Methods of proteomic research. Proteomics techniques (Part 2)

Candidate of bio. Sci. **Meldebekova Aliya A.**

1. Protein Fragmentation and MS detection techniques:

- **Collision-induced dissociation (CID)** – (Диссоциация, вызванная столкновением (CID), или активируемая столкновением диссоциация)
- **Electron-transfer dissociation (ETD)** - Диссоциация с переносом электрона (ETD) - это метод фрагментации многозарядных газовых макромолекул.

2. Protein quantitation:

- **Relative quantitation** or **absolute quantitation**
- **label free quantification (LFQ) and**
- **label based quantification (or label-dependent quantitation) techniques**

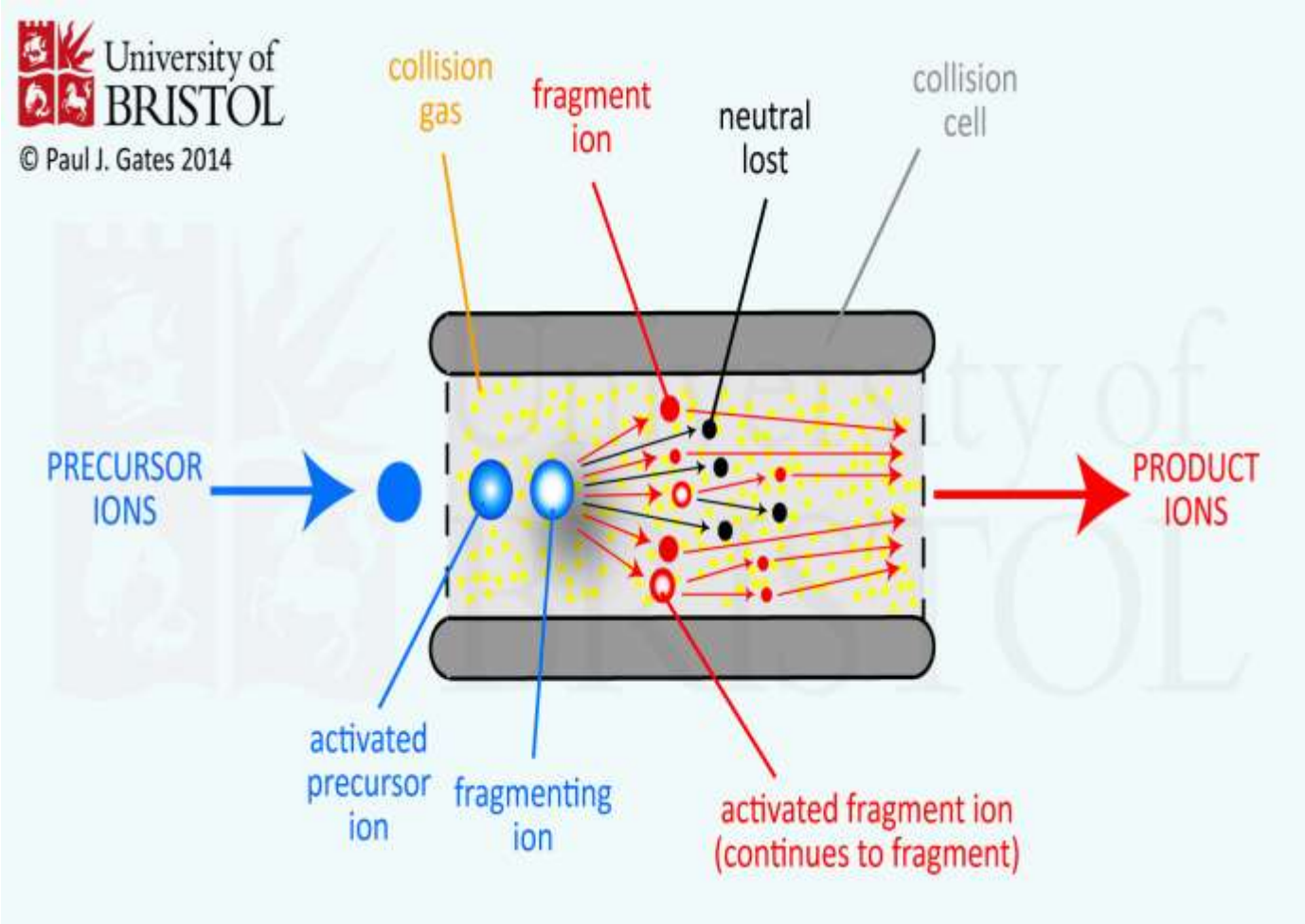
Fragmentation

- Usually measuring the mass of an analyte is not enough to conclusively identify it
- By fragmenting an analyte and measuring the masses of the fragments we can obtain further information to identify the analyte
- There are many types of fragmentation but collision-induced dissociation (CID) is the most common
  - Fastest and most generally successful for the widest variety of proteins and peptides

Fragmentation and MS detection techniques

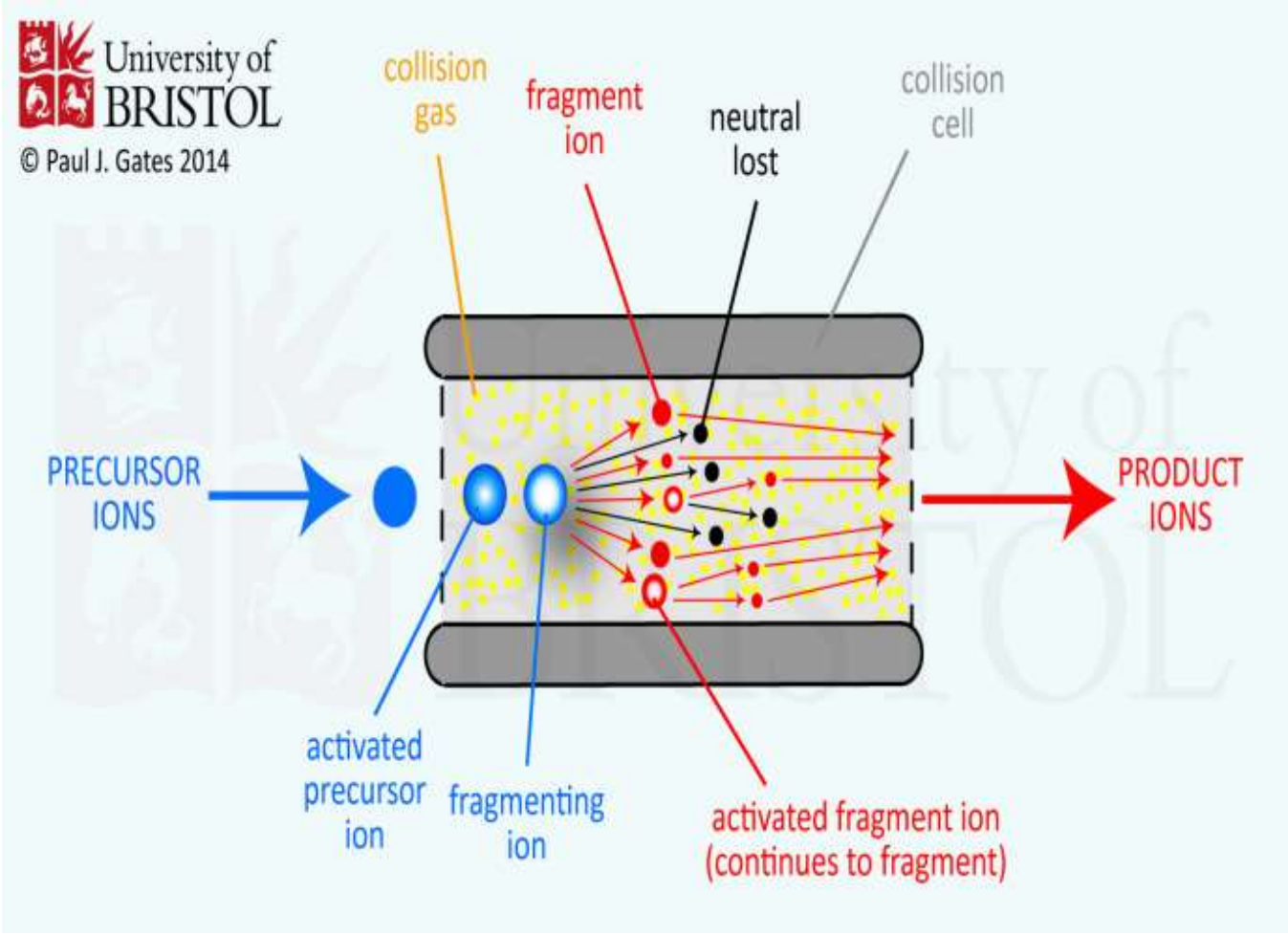
- MS instruments continue to show versatility with various peptide fragmentation and detection configurations.
- **Collision-induced dissociation (CID)** - Диссоциация, вызванная столкновением, или же диссоциация, активируемая столкновением
- Was first introduced in 1981 and is one the most fundamental fragmentation techniques in proteomics .
- The ionized peptides are passed through a vacuum chamber where they collide with a neutral gas such as nitrogen, helium or argon.
- The vibrational energy cleaves the C–N (peptide bonds) to generate **b** and **y** ions series, followed by mass analyzer detection.
- Higher-energy collision dissociation (HCD) which is essentially a vendor-specific term for CID is often used in Orbitraps and hybrid LTQ-Orbitrap mass spectrometers which combine the cycle speed and sensitivity of the linear ion trap with the mass accuracy and resolution of the Orbitrap.
- Precursor ions are shuttled from the C-trap to the collision cell where the ions are similarly fragmented by a neutral gas. The fragments are then transmitted to the Orbitrap analyzer.

- In CID, a strong electric field accelerates ions into a neutral gas (typically He, N2, or A), and each ion collides with gas molecules several times.
- The high kinetic energy of the fast-moving ion is converted to internal energy, thereby increasing the Boltzmann temperature and eventually **breaking the weakest covalent chemical bonds**.
- The resulting fragment ions are extracted from the collision chamber and injected into a mass analyzer.
- Based on the masses (and, with high mass resolution, the elemental compositions) of the various fragment ions, it is possible to identify functional groups and linkages (e.g., for peptides and proteins, the amino acid sequence). A fragmentation pattern, like a fingerprint, can sometimes be used to identify the molecule of interest.

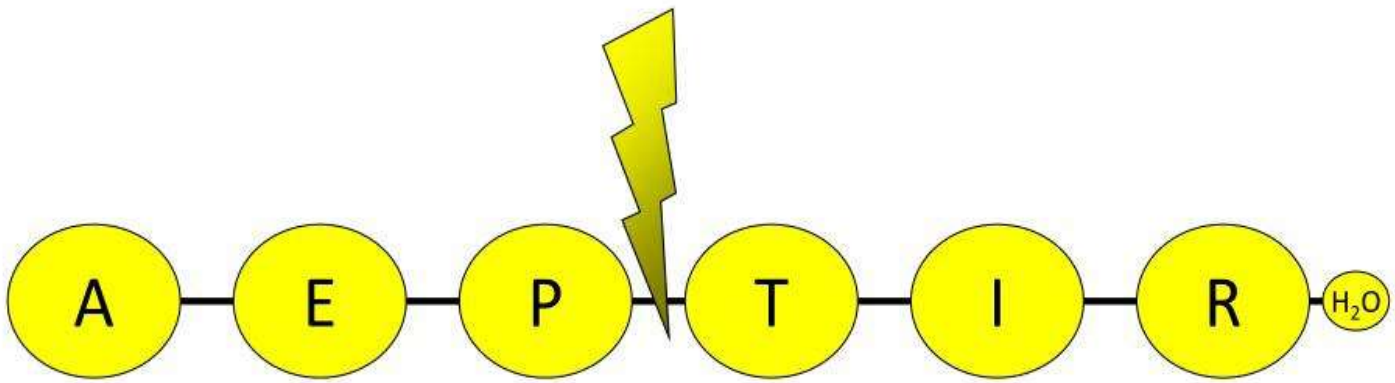


- The figure shows a cartoon schematic of the processes involved in CID. The precursor ion enters the collision cell (or in the case of ion-traps or FT-MS, the precursor ion is isolated in the trap) containing a high pressure of an energised, chemically inert collision gas (e.g. Ar, He, N2, CO2 etc).

The precursor ion undergoes repeated collisions with the collision gas, building up potential energy in the molecule, until eventually the fragmentation threshold is reached and the precursor ion fragments into product ions.

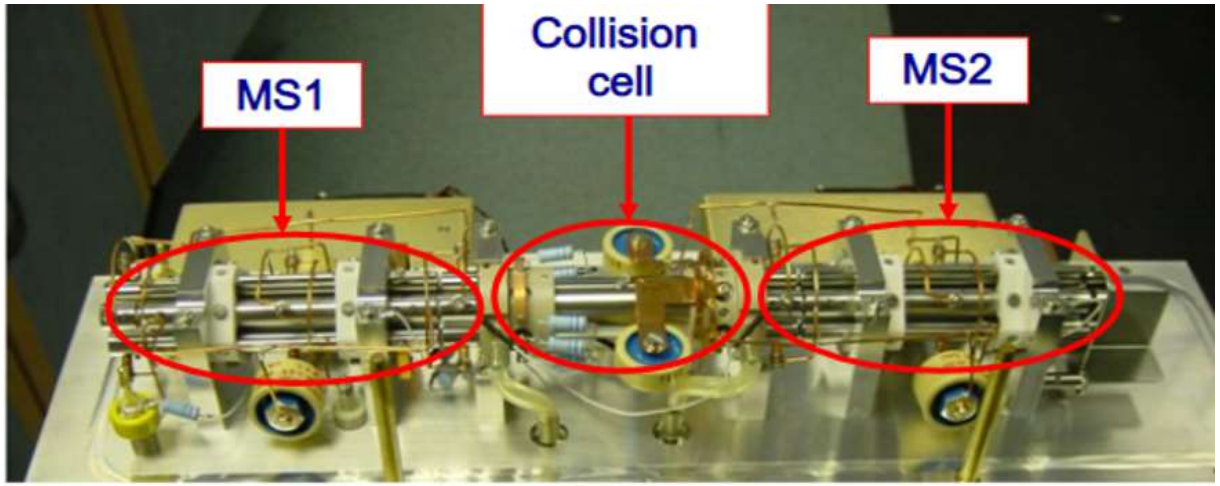


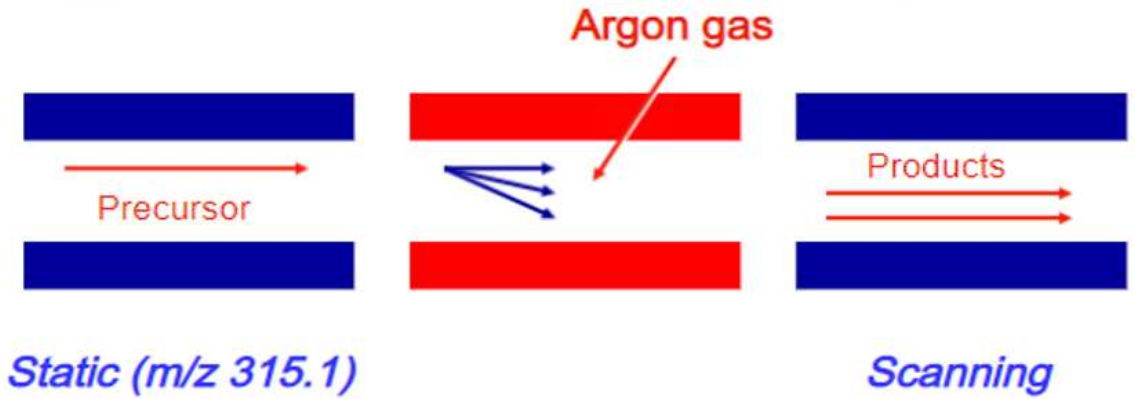
Collision Induced Dissociation



Fragment (somewhat) randomly along the peptide backbone

Product ion scanning

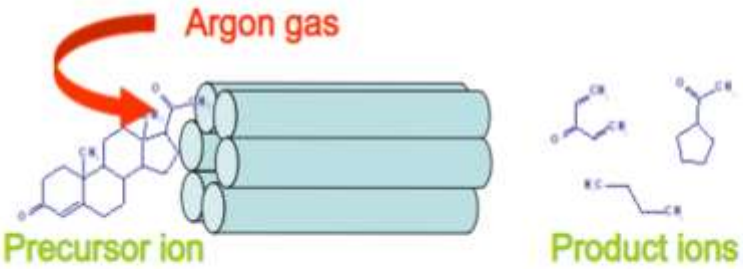




Static (m/z 315.1) *Scanning*

The first quadrupole mass analyzer is fixed at the mass-to-charge ratio (*m/z*) of the precursor ion to be interrogated while the second quadrupole is ***Scanning*** over a user-defined mass range.

Collision induced dissociation



- In the collision cell, the **TRANSLATIONAL ENERGY** of the ions is converted to **INTERNAL ENERGY**.
  - Collision conditions **(FRAGMENTATION)** is controlled by altering:
    - –The collision energy (speed of the ions as they enter the cell)
    - –Number of collisions undertaken (collision gas pressure)

Electron-transfer dissociation (ETD)

It is an ion-ion collision fragmentation-based method where cations (peptides or proteins) collide with charged radical reagent anions.
ETD is particularly useful in the study of diverse modifications because PTM integrity is preserved while still achieving the backbone fragmentation necessary for peptide detection.
Hybrid ETD fragmentation methods like ETD-HCD have been reported where precursors sequentially undergo both types of fragmentation to yield b/y and c/z type fragment ions in a single spectrum.
Additionally, ETD is suitable for the study of top-down proteomics due their high-cationic nature including intact proteins, their PTMs, and protein-protein interactions.
One such application is the study of engineered antibodies which have shown potential to be used as cancer therapies.
ETD-based MS methods have been reported in the literature to characterize monoclonal antibodies (mAb) and antibody drug conjugates (ADC).

- Electron transfer dissociation is a popular peptide fragmentation technique in mass spectrometry. It requires multiply-charged gas-phase cations (z>2) and therefore it is typically limited to an electron spray ion source.

The sample, usually in solution, is first ionized by electrospray, and gas-phase "precursor" ions are then captured in an ion trap.
ETD fragmentation proceeds in two steps.
First, the precursor cation, usually a protein or peptide carrying multiple positive charges, reacts with a pre-ionized anion (e.g., fluoranthene).
Capture of an electron from the anion yields an unstable cation radical (Figure 1).

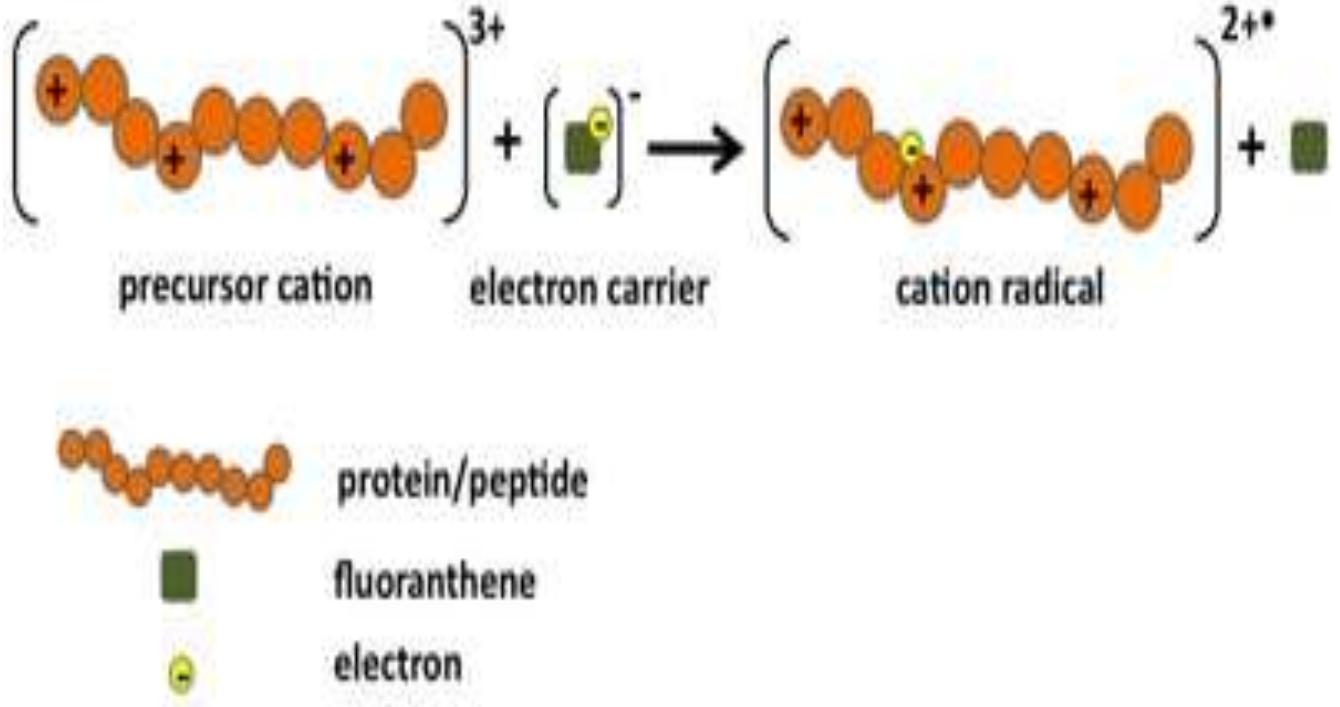


Figure 1. Protein/peptide precursor multiply-charged cation reacts with an anion to form a radical cation

Second, the cation radical then breaks into two fragments, usually one **c** ion and one **z** ion (Figure 2).

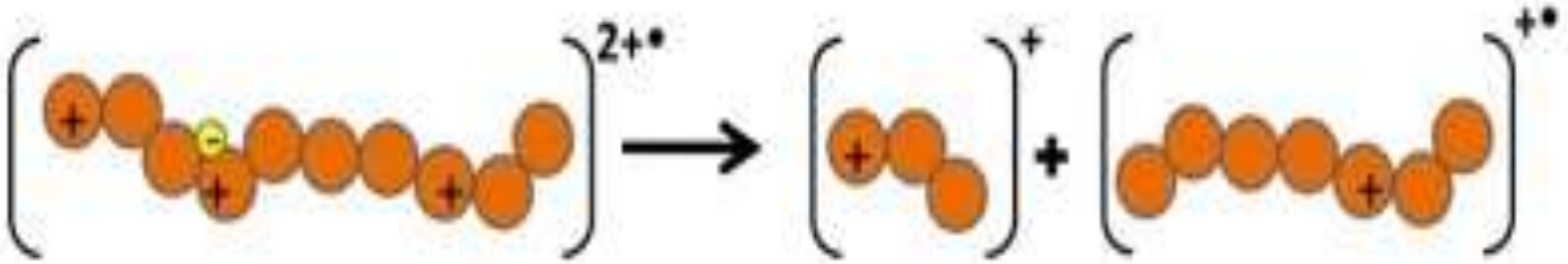


Figure 2. Protein/peptide radical cation fragmentation into a **c** ion and a **z** ion

The cleavage is between **α-carbon** and **nitrogen** as shown in Figure 3.

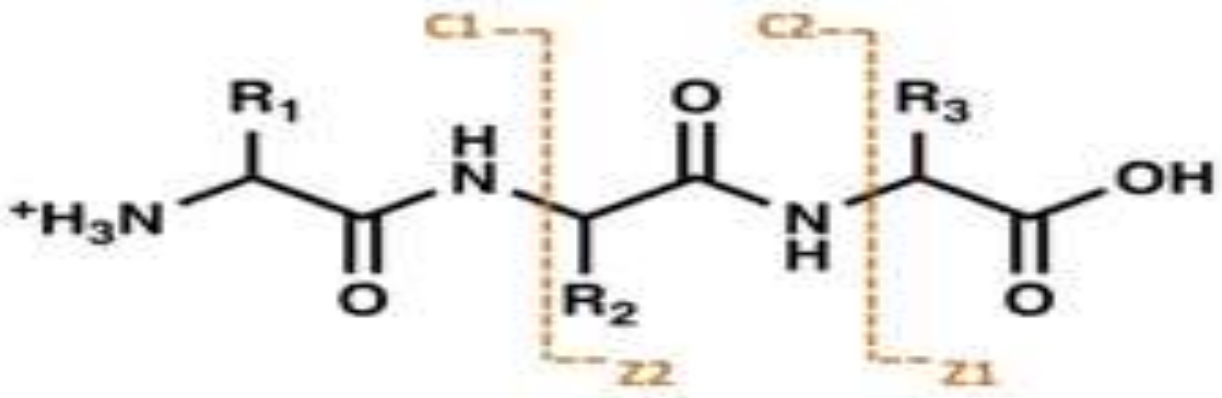


Figure 3. Bond cleavage in ETD and chemical structures of the **c** and **z** ions

The cleavages occur randomly on the protein backbone between any two amino acids (except proline), whereas the side chains can read out a protein amino acid sequence without losing any side chain or chemical modification information

2. Protein quantitation:

There is a large number of global and targeted protein quantitation approaches available, each with their unique sets of advantages and disadvantages.

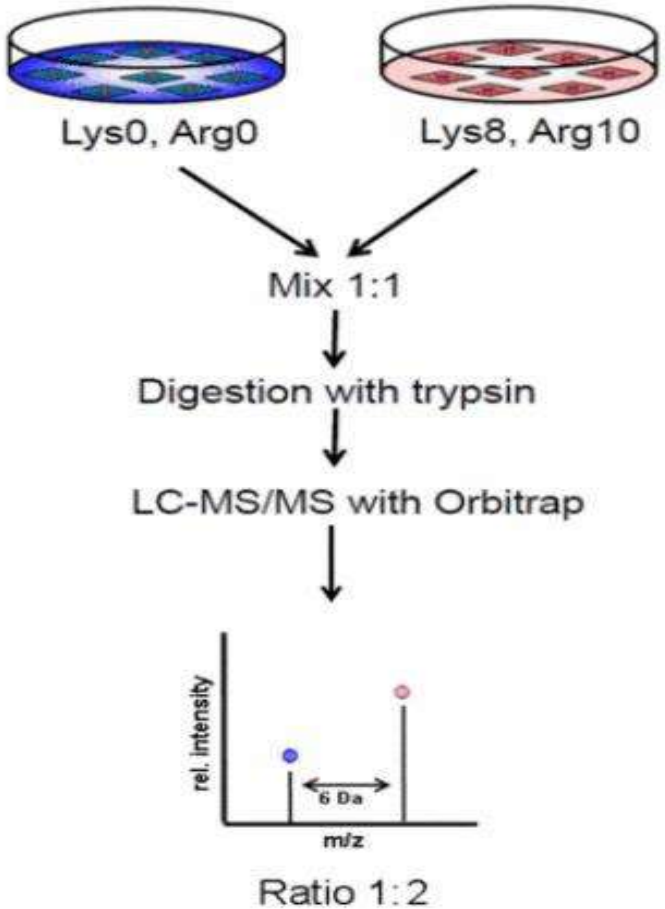
These can be categorized as **relative** or **absolute quantitation**, of which the former can be further divided into **label-dependent** and **label-free techniques (label free quantification - LFQ)**.

Relative quantitation

- **Relative quantitation** often depends on the use of **stable-isotopic labels** that result in covalently derivatized peptides.
- Past methods include metabolic and chemical labelling, such as **dimethylation** and **Stable Isotope Labeling with Amino Acids in Cell Culture (SILAC)** are most applicable to cell line-based studies.
- Additionally, SILAC-based methods have also been adopted to in vivo studies, termed Super-SILAC, where heavy labelled cell line lysates are spiked into tumor tissue lysates as a global control.
- This approach was used for histone PTM quantitation in breast cancer FFPE samples.

In general, relative quantitation strategies are used for **discovery-based** clinical proteomics.
Once proteins of interest have been identified they require further validation.
While antibody-based techniques such as ELISA fulfill this role, MS-based targeted assays are well suited for validation, **especially if no suitable antibodies are available**.

Stable Isotope Labeling of Amino Acids in Culture (SILAC)



- Grow cells in media supplemented with stable isotope-labeled amino acids
- Combine samples at the level of cells and process as one sample
  - Minimize variability between samples for lysis and digestion
- Different samples separated by mass in the MS

Stable Isotope Labeling with Amino Acids in Cell Culture (SILAC)

- The method of working is differentiated into two critical phases**:**

- an adaptation phase and

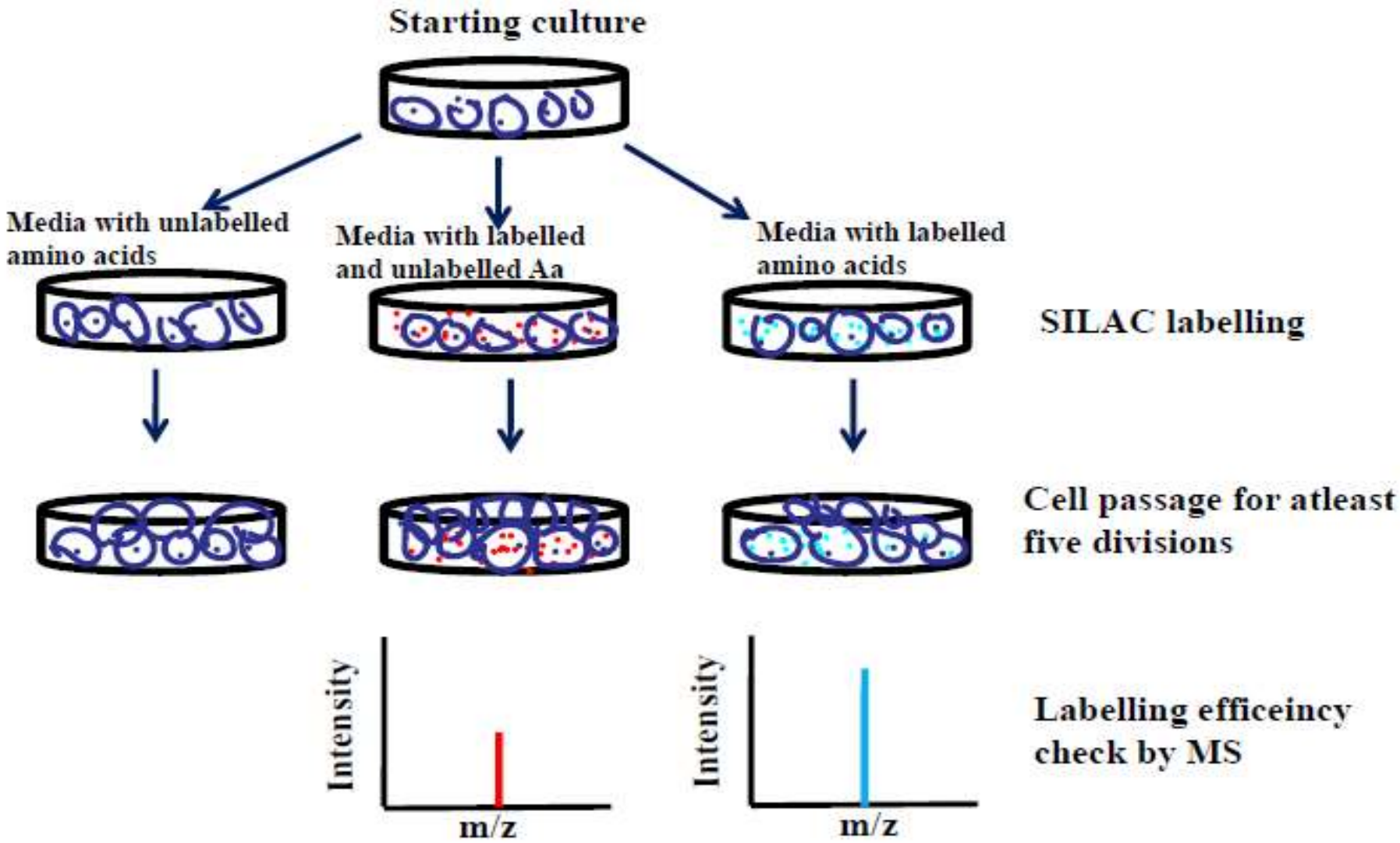
- an experimental phase.

Before initiating the adaptation phase, it is crucial to characterize the cell type to be used for the labeling. In general, dialyzed serum is used to rear the cells in order to negate the availability of free amino acids present in the normal serum. Although, some cell lines do not grow that well in a dialyzed medium due to the absence of some of the growth factors, therefore, supplementation of purified growth factors with the dialyzed media or small percentage of normal serum in the dialyzed media may be helpful.

Adaptation phase

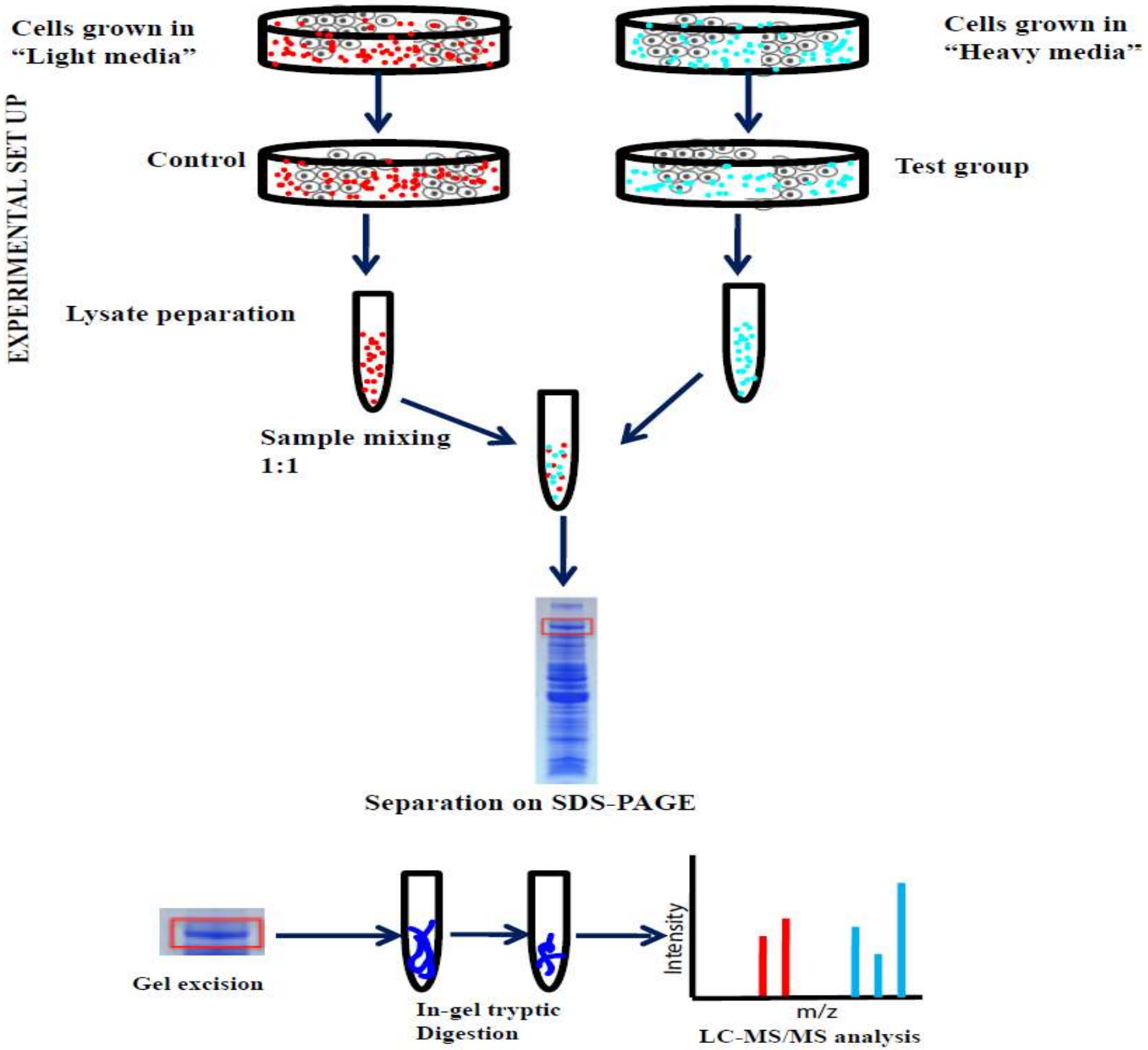
- The selected **cells are acclimatized** in the unlabeled and labelled media until all the cellular **proteins** in the labelled group **are maximally incorporated with the heavy amino acids** (95% or more). The cells are then allowed to divide for at least five cell divisions.
- A small proportion of the heavy labelled cell population is harvested and the cells are digested into small peptides using trypsin followed by evaluation of incorporation of heavy amino acids by LC-MS/MS.
- If the efficiency is not high enough, the cells are allowed to grow further and rechecked for efficiency evaluation.

A. Adaptation phase



Experimental phase

- Once the incorporation of heavy amino acids is ensured, the two cell populations can be subjected to tests according to the experimental plans.
- The cell populations are then combined in equal proportions before subjecting to optional methods, such as subcellular organelle purification, cell lysis, protein extraction, and protein digestion.
- The fragmented samples are further analysed with LC-MS/MS and the quantification is done by comparing the ratios of heavy peptides to light peptides.
- Different proteome discovery **software** such as **MaxQuant, Census, Trans-proteomic pipeline (TPP),** and **pQuant** can be used to delineate the high-quality MS data and quantification of the peptides.
- After all the quantification, in order to develop insight into the proteomic results obtained from MS data, annotation **databases such as GO, KEGG, STRING**, or **bioinformatics tools such** as **GoMiner, cytoscape, DAVID** can be applied.



SILAC can be used to analyse a wide varieties of samples, such as cells, tissues, and body fluids (spike-in SILAC and super-SILAC). It can be implemented easily and is robust in terms of data analysis and result interpretation.

Over other labeling methods, SILAC has more proteome quantitative reliability.

SILAC has only few disadvantages of SILAC as it offers a limited number of cellular states for comparison contrary to iTRAQ or TMT labelling.

Quantification method	Sample processing	Labelling	Level of comparisons	Accuracy level	Proteome coverage	Range of analysis
SILAC (metabolic labelling)	Ex-vivo or in-vivo (Cell cultures, tissues or animal models)	Protein samples	Up to 5 (can be used in combinations)	very high	high	1-2 log
^{15}N (metabolic labelling)	Cells	Protein samples	2	high	high	1-2 log
ICAT (chemical labelling)	in-vitro (cells, animal tissues)	Protein samples	2	high	high	2 log
iTRAQ (chemical labelling)	in-vitro	peptides	up to 4-8	high	high	2 log
TMT (chemical labelling)	in-vitro	peptides	up to 6-10	high	high	2 log
Dimethyl labelling (chemical labelling)	in-vitro	peptides	2-3	high	high	2 log

Table: Comparison of different labelling methods

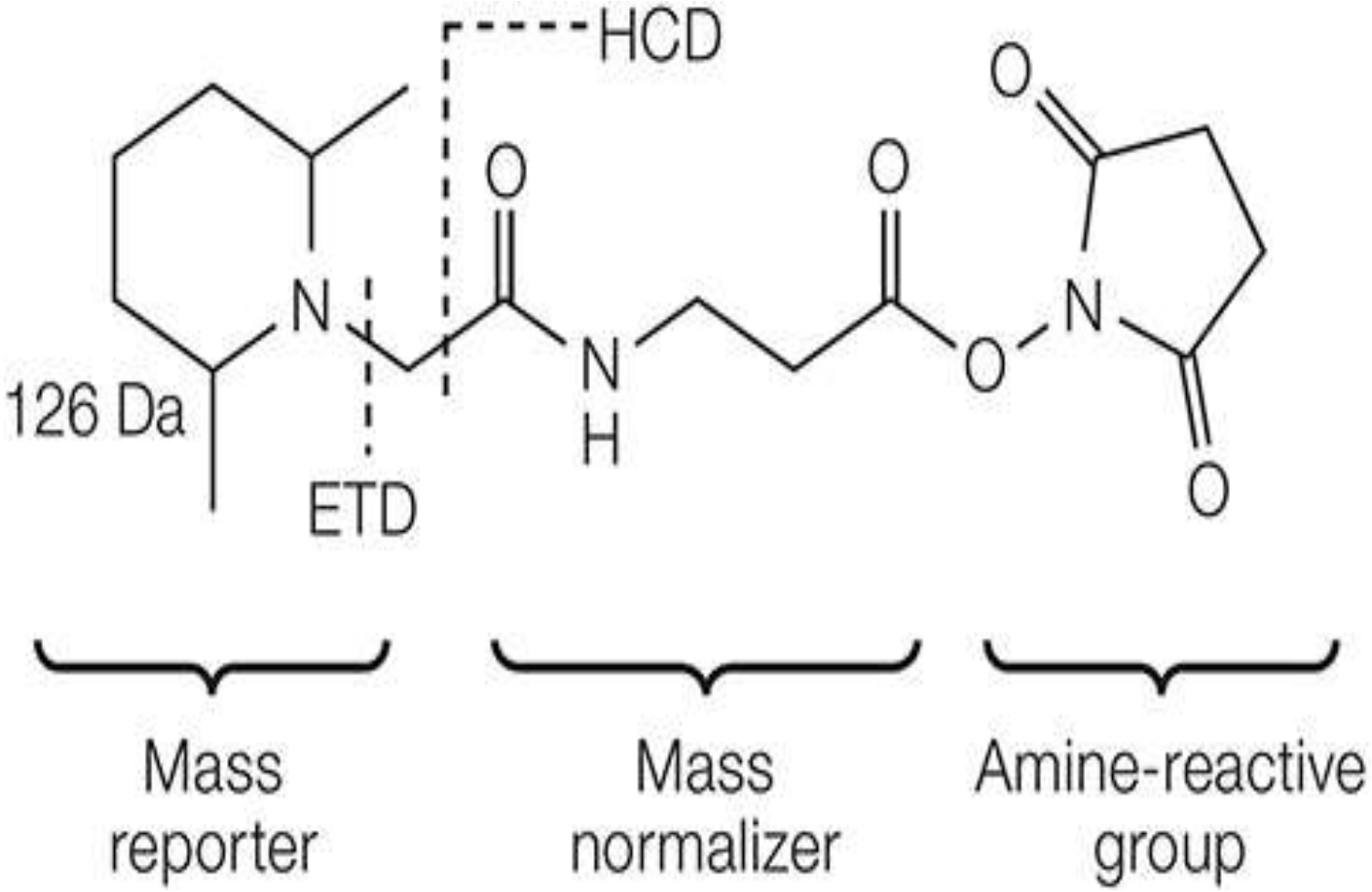
Label-dependent protein quantitation techniques

- More recent labelling strategies, such as **Tandem mass tag (TMT)** and **isobaric tag** for relative and absolute quantitation (iTRAQ), have gained popularity as these techniques apply isobaric tags providing MS2-level relative quantitation while boosting the MS1 signal intensity to improve chances of peptide detection.
- One notable advantage of using TMT in a clinical proteomic setting is the high degree of multiplexing, up to 16-plex, to significantly reduce the LC–MS time requirements of analyzing increasingly large patient cohorts. Potential caveats of TMT -based approaches are that extensive peptide-based fractionation is required to obtain deep proteome profiles and that 1-2 TMT channels are usually dedicated to a global control (i.e. a combined lysate of all clinical lysates analyzed). This reduces the ability of individual projects to be effectively compared against each other. More recently, the ratio compression of TMT labels and thus limited dynamic range of quantitative proteomic data can be overcome with further fragmentation.
- But, the use of stable-isotopic labels adds expenses in clinical proteomic studies.

A **tandem mass tag (TMT)**

- Тандемная масс-метка
- It is a chemical label that facilitates sample multiplexing in mass spectrometry (MS)-based quantification and identification of biological macromolecules such as proteins, peptides.
- **TMT belongs to a family** of reagents referred to as **isobaric mass ta**gs which are a set of molecules with the same mass, but yield reporter ions of differing mass after fragmentation.

A Tandem Mass Tag Reagents are designed to enable identification and quantitation of proteins in different samples using tandem mass spectrometry (MS).
All mass **tagging reagents** within a set **have the same nominal mass** (i.e., are isobaric) and chemical structure composed of **an amine-reactive NHS ester group**, a **spacer arm (mass normalizer)**, and a **mass reporter** (**Figure**)



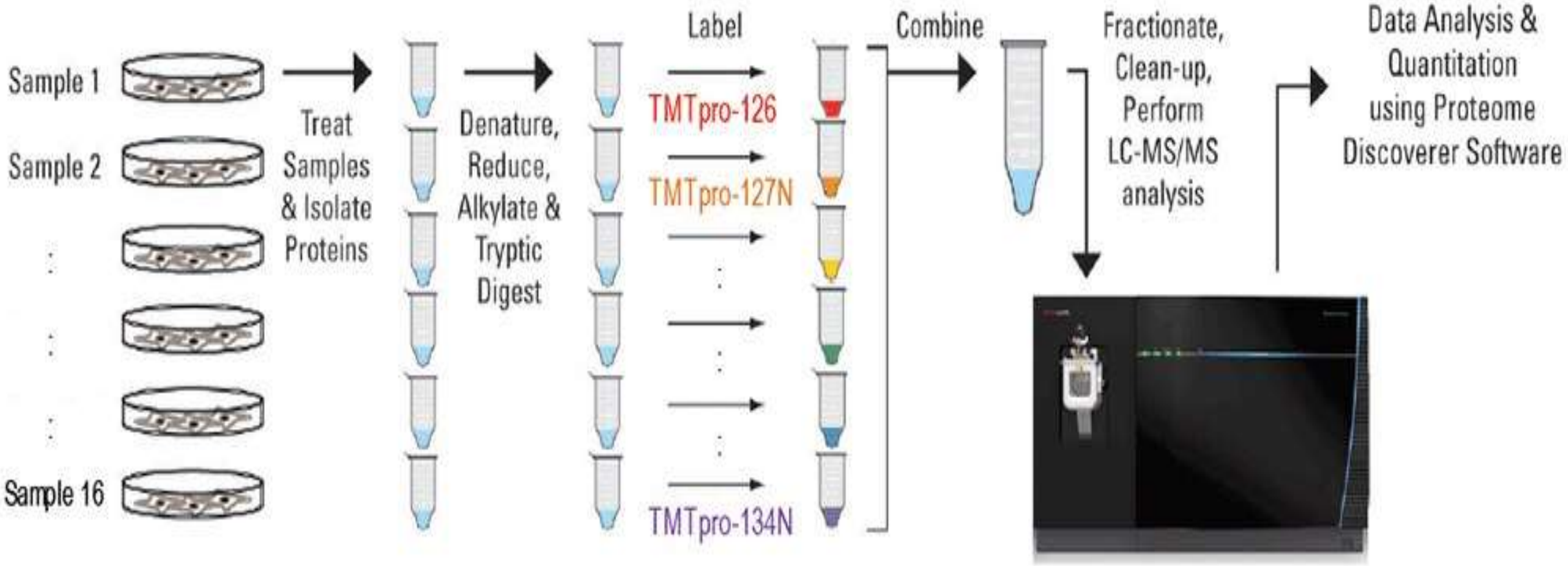


Figure. Procedure summary for MS experiments using TMTpro isobaric mass tagging reagents: Protein extracts isolated from cells or tissues are **reduced, alkylated,** and then **digested** using the EasyPep Mini MS Sample Prep kit or an equivalent method.
Samples are **then labeled with the TMTpro reagents** before sample mixing, fractionation, and cleanup. **Labeled samples are** analyzed on a **high-resolution Orbitrap LC-MS/MS mass spectrometer** before data analysis to identify peptides and quantify relative abundance of reporter ions

Label-free proteomic techniques

- As an alternative, recent advances in data computation have given rise to powerful label-free proteomic techniques.
- Label-free quantitation is more applicable to clinical proteomics due to the inter-patient and intra-patient variability of protein expression.
- Label-free and isobaric tag (iTRAQ) strategies were compared in muscle-invasive and non-muscle-invasive bladder cancer tissues.
- It was concluded that both methods provide comparable proteome coverages and proportional quantitative data, but label-free technique identified a greater number of differentially expressed proteins.
- Latosinska A, Vougas K, Makridakis M, Klein J, Mullen W, Abbas M, et al. Comparative analysis of label-free and 8-plex iTRAQ approach for quantitative tissue proteomic analysis. PLoS ONE. 2015;10(9):1-25.

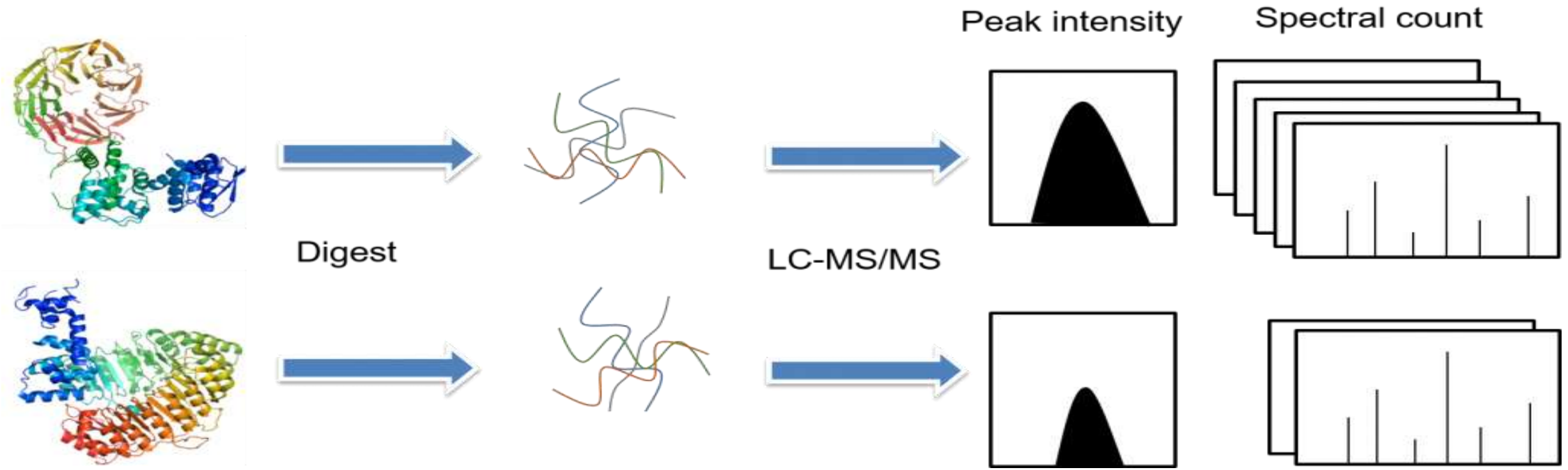
- Label-free protein quantification is a mass spectrometry-based method for identifying and quantifying relative changes in two or more biological samples.
- There is no limit on the number of samples in label-free protein quantification and it's in principle applicable to any kind of sample, including materials that can not be directly metabolically labeled (for instance, many clinical samples).
  - It is ideal for large-sample analyses in clinical screening or biomarker discovery experiments.

Principles

In general, label-free approaches can be divided into two distinct groups according to the method used for data extraction.

On one hand, the quantification can be inferred by counting the number of peptides or spectra assigned to a given protein, and therefore are generically called spectral counting methods.

On the other hand, when liquid chromatography is coupled with mass spectrometry, quantitative values can be measured through the extraction of the area of the precursor ions' chromatographic peaks – the area under the curve (AUC) or MS1signal intensity methods.



Workflow of the label-free method

There are several fundamental steps i in the label-free quantitative proteomics:

- sample preparation (protein extraction, reduction, alkylation, and digestion),
- sample separation by liquid chromatography and
- analysis by tandem mass spectrometry, and data analysis (peptide/protein identification, quantification, and statistical analysis).

In label-free quantitative proteomics methods, each sample is separately prepared and then subjected to individual LC-MS/MS or LC/LC-MS/MS runs. Quantification is based on the comparison of peak intensity of the same peptide or the spectral count of the same protein.