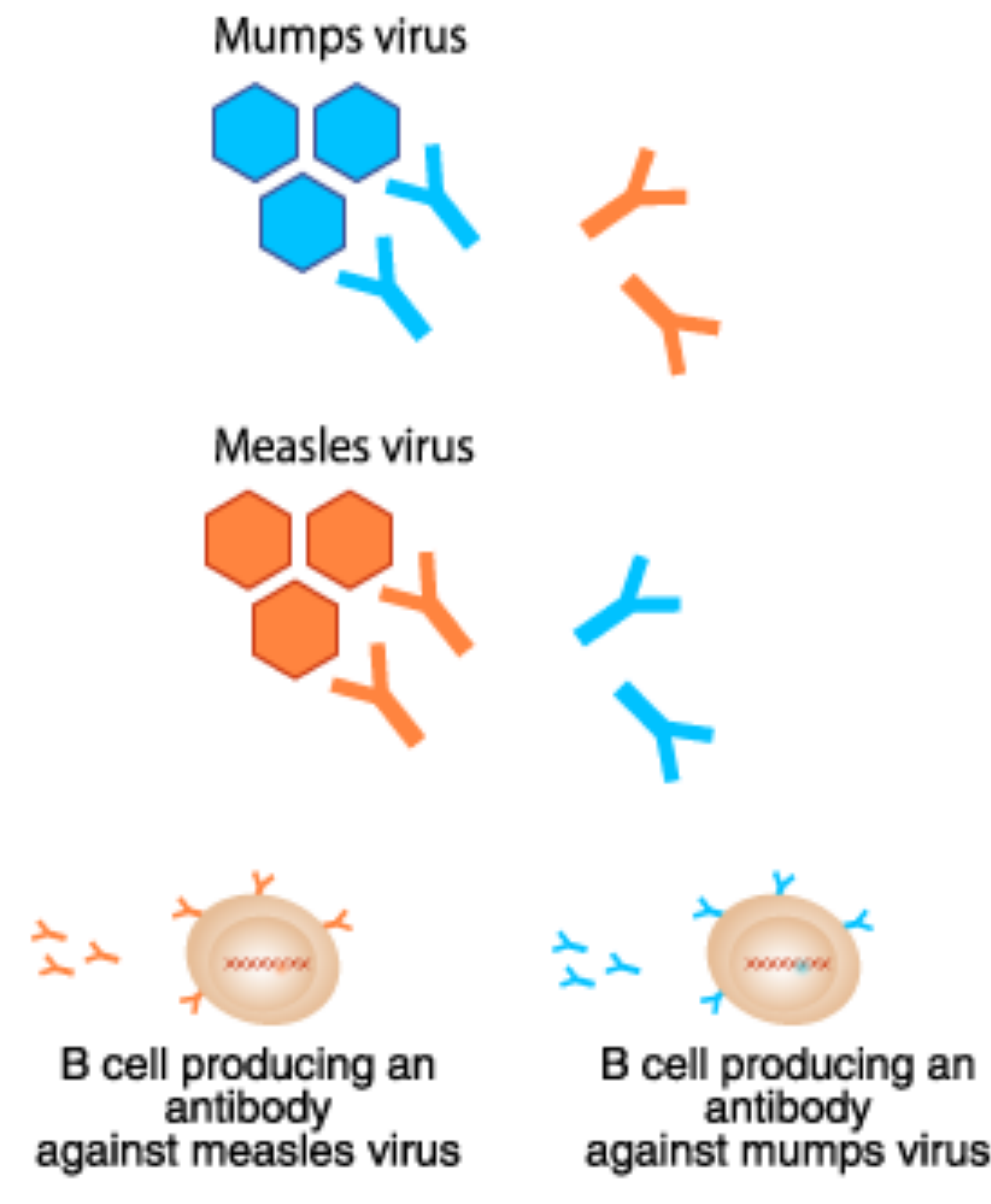
- Cell cultures have one problem: cells **die;** they stop dividing after a number of division (replicative cell senescence)
- **Fibroblasts** divide **25 to 40 times in culture**
- This reflects the **shortening and uncapping of telomeres,** following the **decreased production** of the telomerase
- 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

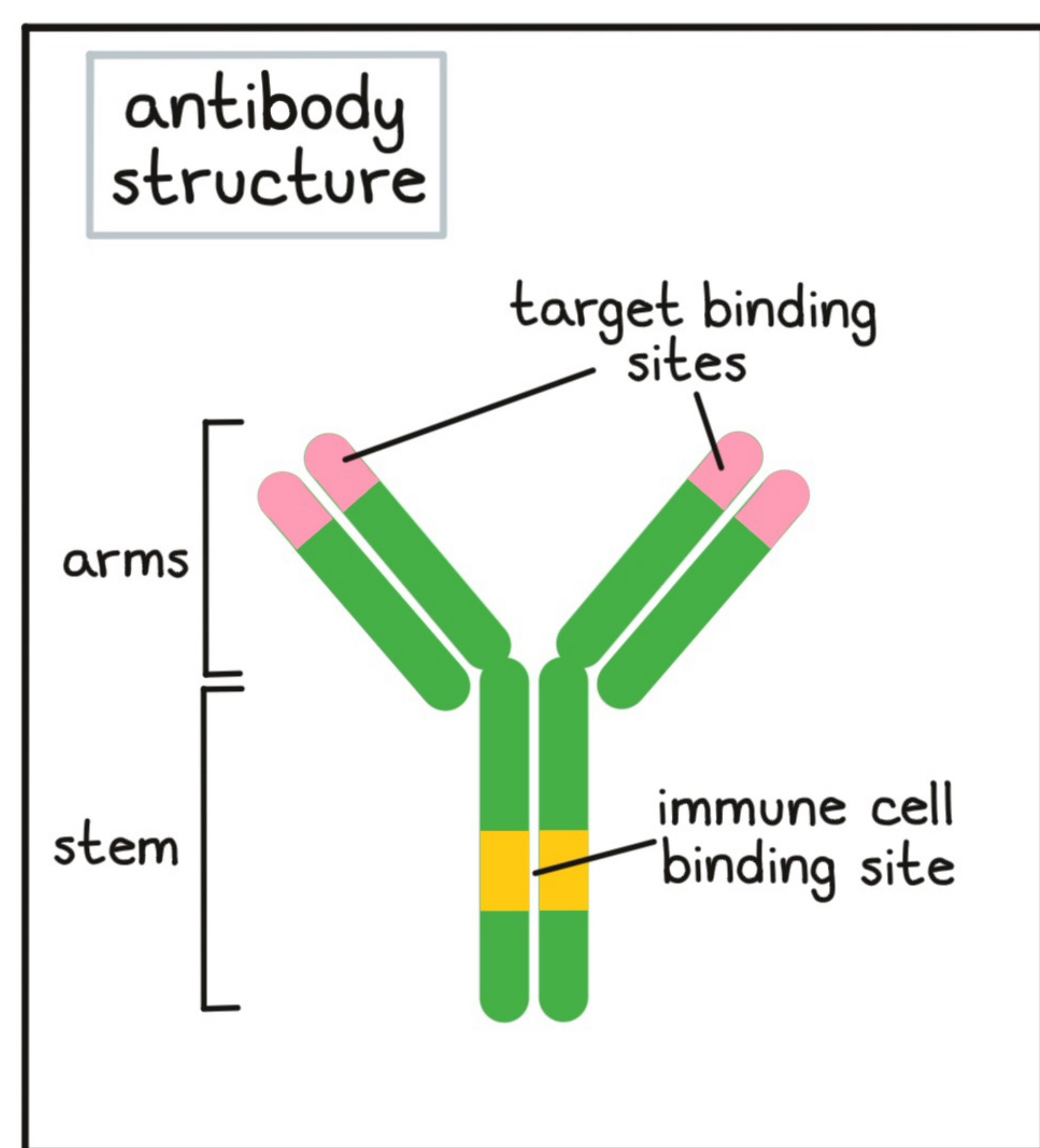
TABLE 8–1 Some Commonly Used Cell Lines

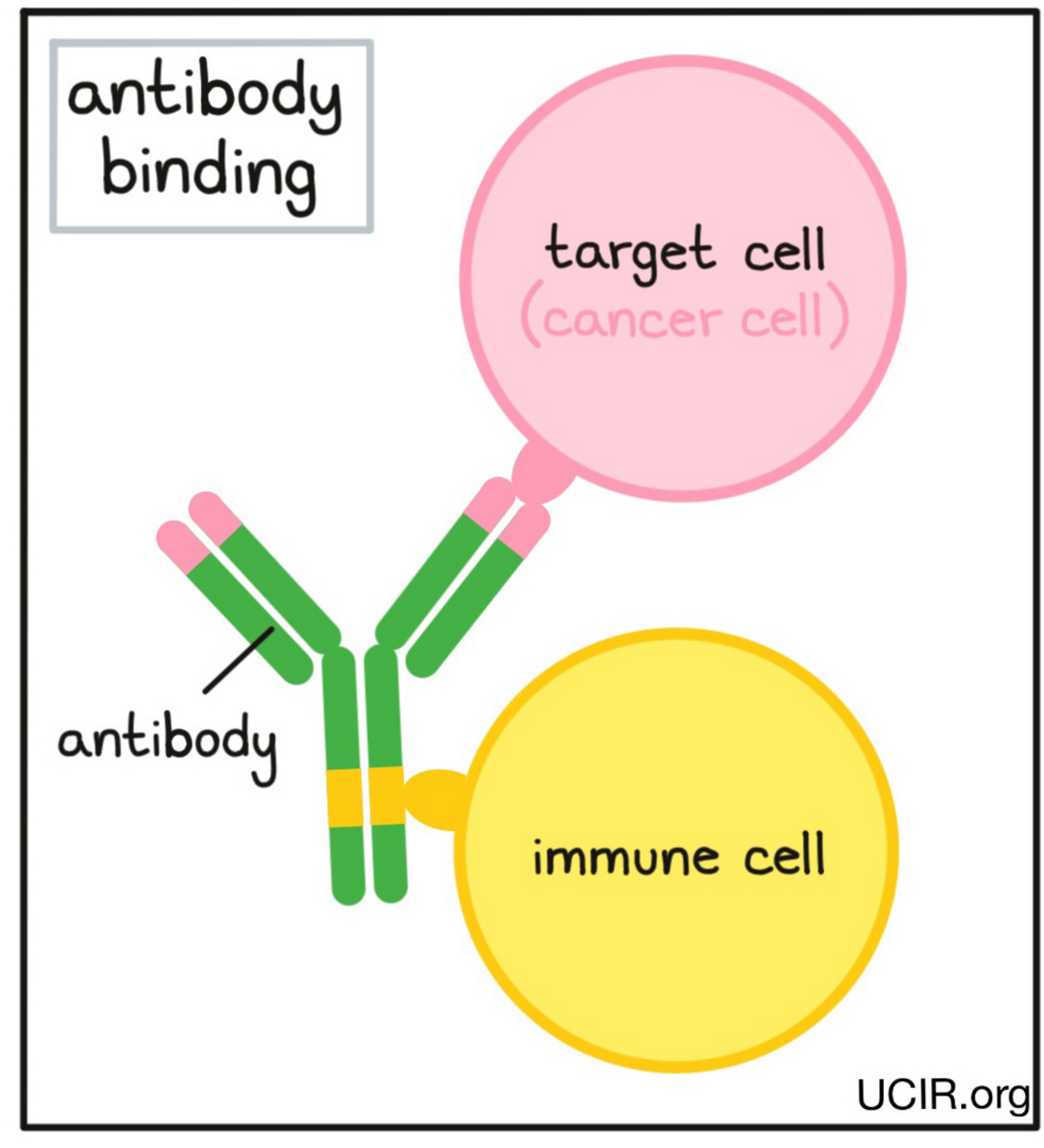
Cell line*	Cell type and origin
3T3	Fibroblast (mouse)
BHK21	Fibroblast (Syrian hamster)
MDCK	Epithelial cell (dog)
HeLa	Epithelial cell (human)
PtK1	Epithelial cell (rat kangaroo)
L6	Myoblast (rat)
PC12	Chromaffin cell (rat)
SP2	Plasma cell (mouse)
COS	Kidney (monkey)
293	Kidney (human); transformed 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







Hybridoma cell lines and monoclonal antibodies

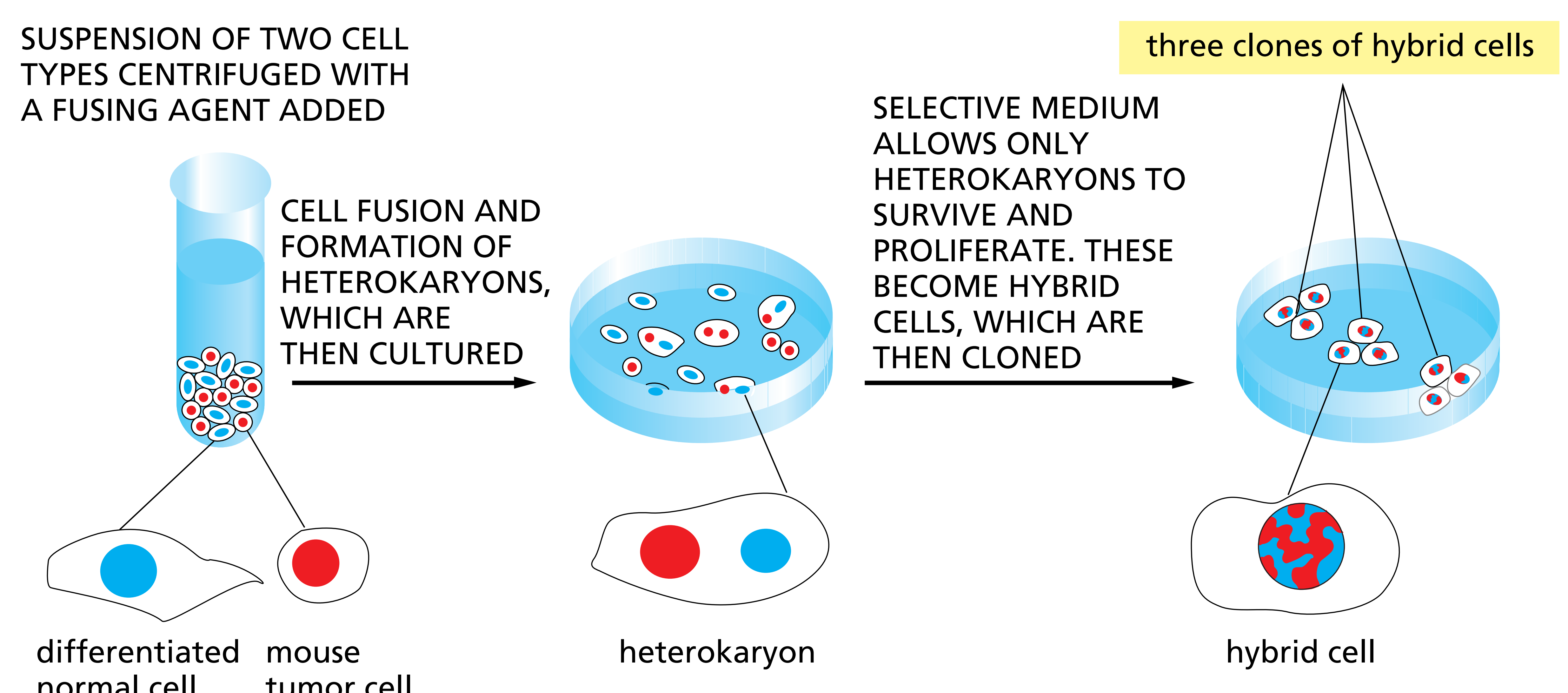
- **Antibodies** are extremely useful tools in the lab and are now also used to treat cancers
- Their **great specificity** allows to **detect selected proteins** among thousands
- They are produced by **inoculating an animal** with the **purified protein** of interest and further isolation of the antibodies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



V. Molecular and Cellular Biology in the lab

1. Model organisms

2. Cell cultures

3. **Studying proteins**

    - Protein sequence
    - Protein purification
    - Protein visualization
    - Protein structure
    - Mass spectrometry

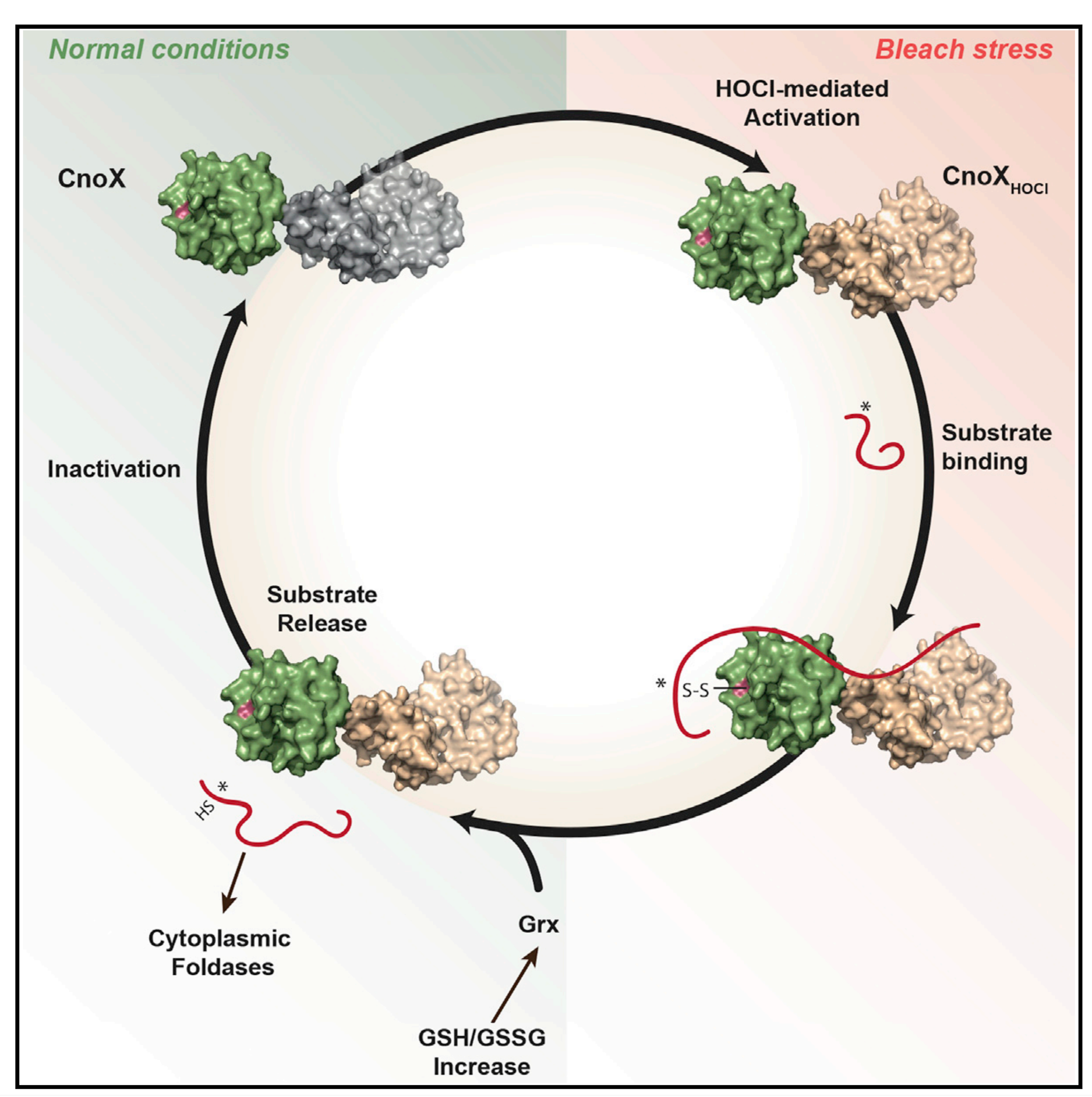
Real-life example: my PhD*

Molecular Cell

Article

CnoX Is a Chaperedoxin: A Holdase that Protects Its Substrates from Irreversible Oxidation

Graphical Abstract



Authors

Camille V. Goemans,
Didier Vertommen, Rym Agrebi,
Jean-François Collet

Correspondence

jfcollet@uclouvain.be

In Brief

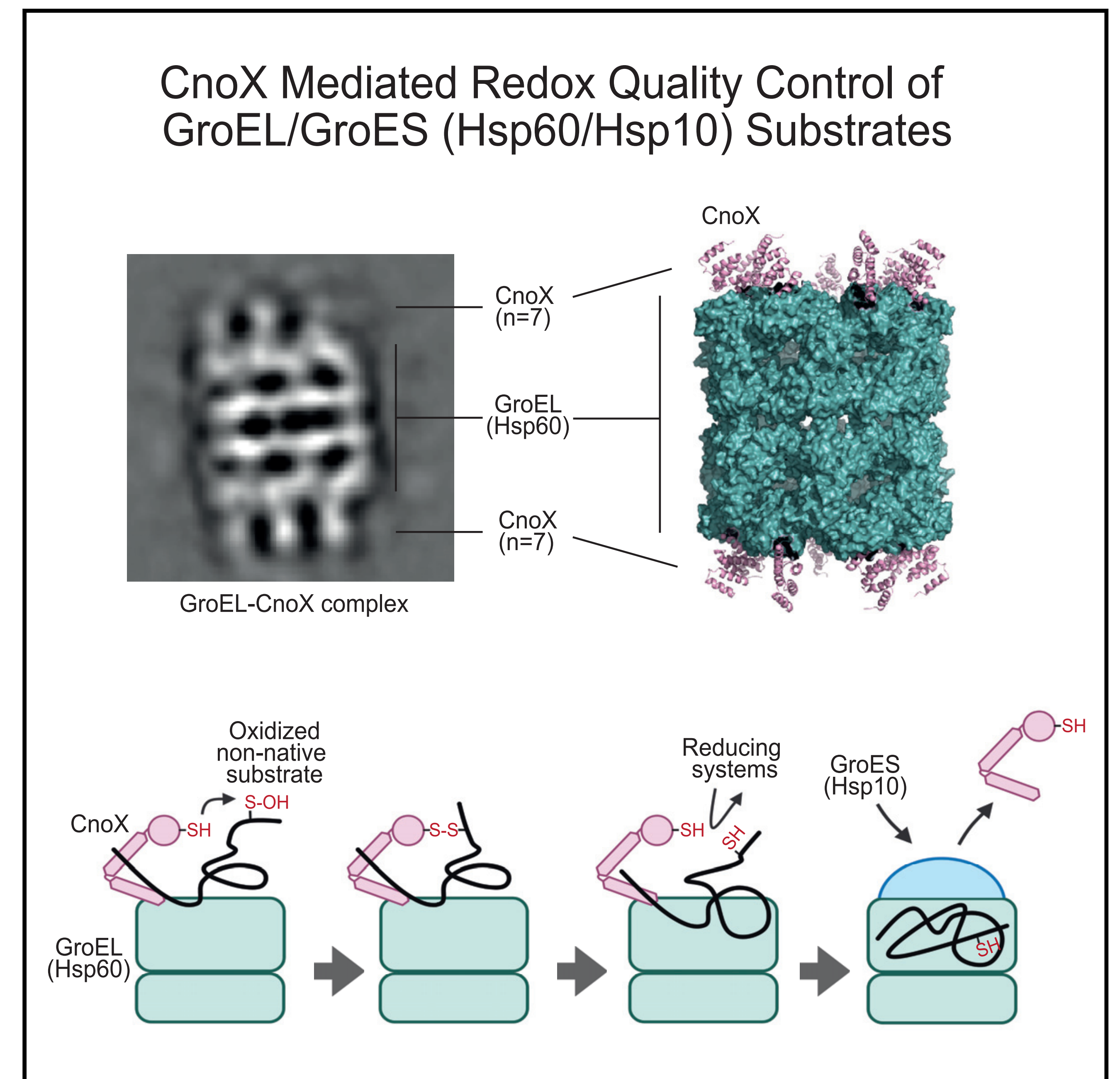
Bleach is a powerful oxidant that kills bacteria by causing protein aggregation. Goemans et al. identified *Escherichia coli* CnoX (YbbN) as a bleach-activated chaperone that uniquely combines holdase activity with the ability to protect its substrates from irreversible oxidation. After bleach stress, CnoX transfers its client proteins to GroEL/ES and DnaK/J/GrpE.

Cell

Article

A molecular device for the redox quality control of GroEL/ES substrates

Graphical abstract



Authors

Emile Dupuy,
Sander Egbert Van der Verren,
Jiusheng Lin, ...,
Camille Véronique Goemans,
Han Remaut, Jean-François Collet

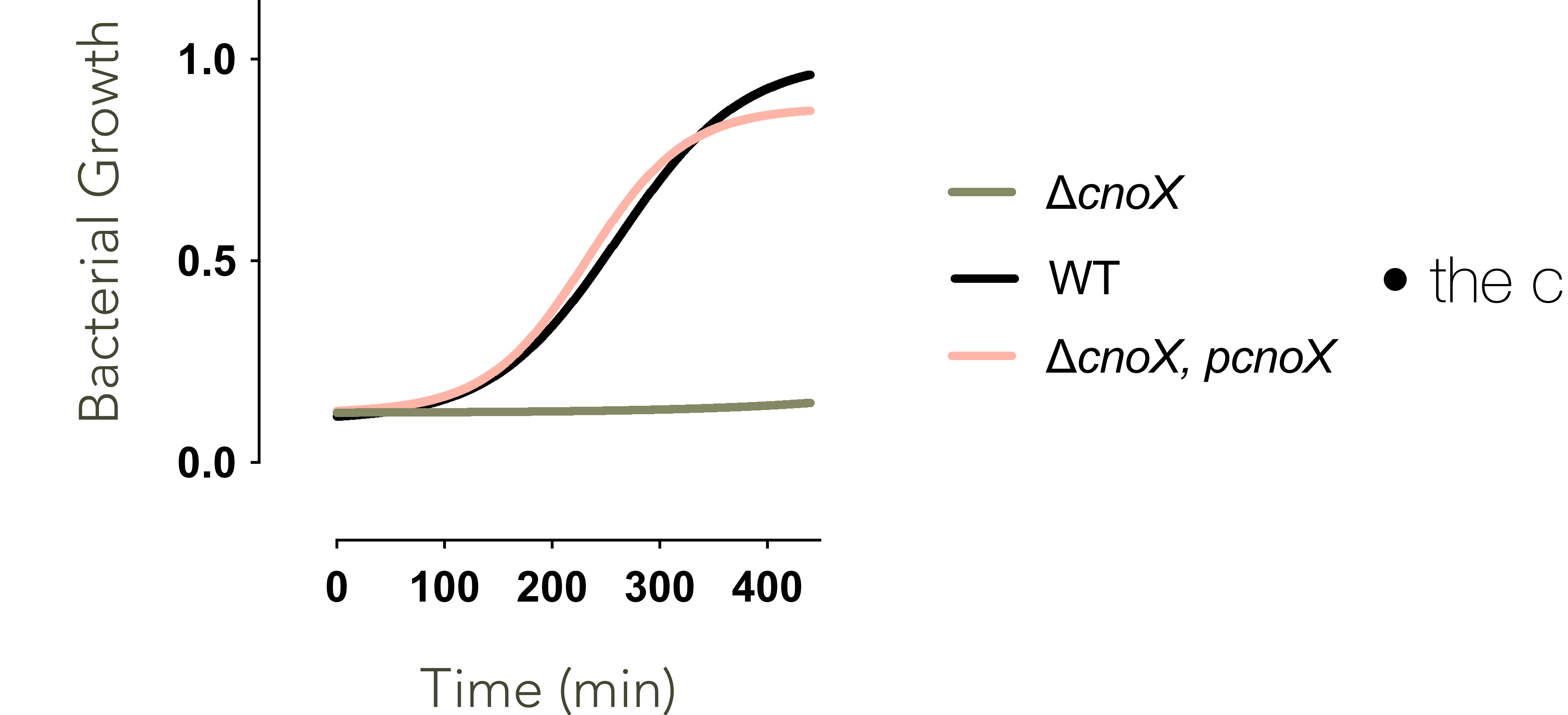
Correspondence

camille.goemans@embl.de (C.V.G.),
han.remaut@vub.be (H.R.),
jfcollet@uclouvain.be (J.-F.C.)

In brief

CnoX is a redox quality-control molecular plugin for an evolutionarily conserved Hsp60 chaperonin complex crucial for protein folding in all living cells.

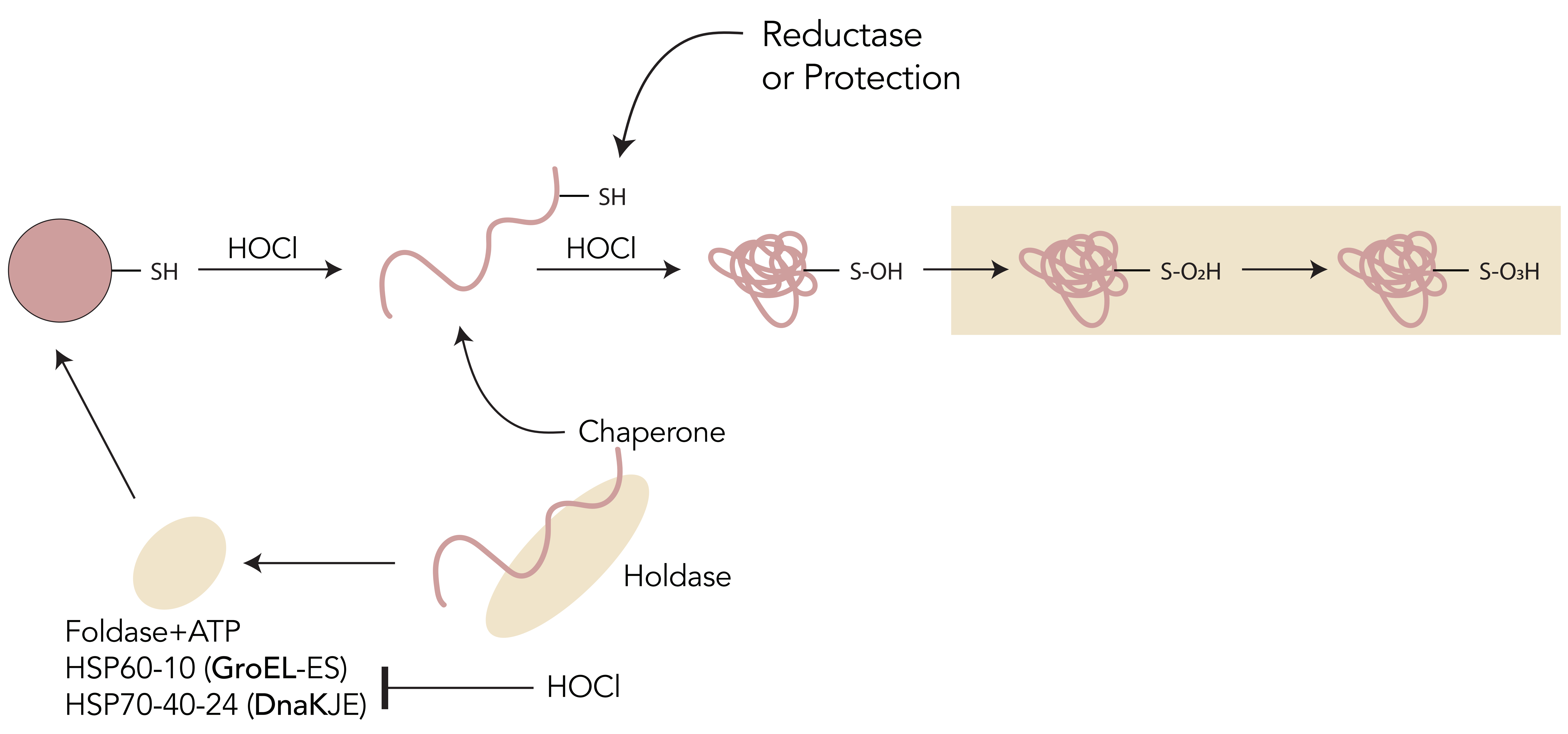
Understanding the function of CnoX*



- the cnoX gene is important for bacteria to survive oxidative stress (HOCl)

How does this work?

Understanding the function of CnoX*



V. Molecular and Cellular Biology in the lab

1. Model organisms
2. Cell cultures
3. **Studying proteins**
    - **Protein sequence**
    - Protein purification
    - Protein visualization
    - Protein structure
    - Mass spectrometry

Getting information from the protein sequence

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

```
 Score =  399 bits (1025), Expect = e-111
 Identities = 198/290 (68%), Positives = 241/290 (82%), Gaps = 1/290

Query: 57   MENFQKVEKIGEGTYGVVYKARNKLTGEVVALKKIRLDTETEGVPSTAIREISLLKELNH 116
            ME ++KVEKIGEGTYGVVYKA +K T E +ALKKIRL+ E EGVPSTAIREISLLKE+NH
Sbjct: 1    MEQYEKVEKIGEGTYGVVYKALDKATNETIALKKIRLEQEDEGVPSTAIREISLLKEMNH 60

Query: 117  PNIVKLLDVIHTENKLYLVFEFLHQDLKKFMDASALTGIPLPLIKSYLFQLLQGLAFCHS 176
             NIV+L DV+H+E ++YLVFE+L  DLKKFMD+         LIKSYL+Q+L G+A+CHS
Sbjct: 61   GNIVRLHDVVHSEKRIYLVFEYLDLDLKKFMDSCPEFAKNPTLIKSYLYQILHGVAYCHS 120

Query: 177  HRVLHRDLKPQNLLINTE-GAIKLADFGLARAFGVPVRTYTHEVVTLWYRAPEILLGCKY 235
            HRVLHRDLKPQNLLI+    A+KLADFGLARAFG+PVRT+THEVVTLWYRAPEILLG +
Sbjct: 121  HRVLHRDLKPQNLLIDRRTNALKLADFGLARAFGIPVRTFTHEVVTLWYRAPEILLGARQ 180

Query: 236  YSTAVDIWSLGCIFAEMVTRRALFPGDSEIDQLFRIFRTLGTPDEVVWPGVTSMPDYKPS 295
            YST VD+WS+GCIFAEMV ++ LFPGDSEID+LF+IFR LGTP+E  WPGV+ +PD+K +
Sbjct: 181  YSTPVDVWSVGCIFAEMVNQKPLFPGDSEIDELFKIFRILGTPNEQSWPGVSCLPDFKTA 240

Query: 296  FPKWARQDFSKVVPPLDEDGRSLLSQMLHYDPNKRISAKAALAHPFFQDV 345
            FP+W  QD + VVP LD  G  LLS+ML Y+P+KRI+A+ AL H +F+D+
Sbjct: 241  FPRWQAQDLATVVPNLDPAGLDLLSKMLRYEPSKRITARQALEHEYFKDL 290
```

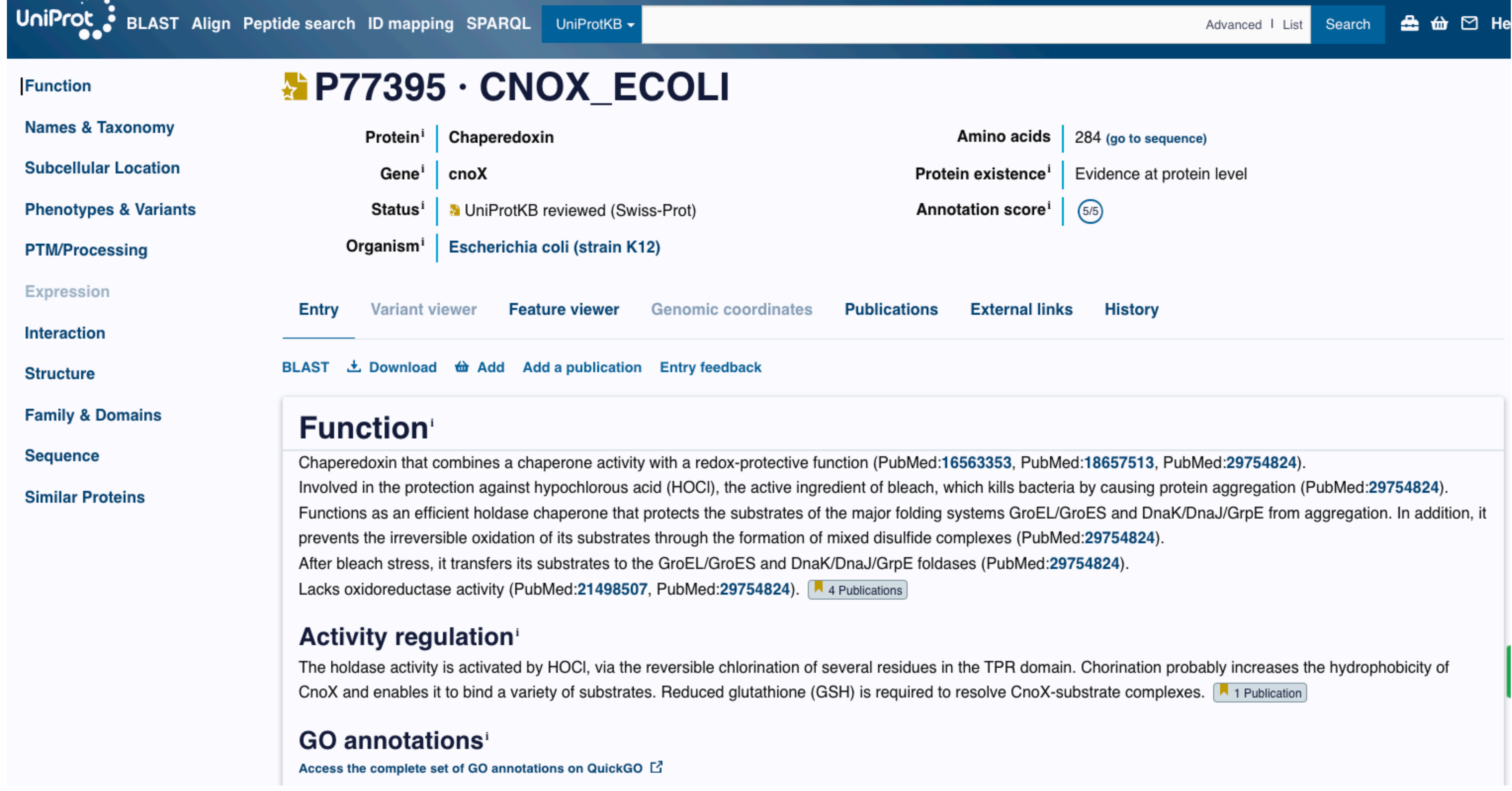
Getting information from the protein sequence

NCBI

Protein sequence

Getting information from the protein sequence

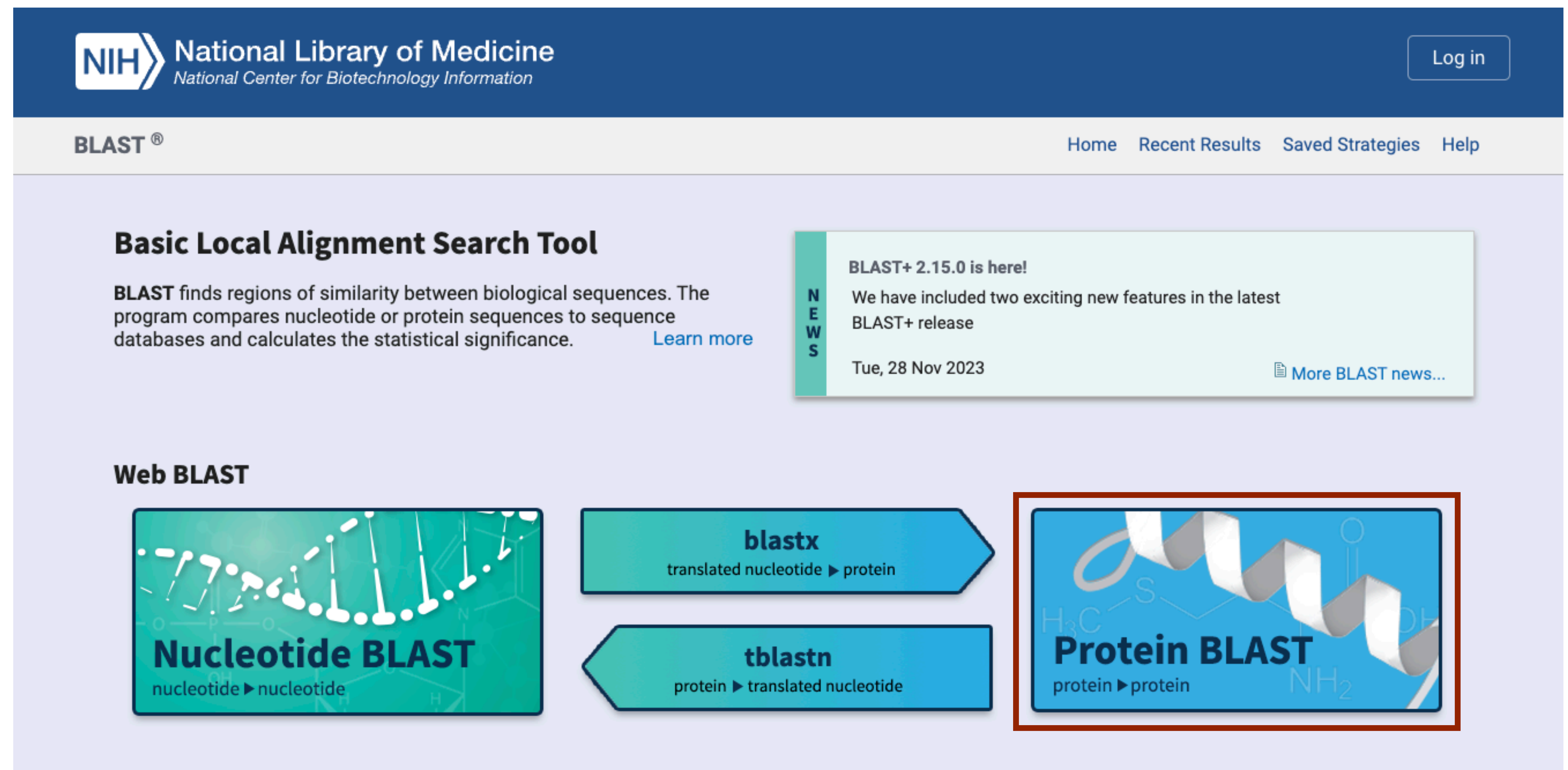
Uniprot



https://www.uniprot.org/uniprotkb/P77395/entry

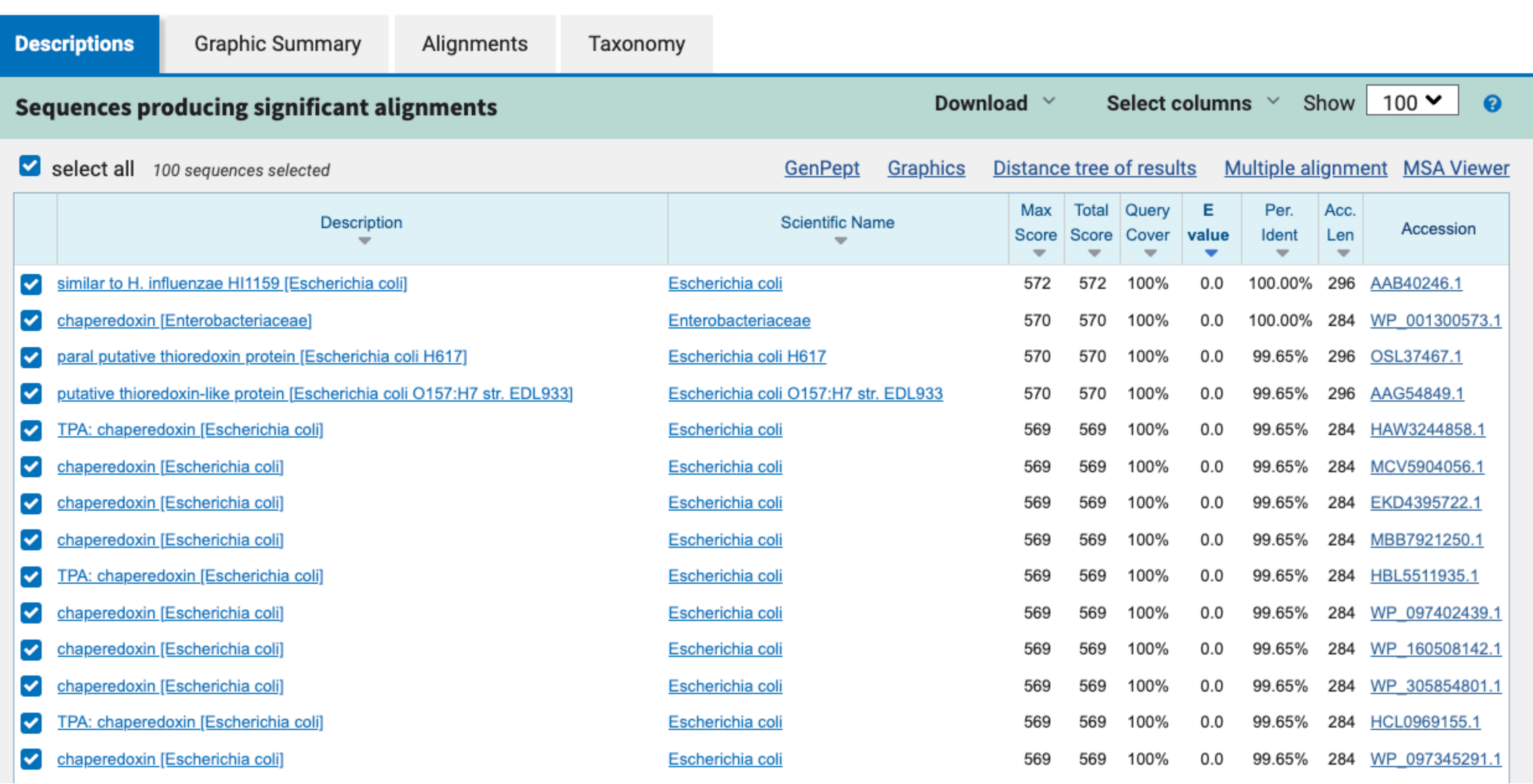
Getting information from the protein sequence

BLAST



Getting information from the protein sequence

BLAST



Descriptions | Graphic Summary | Alignments | Taxonomy

Sequences producing significant alignments Download Select columns Show 100

select all *100 sequences selected* GenPept Graphics Distance tree of results Multiple alignment MSA Viewer

Description	Scientific Name	Max Score	Total Score	Query Cover	E value	Per. Ident	Acc. Len	Accession
similar to H. influenzae HI1159 [Escherichia coli]	Escherichia coli	572	572	100%	0.0	100.00%	296	AAB40246.1
chaperedoxin [Enterobacteriaceae]	Enterobacteriaceae	570	570	100%	0.0	100.00%	284	WP_001300573.1
paral putative thioredoxin protein [Escherichia coli H617]	Escherichia coli H617	570	570	100%	0.0	99.65%	296	OSL37467.1
putative thioredoxin-like protein [Escherichia coli O157:H7 str. EDL933]	Escherichia coli O157:H7 str. EDL933	570	570	100%	0.0	99.65%	296	AAG54849.1
TPA: chaperedoxin [Escherichia coli]	Escherichia coli	569	569	100%	0.0	99.65%	284	HAW3244858.1
chaperedoxin [Escherichia coli]	Escherichia coli	569	569	100%	0.0	99.65%	284	MCV5904056.1
chaperedoxin [Escherichia coli]	Escherichia coli	569	569	100%	0.0	99.65%	284	EKD4395722.1
chaperedoxin [Escherichia coli]	Escherichia coli	569	569	100%	0.0	99.65%	284	MBB7921250.1
TPA: chaperedoxin [Escherichia coli]	Escherichia coli	569	569	100%	0.0	99.65%	284	HBL5511935.1
chaperedoxin [Escherichia coli]	Escherichia coli	569	569	100%	0.0	99.65%	284	WP_097402439.1
chaperedoxin [Escherichia coli]	Escherichia coli	569	569	100%	0.0	99.65%	284	WP_160508142.1
chaperedoxin [Escherichia coli]	Escherichia coli	569	569	100%	0.0	99.65%	284	WP_305854801.1
TPA: chaperedoxin [Escherichia coli]	Escherichia coli	569	569	100%	0.0	99.65%	284	HCL0969155.1
chaperedoxin [Escherichia coli]	Escherichia coli	569	569	100%	0.0	99.65%	284	WP_097345291.1

V. Molecular and Cellular Biology in the lab

1. Model organisms
2. Cell cultures
3. **Studying proteins**
    - Protein sequence
    - **Protein purification**
    - Protein visualization
    - Protein structure
    - Mass spectrometry

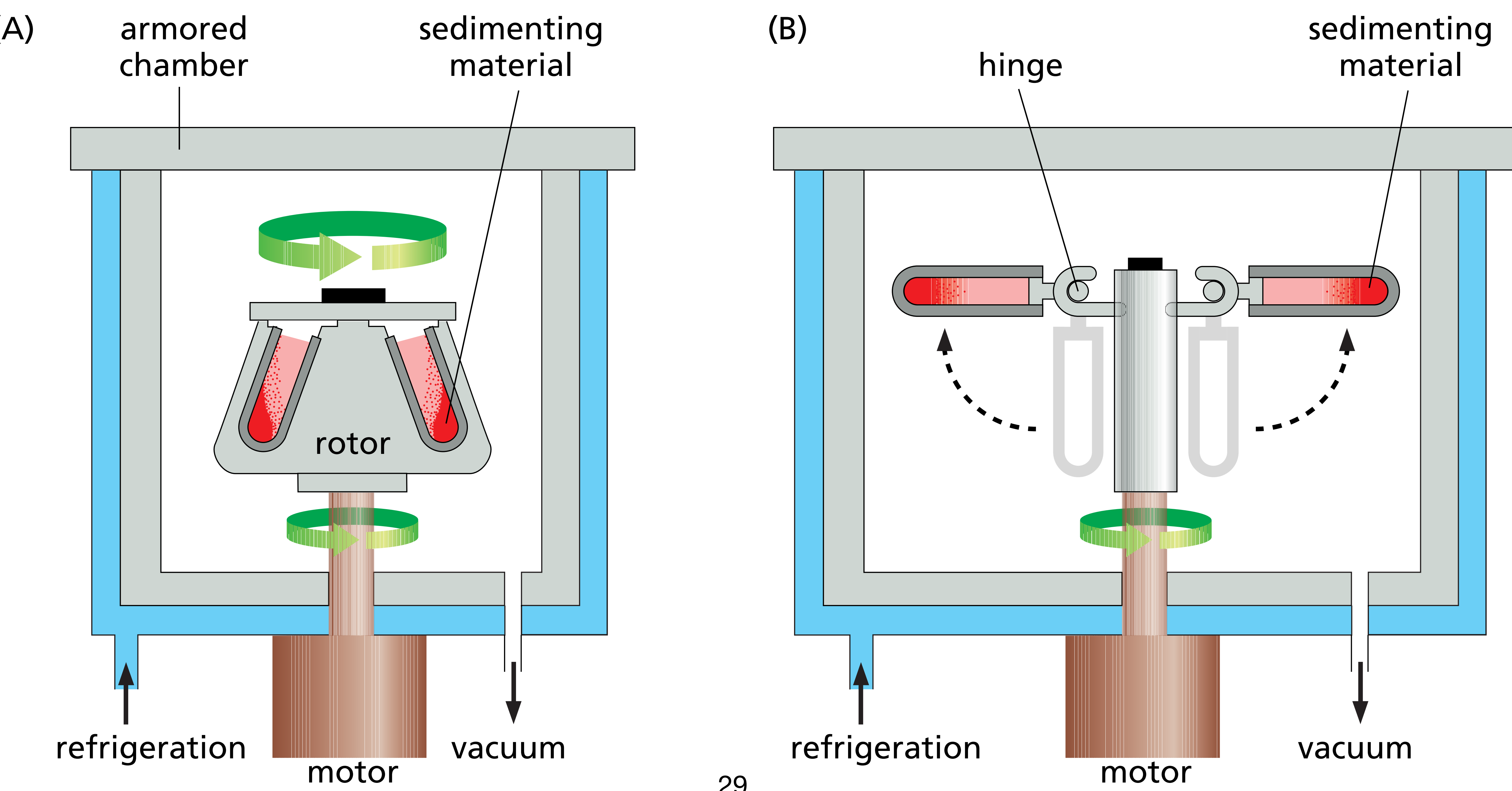
Protein purification

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**

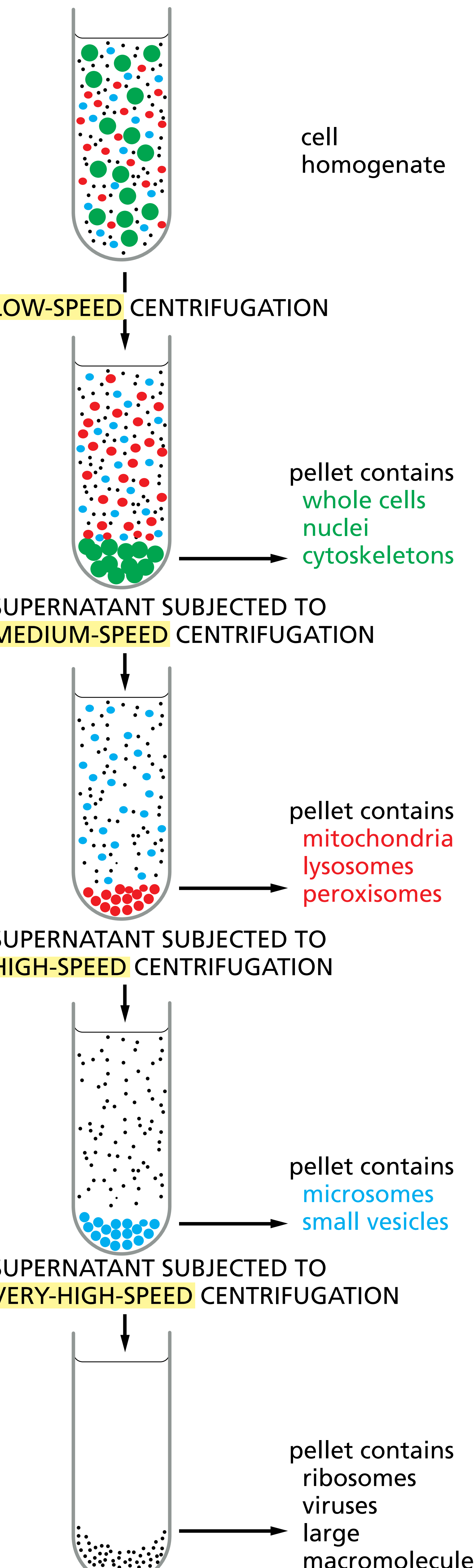
Step 1: opening the cells

- 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**



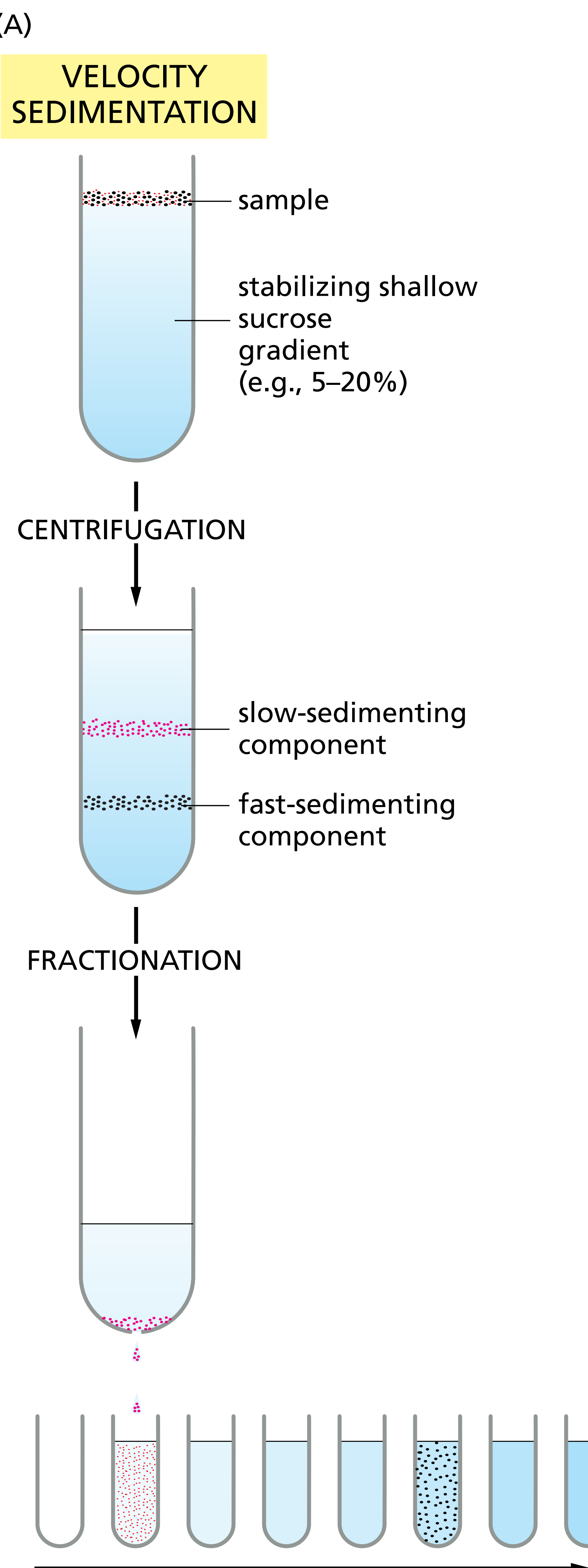
Step 1: opening the cells

MBoC6 m8.09/8.07



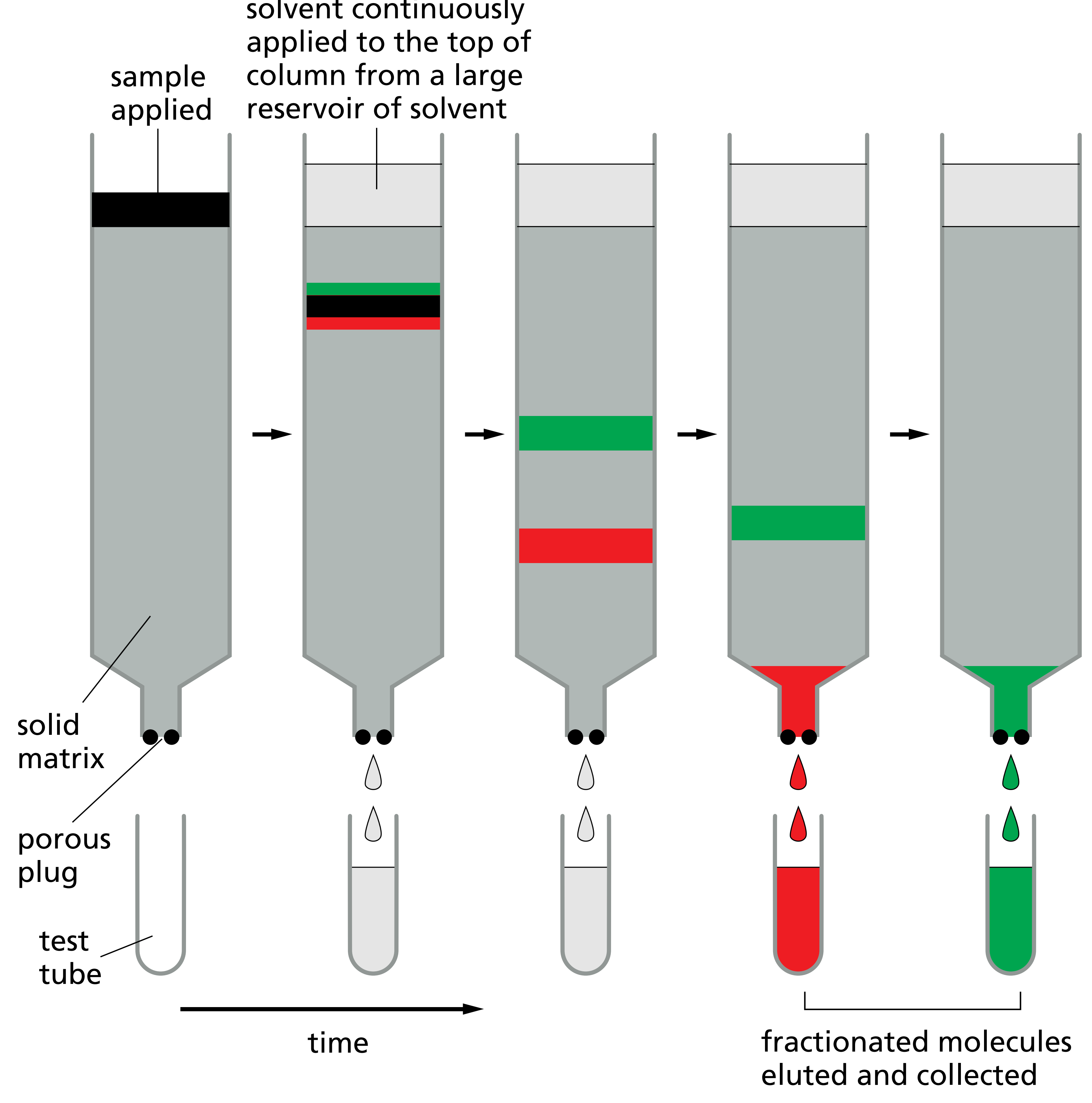
Step 1: opening the cells

- Improving separation using **gradients**



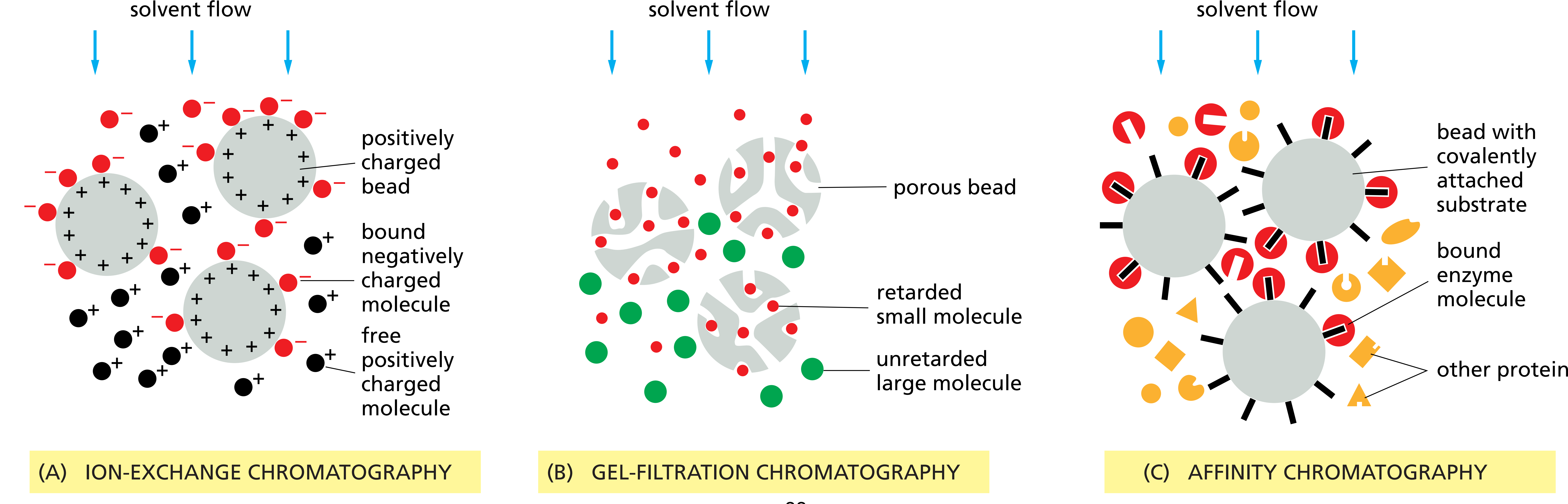
Step 2: separating the proteins

- 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

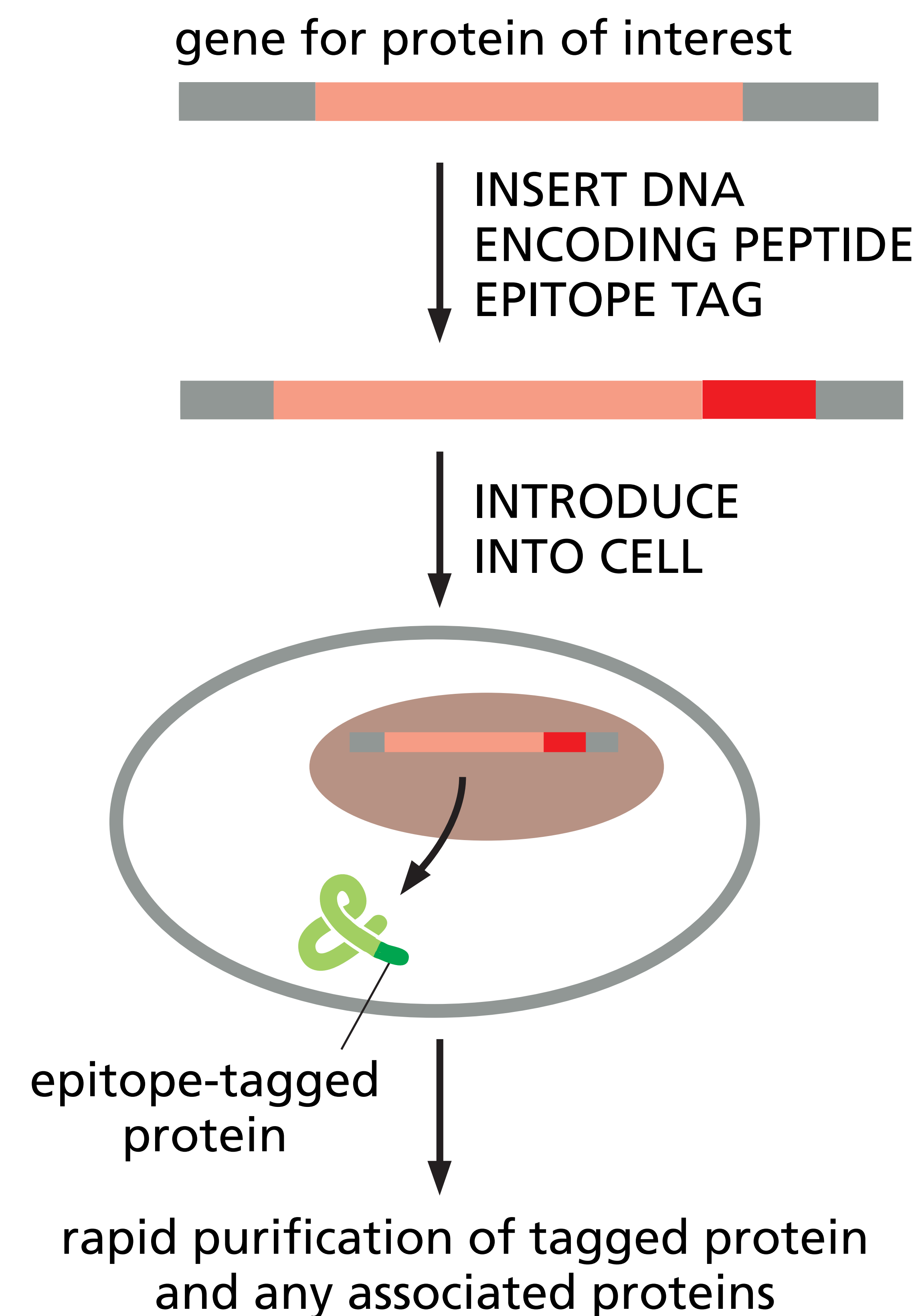
- Depending on the **matrix**, proteins can be separated according to
  - their **charge** (ion-exchange chromatography),
  - their **hydrophobicity** (hydrophobic chromatography)
  - their **molecular weight** (gel filtration = size-exclusion chromatography)
  - their **ability to bind small molecules**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



Step 2: separating the proteins

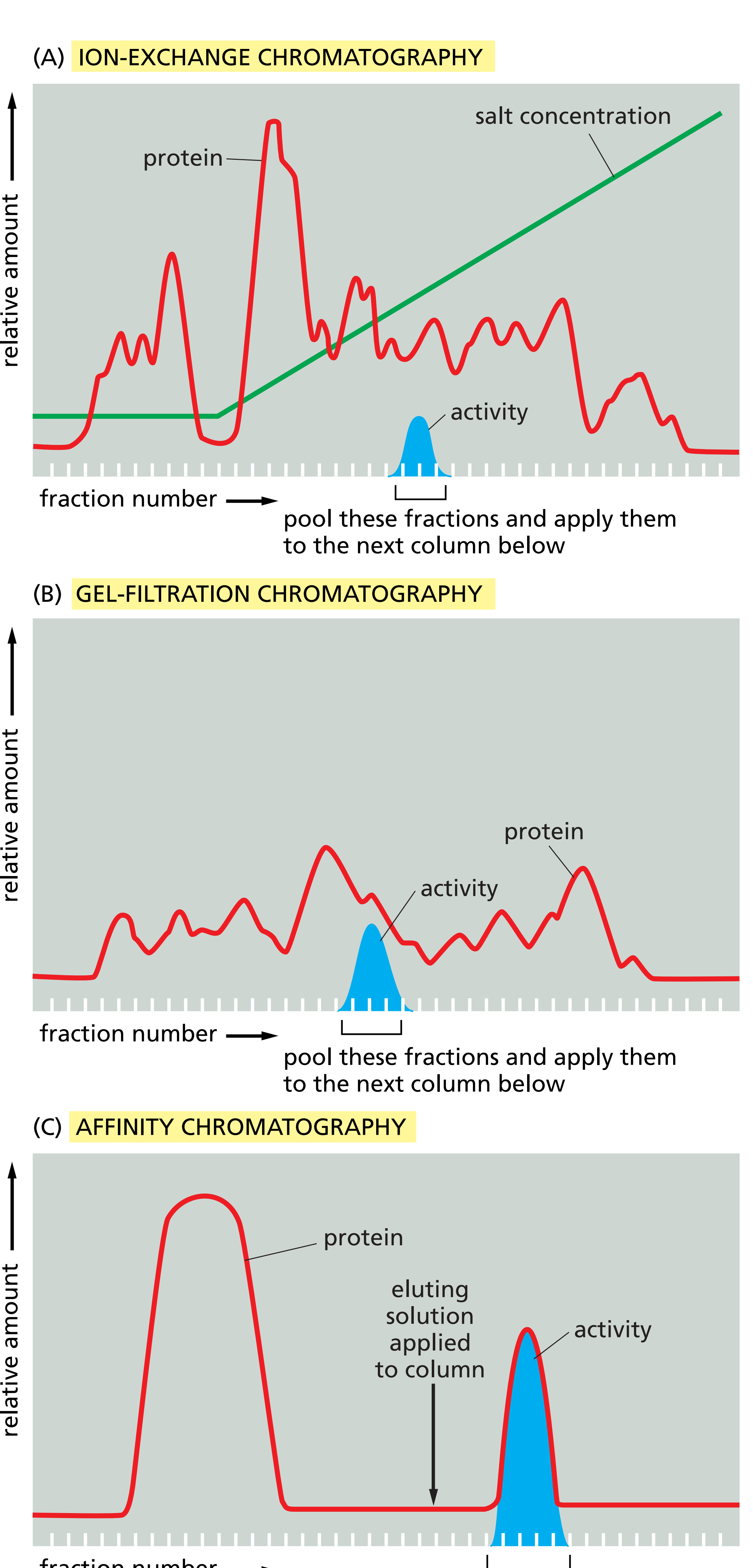
- **Protein tags** provide an easy way to purify proteins
  - His-Tag (addition of 6 His residues) strongly binds to Nickel
  - GST (glutathione-S-transferase, e[...]e protein)
  - **Cleavage site** between the prot[...] and the tag

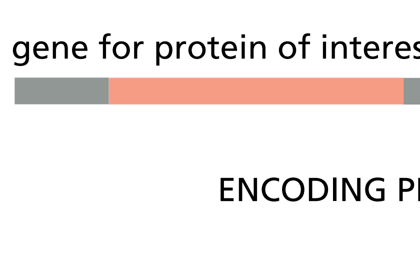
relative amount

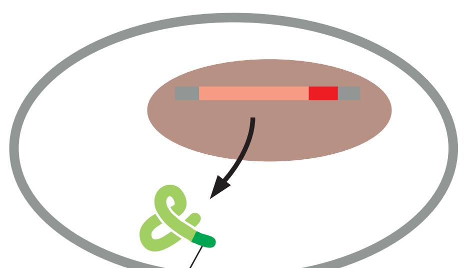


MBoC6 m8.14/8.12

Step 2: separating the proteins

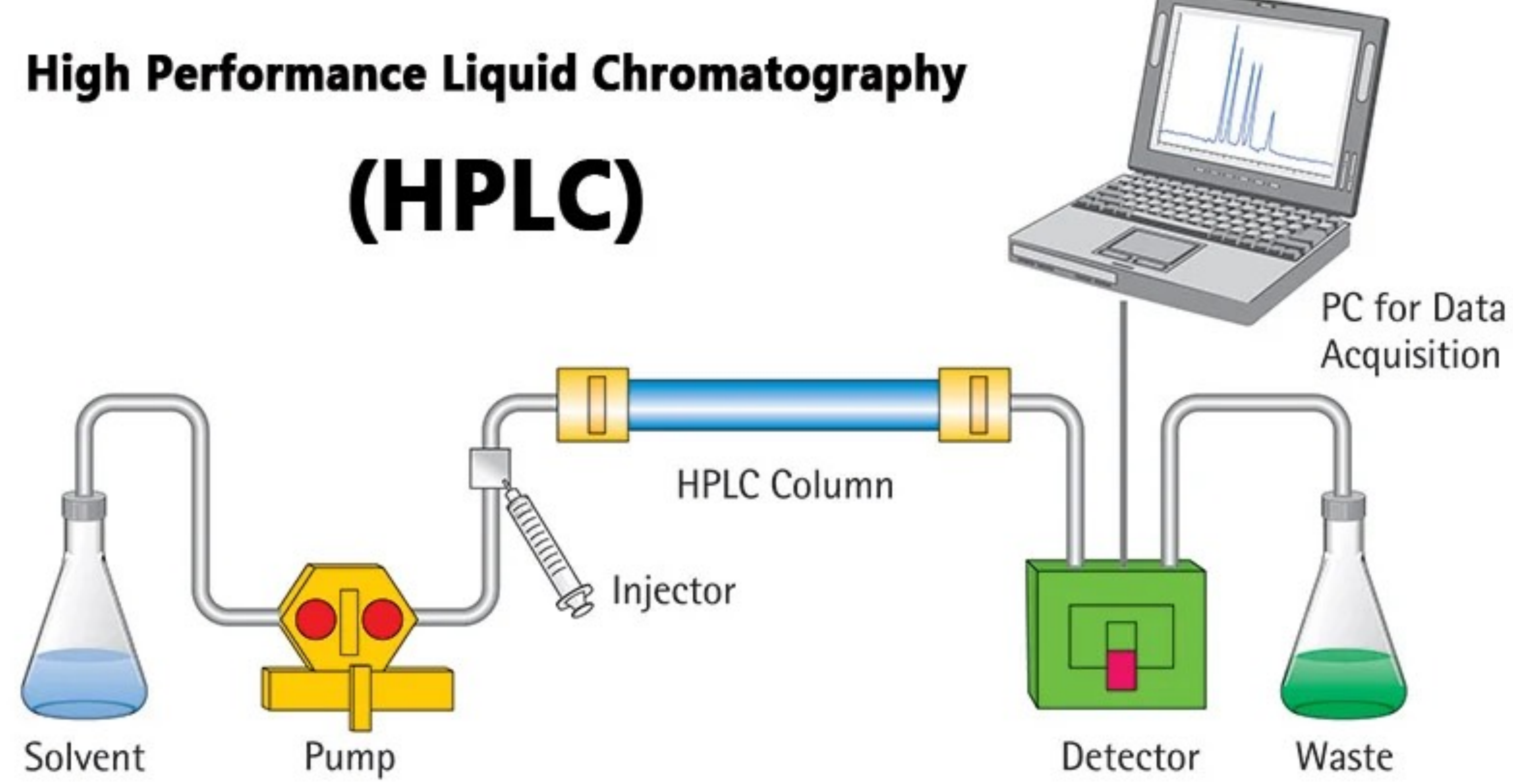






Step 2: separating the proteins

- 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/

V. Molecular and Cellular Biology in the lab

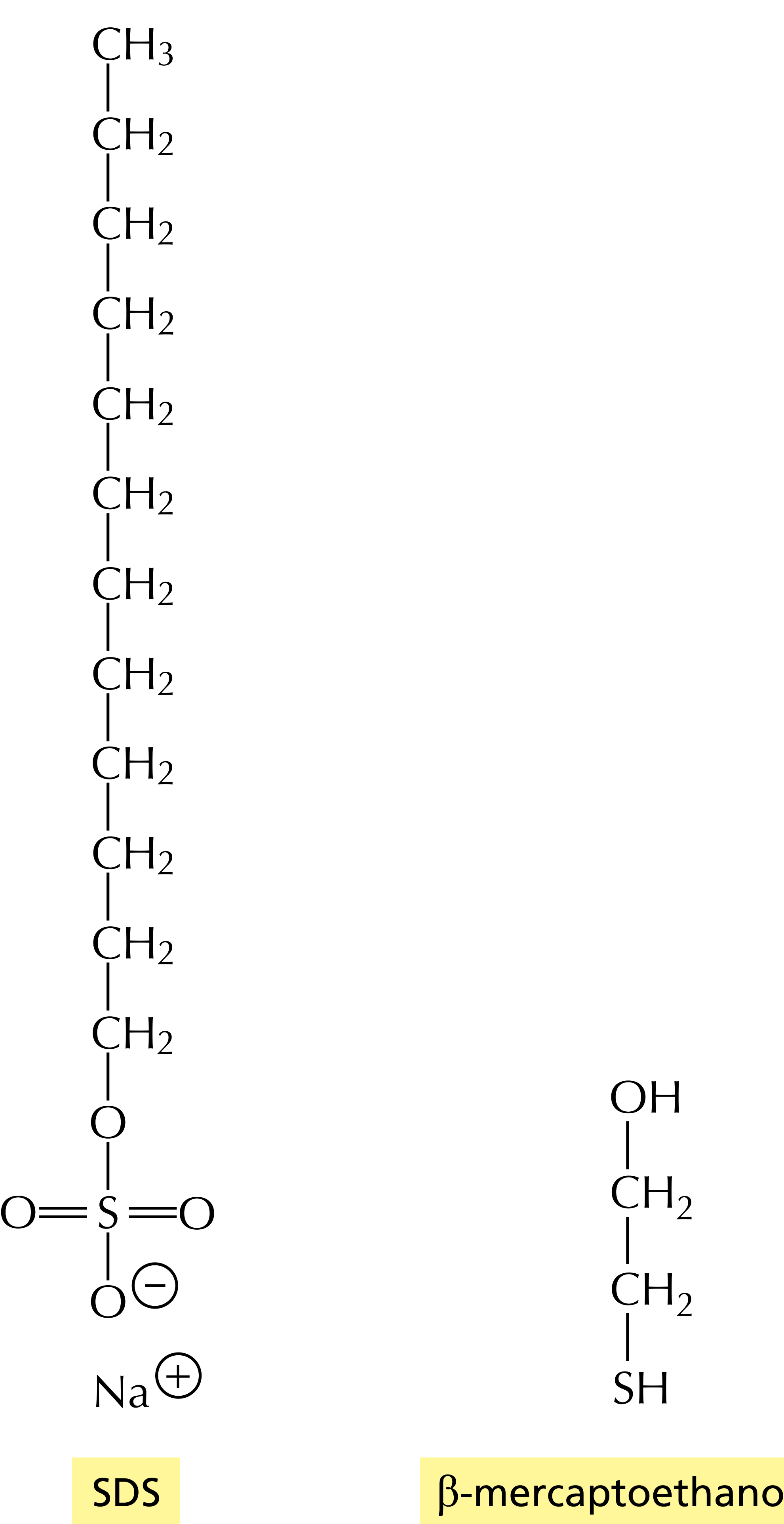
1. Model organisms
2. Cell cultures
3. **Studying proteins**
    - Protein sequence
    - Protein purification
    - **Protein visualization**
    - Protein structure
    - Mass spectrometry

Visualizing/Separating proteins on a gel

- SDS-PAGE = **sodium dodecyl sulfate polyacrylamide gel electrophoresis**
- When an **electric field** is applied to a solution containing proteins, they will **migrate** depending on their 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**



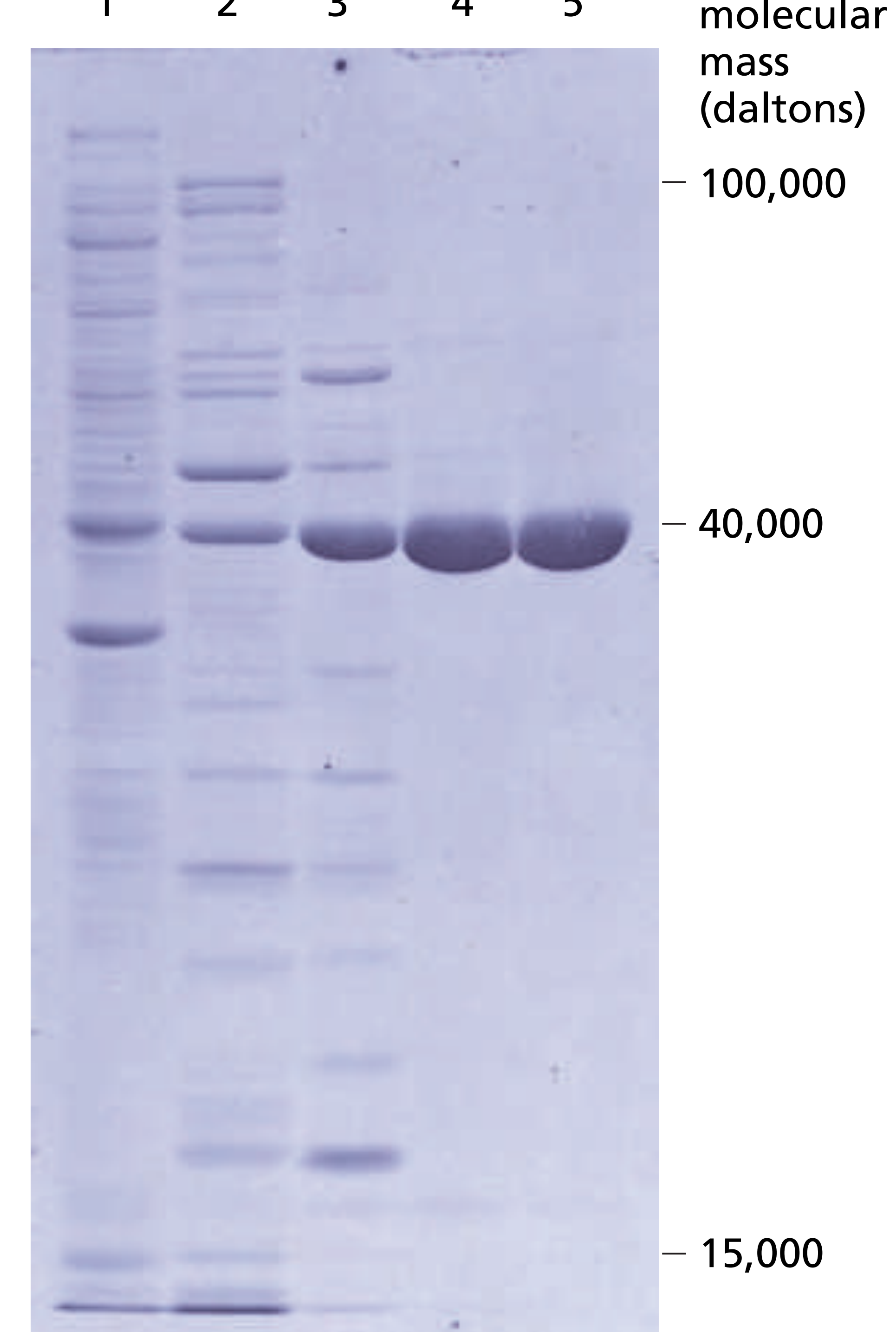
Proteins can be separated on SDS-PAGE

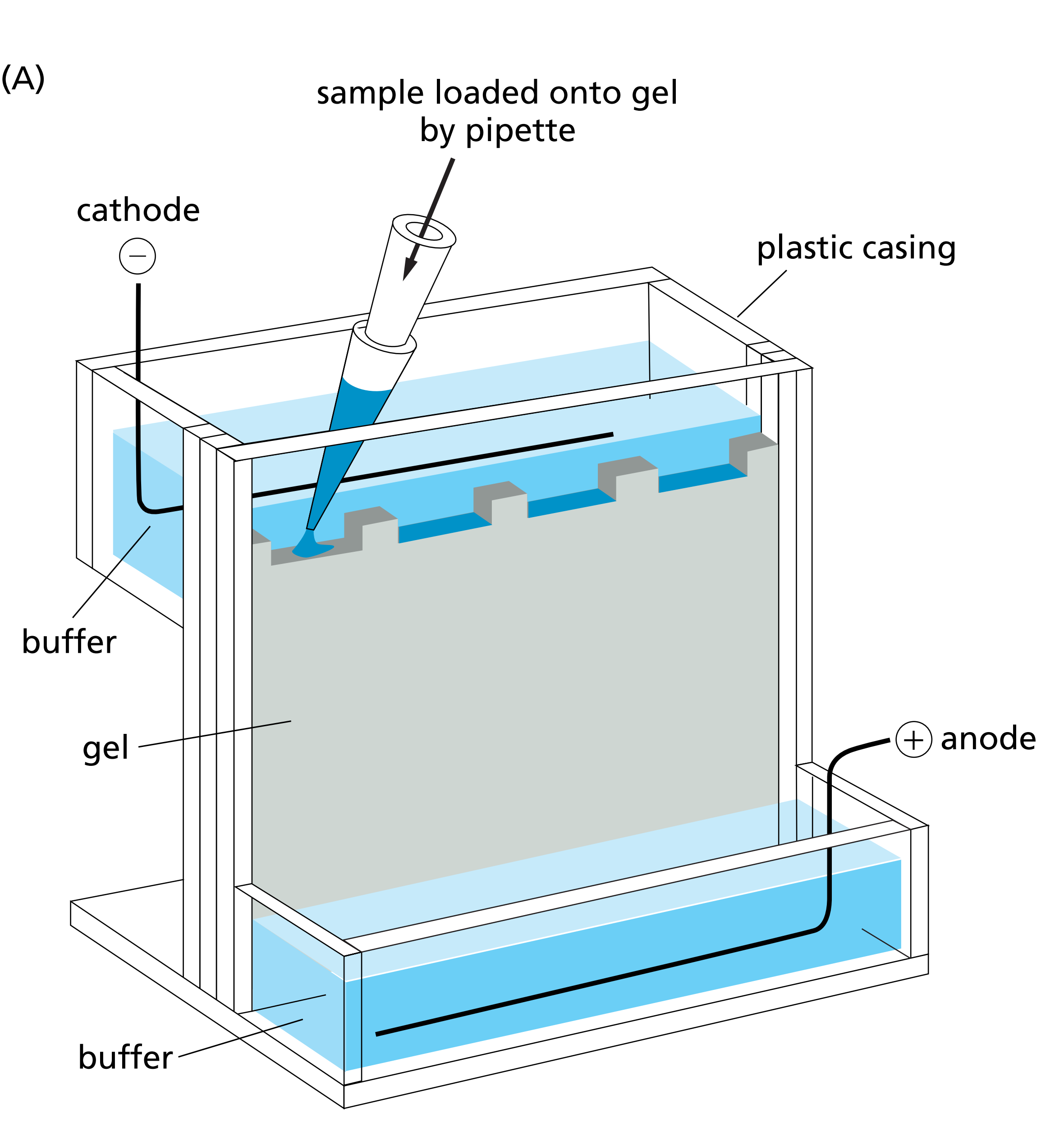
- Binding of SDS masks the **intrinsic charge of the protein** and causes it to migrate towards the **positive electrode** during SDS-PAGE
- Proteins with the **same size** move at the **same speed**
    - they are **unfolded** so their shape is the same
    - they bind the **same amount of SDS** so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

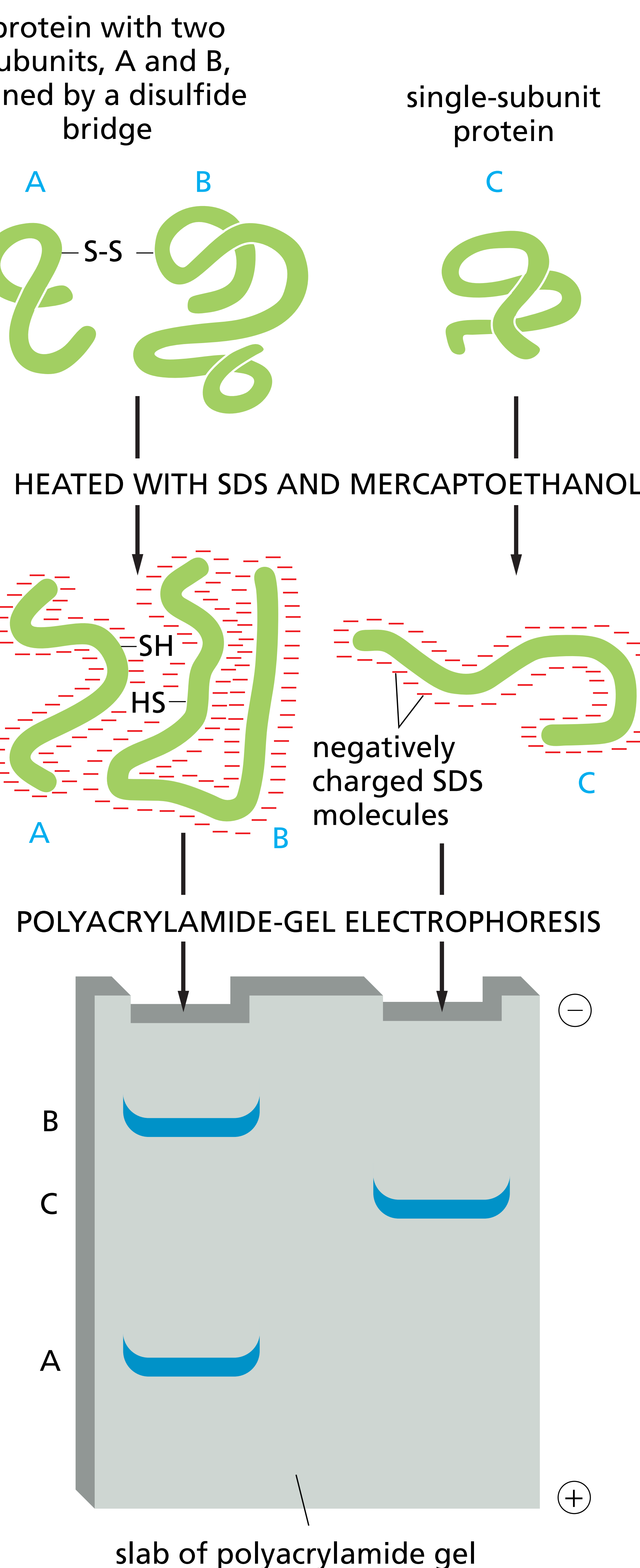
Proteins can be separated on SDS-PAGE

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

MBoC6 m8.18/8.15

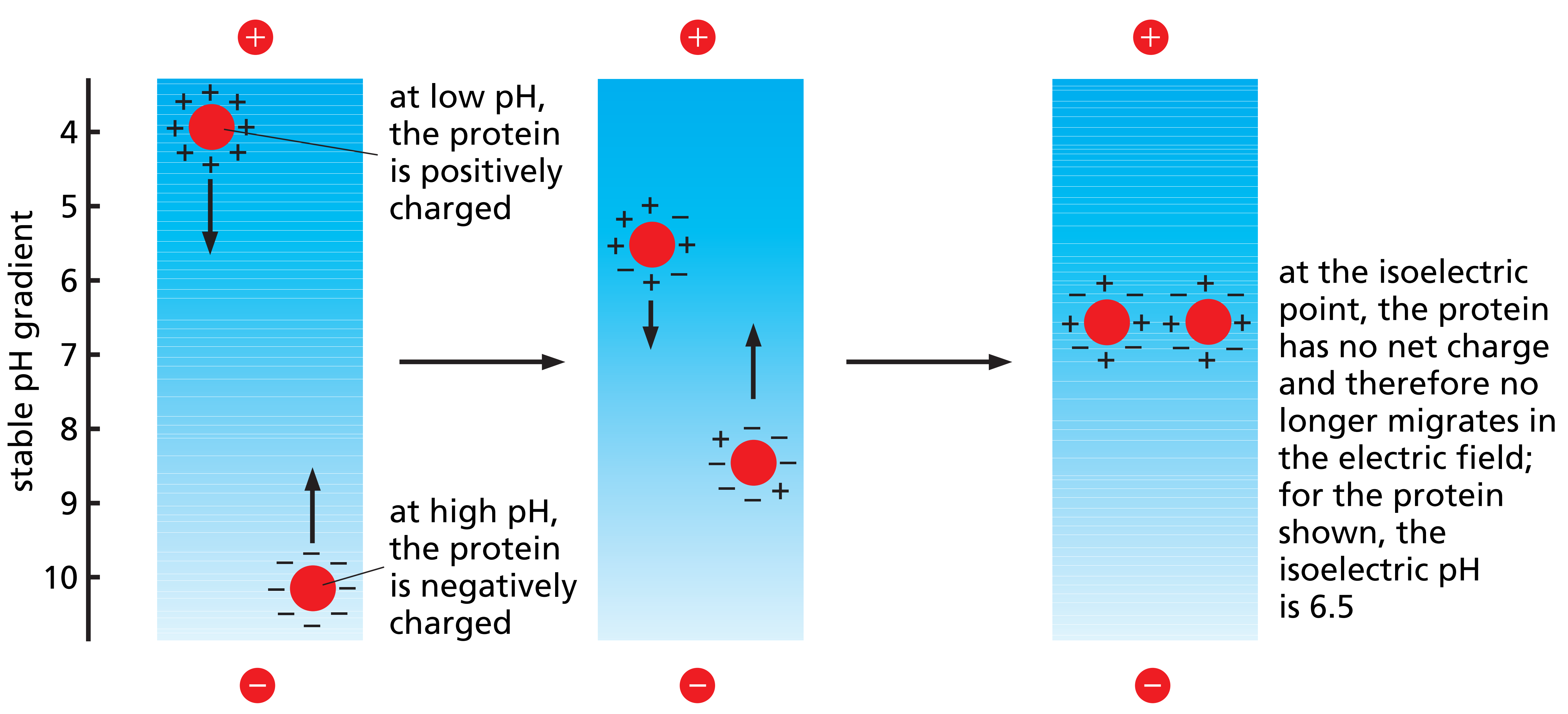






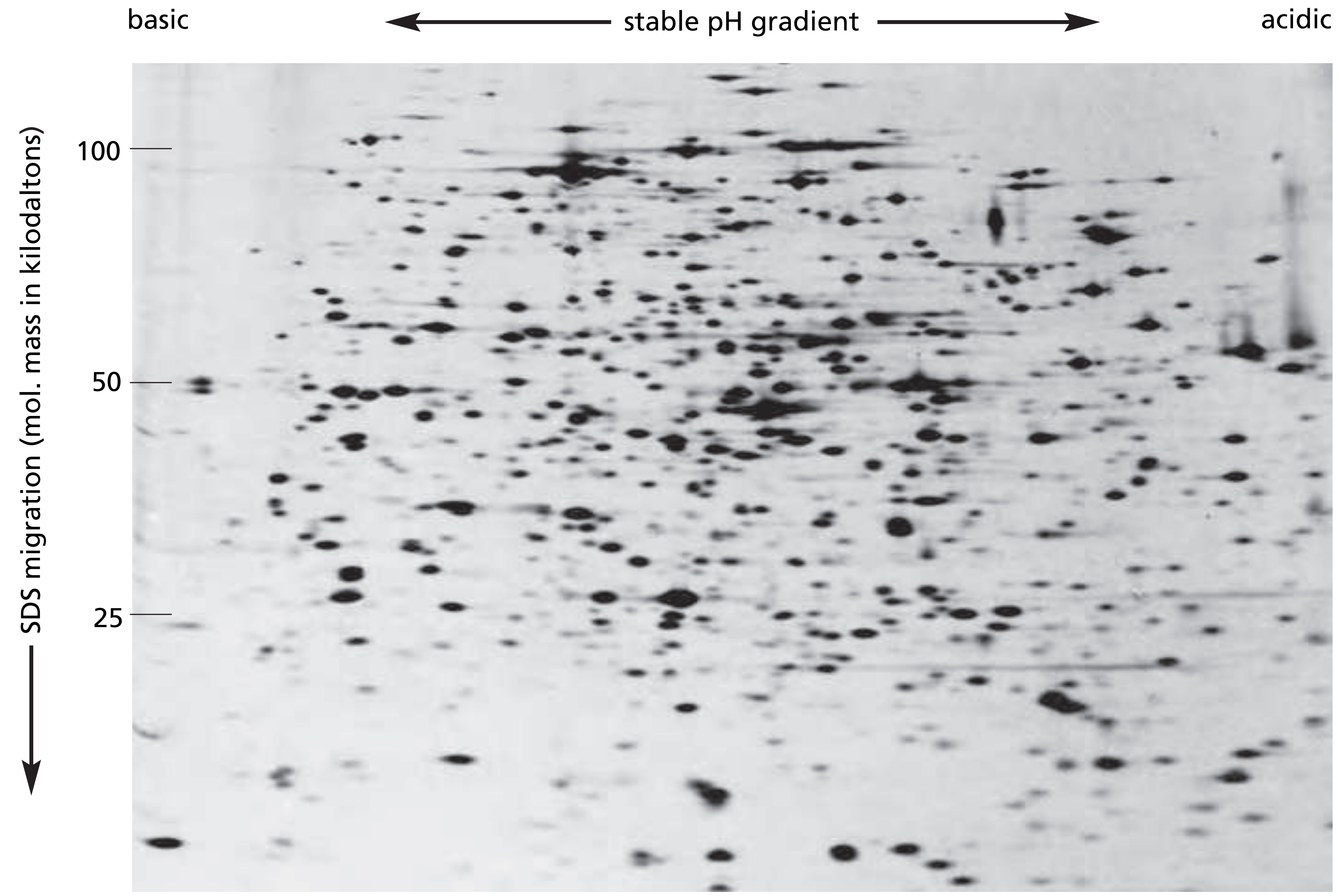
2D-gel electrophoresis

- Two proteins can have the **same size/shape/mass** **but a different intrinsic 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**

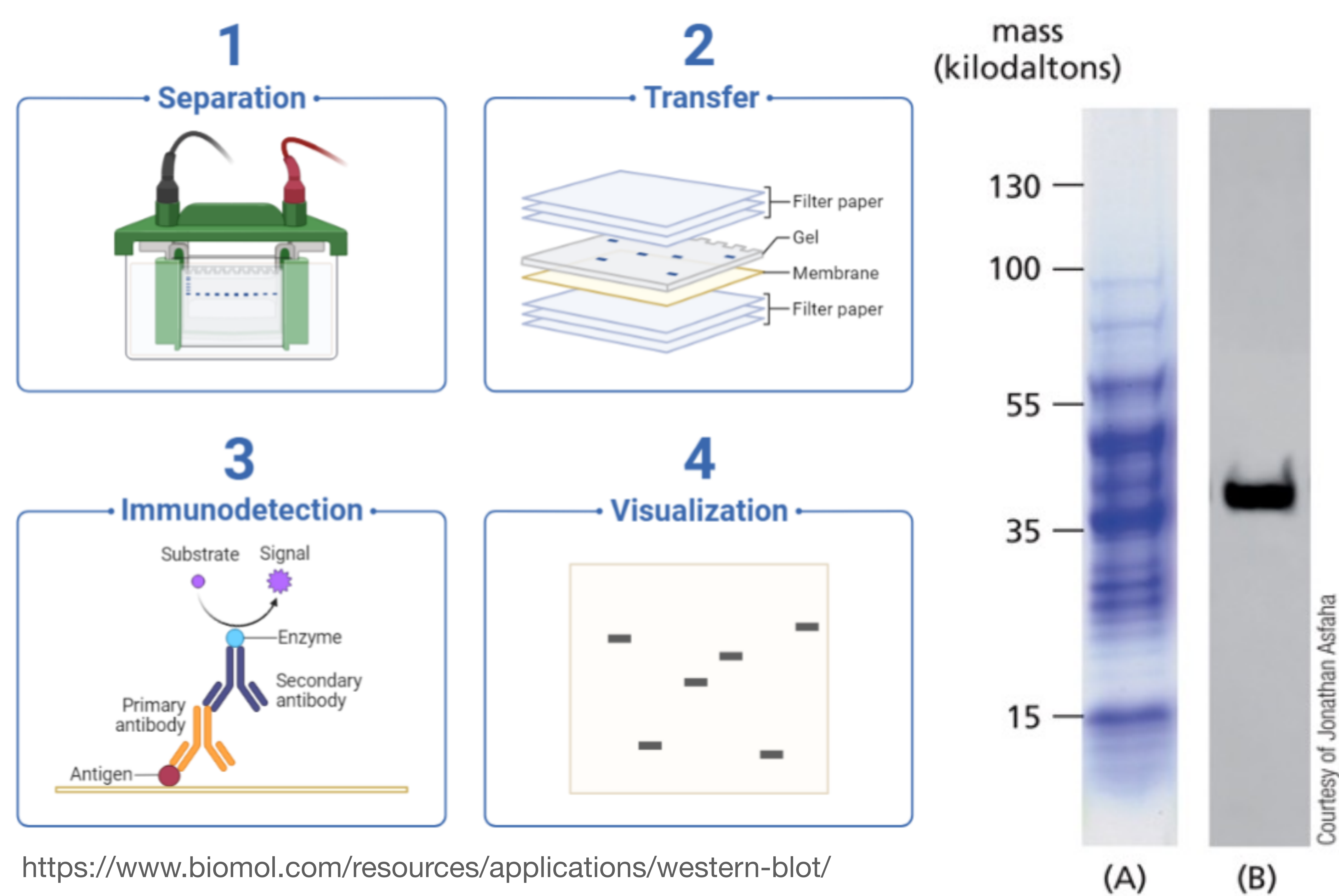


Detecting proteins by Western Blot

- 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.

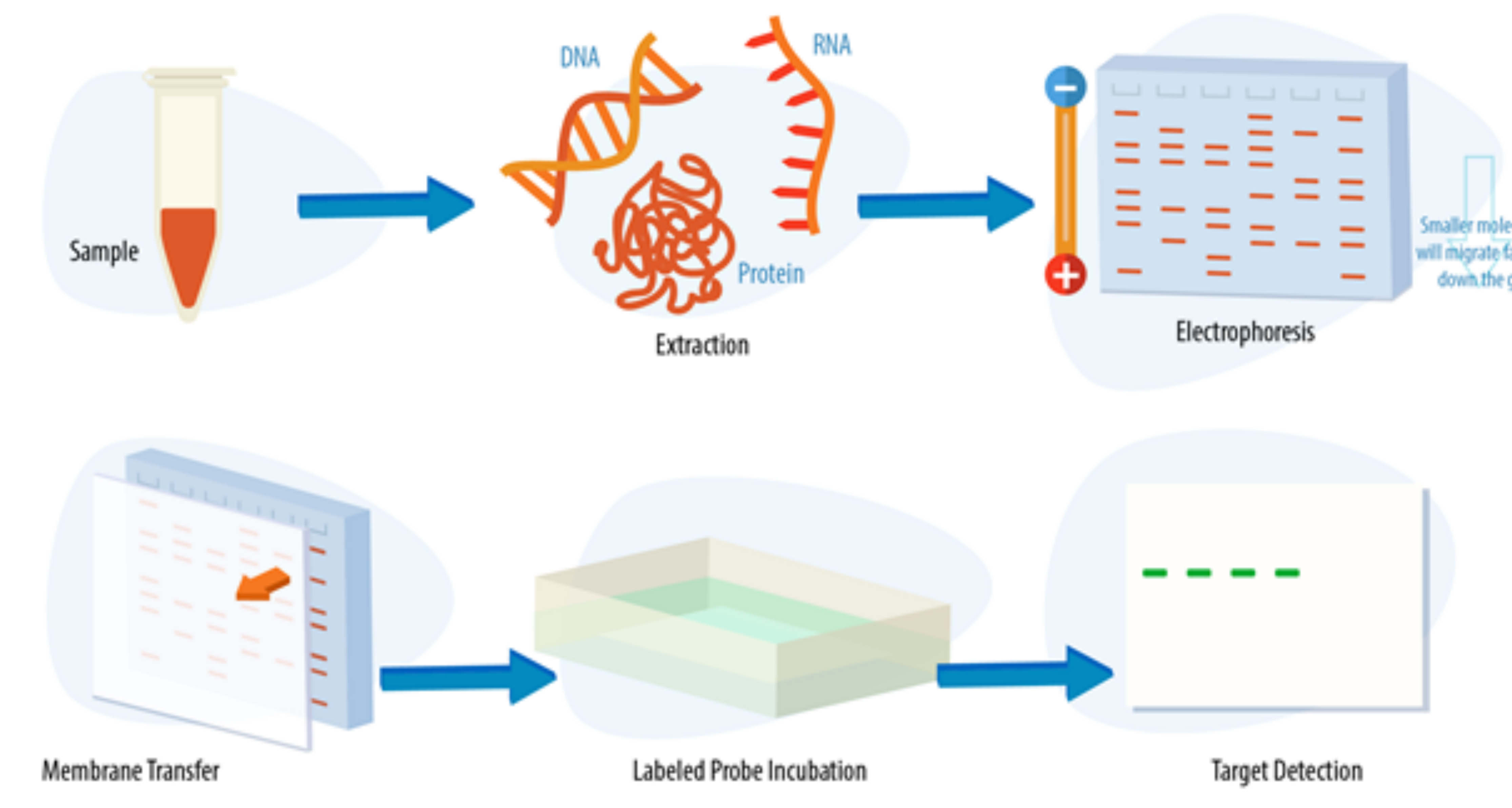
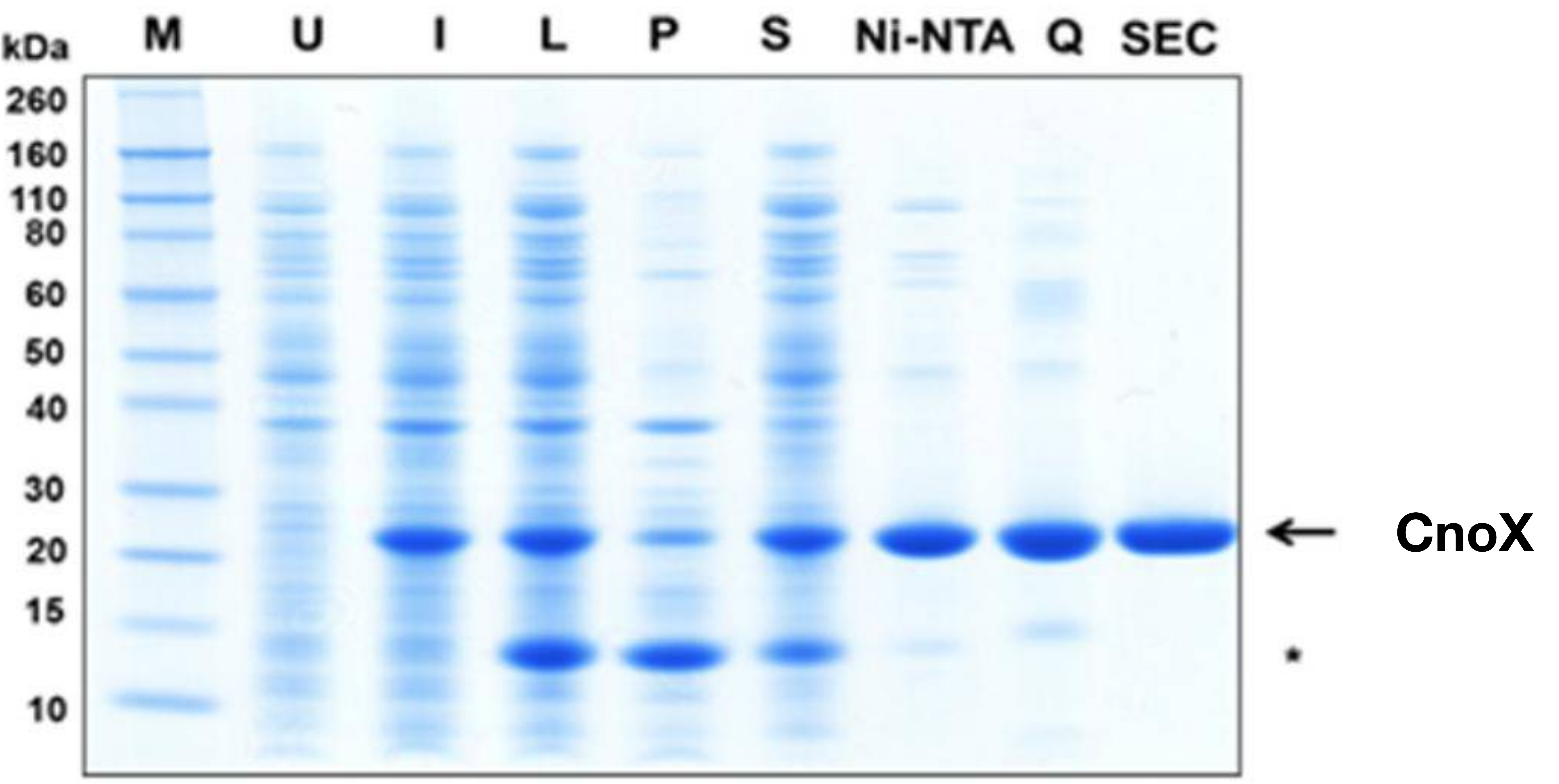


Figure 1: Gel electrophoresis, transfer, incubation, and detection.

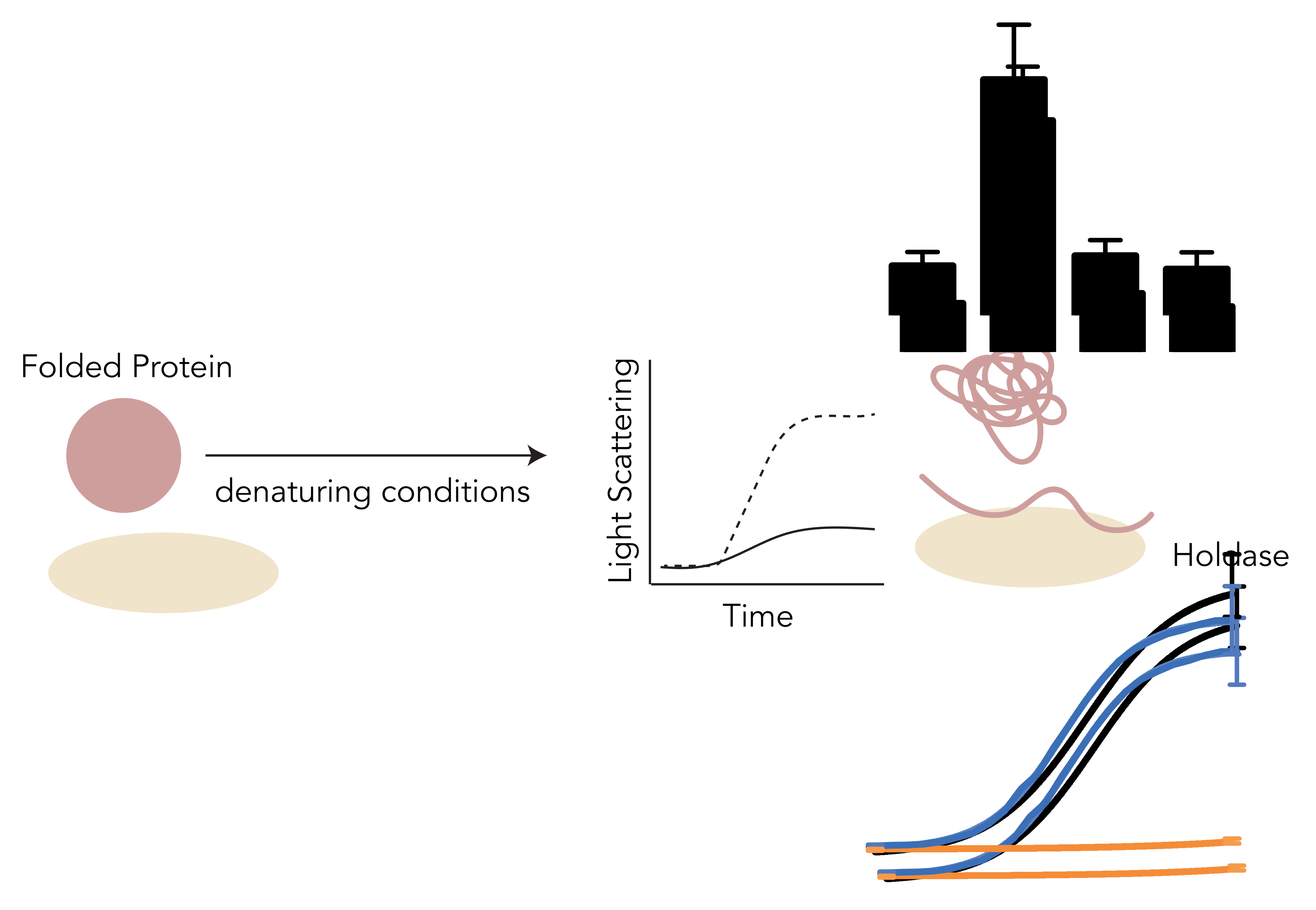
https://www.labmanager.com/southern-vs-northern-vs-western-blotting-techniques-854

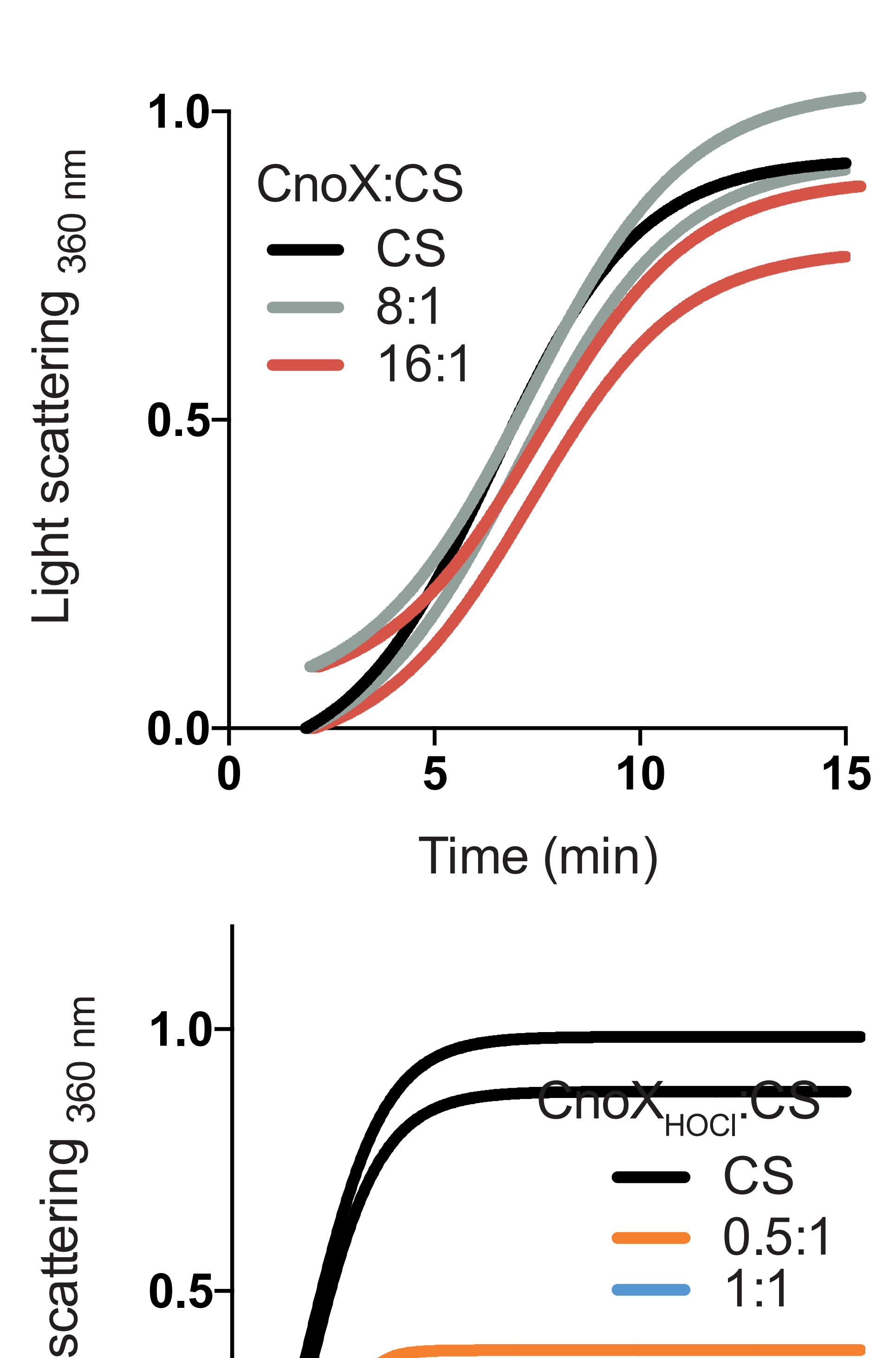
Purifying CnoX*

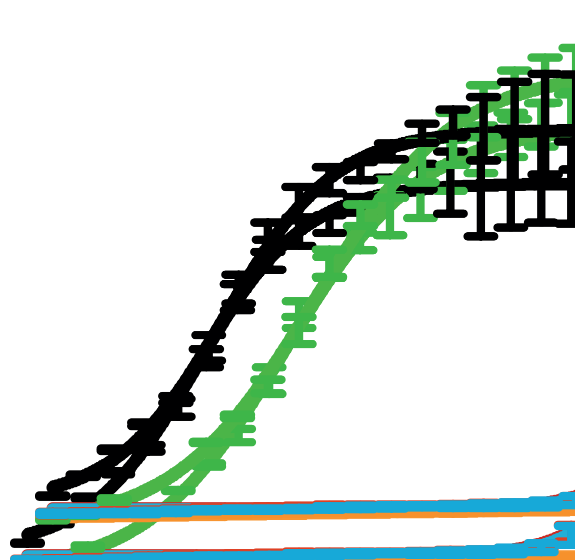


Using purified CnoX for biochemical tests*

Test it biochemically







V. Molecular and Cellular Biology in the lab

1. Model organisms
2. Cell cultures
3. **Studying proteins**
    - Protein sequence
    - Protein purification
    - Protein visualization
    - **Protein structure**
    - Mass spectrometry

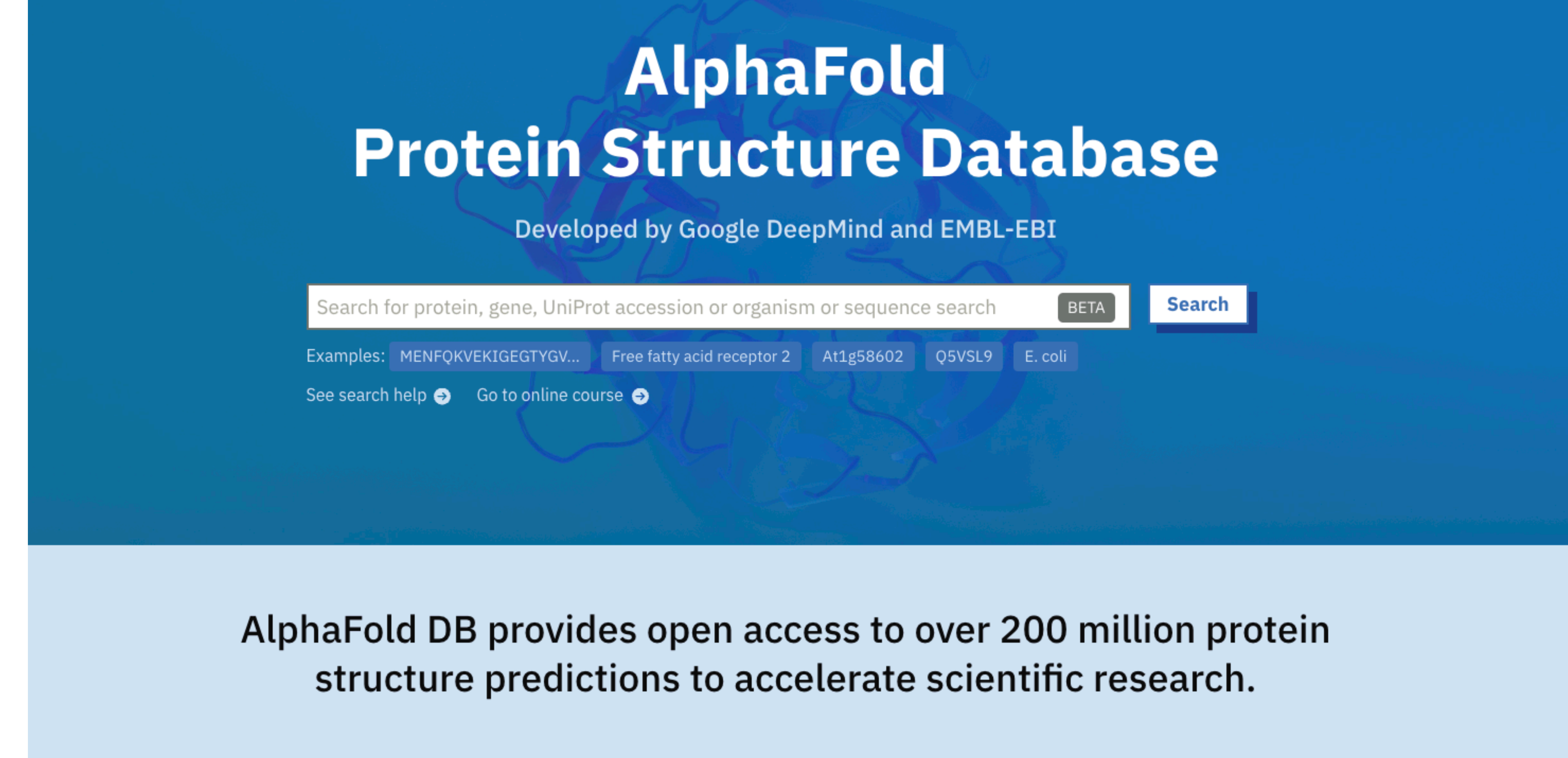
Solving protein structures

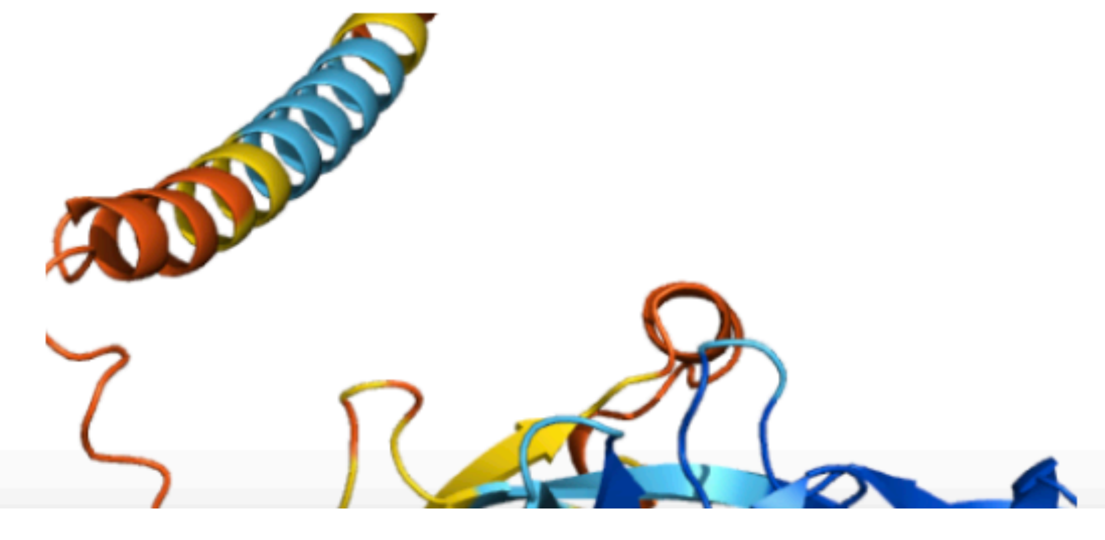
➡ Mapping the **precise position** of each amino-acid within the **3D structure**

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

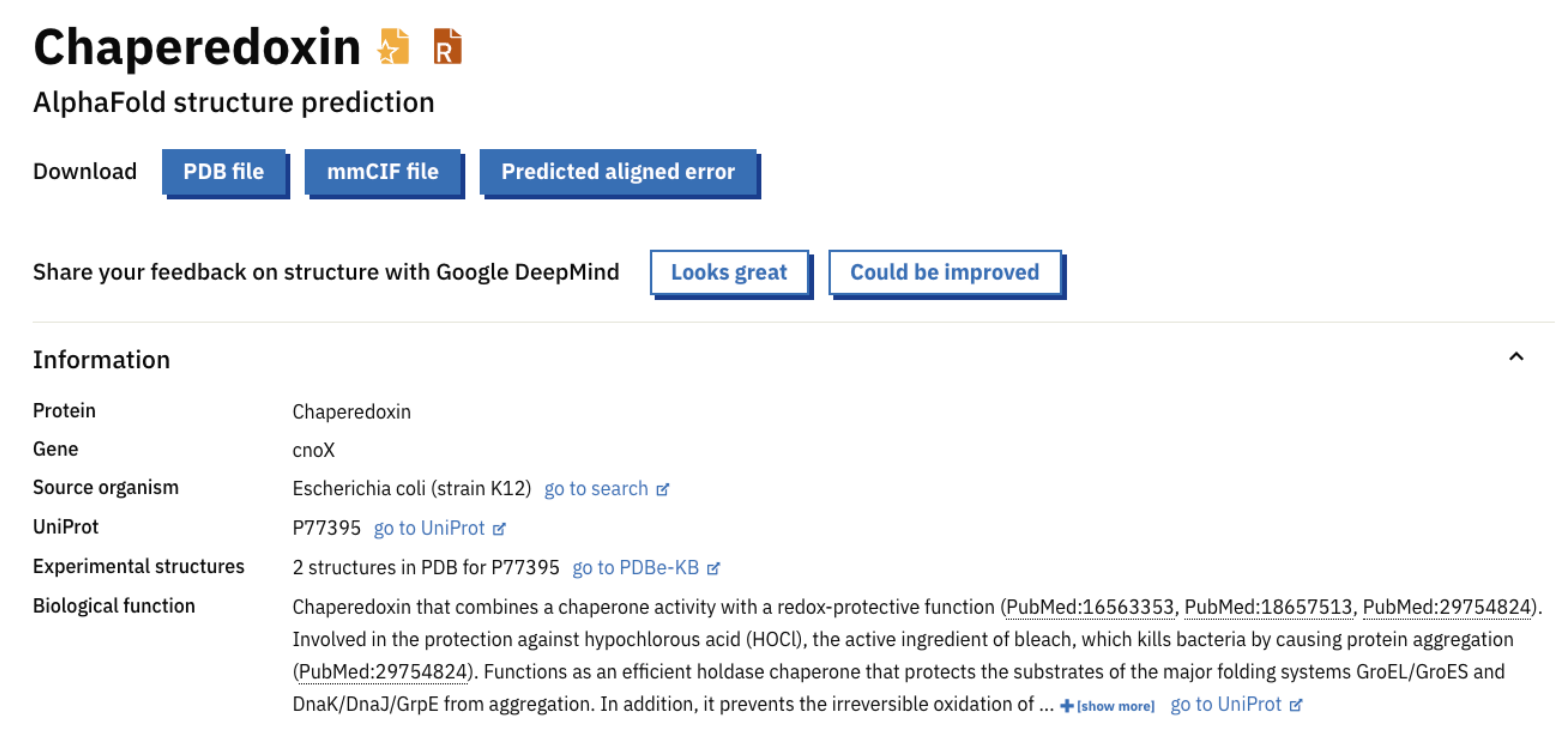
AlphaFold

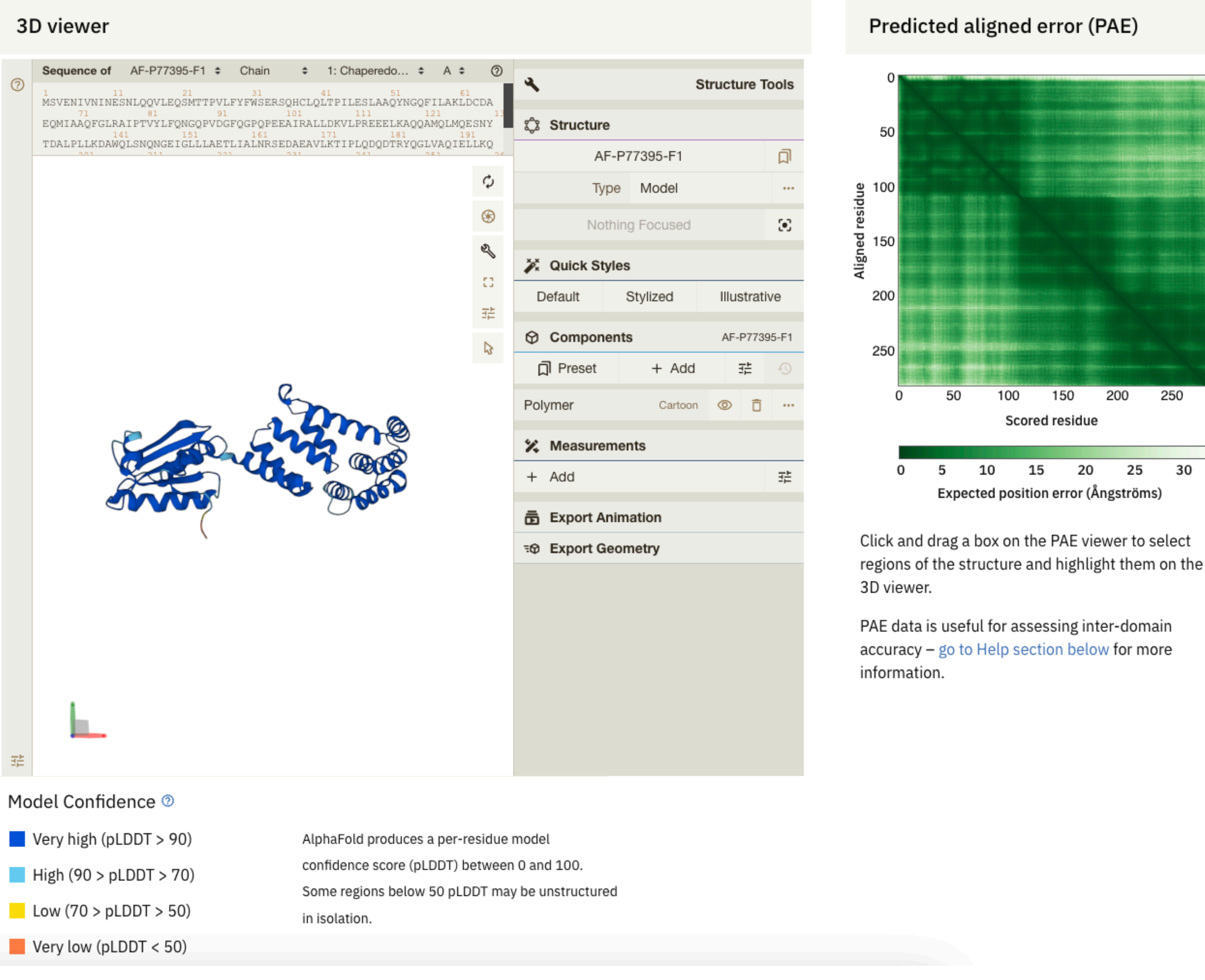
! prediction of the structure





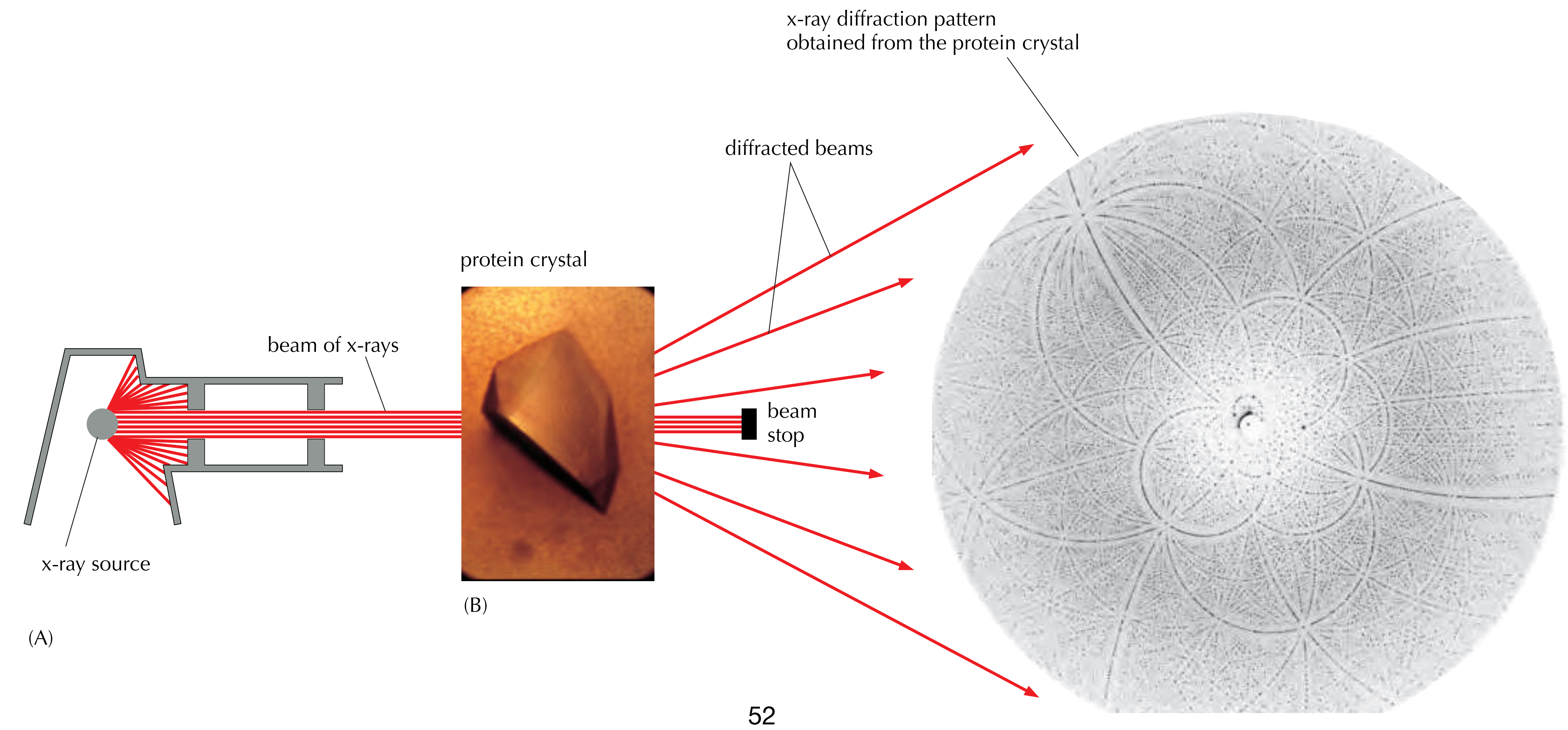
AlphaFold





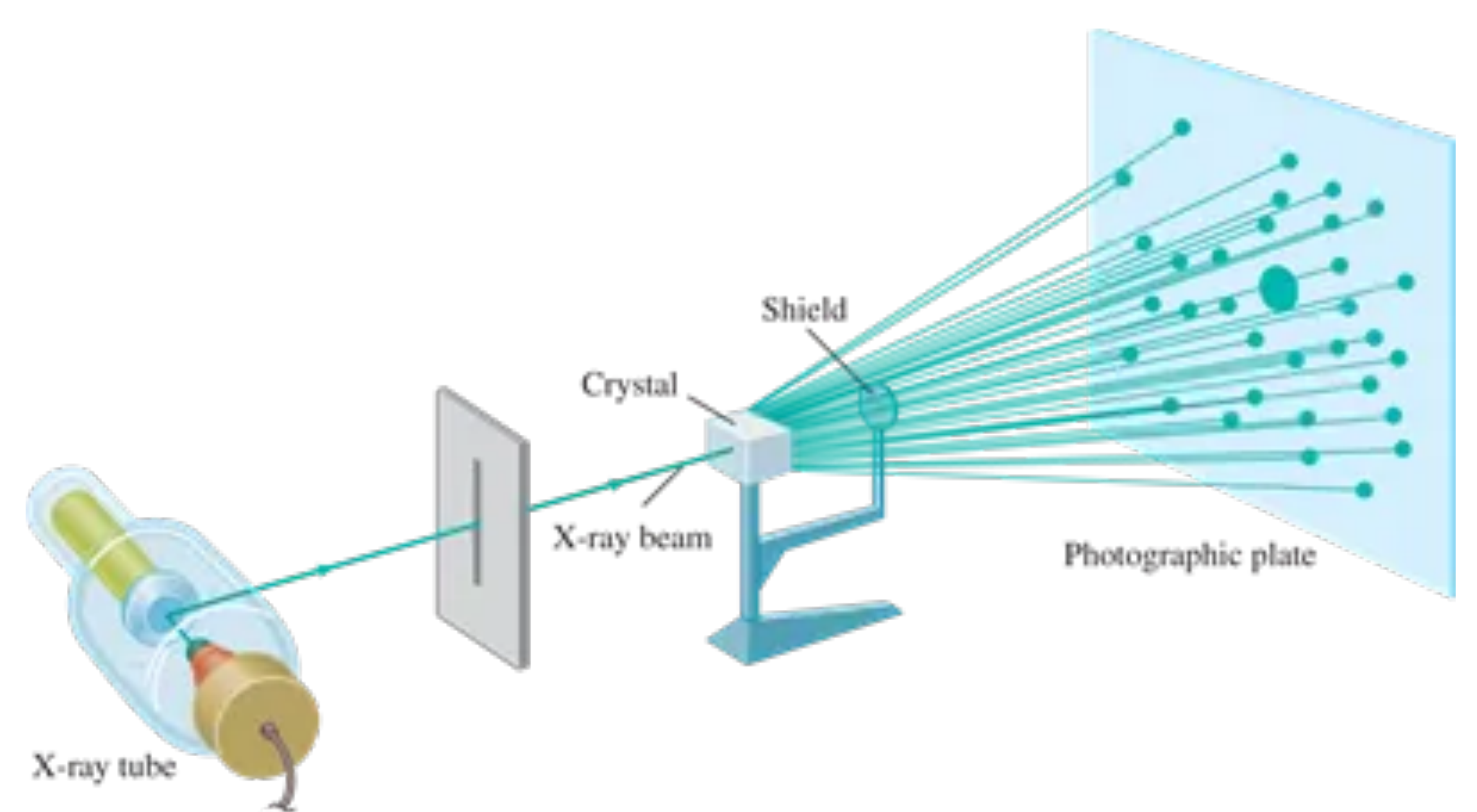
X-ray cristallography

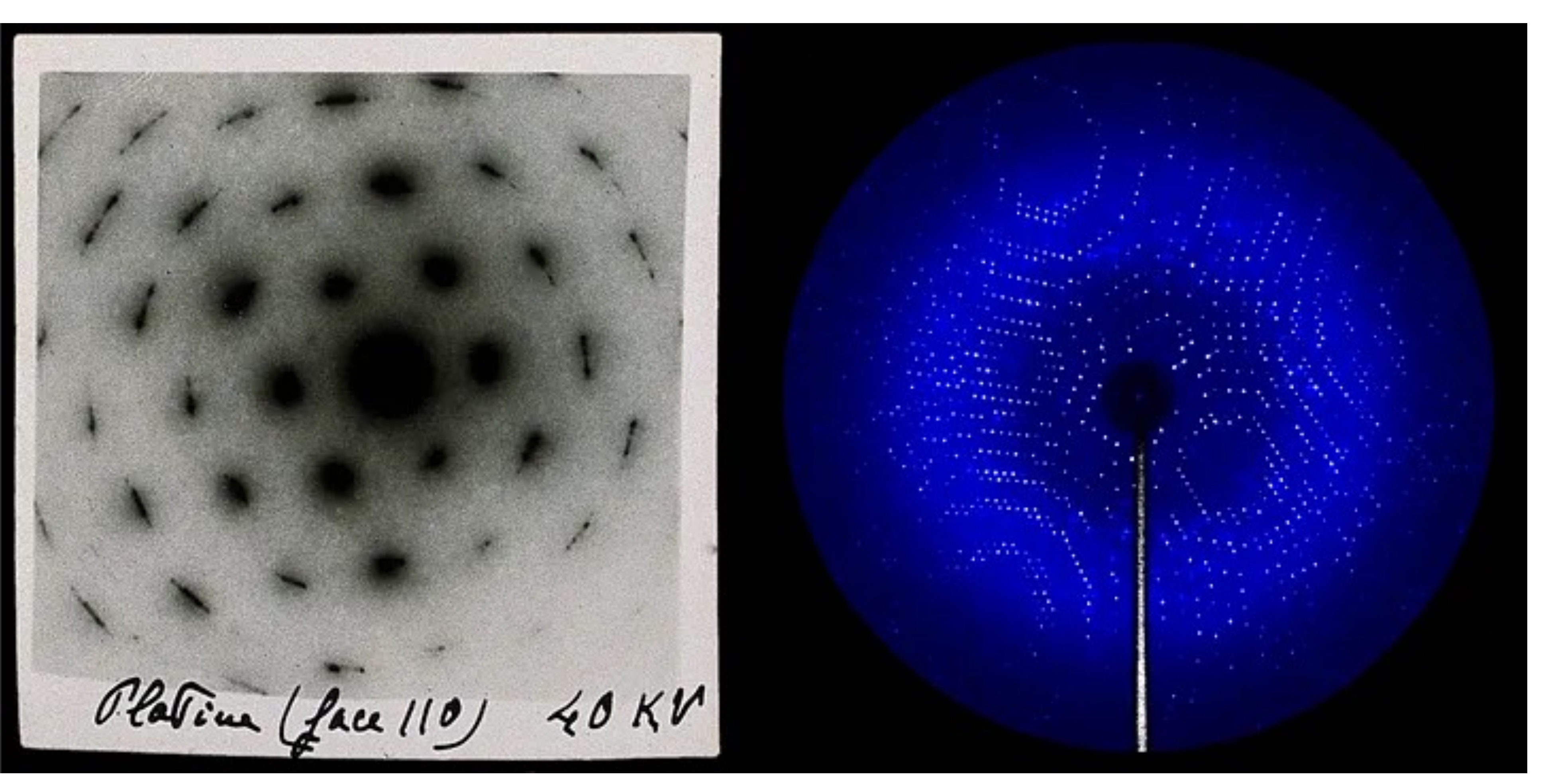
- X-rays are a form of **electromagnetic radiation** with short wavelength (~0,1 nm)
- If a narrow beam of X-rays is directed at a sample of **pure proteins**, most pass through it but a small fraction is **scattered by the atoms of the protein**
- If the sample is a **well-ordered crystal,** the scattered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



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

Using purified CnoX to solve its 3D structure*

Solve its 3D-structure (cristallography)

Nuclear magnetic resonance spectroscopy (NMR)

- Does not depend on a **protein crystal**
- Only suited for **small proteins** (< 50 kDa)
- 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

Cryo-Electron Microscopy (CryoEM)

- Technique that uses an electron microscope to image samples that are rapidly frozen in a thin layer of ice
- Unlike X-ray crystallography, CryoEM does not require crystals, making it ideal for large or flexible complexes

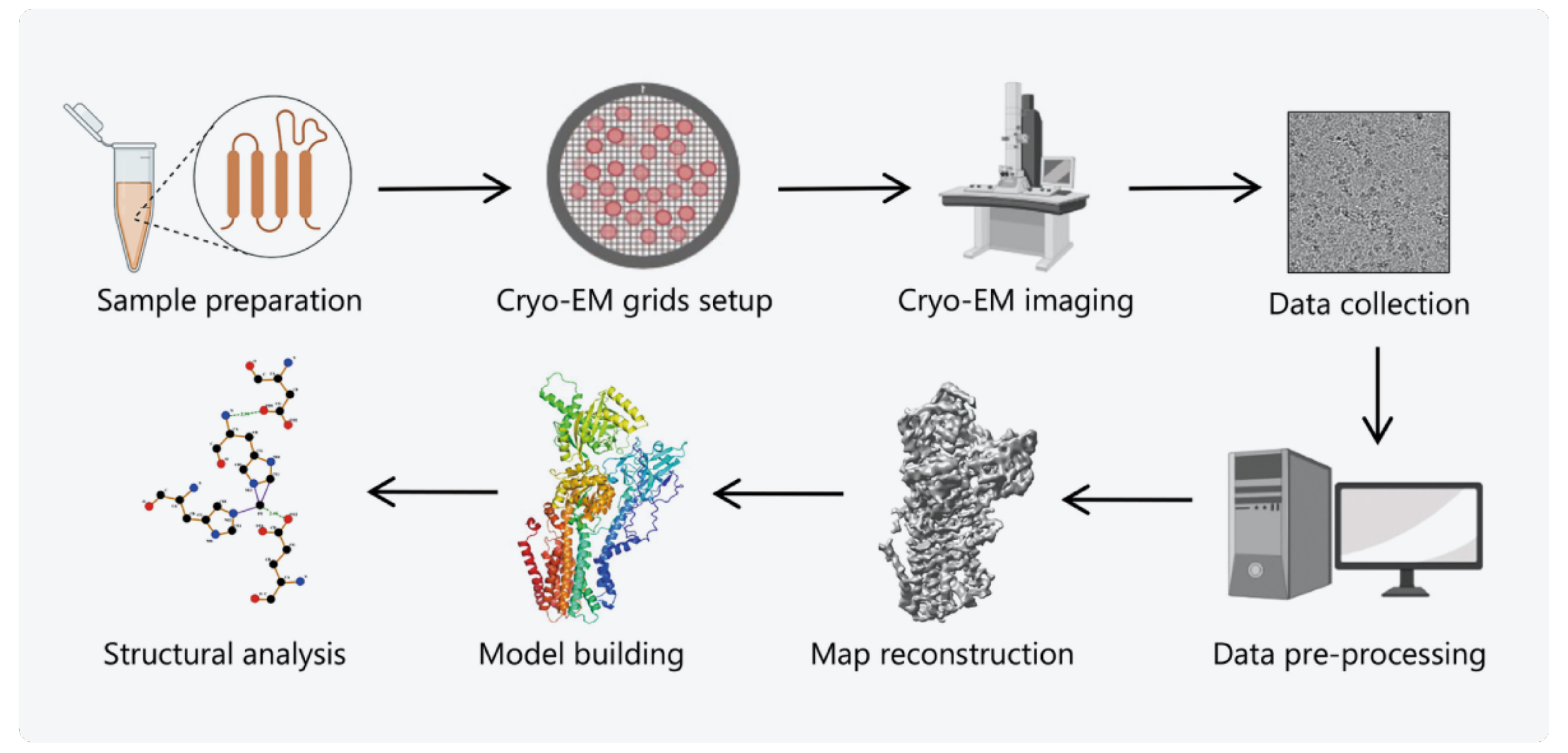
- **Sample prep:**
  1. purified protein or complex is applied to a tiny grid
  2. the grid is plunged into liquid ethane to freeze the sample

  ➡ molecules are trapped in random orientations, this can capture multiple conformations, if any.

- **Data collection:**
  1. The sample is loaded in a cryo electron microscope and an electron beam passes through the sample
  2. An electron detector captures thousands to millions of 2D projections of images in different orientations

Cryo-Electron Microscopy (CryoEM)

- **Image processing**
  1. the software identifies and "cuts-out" the particles (protein projections)
  2. 2D classification: the similar 2D particles are averaged together (2D class averages) to enhance signal and reduce noise
  3. 3D reconstruction using 2D projections
  4. Building of an atomic model into the density



Start the presentation to see live content. For screen share software, share the entire screen. Get help at **pollev.com/app**

Have a nice day!