CHANGES TO STUDY PROGRAMMES / SYLLABUSES IN ADVANCED ANALYTICAL CHEMISTRY IN PROTEOMICS, LIPIDOMICS, AND GLYCOMICS

at University of Aveiro

Contents:

A) Changes to study programme in AAC in proteomics, lipidomics, and glycomics, to be implemented in the PhD programme on Biochemistry of the University of Aveiro

I) Advanced Courses in Biochemistry I

II) Advanced Courses in Biochemistry II

B) Extended contents of the course in lipidomics, to be implemented in Advanced Courses in Biochemistry I of the PhD programme on Biochemistry of the University of Aveiro

C) Extended contents of the course in proteomics, to be implemented in Advanced Courses in Biochemistry II of the PhD programme on Biochemistry of the University of Aveiro

This project has been funded with support from the European Commission.

A) Changes to study programme in AAC in proteomics, lipidomics, and glycomics to be implemented in the PhD programme on Biochemistry of the University of Aveiro

The programme of the PhD programme on Biochemistry of the University of Aveiro has been profoundly changed, in the result of AACLifeSci. The new programme includes now two courses which will be implemented in the scholar year of 2016/17.

I) Advanced Courses in Biochemistry I

1. Learning outcomes of the curricular unit

This course aims to provide students with advanced skills in strategy and methods of modern and specialized biochemical analysis used in research and clinical laboratories and industries. At the end of this course students should:

- Understand the fundamentals and applications of mass spectrometry
- Be able to critically evaluate the methods used to lipidomics and metabolomics studies
- Be able to plan a lipidomics or metabolomics experiments
- Be able to analyse the lipidomics or metabolomics data

2. Syllabus

Module A: LC-MS approaches

- Ions sources, mass analysers and mass spectrometry analysis
- Tandem mass spectrometry
- LC-MS platforms
- Qualitative and quantitative analysis using LC-MS

Module B: Lipidomics

- Lipidomics workflow
  - Lipid extraction from biological samples
  - Lipid analysis by GC and GC-MS
  - Lipid identification and quantification by MS and LC-MS
  - Lipidomics data analysis
  - Lipidomic profile of different cell, tissues and biofluids

Module C: Metabolomics

- Metabolomics workflow
- MS based strategies (GC-MS and LC-MS) for the analysis of metabolome
- Metabolomics resources
- Metabolomic data analysis and metabolomics networks

Note: Due to the characteristics of the UC, offered in modules, it can occur few deviations from the program defined above, depending on the modules attended by the students. This approach however has the advantage

This project has been funded with support from the European Commission.

of allowing a better adaptation of the contents to the formation of the students, and allow better adaptation of content to the formation waiting for students and for their PhD working plan. In either case the adaptation of the contents cannot question the results expected by the frequency of this unit.

3. Demonstration of the syllabus coherence with the curricular unit's learning objectives.

The course aims that students develop advanced skills in most modern strategies and techniques that are used in biomolecular and biochemical analysis, especially with mass spectrometry and mass spectrometry coupled with liquid chromatography and gas chromatography and its application in analysis of metabolites and lipid from samples of different biological origins. The course aims to discuss the correlation between the variations of the metabolome / lipidome and biochemical processes related to disease states which can be useful in developing new therapeutic targets, and identify important markers of a development stage of a pathological condition or disease treatment.

It is expected that students at the end of this course acquire advanced knowledge and skills, theoretical and practical, in mass spectrometry, lipidomics and metabolomics, essential to the biochemistry and related areas as medicine, biology and biotechnology

4. Teaching methodologies (including evaluation)

The course will run for modules to allow specialization. Courses will be based on expository/ interactive lectures with students and demonstration/practical sessions and analysis of practical cases that will allow consolidating specific outcomes.

Students must select two modules from among the modules proposed for this course, in accordance with the conductor of the UC and the advisor. Students can, and in accordance with the teacher and your advisor choose other modules to be offered advanced courses in other third cycles of UA or from another university with which the University of Aveiro has protocol.

This format in modules favors a significant freedom in selecting the acquisition of specific contents, thus offering the opportunity to students to tailor the acquisition of knowledge in a particular method of analysis, particularly important for the situation of each student's doctoral program.

The evaluation will be performed on each individual module and the final grade will be obtained by a weight average of the evaluation obtained in two modules. The implementation in concrete assessment of each module process will be up to the teacher responsible.

5. Demonstration of the coherence between the teaching methodologies and the learning outcomes.

The teaching methodologies favor the practical aspects of learning thereby guaranteeing the achievement of the objectives of the course.

According to the teaching methodologies employed, at the end of the course students will have gained knowledge on advanced analytical strategies based on the MS and LC-MS analysis and in applications in lipidomics and metabolomics.

This project has been funded with support from the European Commission.

6. BIBLIOGRAPHY

1. Leray, C. Introduction to Lipidomics: From Bacteria to Man, , CRC Press, Taylor & Francis Group, Boca Raton FL, 2012.

2. Metabolomics in Practice: Successful Strategies to Generate and Analyze Metabolic Data, Michael Lammerhofer (Editor), Wolfram Weckwerth (Editor) 2013

This project has been funded with support from the European Commission.

II) Advanced Courses in Biochemistry II

1. Learning outcomes of the curricular unit

This course aims to provide students with advanced skills in strategy and methods of modern and specialized biochemical analysis in Omics, used for proteomics, glycomics and bioinformatics, which can be used in research and clinical laboratories and industry.

At the end of the course the student must be able to:

- Critically evaluate which methods should be used to solve specific problems of proteomics and glycomics
- Plan a proteomics and glycomics experiments
- Analyse the data a proteomics and/or glycomics
- Know the most important biological databases
- Learn to navigate these databases and extract information
- Understand the concepts of genome, transcriptome, proteome and metabolome and systems biology
- Know how to select and how to use the most adequate algorithms and tools in biochemistry applications

2. Syllabus

Module A: Proteomics

Principles of protein identification using mass spectrometry

- Sample preparation and pre-fractionation
- Gel-based proteomic analysis
- Gel-Free proteomic analysis
- How Computational tools for Data Analysis in proteomics
- Special Topics in proteomics
  - Top-down vs. bottom up
  - Identification of post translational modifications
  - Quantitative Proteomics

Module B: Glycomics

- Glycomic workflow
- Isolation and separation of glycans and glycoconjugates.
- Strategies hydrolysis and fractionation of glycans and glycosylated derivatives
- Sugar analysis by GC and GC-MS.
- Glycan analysis by mass spectrometry

Module C: Bioinformatics

- Biological and biochemistry data sources: databases, literature
- Bioinformatics applications: programming, storage, tools and algorithms

Note: Due to the characteristics of the UC, offered in modules, it can occur few deviations from the program defined above, depending on the modules attended by the students. This approach however has the advantage

This project has been funded with support from the European Commission.

of allowing a better adaptation of the contents to the formation of the students, and allow better adaptation of content to the formation waiting for students and for their PhD working plan. In either case the adaptation of the contents cannot question the results expected by the frequency of this unit.

3. Demonstration of the syllabus coherence with the curricular unit's learning objectives.

The course aims that students develop advanced skills in advanced strategies and techniques that are used in biomolecular and biochemical analysis, especially with mass spectrometry and mass spectrometry coupled with liquid chromatography and gas chromatography and its application in analysis of proteins and glycans from samples of different biological origins. The course aims to discuss the correlation between the variation of the proteome, glycome, and biochemical processes related to disease states which can be useful in developing new therapeutic targets, and identify important markers of a development stage of a pathological condition or disease treatment.

It is expected that students at the end of this course, acquire advanced knowledge and skills, theoretical and practical, in proteomics, glycomics, and bioinformatics, essential to the biochemistry and related areas as medicine, biology and biotechnology.

4. Teaching methodologies (including evaluation)

The course will run for modules to allow specialization. Courses will be based on expository/ interactive lectures with students and demonstration/practical sessions and analysis of practical cases that will allow consolidating specific outcomes.

Students must select two modules from among the modules proposed for this course, in accordance with the conductor of the UC and the advisor. Students can, and in accordance with the teacher and your advisor choose other modules to be offered advanced courses in other third cycles of UA or from another university with which the University of Aveiro has protocol.

This format in modules favors a significant freedom in selecting the acquisition of specific contents, thus offering the opportunity to students to tailor the acquisition of knowledge in a particular method of analysis, particularly important for the situation of each student's doctoral program.

The evaluation will be performed on each individual module and the final grade will be obtained by a weight average of the evaluation obtained in two modules. The implementation in concrete assessment of each module process will be up to the teacher responsible.

5. Demonstration of the coherence between the teaching methodologies and the learning outcomes.

The teaching methodologies favor the practical aspects of learning thereby guaranteeing the achievement of the objectives of the course.

According to the teaching methodologies employed, at the end of the course students will have gained advanced knowledge on advanced analytical strategies used in proteomics and glycomics and applications of bioinformatics.

This project has been funded with support from the European Commission.

6. Bibliography

1. Glycomics: Methods and Protocols (Methods in Molecular Biology) 2009th Edition, Nicolle H. Packer (Editor), Niclas G. Karlsson (Editor) Human Press (2009)

2. Proteomics: F. Amado, P. Domingues, M. Domingues, R. Ferreira, R. Vitorino. "Análise de proteínas, Guia do Laboratório", 100Luz, 2013

3. Proteomics: Josip Lovric, "Introducing Proteomics: From concepts to sample separation, mass spectrometry and data analysis", Wiley. ISBN: 978-0-470-03523-8.

4. Bioinformatics: Sequence and Genome Analysis, Second Edition David W. Mount. Cold Spring Harbor Press (2004)

5. Bioinformatics: A Practical Guide to the Analysis of Genes and Proteins, 3rd Edition. Edited by Andreas D. Baxevanis

This project has been funded with support from the European Commission.

B) Extended contents of the course in Lipidomics, to be implemented in Advanced Courses in Biochemistry I of the PhD programme on Biochemistry of the University of Aveiro

1. **Introduction to lipidomics**:
    a. Lipidome as a part of metabolome

It is proposed that eukaryotic cells contain up to 200 000 different molecular species and lipidomics aims to quantify and identify lipid molecular species as much as possible and to address differences in lipid profile related to deviation in metabolism and in diseases'. The name "Lipidome" appeared first in the literature in 2001 and described the complete lipid composition within a cell, tissue or organism.

    b. Lipid classes and their functions (membrane regulation, cell signaling): Glycerophospholipids, fatty acids, sterols, eicosanoids and prostanoids.

Lipids are natural molecules that are insoluble or partially soluble in water. These hydrophobic or amphipathic molecules can be either biosynthesized or absorbed from the environment and are vital for the life of all eukaryotic and prokaryotic organisms. Lipids play crucial biological roles through three general mechanisms, they affect the cellular membrane structures and protein–membrane interactions, they provide a source of energy through processes of oxidation, and they serve as signaling molecules, binding to plasma membrane or nuclear receptors mediating transmembrane signaling and cell-to-cell communication.

    c. Lipidomic pathways.

In order to understand how lipid homeostasis regulates biological processes it will be necessary to combine pathway analysis with MS-based lipidomics. It is very important to investigate how the perturbance of one pathway influences the lipidome of cells. Pathway examination in the context of lipid analysis will lead to an increase or decrease in certain lipid metabolites, either directly or indirectly through a cascade of signaling events.

    d. Lipidomics.

The term lipidomics was introduced in the literature in 2001, however it did not refer to large-scale MS-based analyses. Lipidomics was seen mainly in the context of other functional 'omics' technologies such as genomics. Nowadays, lipidomics is seen in a broader context as a field of systems level analysis. Thus lipidomics is defined as the discipline that studies the largescale changes in lipid composition accompanying perturbations of biological systems. It involves the comprehensive identification and quantification of all individual cellular lipid species and the characterization of their interactions with neighboring lipids and proteins. The ultimate goal of lipidomics is to understand the role of lipids in the biology of living organisms.

    e. Oxidized lipids.

The early products of phospholipid peroxidation are hydroperoxides, which can rearrange to compounds containing epoxides, hydroxides, or ketones and isoprostane-like structures. Further oxidation and destabilization can result in cleavage of the oxidized fatty acyl chain to yield chain-shortened phospholipids and a corresponding nonesterified molecule.

Several families of nonenzymatic polyunsaturated fatty acids (PUFA) oxidation products were shown. These include PUFA fragmentation products (e.g., 4-hydroxynonenal), hydroxides and hydroperoxides of PUFAs,

This project has been funded with support from the European Commission.

and isoprostanes. In addition to free PUFAs, esterified oxidized residues the majority of PUFAs are esterified in phospholipids (PLs). Nonenzymatic peroxidation generates PLs containing a mixture of oxidized residues (OxPLs). OxPLs were characterized as the major component of minimally oxidized low-density lipoprotein responsible for its ability to induce synthesis of cytokines and chemokines (IL-6, IL-8, MCP-1, etc.) and stimulate adhesion of monocytes to endothelial cells.

2. **Analytical techniques in lipidomics.**

   a. Sample preparation: lipid extraction methods (Folch, Bligh and Dyer, MTBE - methyl-tert-butyl ether extraction); separation of lipid classes (TLC and HPLC).

Commonly, lipids are extracted using the chloroform- and methanol-based protocols of Folch, or Bligh and Dyer. Additionally, another type of lipid extraction, the so-called methyl-tert-butyl ether extraction has emerged which extracts lipid very well but has the advantage that lipids partition into the upper layer of a two-phase extraction system.

   b. Separation techniques coupled to mass spectrometry: liquid chromatography (HPLC, UPLC, nano-LC, 2D-LC); gas chromatography (GC, 2D-GC);

Traditionally, lipid analyses are based on analytical techniques with low resolution and sensitivity, such as thin-layer chromatography (TLC). This method is still commonly used today because of its relative ease of use and low costs. The development of gas chromatography (GC) and liquid chromatography (LC) allow the separation of different fatty acid components from simple mixtures. The identification and quantification of individual lipids from complex matrices such as a lipid extracts from biological samples is a real challenge. Therefore, GC and LC, as major separation techniques are combined with mass spectrometry (MS).

   c. Mass spectrometry for lipidomics, ionization methods (ESI and MALDI), untargeted and targeted approaches: shotgun and global lipidomics, MS fingerprinting, data-dependent acquisition, lipid class specific scanning (precursor ion scanning, neutral-loss scanning)

In past decades, the field of lipid research has made great progress by impressive advances in MS, in particular, soft ionization techniques such as electrospray ionization (ESI) and MALDI. Other ionization technique such as atmospheric pressure ionization has also been used.

Analysis of lipids by ESI-MS is based on the ability of each class of lipids to acquire positive or negative charges when in solution during ionization. in addition, electron ionization (EI) is only used in combination with GC-MS to analyze neutral lipids such as cholesterol, cholesterol esters, and derivatized fatty acids.

Lipids can be directly identified from crude lipid extracts solely by their mass. This approach is called top-down lipidomics and requires high-resolution mass spectrometers such as FT-Orbitrap, linear ion trap hybrids (LTQ-FT and LTQ-Orbitrap), triple quadrupole or Q-TOF instruments. In top-down lipidomics, the goal is mainly to distinguish differences in the lipid pattern rather than quantifying individual lipid species although it can also be used for this purpose.

In a typical lipidomics experiment, the lipid sample is delivered to the mass spectrometer, ionized and vaporized, and the resulting ions are sorted according to their *m/z* ratio in the mass analyzer. Tandem mass spectrometry allowed the confirmation of the identity of each lipid molecular specie. If a crude lipid extract

This project has been funded with support from the European Commission.

is introduced directly into the mass spectrometer without previous chromatographic separation and analyzed and quantified using neutral loss or precursors ion scan experiments, it is called shotgun lipidomics.

Alternatively, the lipid extract is introduced to the mass spectrometer after liquid chromatography (LC-MS or HPLC-MS), which increases the number of detectable lipids due to reduced ion suppression. Thus actually most lipidomics experiments are performed using LC-MS and MS/MS. This approach has been termed bottom-up lipidomics.

d. Lipid class targeted profiling

Another lipidomic approach that main goal is to achieve highest sensitivity is targeted lipidomics. This approach is most useful for quantification of a specific lipid class or subclass. Commonly it refers to MS/MS experiments, mainly to selected reaction monitoring (SRM) or multiple reactions monitoring (MRM). In this approach the signals are only recorded when a specified precursor/fragment mass pair is detected. In a targeted lipidomics approach the different lipid classes are analyzed separately using extraction and analytical protocols designed specifically for each lipid species.

e. Targeted quantitative analysis of lipid metabolites with LC-QQQ.

3. **Compound identification.**

a. Isotopic distribution,

b. Fragmentation rules for lipid species,

c. Lipidomic tools: LIPID MAPS and LipidBank.**,** libraries, databases: Human Metabolome Database and MassBank database.

Currently available lipidomics tools for annotation (LIPID MAPS, LipidBank), shotgun MS/MS software tools such as the publicly available LIPIDXPLORER, LC-MS/MS tools and databases containing reference spectra such as the Human Metabolome Database or the MassBank database) demonstrate the growing need for detailed bioinformatics resources to facilitate lipid research.

4. **Pathway analysis of lipidomics data**

a. Selected lipid pathway of interest

b. Large scale pathway analysis

5. **Clinical and biomedical applications.**

Lipids are involved in the pathogenesis of various diseases and defects in lipid metabolism are central to many devastating diseases such as insulin-resistant diabetes, Alzheimer's disease, cancer, atherosclerosis, steatohepatitis and obesity. A comprehensive analysis of the lipid metabolic changes is therefore crucial for the understanding of cellular pathologies and lipidomics will be a useful tool in the elucidation and characterization of defects in lipid homeostasis.

This project has been funded with support from the European Commission.

PRACTICAL Subjects

1. **Analysis of plasma phospholipid profile**:
    a. Lipid extraction from plasma
    b. Quantification of total phospholipids content by phosphorous assay
    c. Separation of majority of the phospholipid classes with use of thin layer chromatography (TLC)
    d. Quantification of phospholipids content in each class by phosphorous assay
    e. QTOF spectrometer tuning and calibration for high mass accuracy and resolution
    f. Direct plasma phospholipid species profiling by LC-QTOF analysis of the total lipid extract (data dependent scanning mode).
    g. Analysis of obtained data: Qualitative Mass Hunter
    h. Phospholipids identification based on: structural elucidation, chromatographic properties, fragmentation pattern, MS/MS spectra (Lipid MS predictor).
2. **Targeted lipidomics.**
    a. Sample preparation: protein precipitation, SPE extraction and cleaning
    b. Quantitative analysis of isoprostanes (8-isoPGF$_{2\alpha}$) in plasma - LC-MS/MS method (MRM mode).
    c. Analysis of obtained data: quantitative procedure with Quant Mass Hunter (calibration curve construction, ESTD or ISTD method).

This project has been funded with support from the European Commission.

C Extended contents of the course in Proteomics, to be implemented in Advanced Courses in Biochemistry II of the PhD programme on Biochemistry of the University of Aveiro

Introduction: Mass spectrometry principles for the identification of proteins

In recent years, mass spectrometry has evolved dramatically and is now regarded as a key technology for the identification, characterization and quantification of proteins and post-translational modifications. The role of mass spectrometry in biochemistry is the result of development of new ionization techniques, the creation and development of more sensitive and accurate analyzers, development and developing more effective fragmentation methods, important advances in separation technologies and rapid development of protein analysis methods using mass spectrometry. In this course we will briefly discuss these characteristics of modern mass spectrometry and the way it is used to identify proteins, in a global perspective.

Mass spectrometry (MS) is an analytical technique that involves the study of ionized gas phase molecules, in order to determine the molecular weight of the compounds, to characterize their structure and to determine its abundance. Briefly, mass spectrometry workflows consist generally in ionization of a sample in an ion source, the separation of ionised molecules according to their ratio of mass/charge (m/z) using an analyzer, detection of molecules an ionized detector and analyzing and interpreting the resulting mass spectrum. The tandem mass spectrometers have the additional capability of selecting an ion and induce their fragmentation in order to obtain detailed structural information of the selected species (MS / MS).

There are various methods of ionization that allow the analysis of organic compounds. However, the development of two soft ionisation techniques, electrospray ionization (ESI) desorption and ionization by matrix assisted laser (MALDI), able to ionize organic compounds not volatile and unstable term enabled the development of mass spectrometry as an analytical technique for the analysis of biological molecules. The method of ionisation by ESI is based on applying a high potential to a metallic capillary located close to a counter electrode. The ionization process occurs in four major steps: the formation of electrically charged droplets from a dissolved electrolyte; the reduction of the drop volume by evaporation of solvent and repeated disintegration leading to formation of highly charged droplets capable of producing ions in gas phase; ion production in the gas phase; and secondary processes suffered by the ions in the gas phase. The method of MALDI ionization is based on irradiation with a laser sample, previously dissolved in a matrix which absorbs energy at the laser wavelength used. The ionization mechanism is still under discussion but involves desorption and ionization of analyte proton transfer array.

Most researchers in proteomics proceeds to fractionation and preconcentration of the sample using high-performance liquid chromatography (HPLC), typically using a capillary column (nano-LC). This approach is used to increase protein identification and coverage rates observed protein sequence. The existing discrimination in the analysis of different peptides in MALDI and ESI also contributes to different protein coverage rates. Several authors have reported different results when using these two ionization methods. However, the results indicate that, in general, the ionization ESI allows a slightly better coverage rate. These studies also suggest that to increase the coverage rate should be used both ionization methods.

In mass spectrometry, mass analyzers are the components which separate ions according to their ratio of mass

This project has been funded with support from the European Commission.

/ charge. In fact, when subjected to a constant acceleration voltage, the separation of ions according to their moment or speed is equivalent to separation based on a ratio of mass / charge. The separation of ions according to the present is the basis of the magnetic deflection instruments. The separation according to the speed of ions is the basis of the time of flight instruments (TOF). The quadrupoles (Q) and the quadrupole ion traps (IT) separate the ions directly in accordance with m / z utilizing RF and DC voltages. In Orbitrap, the analysis of ions is performed by measuring the frequency of harmonic oscillations induced by a linear electric fields, whereas in ion-cyclotron resonance instruments Fourier transforms (FT-ICR), the ions are separated according to the cyclotron frequency when subjected to a magnetic field.

Tandem mass spectrometry or mass spectrometry/mass spectrometry (MS/MS) can be defined as the ion detection that after forming the source, suffer load changes and / or mass during the analysis with the spectrometer pasta. The most common way to provide additional energy to the precursor ion to fragment for this consists in introducing a gas in a collision cell. In this case, the activation process is the translational energy transfer from the parent ion into internal energy by collision with a neutral target gas, typically introduced into a cell collision. This process is most often referred to as collision induced dissociation (CID). In peptide analysis using tandem mass spectrometry, the fragment ions, also called product ions, provide information on the composition and amino acid sequence, since the cleavage of amide linkages are favored. However, it is important to bear in mind that the peptides containing PTMs may have different preferred fragmentation pathways. In fact, several PTMs are known as such occurs, such as sulphation, glycosylation and phosphorylation. In these cases, fragmentation by loss of the modification in the form of a neutral molecule may be the dominant fragmentation pathway. As a result, the mass spectrum MS / MS of peptides with PTMs may contain sufficient information for the location of the site of modification. In the electron transfer dissociation method (ETD), fragmentation is induced by an electron transfer process. In this method, the MTPs labile neutral loss is not preferred fragmentation pathway and phosphorylated particularly glycated peptides. Several studies have shown that the ETD is a more informative fragmentation method than the CID, particularly for peptides with higher load than 2, despite having lower sensitivity, generating information complete. There are thus advantages in using both fragmentation methods when intended high coverage rates, such as the discovery of MTPs.

The methods used to identify proteins and are designated generally by "bottom-up", "top-down" approaches and "middle-down". These approaches can also be used for the identification of MTPs using mass spectrometry. The approach "bottom-up" is the traditional proteomics approach and is what we will discuss in this lesson. In this approach, proteins are enzymatically digested and the resulting peptides are analyzed, either directly or after separation by HPLC, using mass spectrometry and tandem mass spectrometry. The MTPs using this detection method, based on the change in weight observed in the mass spectra of the tryptic peptides resulting from chemical modification that occurs in the side chain of amino acid residues. However, for the unambiguous allocation of a local data modification experiments are necessary for using tandem mass spectrometry. In these MS / MS spectra attempt to identify the mass change detected in the precursor ion (peptide obtained by tryptic digestion of the modified protein) in ion fragments having the modified amino acid residue. However, as mentioned above, the modified peptides may have other fragmentation pathways that hinder the identification of the site of modification. These pathways are the loss of modification in the form of neutral or charged in the

This project has been funded with support from the European Commission.

form of residue. The latter fragmentation pathways occur predominantly in more labile modifications.

In general, the approach of "bottom-up", the proteins are first separated. These purification methods include electrophoretic methods generally mono or bidimensional gel, but may include other chromatography techniques, including affinity chromatography and immunoprecipitation. These latter methods are very important to previously concentrate the sample with proteins or peptides containing the change to be identified. In fact, when performing an experiment "bottom-up" usually only detect a small number of PTMs due to the complexity of the proteome found in most biological samples. It is therefore very important to use enrichment methods proteins and peptides modified in order to decrease the complexity of the sample and increasing the number of PTMs identified. Thus, there have been developed a variety of methods for selective enrichment of modified proteins for identification and quantification of MTPs.

The use of mass spectrometry in proteomics and MTPs ID allows simultaneous identification of thousands of modified peptides. Of all the tasks workflow, data processing and analysis and presentation of results are probably the most challenging aspects of this method. The characterization of MTPs using data obtained using the tandem mass spectrometry involves exhaustive search of all possible combinations of mass for each identified peptide from a protein database. (29,30) Although the development of new bioinformatics tools for analysis of large number of MTPs have allowed significant advances in this area, most investigators consider that it still requires manual validation of MS / MS data. This poses enormous difficulties in validating large data sets and obtaining conveniently validated databases (cured).

1. **Introduction to proteomics**:
   a. Proteome
   - Definition of proteome as the entire set of proteins expressed by genome, cell, tissue or organism at a certain time point;
   - Determinants of proteome: gene expression regulation (open reading frames ORF, single nucleotide polymorphism SNPs), post-translational events (alternative splicing), protein post-translational modifications (PTMs);

   b. Proteomics
   - Definition of proteomics as the large scale study of proteins, there structure and function in certain cell/tissue/organism type or under certain conditions;
   - Proteomics studies not only protein expressions, but also the interplay of proteins, protein complexes, signaling pathways, and network modules;
   - Limitations of proteomics studies;

   c. Proteomic pathways
   - Proteomic studies relate to various areas of the proteome analysis, such as: functional proteomics, structural proteomics, quantitative proteomics, clinical proteomics, post-translational modifications (containing redox proteomics);

2. **Types of proteomics research:**

This project has been funded with support from the European Commission.

a. Protein profiling

- Protein profiling is defined as identifying the character and quantity of specific sets of proteins in a particular tissue, under a specified set of conditions and at a particular time, usually compared to expression in reference samples.

b. Predicting protein structure

- Protein structure prediction is the prediction of the three-dimensional structure of a protein from its amino acid sequence - that is obtained during MS/MS experiments. Knowledge about amino acid sequence allows for the prediction of protein folding and its secondary, tertiary, and quaternary structure.

c. Protein networks

- The majority of proteins are in highly interactive networks, therefore the goal of the protein network analyzes is to understand how protein complexes regulate cellular behavior (robustness and adaptability of living cells, perturbations in these networks in pathological conditions such as cancer or neurological disorders).

3. **Proteomic methodologic approaches:**

a. Bottom up

- Bottom-up proteomics is a common method to identify proteins and characterize their amino acid sequences and post-translational modifications by proteolytic digestion of proteins prior to analysis by mass spectrometry.
- The proteins are purified and separated by electrophoresis, what resulting in one or a few proteins in each proteolytic digest.
- Peptides are identified by comparing the masses of the proteolytic peptides with those predicted from a sequence database.

b. Top-down

- Top-down proteomics is a method of protein identification that uses MS/MS experiments to provide the information about native protein molecular weight and can be performed to sequence the protein without digestion.
- Before analysis proteins should be separated and purified.
- Proteins are typically ionized by electrospray ionization, and the most abundant ions are fragmented in MS/MS experiments.
- Fragmentation for tandem mass spectrometry is usually accomplished by electron-capture dissociation or electron-transfer dissociation.
- Peptides are identified by comparing the masses of the proteolytic peptides with those predicted from a sequence database.

c. Shotgun

- Shotgun proteomics refers to the use of bottom-up proteomics techniques in identifying proteins in complex mixtures using a combination of high performance liquid chromatography combined with mass spectrometry.

This project has been funded with support from the European Commission.

- The proteins in the mixture are digested directly, followed by one or more dimensions of separation of the peptides by liquid chromatography. Tandem mass spectrometry is used to identify the peptides.

4. **Analytical techniques in proteomics:**

a. Sample preparation: SDS-PAGE, 2DG, HPLC, ProteinChips, protein digestion techniques

- Because the proteome is so complex, there is no one standard method for preparing protein samples for analysis by mass spectrometry. Protein preparation for MS analysis can be accomplished by many methods, so it is important to understand the steps leading to analysis. The most common isolation methods of selected protein are electrophoreticaly separation, such as SDS-PAGE and 2DG.
- Before digestion proteins have to be denatured. For in-solution digestion strategies, proteins are denatured with strong chaotropic agents such as urea. This step is either followed by or combined with disulphide reduction. The free sulfhydryl groups on the cysteine residues are then alkylated with reagents such as iodoacetamide.
- The denatured, reduced and alkylated proteins are then digested by endoproteinases, (e.g., trypsin, chymotrypsin, Glu-C and Lys-C), which hydrolytically break peptide bonds to fragment proteins into peptides.
- Protein separation by 1- or 2-dimensional gel electrophoresis (1DE and 2DE, respectively) is an alternative to in-solution protein denaturation, reduction, alkylation and digestion. The peptides are then extracted from the gel matrix and prepared for MS analysis.
- In the case of top down approach to separation digested peptides also liquid chromatography is used. To do the preliminary analysis of the protein content in the test samples also can be used specific ProteinChips.

b. Separation techniques coupled to mass spectrometry: liquid chromatography (HPLC, UPLC, nano-LC), capillary electrophoresis CE, gas chromatography (GC/MS)

c. Types of ionization: MALDI/ESI

- Ionization in peptides analysis is a key to determine what types of samples can be analysed by mass spectrometry.

d. Mass analyzers: types, comparison: sensitivity, resolution, mass accuracy

e. Mass spectrometry for proteomics:

- targeted and untargeted (discovery) approaches,
- data-dependent and data-independent acquisition,
- Selected/Multiple Reaction Monitoring (SRM or MRM).

f. Stable Isotope Labelling of Amino acids (SILAC), Isotope-Coded Affinity Tag (ICAT), isobaric Tags for Relative and Absolute Quantitation (iTRAQ)

- Description of quantitative techniques used in proteomics, description of markings, sample preparation, its advantages/disadvantages.

5. **Clinical and biomedical applications**

Proteomics has already many applications in the discovery of new diagnostic, prognostic and therapeutic targets, a team of international specialists from research institutions, hospitals and companies contribute with their specific expertise. The example of that can be its applications for the diagnosis the most important diseases, such as heart and cardiovascular disorders, cancer, infectious diseases and diseases of the nervous system.

This project has been funded with support from the European Commission.

PRACTICAL PART

1. **Bottom-up proteomics:**
   a. 2D electrophoresis
   b. In-gel trypsin digestion
   c. Separation of digested peptides
   d. Data-dependent acquisition of digested peptides
   e. Analysis of the obtained spectra

This project has been funded with support from the European Commission.

University of Aveiro – AACLifeSci Syllabus

The PhD program in Biochemistry of the University of Aveiro is organized in four years. The first year includes 7 curricular units, four of which are optional. Two of these optional Curricular Units (Advanced courses in biochemistry and Advanced courses in biochemistry II) will be based on the contents of AACLifeSci. These UCs are composed of three optional modules corresponding each to 3 ECTS. To complete each of these UC´s students must successfully complete two of these modules. The evaluation will be performed on each individual module and the final grade will be obtained by a weight average of the evaluation obtained in two modules. The courses and the respective modules are:

Advanced courses in biochemistry I:

This course aims to provide students with advanced skills in strategy and methods of modern and specialized biochemical analysis used in research and clinical laboratories and industries. At the end of this course students should:

-Understand the fundamentals and applications of mass spectrometry and hyphenated chromatographic techniques;

- Be able to critically evaluate the methods used to lipidomics and metabolomics studies;

- Be able to plan a lipidomics or metabolomics experiments;

- Be able to analyze the lipidomics or metabolomics data.

	UA	**Chromatography and Mass Spectrometry for the for Life Sciences**	**Lipidomics**	**Metabolomics**
Lectures and workshops	20 hours	10	10	10
Practical Lessons		5	5	5
Student centred learning	150 hours	75	75	75
Total student effort	180 hours	90	90	90
ECTS	6	3	3	3

This project has been founded with support from the European Commission.

Advanced courses in biochemistry II

This course aims to provide students with advanced skills in strategy and methods of modern and specialized biochemical analysis in Omics, used for proteomics, glycomics and bioinformatics.

At the end of the course the student must be able to:

- Critically evaluate which methods should be used to solve specific problems of proteomics and glycomics

- Plan a proteomics and glycomics experiments

- Analyze the data a proteomics and /or glycomics

-Know how to select and how to use the most adequate algorithms and tools in these biochemistry applications

	UA	Proteomics	Glycomics	Bioinformatics
Lectures and workshops	20 hours	10	10	10
Practical Lessons		5	5	5
Student centred learning	150 hours	75	75	75
Total student effort	180 hours	90	90	90
ECTS	6	3	3	3

Overall, AACLifeSci will be fully integrated in 4 modules of the PhD program in Biochemistry of the University of Aveiro, each corresponding to 3 ECTS, to a total of 12 ECTS.

Module 1- Chromatography and Mass Spectrometry for the for Life Sciences (3 ECTS)

Module 2- Metabolomics (3 ECTS)

Module 3- Lipidomics (3 ECTS)

Module 4- Proteomics (3 ECTS)

This project has been founded with support from the European Commission.