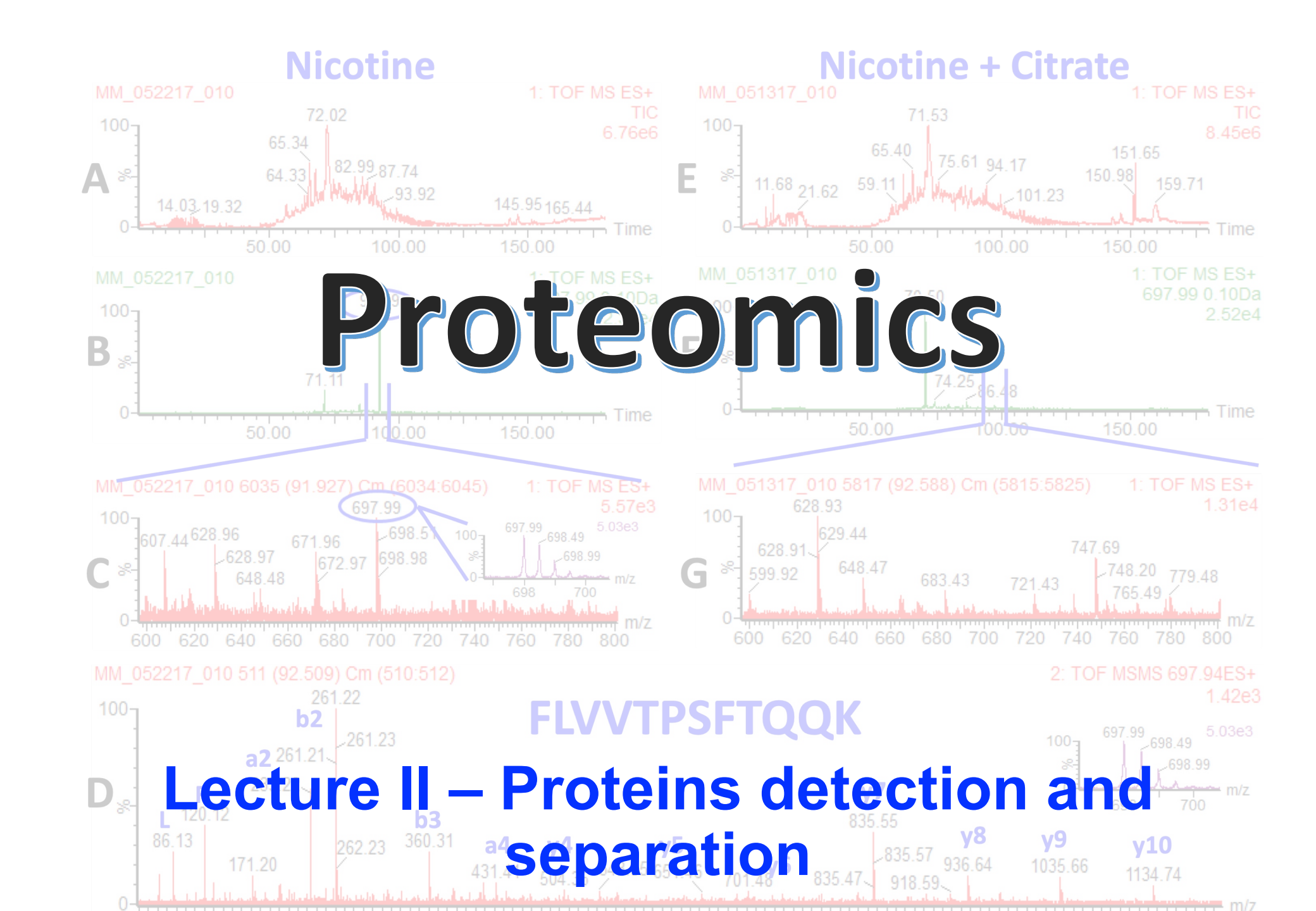
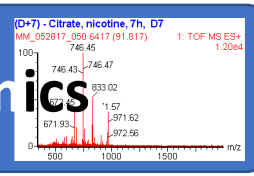
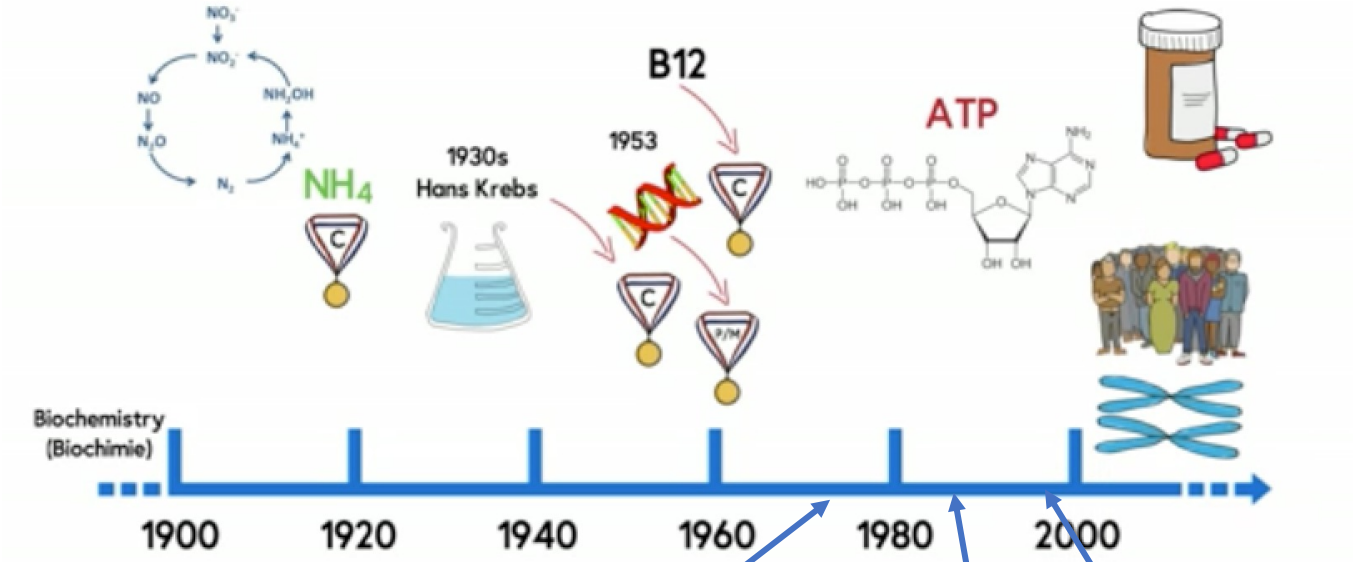

Nicotine
Nicotine + Citrate
Proteomics
Lecture II – Proteins detection and separation
FLVVTPSFTQQK

Key technologies that have enabled the development of proteomics





1975 – first proteomics study:
2D-SDS PAGE of *E. coli* proteins

1988 , Hillenkam develops the MALDI-MS

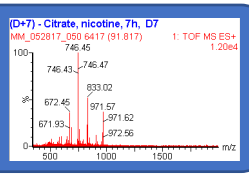
1988, Fenn develops the ESI-MS

1995, Wilkins coins the term proteomics

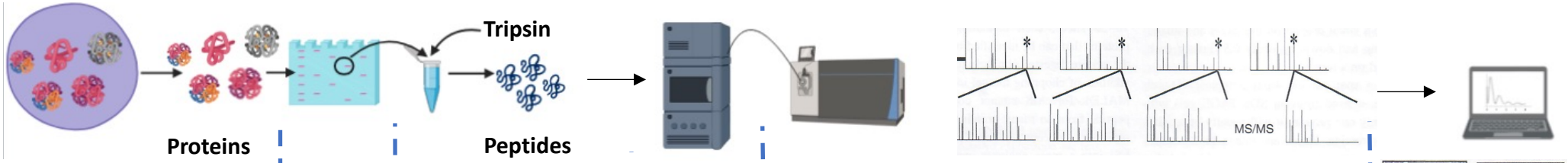
The emergence and recent development of proteomics as a branch of life sciences have been made possible due to significant discoveries in four major directions:

1. Development of methods for **separation, fractionation, and detection** of proteins and peptides.
2. Development of accurate **mass spectrometry instruments** and efficient **peptide ionization methods**.
3. Sequencing of nucleic acids and deposition of **sequences** in public **databases**.
3. Development of bioinformatics methods for **computer-based processing** of mass spectra.

Steps in a proteomics experiment



For the simultaneous study of the entire set of proteins produced at a given moment by a cell, tissue, or organism, various proteomic methods have emerged, but they commonly involve the following major steps:



A. **Cell lysis** and **protein extraction** from a given biological sample

B. **Protein fractionation** most commonly performed using electrophoresis (not a mandatory step).

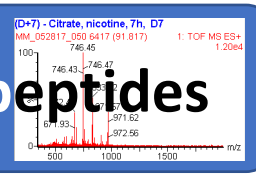
C. **Protein digestion** in order to generate shorter peptides.

D. **Peptides fractionation** using liquid chromatograpgy

E. **Sequencing of the resulting peptides** using mass spectrometry.

F. Raw data processing and identification of proteins present in the sample using bioinformatics tools.

Chapter I. Separation, fractionation, and detection of proteins/peptides



One of the problems that proteomic studies must solve is the complexity of the sample to be analyzed, a complexity that refers to:

A. **The presence of a large number of non-protein contaminants**. In any study, one starts from a biological sample that is disaggregated (by pulverization, lysis, sonication, homogenization) to release the protein molecules. However, along with these, a large number of molecules that are not of interest in this case are also released - lipids, carbohydrates, nucleic acids - and thus become **contaminants**.

TABLE 1–1	Molecular Components of an *E. coli* Cell	
	Percentage of total weight of +1cell	Approximate number of different
Water	70	1
Proteins	15	3,000
Nucleic acids		
DNA	1	1
RNA	6	>3,000
Polysaccharides	3	5
Lipids	2	20
Monomeric subunits and intermediates	2	500
Inorganic ions	1	20

Table 1-1
Lehninger Principles of Biochemistry, Fifth Edition

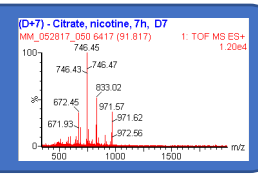
B. **The large number and varied concentration of target molecules**. All proteomic studies are currently based on a single type of analytical instrument - the **mass spectrometer** - which sequences the peptides resulting from the protein digestion from the sample being analysed. If we consider a human cell that produces around **30,000** proteins with an average mass of 50 kDa, each protein will generate.

approximately 30 different peptides. In the analysis of the proteome of a human cell, the instrument must therefore **analyze approximately 6,000,000 peptides**. This huge number of peptides cannot be fully sequenced by any mass spectrometer currently available.

which sequences the peptides resulting from the digestion of proteins from the sample being analysed
All proteomic studies are currently based
on a single type of analytical instrument - the mass spectrometer

All proteomic studies are currently based | on a single type of analytical instrument - the mass spectrometer | which sequences the peptides resulting from the digestion of proteins from the sample being analyzed

I. 1. Protein extraction and separation



For this reason, in one form or another, all proteomic studies include steps of:

1. **Extraction of proteins** from the biological matrix and separation of these from non-protein contaminant molecules;
2. **Fractionation of proteins from the sample being analyzed, of the resulting peptides, or both**. Through this, the peptides to be sequenced are divided into fractions and gradually delivered, in smaller quantities, to the analytical instrument so that it can analyze them.

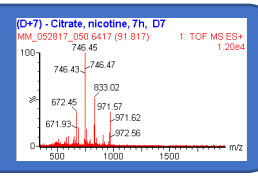
I. 1. Proteins extraction and separation

The purpose of this step is to extract protein molecules from the sample being analyzed and solubilize them for subsequent analysis. The step must be carried out **thoroughly**, as much as possible to ensure that the extract contains **all the proteins of interest from the sample being analyzed**, while minimizing the presence of contaminant molecules. Additionally, it is crucial for this stage to **preserve the naticve post-translational modifications of proteins**. The step typically begins with the operation of lysing the sample and releasing the cellular contents into a **suitable extraction buffer**.

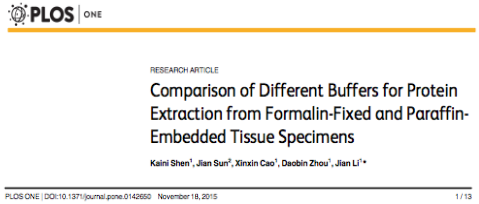
The lysis of the biological sample is performed according to the specific characteristics of the sample being analyzed, respecting a few principles:

- The cells must be dislodged from tissues through mechanical action - homogenization, ultrasonication;
- **Bacterial cells** must be treated with **lysozyme** to disrupt the **peptidoglycan cell wall** and then resuspended in a lysis buffer solution;
- **Yeast** are treated with **lyticase**, which destroys the **chitinous** cell wall, and then resuspended in a lysis buffer solution;
- **Plant** cells are treated with **cellulase**, which destroys the cellulose cell wall, and then resuspended in a lysis buffer solution;
- **Eukaryotic** cells from cell cultures are directly resuspended in a buffer solution that acts as a lysis solution and contains detergents.

I. 1. Protein extraction and separation



Although it is considered that there is no universally applicable extraction buffer solution capable of preserving all proteins soluble, regardless of their origin, various recipes specific to certain cells or tissues are currently available and used.

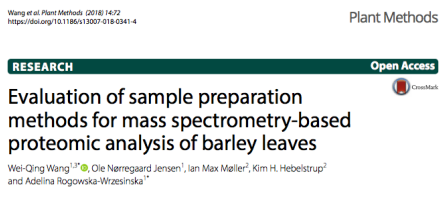

PLOS ONE

RESEARCH ARTICLE

Comparison of Different Buffers for Protein Extraction from Formalin-Fixed and Paraffin-Embedded Tissue Specimens

Kaini Shen[1], Jian Sun[2], Xinxin Cao[1], Daobin Zhou[1], Jian Li[1]*

PLOS ONE | DOI:10.1371/journal.pone.0142660 November 18, 2015 1 / 13

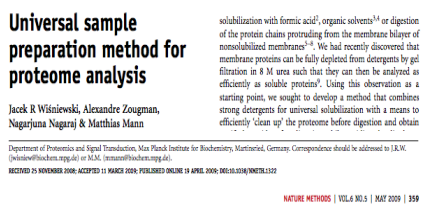

Wang et al. Plant Methods (2018) 14:72
https://doi.org/10.1186/s13007-018-0341-4

Plant Methods

RESEARCH Open Access

Evaluation of sample preparation methods for mass spectrometry-based proteomic analysis of barley leaves

Wei-Qing Wang[1,3*], Ole Nørregaard Jensen[1], Ian Max Møller[2], Kim H. Hebelstrup[2] and Adelina Rogowska-Wrzesinska[1*]


Universal sample preparation method for proteome analysis

Jacek R Wiśniewski, Alexandre Zougman, Nagarjuna Nagaraj & Matthias Mann

solubilization with formic acid[2], organic solvents[3,4] or digestion of the protein chains protruding from the membrane bilayer of nonsolubilized membranes[5–8]. We had recently discovered that membrane proteins can be fully depleted from detergents by gel filtration in 8 M urea such that they can then be analyzed as efficiently as soluble proteins[9]. Using this observation as a starting point, we sought to develop a method that combines strong detergents for universal solubilization with a means to efficiently 'clean up' the proteome before digestion and obtain

Department of Proteomics and Signal Transduction, Max Planck Institute for Biochemistry, Martinsried, Germany. Correspondence should be addressed to J.R.W. (jwisniew@biochem.mpg.de) or M.M. (mmann@biochem.mpg.de).

RECEIVED 25 NOVEMBER 2008; ACCEPTED 11 MARCH 2009; PUBLISHED ONLINE 19 APRIL 2009; DOI:10.1038/NMETH.1322

NATURE METHODS | VOL.6 NO.5 | MAY 2009 | 359


PLOS ONE

RESEARCH ARTICLE

Evaluation of sodium deoxycholate as solubilization buffer for oil palm proteomics analysis

Benjamin Yii Chung Lau*, Abrizah Othman

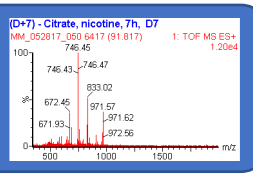
Malaysian Palm Oil Board, No 6, Persiaran Institusi, Bandar Baru Bangi, Kajang, Selangor, Malaysia

PLOS ONE | https://doi.org/10.1371/journal.pone.0221062 August 15, 2019 1 / 23

Generally speaking, any method of protein extraction from a biological matrix for proteomic analysis uses an extraction buffer solution containing:

1. A **strong detergent** such as SDS (Sodium dodecyl sulfate), CHAPS (3-([3-cholamidopropyl]dimethylammonio)-1-propane sulfonate), or Tween (commercial name for non-ionic detergents from the polysorbate class - derivatives of sorbitol esterified with fatty acids) which solubilize membrane proteins, extract them from the lipid bilayer, and separate them from lipid molecules;
2. **Denaturing agents**, usually in high concentrations - most commonly urea or strong acids - that modify the ionic strength and pH of the solution, thereby disrupting interactions between proteins and destabilizing secondary and tertiary structures;
3. **Reducing agents** - most commonly dithiothreitol (DTT), mercaptoethanol, or thiourea - that break disulfide bonds and prevent oxidation of functional groups;
4. Various **enzymes** - (most commonly used DNase, RNase) – for the removal of contaminating macromolecules.

I. 2. Protein fractionation



The purpose of fractionation methods is to separate complex mixtures of proteins into fractions that group molecules with similar properties. All proteins have molecular weight and electric charge, with the values of these two properties being characteristic to each molecule and thus can be used to differentiate one protein from another. Protein fractionation in proteomic studies is most commonly achieved through two types of methods:

A. **Electrophoretic methods** - SDS-PAGE, Isoelectric focusing, 2D-PAGE
B. **Chromatographic methods** - High Performance Liquid Chromatography (HPLC)

A. Electrophoretic methods

Electrophoresis is a separation method **based on the ability of colloidal particles or electrically charged molecules to migrate, under the influence of an electric field generated by two electrodes, towards the electrode carrying the opposite charge**. The migration velocity mainly depends on the size of the electric charge, and for this reason, molecules with different charges will have different electrophoretic mobilities, separating into fractions. In principle, this method can be applied to all electrically charged macromolecules.

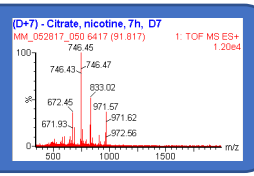
Arne Wilhelm Kaurin Tiselius
1902 - 1971

Depending on the medium in which the separation is carried out, three technical solutions can be distinguished for electrophoretic separation of molecules:

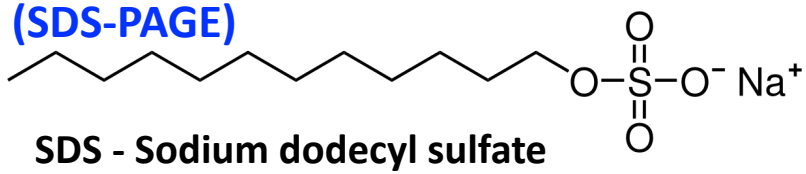
1. **Separation in buffers** - very rarely used, of historical importance. It represents the first method of electrophoretic separation and was developed by Arne Tiselius in 1937.
2. **Separation in stabilizing media** (silica gel plates, films, gels). Depending on the physical location of the stabilizing medium, we can distinguish:

2.A. Horizontal electrophoresis - for supports inert to oxygen (paper, silica gel plates, starch, agarose).
2.B. Vertical electrophoresis - supports that need to be shielded from oxygen, such as polyacrylamide. It represents the standard type of electrophoresis for protein separation in proteomic studies.

I. 2. Protein fractionation



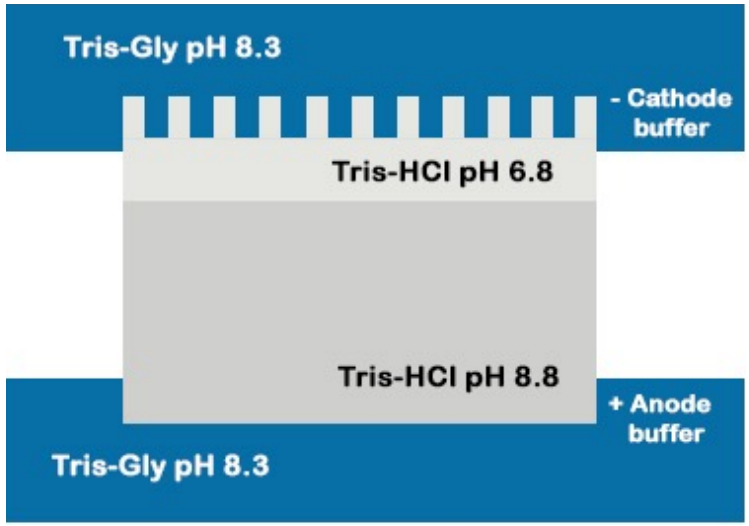
A.I. Sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS-PAGE)



SDS - Sodium dodecyl sulfate

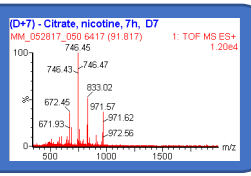
The name of this method refers to the fact that protein molecules are not separated in their native form, but are **unfolded under the action of temperature, SDS detergent, and reducing agent such as β-mercaptoethanol.** The three-dimensional folding of protein molecules is removed, proteins being transformed into linear chains of amino acids. Additionally, **SDS binds to the polypeptide chains in an amount proportional to their length and screens the charges due to the amino acid residues**. Therefore, all proteins will be negatively charged, the magnitude of the charge being directly proportional to the amount of SDS bound, i.e., to the length of the chain. In this type of electrophoresis, separation is based on a single parameter, namely **the length of the polypeptide chains, i.e., the molecular weight.**

Proteins are separated on polyacrylamide gels (PAGE - polyacrylamide gel electrophoresis) using a discontinuous buffer system. Discontinuity refers to four parameters:



1. **Gel porosity** - the gel used for electrophoretic separation contains two regions with different concentrations, one in the upper region with **low concentration (large pores, stacking gel**), and one in the lower region with high concentration (small pores, separating gel).
2. **Gel pH** - the **stacking gel has a pH of 6.8**, while the **separating gel has a pH of 8.8.**
3. **Ionic strength** of the buffer solutions within the gel - the stacking gel contains 0.125 mol/L TRIS-HCl, while the separating gel contains 0.375 mol/L TRIS-HCl.
4. **Ions in the gel and in the electrode buffer**; the gel contains Cl^-, while the migration buffer contains glycine.

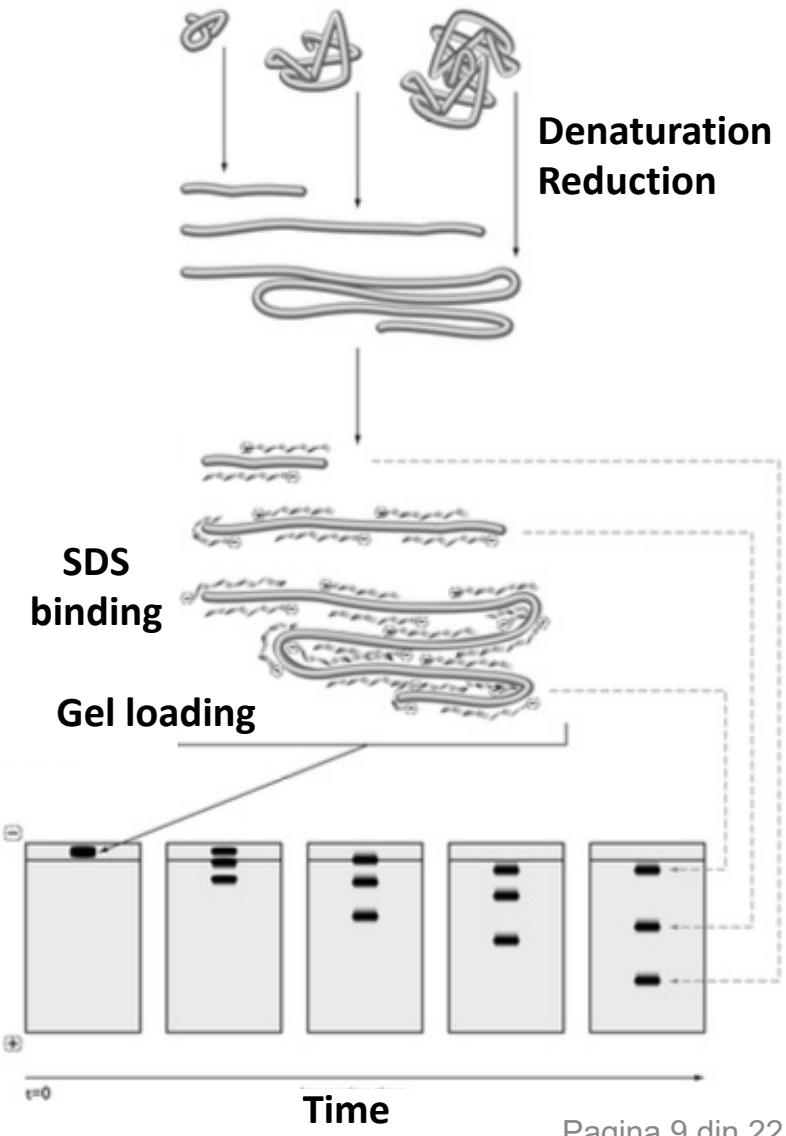
I. 2. Protein fractionation



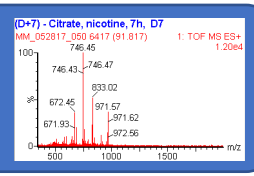
The separation of proteins in the discontinuous system occurs in two steps, corresponding to the two gels:

1. **Protein stacking**: The stacking effect happens because the proteins encounter less resistance as they move from the sample wells into the stacking gel due to the lower pH. As they move further into the stacking gel, the pH becomes more basic, causing the proteins to lose their positive charges. This process allows the proteins to concentrate into a tight band. Once the proteins reach the interface between the stacking gel and the separating gel, the pH abruptly changes to a higher value (usually around pH 8.8) in the separating gel. This change in pH causes the proteins to regain their negative charges, but now they are concentrated into a narrow zone.

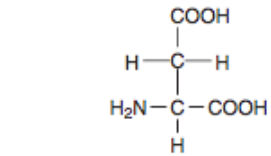
2. **Separation proteins into fractions:** In the separating gel, the proteins are separated based on their size and charge. Because they are now concentrated into a narrow band, the proteins can migrate more uniformly and efficiently through the gel, resulting in better resolution and clearer bands.



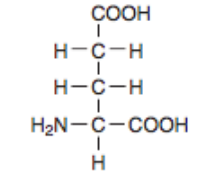
I. 2. Protein fractionation



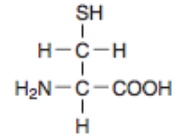
A.II. Isoelectric focusing



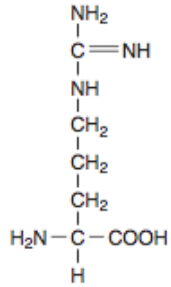
Aspartic acid (ASP/D)
$pK_a = 3.86$



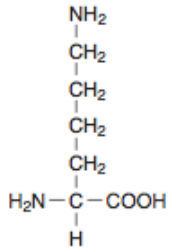
Glutamic acid (Glu/G)
$pK_a = 4.07$



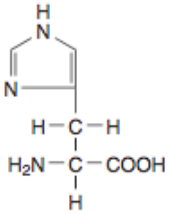
Cysteine (Cys/C)
$pK_a = 8.00$



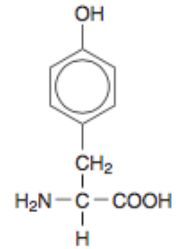
Arginine (Arg/R)
$pK_a = 12.48$



Lysine (Lys/K)
$pK_a = 10.53$



Histidine (His/H)
$pK_a = 6.10$



Tyrosine (Tyr/Y)
$pK_a = 10.07$

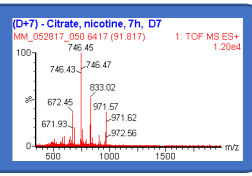
A significant portion of proteinogenic amino acids contain functional groups that, under certain conditions, can ionize, thus generating electric charges (COOH into COO^-, $-NH_3$ into $-NH_4^+$). Additionally, post-translational modifications can introduce ionizable functional groups onto the surface of protein molecules (PO_4^-) and thus generate charges. The ionization of these groups is influenced by the pH of the solution. **At low pH values**, most functional groups capable of accepting H^+ will do so and become **neutral or positively charged**. At **high pH values**, most functional groups capable of donating H^+ will do so and become **neutral or negatively charged**. Each functional group is thus characterized in this manner Through a specific pH value where ionization occurs, and this value is the **pK_A**. The ionizable functional groups within the structure of amino acids influence each other, resulting in **each amino acid being characterized by a unique pH value at which none of the functional groups are electrically charged, rendering the amino acid uncharged – the isoelectric point, pI.**

The same is true for proteins. In the case of proteins, the functional groups of amino acids on the surface of protein molecules can ionize and thus generate charges. The interaction between the ionizable functional groups on the surface of the protein molecule causes each protein to be characterized by a unique pI. Depending on its value, proteins can be divided into:

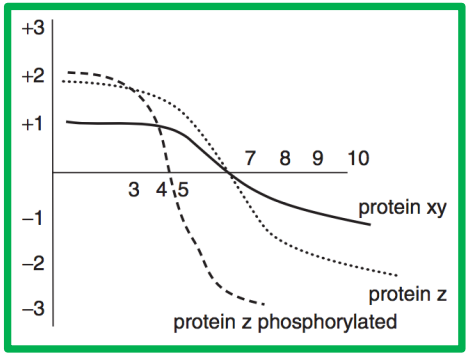
acidic proteins, pI < 7,

basic proteins, pI > 7.

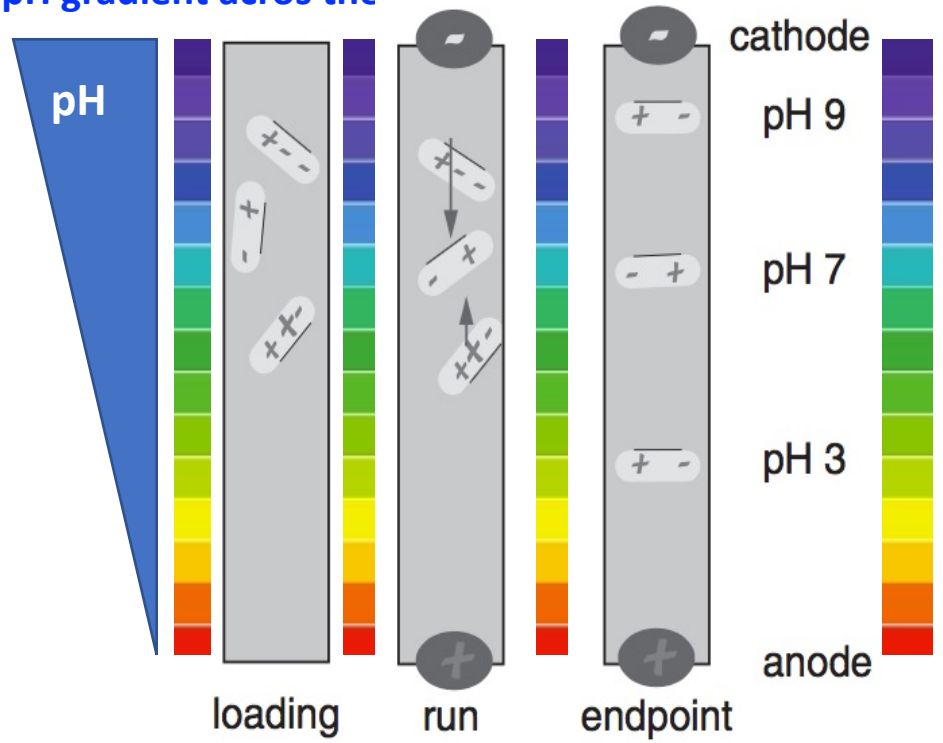
I. 2. Protein fractionation



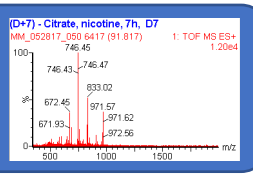
Regardless of their type, **proteins will be positively charged at pH values lower than their pI**, **negatively charged at pH values higher than their pI**, and **neutral at their pI**. This variability in protein charge depending on the pH underlines the principle of a fractionation method called **isoelectric focusing**.



Isoelectric focusing is a method for fractionating and separating proteins based on their pI values. Specifically, **proteins are placed into gels that have a pH gradient acros their full length.** By applying an electric current, the **proteins move towards one of the electrodes based on their charge.** As they migrate towards the electrodes, proteins pass through gel regions with different pH values, and their charge changes accordingly. **When a protein molecule reaches the gel region with a pH equal to its pI, the protein's charge becomes neutral, and it stops moving**. Regardless of the duration of the electric current application, a protein molecule reaching its pI will no longer move within the gel. **Through prolonged application of the electric current, all protein molecules with the same pI, regardless of sequence and initial location in the gel at the beginning of separation, will migrate and concentrate in the same zone** – hence the term isoelectric focusing.



I. 2. Protein fractionation



A.III. 2D-PAGE

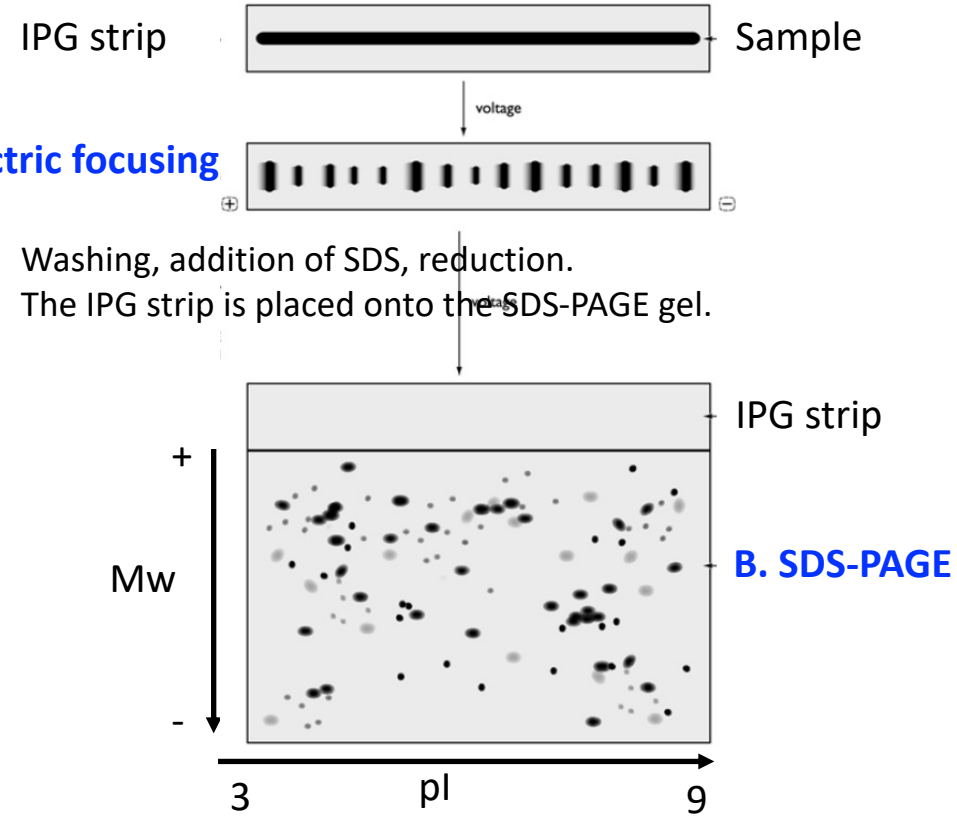
2D-PAGE, or two-dimensional polyacrylamide gel electrophoresis, is a method for separating proteins in polyacrylamide gels that combines the principles described earlier to separate protein mixtures based on two parameters or dimensions (2D):

1. **Isoelectric point** - through an initial step of isoelectric focusing.
2. **Molecular weight** - through an SDS-PAGE step.

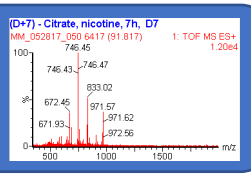
Although the principles and the method itself were introduced as early as the 1970s, it wasn't extensively utilized due to technical difficulties encountered in:

A. Achieving a reproducible pH gradient and
B. Attaching the gel for isoelectric focusing onto the SDS-PAGE gel and transferring proteins from one gel to another.

Currently, by utilizing gel strips in the isoelectric focusing step where compounds generating the **pH gradient are immobilized on an acrylamide gel** (**IPG strips - immobilized pH Gradient**), 2D-PAGE has become the gold standard method for proteomic studies. It allows the separation of protein molecules with the same sequence but different post-translational modifications (e.g., various levels of phosphorylation).



I. 2. Protein fractionation



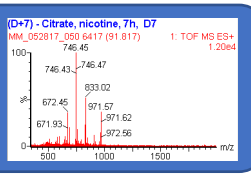
A.IV. Detecting proteins in gels

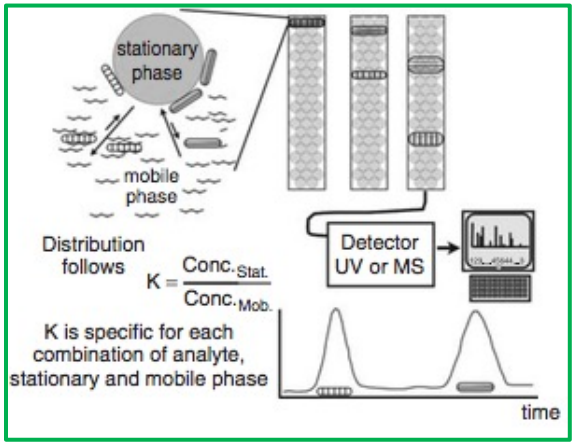
For the visualization of proteins separated on polyacrylamide gels, numerous protocols have been described, which can be divided into two main categories:

1. **Specific staining** methods that highlight only certain fractions of all separated proteins; these stains are used especially for the visualization of enzymes.
2. **Non-specific staining methods** that achieve uniform staining of all separated proteins. The various non-specific staining methods differ in their sensitivity level and complexity. Regardless of the staining method used, they are all based on the reaction of a color reagent with specific amino acid residues within the polypeptide chain (e.g., aromatics). For this reason, the color intensity depends not only on the quantity of protein in the gel but also on the amino acid sequence of the stained protein.

Dye method	Sensibility	Observations
Coomassie brilliant blue R-250	0,3-1 µg/fraction	Simple, traditional, with reduced sensitivity. Recommended for SDS-PAGE gels and 2D-PAGE.
Coomassie brilliant blue G-250 (coloidal)	100 ng/fraction	Recommended for all types of gels, good sensitivity but requires a longer staining time.
Silver staining	10 ng/fraction	Very good sensitivity, rather difficult to perform, requires exceptionally high-quality reagents.
Copper staining	10-100 ng/fraction	Fast and easy to perform, can only be applied to SDS gels.
Zinc staining	10-100 ng/fraction	Quick and easy to perform, can only be applied to SDS gels.
SYPRO	1-8 ng/fraction	Fast and easy to perform, applicable to all types of gels, requires a UV imaging system.

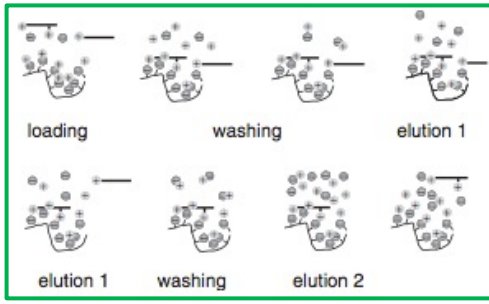
I. 2. Protein fractionation





B.I. Chromatographic methods for protein fractionation

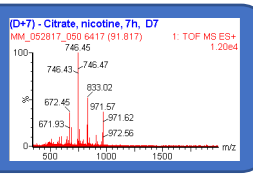
Chromatography is a physical method of separating the components of a mixture based on the different interactions they have with two phases: a mobile phase and a stationary phase. Depending on the interactions on which it is based, the chromatographic methods applicable to proteins are:



1. Ion exchange chromatography (IEX) is based on the presence of electric charges on the surface of protein molecules. The stationary phase is also electrically charged (positive or negative), so proteins will have a greater or lesser affinity depending on this charge. The name of this type of separation comes from the fact that protein elution from the column is achieved through increased concentrations of ionized salts, with the stationary phase exchanging the proteins bound to it with the salt ions.

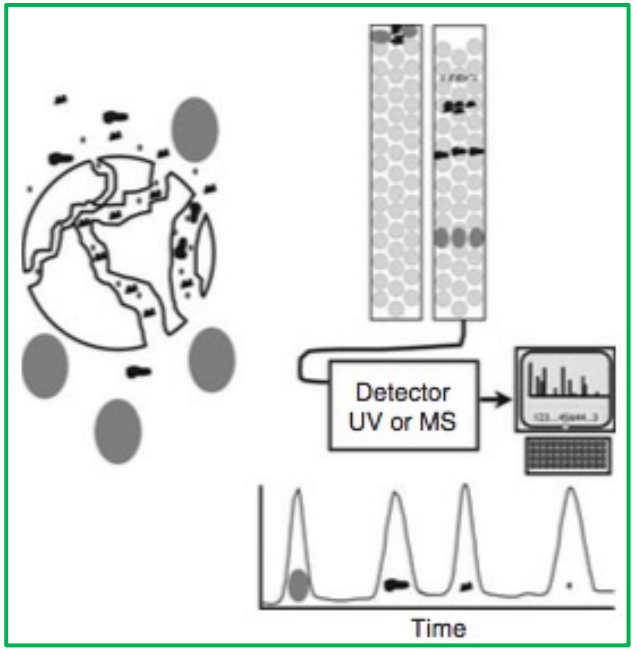
2. Hydrophobic Interaction Chromatography , (HIC) Hydrophobic Interaction Chromatography (HIC) is a chromatographic technique used for the separation of molecules, including proteins, based on their hydrophobic interactions. In HIC, separation occurs at high salt concentrations, which strongly retain water in the hydration sphere. Under these conditions, protein molecules are deprived of hydration water and interact with a mildly hydrophobic stationary phase. The strength of the interaction depends on the size of the hydrophobic patches on the surface of the protein molecules. As the salt concentration decreases, hydration water becomes available, and the protein molecules interact with it, causing them to elute from the stationary phase. HIC is commonly used in protein purification processes due to its ability to separate proteins based on their hydrophobicity.

I. 2. Fracționarea proteinelor



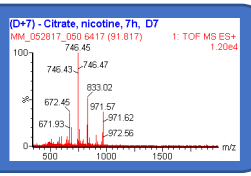
3. Hydrophilic interaction chromatography, HILIC is a chromatographic technique based on interactions with a strongly hydrophilic stationary phase. However, the applicability of this method in protein separation is limited to membrane proteins and apolipoproteins.

4. **Size-exclusion chromatography (SEC),** also known as gel gel filtration, is based on the principle of separation by exclusion according to size. Depending on their mass, molecules of different sizes are differentially retained as they pass through a gel with fixed-size pores. Larger molecules, which cannot penetrate the pores, will have a lower degree of retention, while smaller molecules will penetrate the gel pores and be "lost," resulting in significant delay.



5. **Reversed-phase liquid chromatography** (RPLC) is traditionally used for the separation of peptides, but it has recently found applications in protein separation as well. For pedagogical reasons, the principles and additional details will be discussed in the corresponding chapter on post-digestion peptide separation.

I. 2. Fracționarea proteinelor

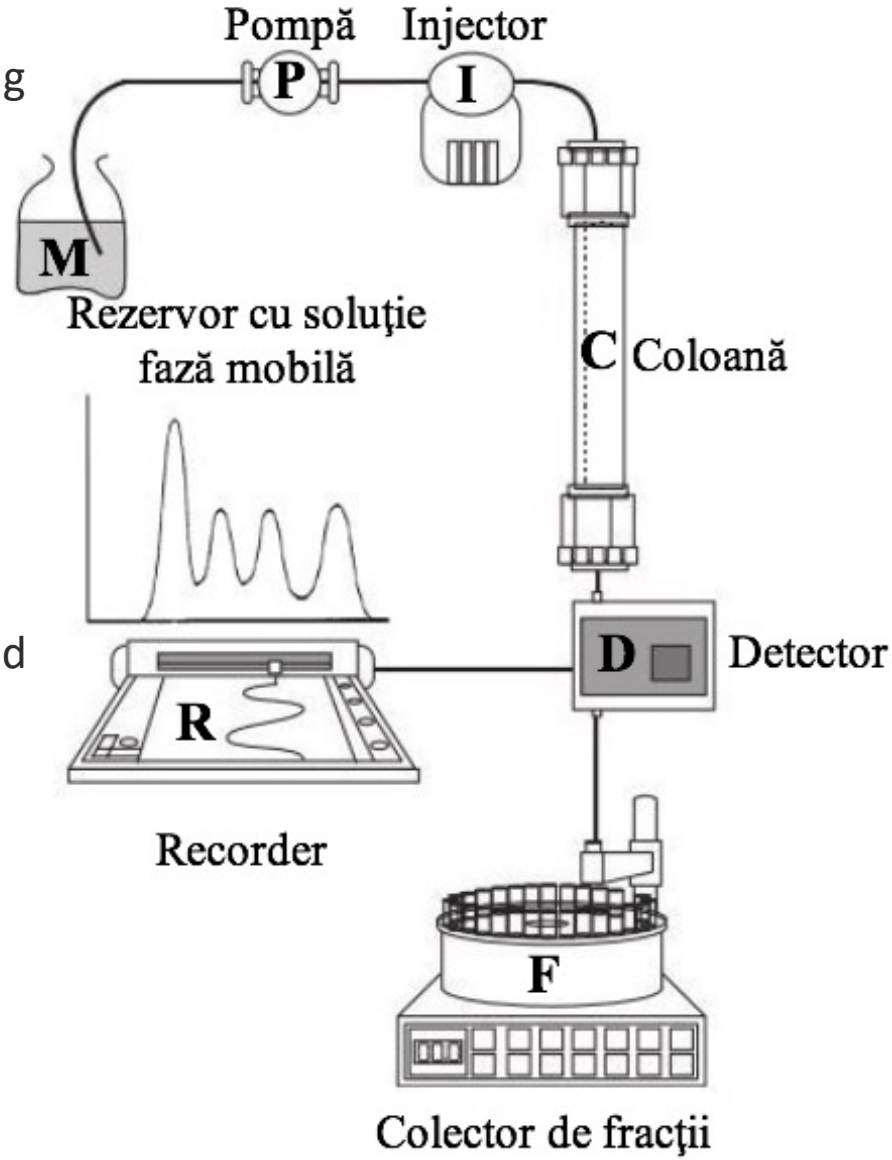


Fractionation of proteins using chromatographic techniques is carried out in **chromatographic systems** typically consisting of:

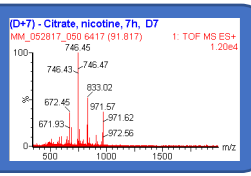
1 **reservoir containing the mobile phase solution** (M) - The minimum capacity of the reservoir is determined by the dimensions of the chromatographic column, as it needs to contain enough mobile phase to allow for column equilibration and actual chromatographic separation.

2 **pump (P)** - It is responsible for ensuring a constant flow rate during the separation process. For gravity flow (GF), a simple peristaltic pump is sufficient, as the pressures reached in the column are low. More complex applications may require a piston pump or even a system consisting of two or four piston pumps.

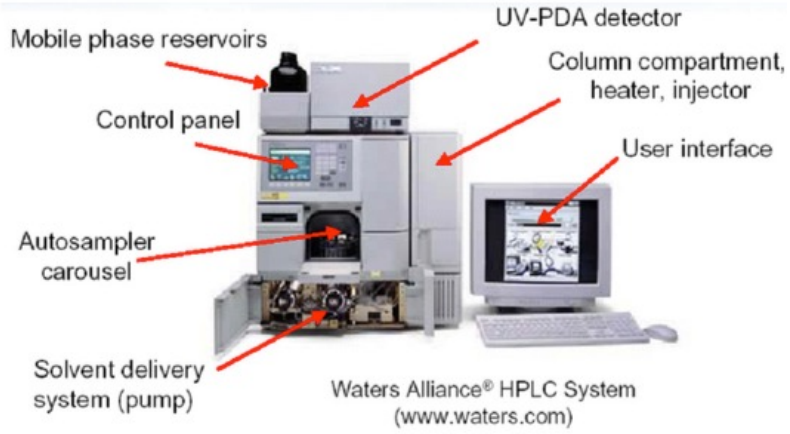
3 **injector (I)** - It allows for the injection of the sample into the column at a user-defined time. In modern instruments, the injector is often replaced by an autosampler that automatically loads the samples for analysis.



I. 2. Fracționarea proteinelor



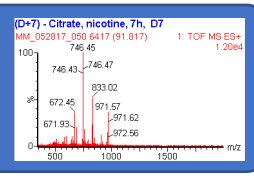
4. Chromatographic column (C) - This is the element where the actual separation takes place. Chromatographic columns contain various types of stationary phase, depending on the type of separation (SEC; HILIC; RPLC, HIC, IEX).



5. Detector (D) - It allows for the real-time monitoring of protein concentration in the mobile phase as it exits the chromatographic column and detects the moment when the investigated protein is eluted.

6. Fraction collector (F) - It allows for the automatic fractionated collection of eluent into separate tubes from the column. It is not mandatory, as chromatographic systems can be directly connected to a mass spectrometer.

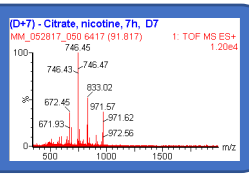
7. Recorder (R) - Its role is to record and present graphically the values detected by the detector. The information is presented in the form of a graph called a chromatogram, with time (or volume) represented on the abscissa (X-axis) and the mobile phase absorbance on the ordinate (Y-axis). In modern chromatography systems, the recording system is replaced by control software for the entire chromatographic system. This software allows not only monitoring but also real-time adjustment of separation parameters.

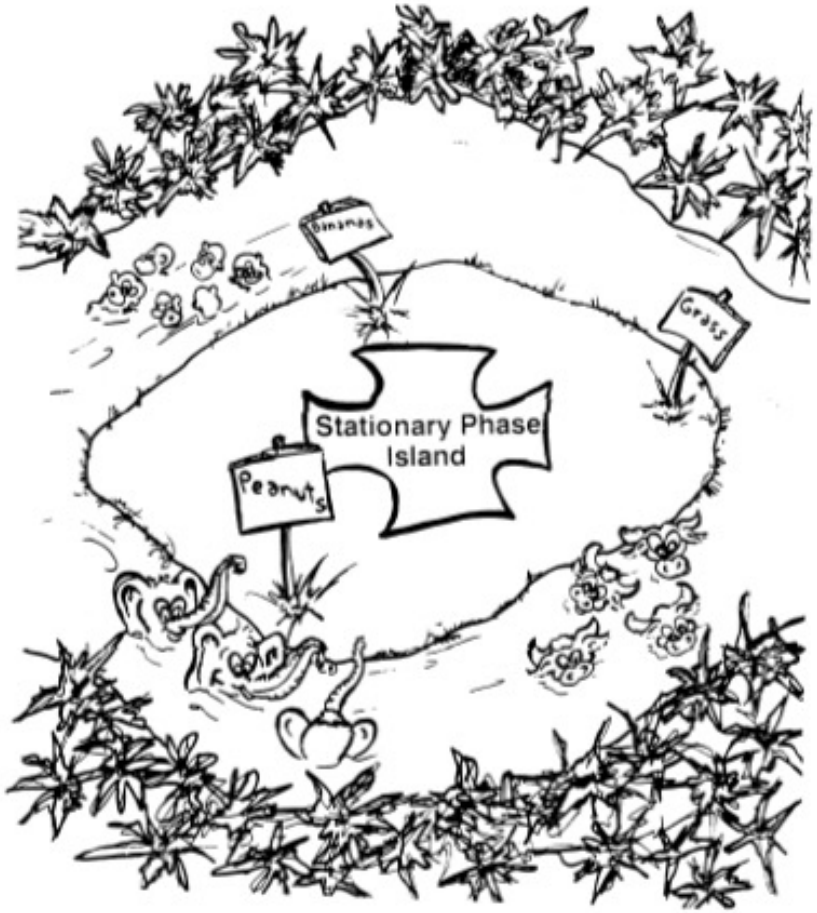


A Simple Analogy

A zoo train derailed over a river. Monkeys, elephants and cows were free!

SIGMA-ALDRICH®

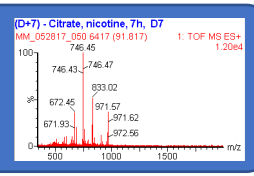


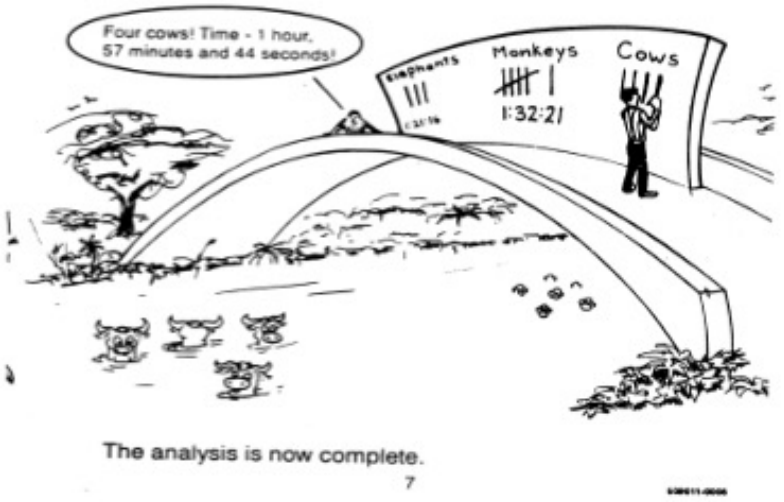


A Simple Analogy

The animals drifted by an island that had bananas, grass and peanuts. Each animal stopped at their favorite food. They each took their time eating before setting off.

SIGMA-ALDRICH®

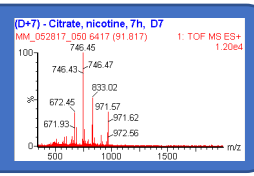


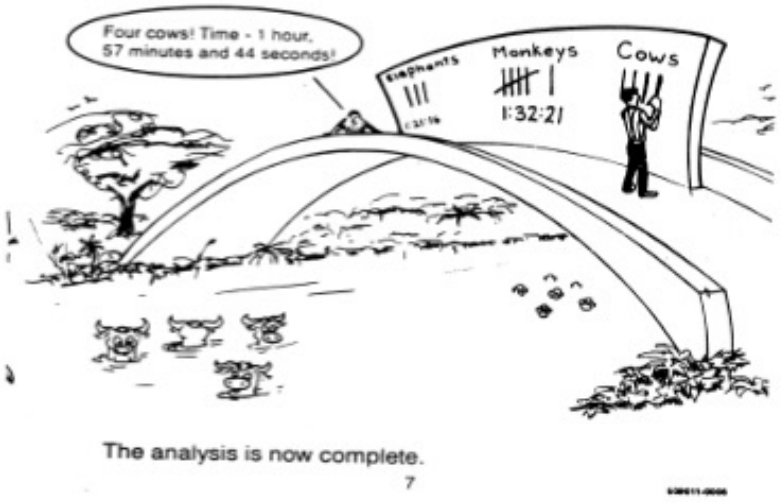


A Simple Analogy

Eventually, the animals floated under a bridge. As the animals went under the bridge, the "detector" called out the time and the number of each animal.

SIGMA-ALDRICH®

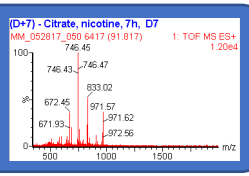




A Simple Analogy

Eventually, the animals floated under a bridge. As the animals went under the bridge, the "detector" called out the time and the number of each animal.

SIGMA-ALDRICH®



A Simple Analogy

The caretakers graphed the detector's data. They plotted the number of each animal on the y-axis and the time it passed under the bridge on the x-axis.

SIGMA-ALDRICH®