Gonçalo Nuno Gouveia Martins Pinto

Licenciado em Ciências de Engenharia Química e Bioquímica

Engineered design of new nano-micro materials for proteomics and phosphoproteomics applications

Dissertação para obtenção do Grau de Mestre em

Engenharia Química e Bioquímica

Orientador: Dr. Hugo Santos, Investigador Auxiliar, FCT-UNL

Coorientador: Dr. José Paulo Mota, Professor Catedrático, FCT-UNL

Júri

Presidente: Professor Dr. Mário Fernando José Eusébio

Arguente: Professor Dr. Luís Fonseca

Vogal: Dr. Hugo Miguel Santos

FACULDADE DE CIÊNCIAS E TECNOLOGIA
UNIVERSIDADE NOVA DE LISBOA

Setembro 2017

Gonçalo Nuno Gouveia Martins Pinto

Licenciado em Ciências de Engenharia Química e Bioquímica

Engineered design of new nano-micro materials for proteomics and phosphoproteomics applications

Dissertação para obtenção do Grau de Mestre em Engenharia Química e Bioquímica

Orientador: Dr. Hugo Santos, Investigador Auxiliar, FCT-UNL
Coorientador: Dr. José Paulo Mota, Professor Catedrático, FCT-UNL

Júri

Presidente: Professor Dr. Mário Fernando José Eusébio

Arguente: Professor Dr. Luís Fonseca

Vogal: Dr. Hugo Miguel Santos

Acknowledgments:

It has been quite a journey since I join the university and, therefore, I would like to express thanks to all my colleagues for all the support, academical and not academical, through this phase of my life.

I want to start my special acknowledgments by thanking to Prof. Carlos Lodeiro and to Prof. José Luís Capelo for letting me enter this brilliant scientific group (Bioscope), for all the support given and for believe in my abilities.

To Prof. José Paulo Mota for accepting to be my co-supervisor and for his time and knowledge.

A thank you to Mr. Eduardo Araújo and Ms. Susana Jorge for being available to help me every time I needed and for passing techniques for my laboratorial performance.

To Mr. João Prates for the support given in the laboratory and for turning lab days less monotonous.

I want to thank to Dr. Elisabete Oliveira for all the motivation given before and during my master thesis and all the advice given.

A big thank you for Mr. Gonçalo Marcelo, for being my “brother in arms” during my staying in the university. Thank you for all the support, laughs and carpools.

To Ms. Inês Ferreira for always believe in me, for always been capable to motivate me increasing my self-esteem and for helping me to push away my doubts about my skills.

To Ms. Ana Patrícia Martins, for being there in good and tough times and for all the shares during these last five years.

To all the Bioscope group, thank you, for all the support and for receiving me.

To my family, in special to my mother, for all the support on the pursuit of my dreams, patience and for forgiving me my often absence.

My last but no least acknowledgement is to Dr. Hugo Santos for the immeasurable patience, for all the academical and social wisdom transmitted, for all the support and motivation, all the time spent teaching me and for being my friend during these last months, a huge and unforgettable thank you.

Resumo

De modo a obter um conhecimento aprofundado dos mecanismos biológicos dos seres vivos é necessário estudar o seu proteoma. De modo a obter este conhecimento é imprescindível a identificação das fosfoproteínas. Para este propósito tanto a digestão de proteínas como a pré-concentração de fosfopeptídeos são etapas fundamentais.

Para a digestão das proteínas recorreu-se a tripsina imobilizada de modo a evitar problemas associados com a sua autodigestão. A tripsina utilizada neste trabalho é imobilizada num suporte que tem um núcleo de ferro, que lhe confere propriedades magnéticas. Tem também um tamanho de 80nm que tanto quanto sabemos é o menor existente; aumentando assim a área de superfície de contacto para o mesmo volume. Assim, nestas condições, foi possível identificar quatro vezes mais proteínas com o sistema desenvolvido no grupo Bioscope que com o sistema comercial.

As fosfoproteínas representam cerca de 30% do proteoma, mas a dificuldade da identificação destas proteínas prende-se com a sua baixa concentração em relação as outras. Para eliminar este problema, recorre-se a materiais capazes de isolar e pré-concentrar os fosfopeptídeos presentes na amostra. Para o efeito, neste trabalho foi utilizada uma cromatografia de afinidade de metal imobilizado (IMAC). Os IMACs criados neste trabalho têm por base uma matriz de poliestireno. Para obter uma matriz de poliestireno dentro do tamanho pretendido (20-50nm), um desenho de experiências 2^3 foi feito de modo a identificar os valores ótimos para as variáveis estudadas. Partindo de uma matriz de tamanho nanométrico (40nm) foi possível obter também um IMAC na mesma escala (250nm). Os metais utilizados nos IMACs foram titânio e lantânio, e a capacidade de pré-concentração e isolamento de fosfopeptídeos foi demonstrada. Como prova de conceito foram identificados 99 fosfopeptídeos, um número muito superior ao descrito na literatura (20).

Abstract:

For a deep understanding of the biological mechanisms of the living organisms, a detailed study of phosphoproteins is vital. Digestion and pre-concentration of phosphoproteins are critical steps in phosphoproteomics. analysis. In this work, we used a new nano-sized system in both steps.

For digestion of phosphoproteins the standard used trypsin enzyme was selected. This trypsin was immobilized in an iron core to avoid trypsin auto-lysis. This new system was found to be highly effective for protein digestion and when compared with the available commercial systems, it performed better. Thus, it was possible to identify four times more proteins than using the standard procedures.

Phosphoproteins are estimated to be 30% of the entire proteome, however, exist some issues regarding its identification such as low levels of concentration. To overcome these obstacles, ion metal affinity chromatography, IMAC, is currently used. In this work, we synthesized new nano-sized IMACs from a polystyrene matrix. These polystyrene matrices were also synthesized using a 2^3 experimental design to unravel the conditions to create them in a certain range of size. Using a polystyrene matrix of 40nm, it was possible to create an IMAC of 250nm, in the nano scale range. Metals used for the IMACs were titanium and lanthanum. Both IMACs proved to be efficient on the phosphopeptide enrichment having superior capacity than others described in the literature. The number of phosphopeptides identified from a simple sample of α-casein was 99, five times more than the best results described in literature.

Oral communications and posters presentation obtained from present master thesis:

"Expanding the toolbox for robust and reproducible protein digestion in mass spectrometry-based proteomics: the new nano-immobilized magnetics trypsin"

Gonçalo Martins, Javier Fernández-Lodeiro, Jamila Djafari, João Prates, Susana Jorge, Eduardo Araújo, Adrián Fernández-Lodeiro, Elisabete Oliveira, José P. Mota, Carlos Lodeiro, José L. Capelo, Hugo M. Santos

V International Congress on Analytical Proteomics 2017, Costa da Caparica, Portugal – 3rd-6th July 2017.

Type of contribution: Poster - http://www.icap2017.com

"Expanding the toolbox for robust and reproducible protein digestion in mass spectrometry-based proteomics: the new nano-immobilized magnetics trypsin"

Gonçalo Martins, Javier Fernández-Lodeiro, Jamila Djafari, João Prates, Susana Jorge, Eduardo Araújo, Adrián Fernández-Lodeiro, Elisabete Oliveira, José P. Mota, Carlos Lodeiro, José L. Capelo, Hugo M. Santos

V International Congress on Analytical Proteomics 2017, Costa da Caparica, Portugal – 3rd-6th July 2017.

Type of contribution: Shotgun Presentation - http://www.icap2017.com

Winner Award: Excellent Shotgun Presentation

"Trypsin goes nano! New nano-sized magnetic trypsin for ultra-high effective protein digestion"

Gonçalo Martins, Javier Fernández-Lodeiro, Jamila Djafari, João Prates, Susana Jorge, Eduardo Araújo, Adrián Fernández-Lodeiro, Elisabete Oliveira, José P. Mota, Carlos Lodeiro, José L. Capelo, Hugo M. Santos

XXV Encontro Nacional da SPQ, Lisboa, Portugal – 16th-19th July 2017.

Type of contribution: Poster - http://www.spq.pt/agenda/event/306

"New insights on nano-magnetic immobilized trypsin for mass spectrometry-based proteomics"

Gonçalo Martins, Javier Fernández-Lodeiro, Jamila Djafari, João Prates, Susana Jorge, Eduardo Araújo, Adrián Fernández-Lodeiro, Elisabete Oliveira, José P. Mota, Carlos Lodeiro, José L. Capelo, Hugo M. Santos

III International Caparica Symposium on Profiling 2017, Costa da Caparica, Portugal – 4th-7th September 2017.

Type of contribution: Poster - http://www.isprof2017.com

Future Prospects:

“The last frontier in proteomics: The race towards nano” Publication

Gonçalo Martins, Jamila Djafari, Javier Fernández-Lodeiro, Carlos Lodeiro, José P. Mota, José L. Capelo e Hugo M. Santos

Manuscript in preparation

General Index

Figure Index:

Table Index:

Abbreviations

1D-GE	One dimensional gel electrophoresis
2D-GE	Two dimensional gel electrophoresis
2D-DIGE	Two-dimensional difference gel electrophoresis
ACN	Acetonitrile
ADP	Ammonium dihydrogen phosphate
AIBN	2,2'-Azobis(2-methyl-propionitrile)
AmBic	Ammonium Bicarbonate
APS	Ammonium Persulfate
BSA	Bovine Serum Albumin
CHCA	α-cyano-4-hydroxycinnamic acid
DLS	Dynamic Light Scattering
DTT	Dithiothreitol
ESI	Electrospray Ionization
FA	Formic Acid
GMA	Glycidyl Methacrylate
IAA	Iodoacetamide
IMAC	Immobilized metal affinity Chromatography
LC	Liquid Chromatography
MALDI	Matrix assisted laser desorption/ionization
MOAC	Metal oxide affinity chromatography
MQH_2O	MilliQ Water
MS	Mass Spectrometry
PDI	Polydispersion
PS	Polystyrene
PTMS	Post-Translational Modifications
PVA	Polyvinyl alcohol
Q	Quadrupole
SCX	Strong Cation Exchange
SDS	Sodium Dodecyl Sulfate
TFA	Trifluoracetic acid
THF	Tetrahydrofuran
TMPTMA	Trimethylolpropane Trimethacrylate
TOF	Time of Flight

1 Introduction

1.1 What is Proteomics?

Proteomics is defined as the systematic and large-scale analysis of proteomes. A proteome is a set of proteins encoded by the genome of a given cell, tissue or organism, and it differs from cell to cell and changes over time. [1, 2]

In 1938, the term protein was introduced by Jöns Jakob Berzelius when he wanted to describe a class of macromolecules that are abundant in living beings and made up of amino acids. However, the first protein studies that can be called proteomics began in 1975. In this study proteins from *E.coli*, guinea pig and mouse could be separated and visualized but not identified. Proteomics and proteome terms were coined around the nineties during the genomics revolution. Since there the field of Proteomics had evolved from a concept to a mainstream technology with a global market value of more than six billion dollars in 2015[1]

Proteins are involved in almost every biological activity. Therefore, an exhaustive analysis and comprehension of the proteins in a certain organism provide us perspective of how these molecules interact and corporate to assure a working biological system. Organisms respond to internal and external changes by regulating the level and activity of its proteins so changes occur in the proteome that can be of interest proving that the proteome is a dynamic and complex entity.[2]

In a broader scope, proteomics is used to investigate:[3]

- when and where proteins are expressed;
- rates of protein production, degradation, and steady-state abundance;
- how proteins are modified;
- the movement of proteins between subcellular compartments;
- the involvement of proteins in metabolic pathways;
- how proteins interact with one another.

Proteomics has applications to medicine through identification of protein markers of a disease or identification of targets of new drugs.[2]

As stated before Proteomics has grown to a mainstream technology. To achieve their goals, Proteomics will require the involvement of different disciplines such as biochemistry, biochemical engineering, and bioinformatics. Figure 1.1 indicate various types of proteomics fields and their applications.[4]

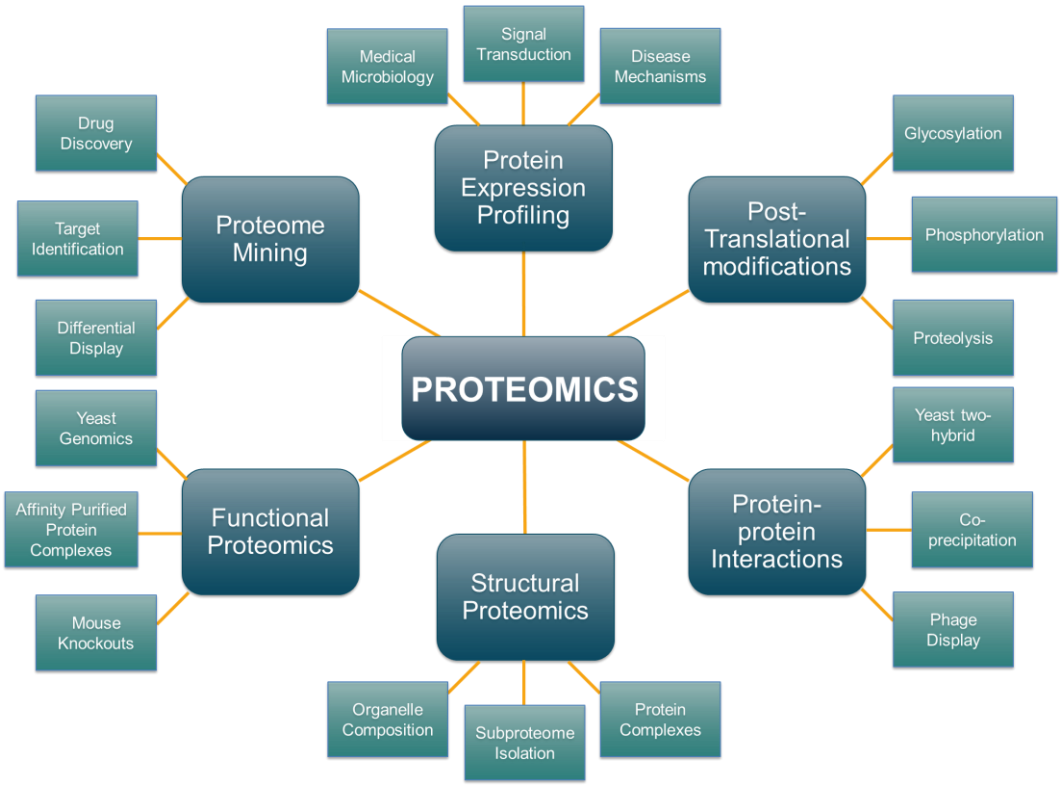


Figure 1.1: Fields of Proteomics and their applications.(adapted)[4]

One of the most challenging tasks in Proteomics is the post-translational modifications (PTMS) analysis is known that proteins modified post-translationally in response to several extracellular and intracellular signals. As an example of this PTMS, protein phosphorylation is an important signalling mechanism and dysregulation of protein kinases can result in oncogenesis.[4]

1.2 Post-Translational Modifications:

An understanding of human complexity demands not only the knowledge of all genes but also the knowledge of the proteome generated, post-translational modifications and the release of active products after biological activation.

Figure 1.2 illustrates that many things can occur during the process of protein creation.

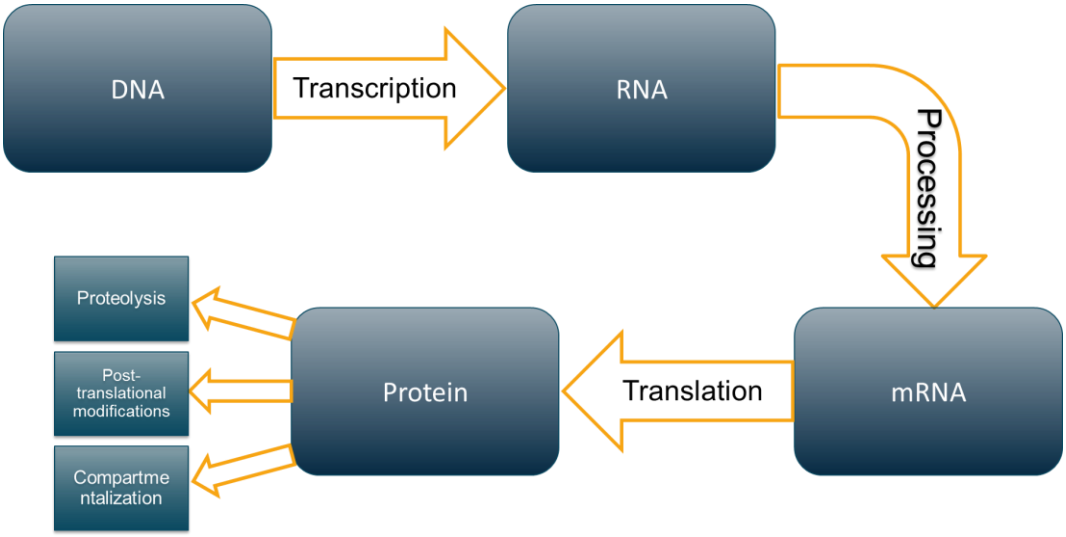


Figure 1.2: Pathway from gene to protein. (adapted)[4]

It is easy to understand by Figure 1.3 that the complexity of the proteome is bigger than the complexity of the genome. One gene can provide more than one protein. It is estimated that human integrity relies on the existence and action of 10^6 individual molecular species.

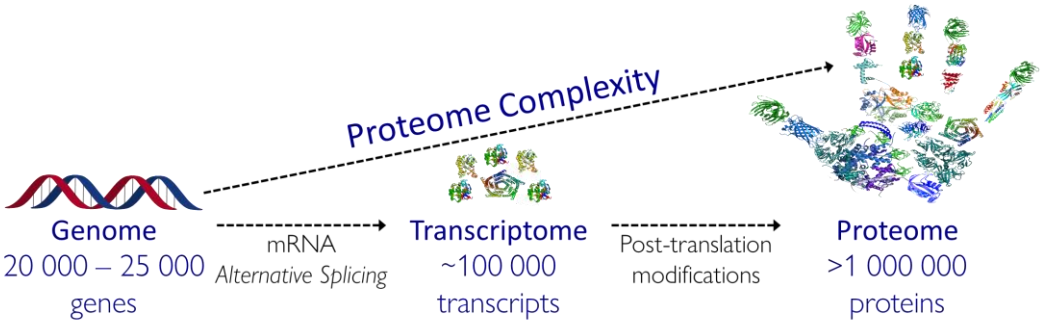


Figure 1.3: Level of complexity of genome and proteome.(adapted)[5]

PTMS are covalent processing events that modified the properties of the protein. This will occur by addition of a modifying group or by proteolytic cleavage on one or more amino acids. PTMS determines the activity state and localization of protein as well as protein interactions.[6]

Phosphorylation, Glycosylation, Ubiquitylation, Acetylation, Methylation and Sumoylation are some of the common and more important PTMS.[6,7]

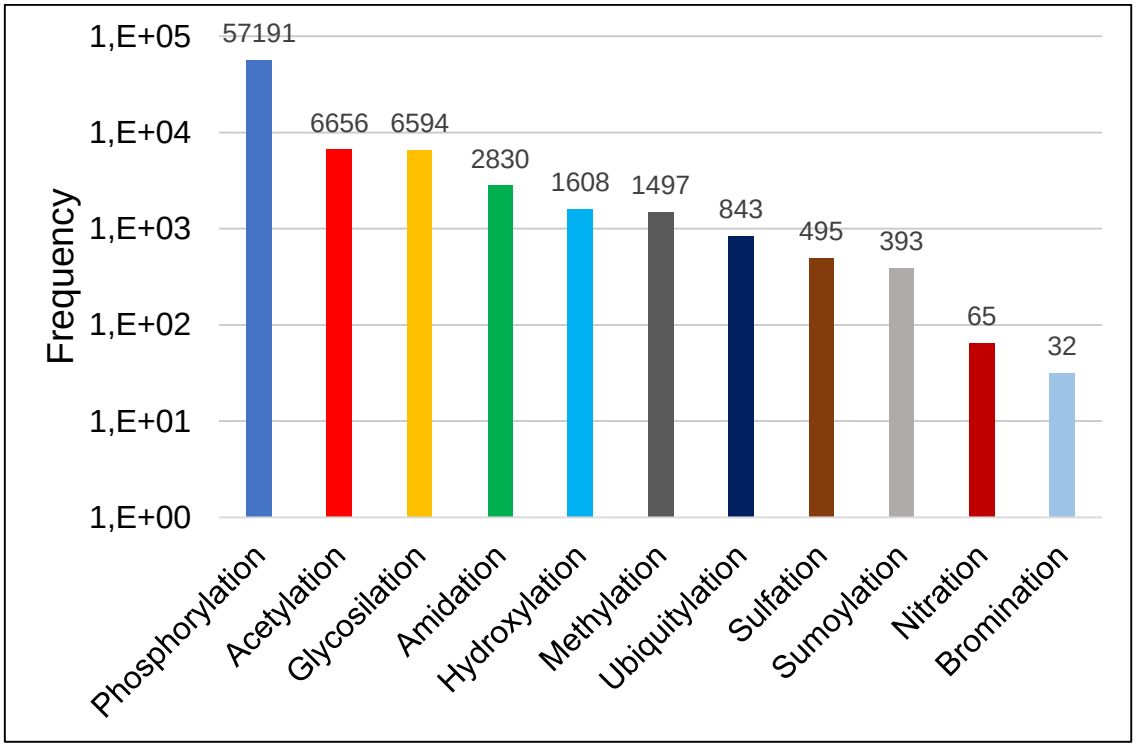


Figure 1.4: Frequency of PTMS(adapted)[7]

Figure 1.4 exemplifies how abundant phosphorylations are. Phosphorylation is not only important by its frequency but also by its biochemical process of supreme biological relevance.

1.3 Importance of Phosphorylation:

Phosphorylation is the covalent addition of a phosphate group to an amino acid, Figure 5. Phosphorylation generally occurs in serine (SER), threonine (Thr) and tyrosine (Tyr) although it may also occur in histidine (His), aspartate (Asp), cysteine (Cys), lysine (Lys) and arginine (Arg).[8]

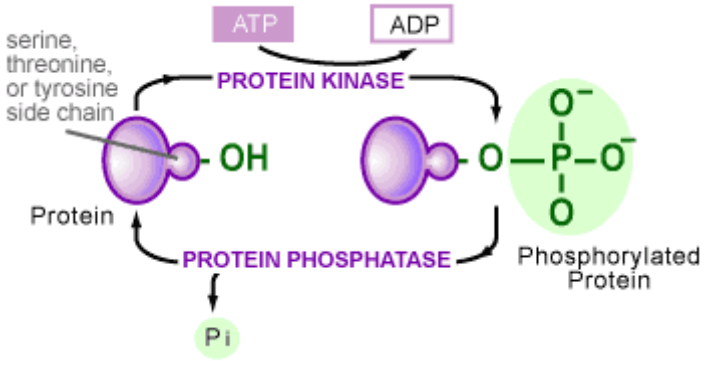


Figure 1.5: Addition of a phosphate group to an amino acid.[9]

Phosphorylation is a reversible modification that adjusts the function of a certain protein. In these functions, are include enzymatic activities, protein localization, formation/degradation of proteins, cell signalling and more. Due to this functions protein phosphorylation is one of the key regulatory

mechanisms present in many important cellular processes. Abnormal phosphorylation may result in major diseases such as cancer, diabetes and rheumatoid arthritis.[8,10,11,12]

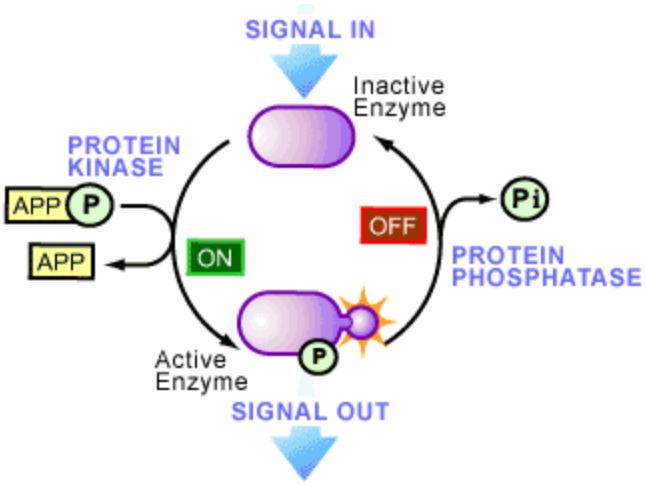


Figure 1.6: Protein Signalling by Phosphorylation.[9]

Phosphorylation is considered to affect around 30% of a proteome thus is indispensable to understand why, when and where phosphorylation occurs to prevent diseases and discover biomarkers.[13]

Phosphoproteomics is more complicated than simply measuring protein expression because the stoichiometry of phosphorylation is low and the determination of the site where phosphorylation occurs needs to be accurate.

1.4 Proteomics Techniques:

Processing and analysis of Proteomics are in fact a very complex and multistage process. Liquid Chromatography-mass spectrometry (LC-MS) and LC-MS/MS analysis of data requires multistages remaining this process as the main bottleneck for many larger proteomics studies. To overcome these issues, highly efficient sample preparation, state-of-the-art in mass spectrometry instrumentation and extensive data processing and analysis are demanded.[14]

Before exploring proteomics techniques, is important to understand the pathway required to do from sample to protein identification. Figure 1.7 shows the workflow for a MS-based proteomics.

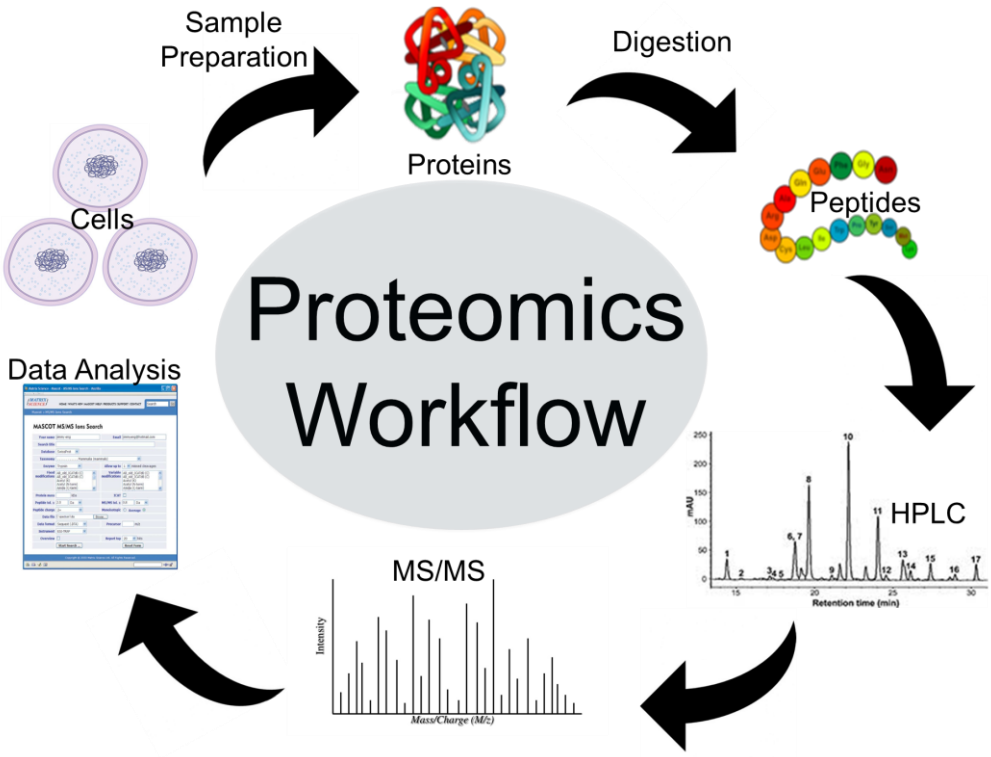


Figure 1.7: MS-based Proteomics workflow.(adapted)[15]

Being Proteomics currently such an enormous field, there is an array of techniques that can be used where some can depend on the final objective of the study.

Figure 1.8 illustrates a group of techniques used currently in Proteomics.

MS analysis is an important technique in Proteomics, and it has become the method of choice for complex protein sample analysis. This discipline was made possible by the existence of genome sequence databases and technical advances in many areas.[16]

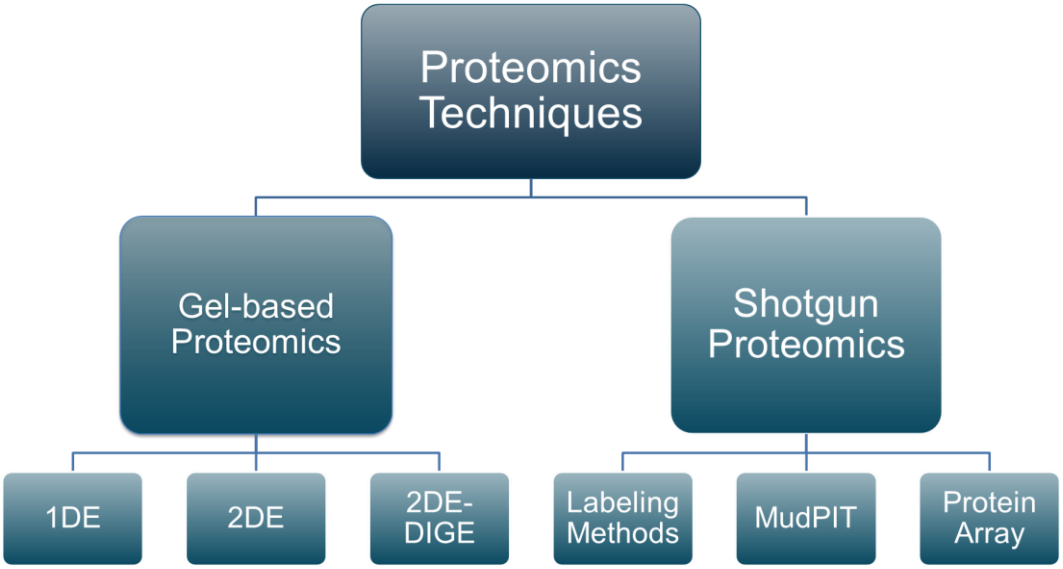


Figure 1.8: Techniques in Proteomics.(adapted)[14]

1.5 Gel-based Proteomics:

There are three gel-based techniques in Proteomics. These techniques involve the use of intact proteins during all stages of analyses. As suggested by the name this technique is performed on a gel made of polyacrylamide. These three techniques are very similar, but they have an increased power of separation. 1-dimensional electrophoresis (1D-GE), 2-dimensional electrophoresis (2D-GE) and 2 dimensional electrophoresis difference gel (2D-DIGE) are the gel-based techniques (Table 1.1).[14,17]

1D-GE separates proteins in groups by molecular weight, 2D-GE separates proteins not only by molecular weight but also by isoelectric point.[17,18,19]

2D-DIGE has a different form of gel that is capable of separating up to three protein samples, reducing intergel variability. Fluorophores are needed for visualization of proteins.[14,17]

Table 1.1: Gel-based techniques strengths and limitations.[14]

Technology	Application	Strengths	Limitations
2D-GE	Protein separation Quantitative expression profiling	Relative quantitative PTM information	Poor separation of acidic, basis, hydrophobic and low abundant proteins
2D-DIGE	Protein separation Quantitative expression profiling	Relative quantitative PTM information High sensitivity Reduction of variability	Proteins without lysine cannot be labelled. Expensive

1.6 Shotgun Proteomics:

Shotgun proteomics approaches use multi-dimensional capillary liquid chromatography combined with tandem mass spectrometry (MS/MS) to separate and identify the obtained peptides from the enzymatic digestion. It is important to notice that in shotgun proteomics it is not the protein that is separated. Instead of that protein are transformed into peptides by enzymatic digestion, and those peptides are separated and expose to MS/MS analysis. Once peptides are easier to separate by LC than proteins, shotgun proteomics are faster and cheaper than gel-based analysis.[20]

Shotgun proteomics has also isotopic labelling methods such as O18.[14]

1.7 Immobilized Trypsin:

Digestion is by far the most crucial step in protein identification and quantification.[21]

This concept refers to the enzymatic transformation of proteins into peptides. Although many enzymes can be used to perform protein digestion, trypsin is, by far, widely used.[22]

Trypsin cleaves peptide bonds at the carboxyl side of arginine and lysine. Therefore, this enzyme can produce a reproducible pool of peptides with an average size range between 600-2500 Da.[23]

Tryptic digestion can be performed in a heterogeneous or homogeneous phase. In the homogeneous phase, trypsin is in solution with the proteins. By contrast, the heterogeneous phase has the protein in solution, and the trypsin is immobilized onto a solid support. Immobilized trypsin prevents auto digestion of the enzyme avoiding downstream conflicts with MS analysis

and posterior protein identification. This technique also increases effective trypsin concentration which may result in shorter digestion times.[21,24]

Trypsin immobilization can occur in varied materials such as polymeric or metallic materials. Currently and independent of the material, these reactors are micro-sized.[25]

In this work, will be used a new and revolutionary type of immobilized trypsin. We will use iron nanoparticles with 80nm and magnetic properties where trypsin will be covalently attached to the nanoparticle. TEM image of the immobilized trypsin magnetic nanoparticles in Figure 1.9.

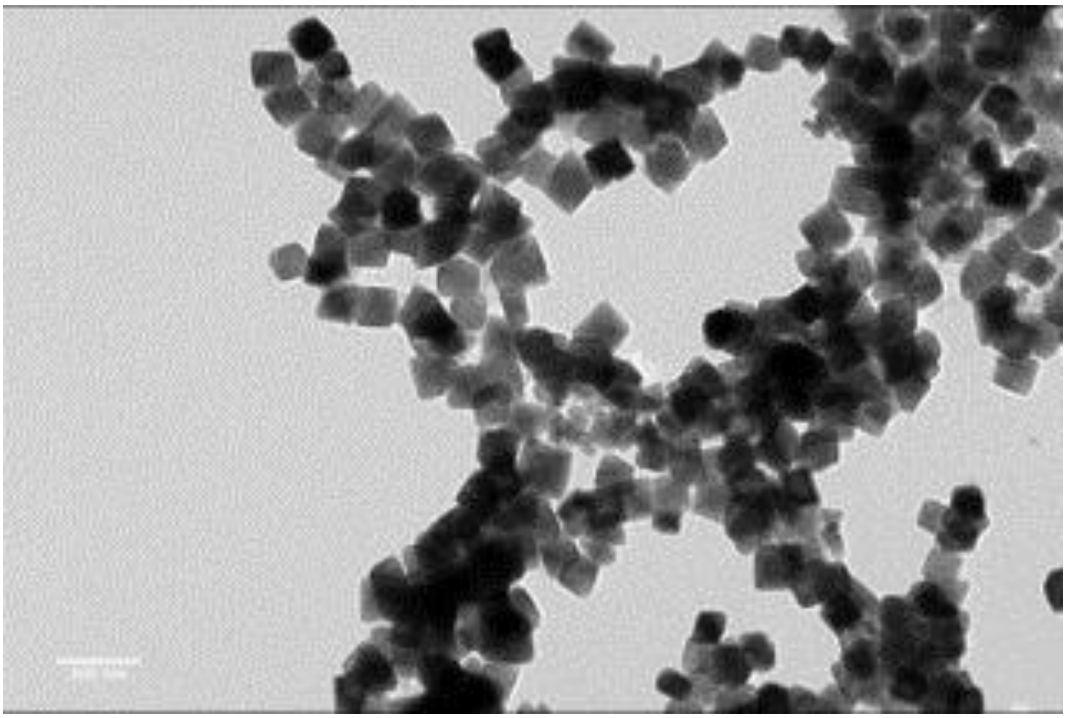


Figure 1.9: TEM picture of Immobilized Trypsin magnetic nanoparticles

1.8 Preconcentration of Phosphopeptides:

MS is currently the method of choice to detect changes in protein phosphorylation and to identify the position of specific phosphorylation events. However, even with the most recent advances in MS instrumentation, the detection and identification of phosphoproteins are compromised by a low ratio of phosphorylated and non-phosphorylated proteins. Only 1 to 2% of the entire protein amount is phosphorylated. Another issue is the phosphorylation cycles that may occur on a very short timescale.[12,26]

To be successful in such an endeavour, there is a prerequisite for an effective enrichment of phosphopeptides. New alternatives have been developed to overcome these issues.[13] Figure 1.10 shows the most common techniques employed in the enrichment of phosphopeptides.

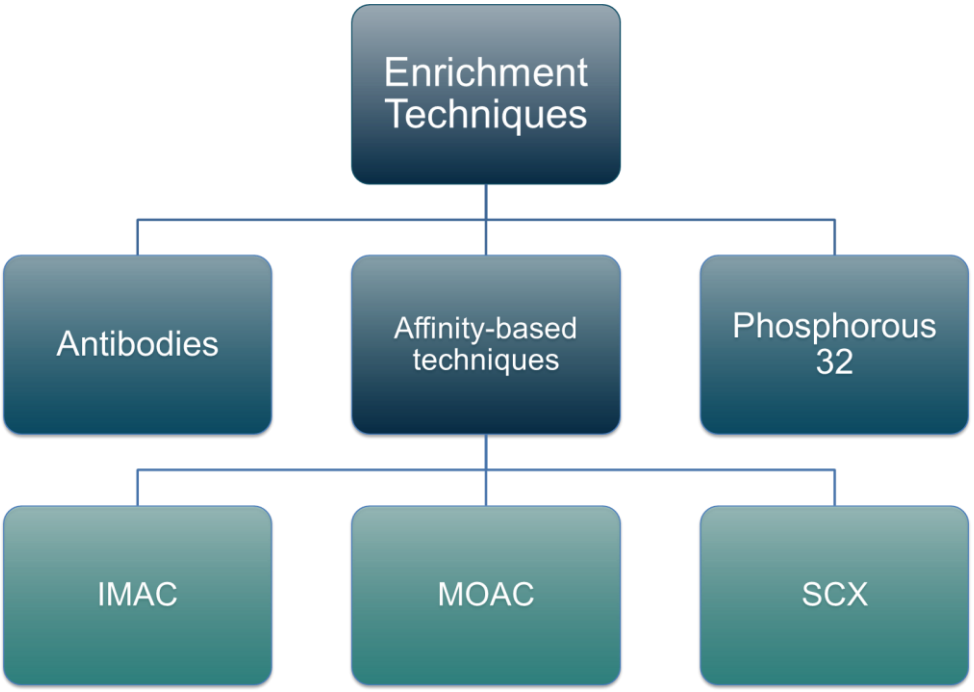


Figure 1.10: Common enrichment of phosphopeptides techniques.(adapted)[27]

Some alternative approaches have been performed to determine the phosphorylation stoichiometry. Conventionally, this has involved assessing the amount of 32P incorporation. However, this method is becoming less used nowadays due to safety constraints. Alternatively, protein phosphorylation can also be measured by MS. Typically, an enrichment step to pre-concentrate the low abundance phosphopeptides using immunoprecipitation, affinity purification, and strong cation exchange chromatography is required before MS analysis. However, these techniques are expensive, time-consuming, and a skilled operator is demanded to operate them. Immunoprecipitation techniques perform a decent enrichment however it has been proven to be highly applicable to samples containing peptides with phosphotyrosine.[28]

Immobilized metal affinity chromatography (IMAC) and metal oxide affinity chromatography (MOAC), were the first successful strategies developed for phosphopeptide enrichment. Both involve an immobilized metal ion or metal oxide, which is capable of coordination and has a high preference for phosphate groups. IMAC with Fe^{3+}, Ga^{3+}, or Ti^{4+}, and MOAC mostly with TiO_2 are nowadays the most-used enrichment methods for phosphopeptides.

Chemical coupling is a different approach for enrichment, in this technique phosphopeptides are covalently attached to a polymeric support. This requires multiple reactions steps and purification increasing the complexity of the technique and resulting in sample loss.[10,13]

Strong cation exchange (SCX) chromatography and IMAC are the most popular strategies. SCX is performed at very low pH (≈2.7) while phosphopeptides can remain negatively charged at this conditions allowing a major separation between phosphopeptides and nonphosphopeptides.[13] Other authors consider this technique as a prefractionation technique rather than a enrichment

technique.[27,29] While other authors refer that a combination of SCX followed by IMAC is the most successful technique.[10]

IMAC is considered to be the first truly successful technique for enrichment of phosphopeptides. This strategy involves an immobilized metal ion capable of coordinating specifically with phosphate groups due to its high preference for these groups. Iron and gallium metal ions are widely used. However, one of IMAC's limitations is the nonspecific adsorption resulting of nonphosphopeptides containing multiple acidic amino acids such as glutamate and aspartate.[13]

1.9 Mass Spectrometry:

MS uses mass analysis for protein identification, and it is the most popular and versatile technique for large-scale proteomics. MS measures the mass to charge ratio (m/z) of gas-phase ions. Thus a mass spectrometer equipment consists of an ion source, converting molecules into gas-phase ions, a mass analyser, separating ions based on m/z, and a detector recording the number of ions of each m/z.[30,31] Simple diagram in Figure 1.11.

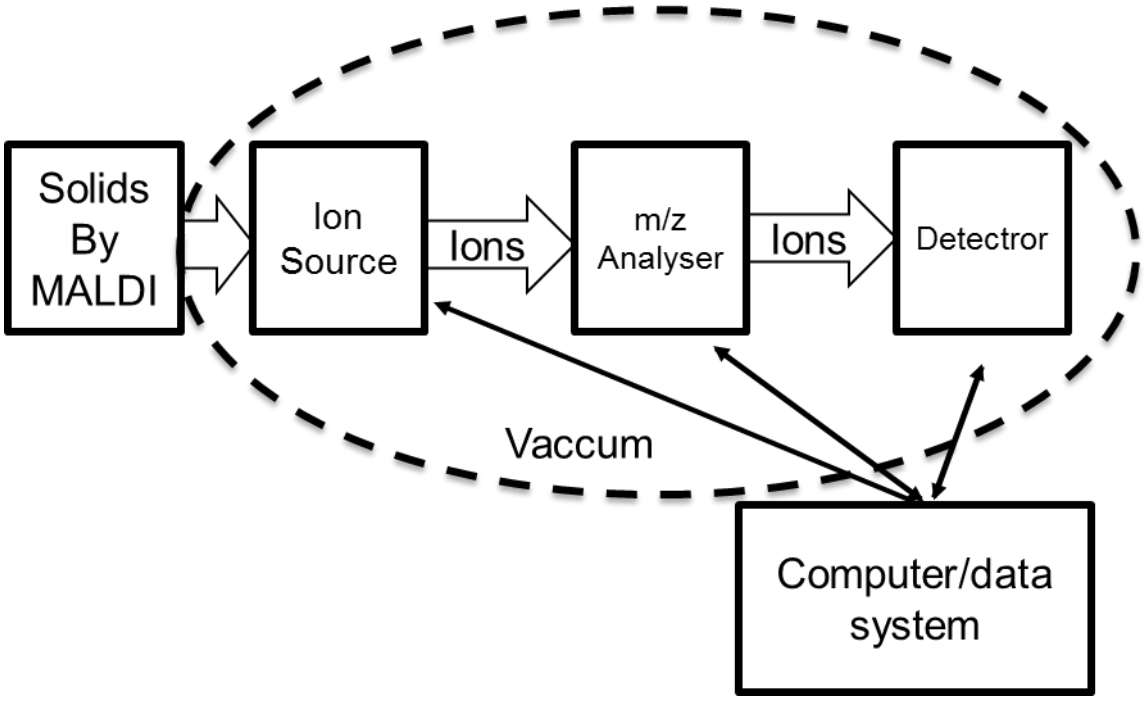


Figure 1.11: Simple diagram of a mass spectrometer.(adapted)[32]

In a mass spectrometer, a form of energy ionizes and fragments a molecule; next, this molecule is accelerated by an electromagnetic field separating the fragments according to their m/z after that a detector counts the number of fragments of each m/z. Using a proper software, a graphic of abundance vs m/z is presented.[32]

1.10 Soft Ionization Techniques:

Matrix-assisted laser desorption/ionization (MALDI) and Electrospray ionization (ESI) are the two preferred techniques to ionize molecules for prior mass spectrometric analysis. Before soft ionization techniques become available, MS was not considered to perform in biological sciences.[33] Common MALDI analytes are peptides, proteins and nucleotides; sample introduction is in a solid matrix. MALDI sublimates the dried samples out of a crystalline matrix through a laser beam, Figure 1.12.[16,30,34]

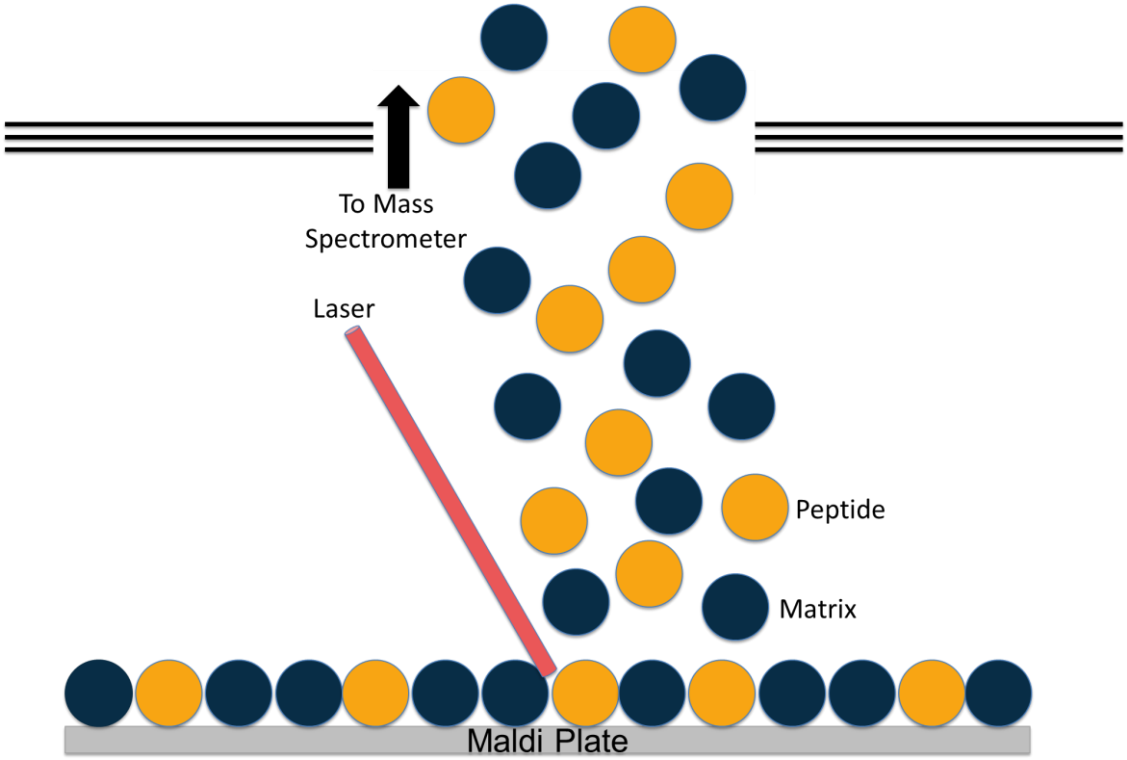


Figure 1.12: MALDI ionization technique.(adapted)[35]

Currently, MALDI is crucial in proteomics due to its sensitivity and simplicity besides this MALDI-TOF-MS (time of flight – TOF) accomplishes fast analysis generating amounts of date in a brief period. As stated before MALDI samples must be in matrixes, Figure 1.13 shows common matrixes used in MALDI.[36]

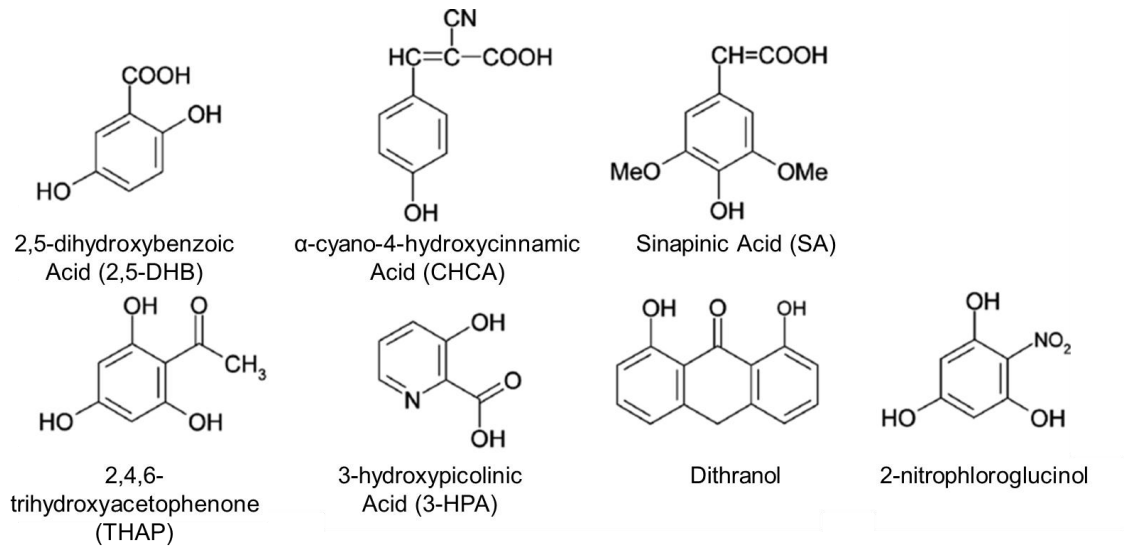


Figure 1.13: Common MALDI matrixes[36]

ESI is driven by high voltage (between 2 and 6 kV) applied between the emitter and the inlet of the mass spectrometer. ESI processes involve creation of electrically charged spray, trailed by formation and desolvation of sample droplets.[30,33] Samples are insert in a liquid form.[16,31]

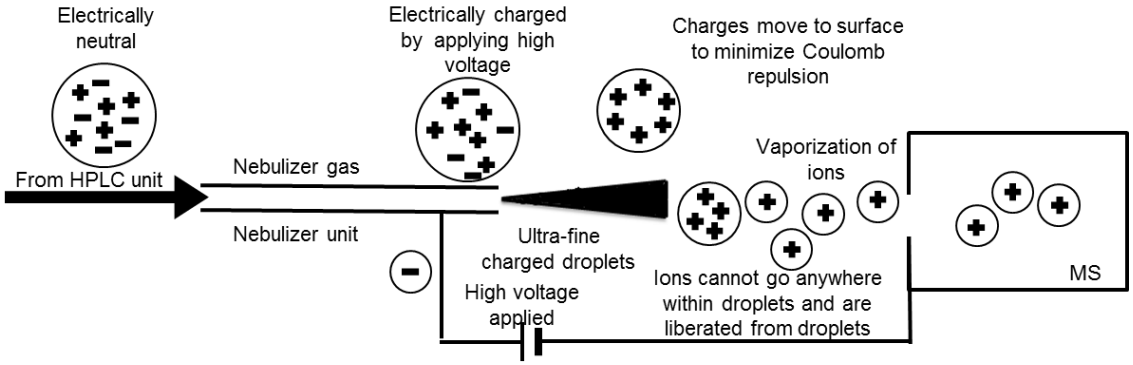


Figure 1.14: ESI ionization technique.[35]

1.11 Mass spectrometer analysers:

Mass analysers are a critical technology for MS. MS-based proteomics analysers are required to have high resolution, sensitivity and mass accuracy. There are four analysers commonly used in proteomics, ion trap, time of flight (TOF), quadrupole (Q) and orbitrap. These analysers can perform alone or together in MS/MS to take advantages of each analyser strength. Usually, MALDI is coupled with TOF, more recently TOF TOF or even triple quadrupole. ESI is coupled with ion trap or triple quadrupoles.[36]

TOF-TOF instruments incorporate a collision cell between the two TOF section. In the first TOF ions with a specific m/z are selected, next this selected are fragmented in the collision cell, in the second TOF the fragments created are separated.[16] Illustration in Figure 1.15.

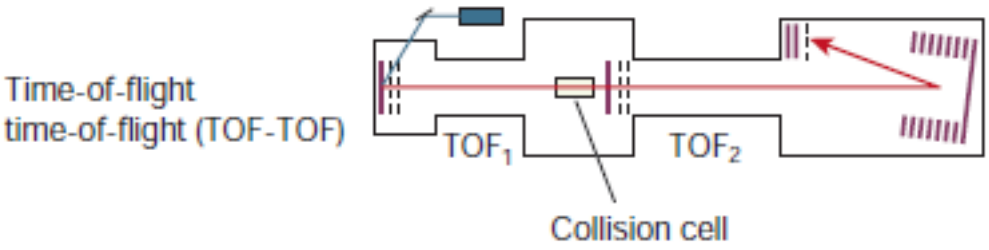


Figure 1.15: TOF-TOF analyser[16]

Ions of a selected m/z in a quadrupole mass spectrometer have a stable trajectory due to time-varying electric fields between four rods present in the equipment. Similar to TOF-TOF, in triple quadrupoles, ions of a specific m/z are selected in the first quadrupole, fragmented is the second quadrupole and separated on the third.[16] Illustration in Figure 1.16.

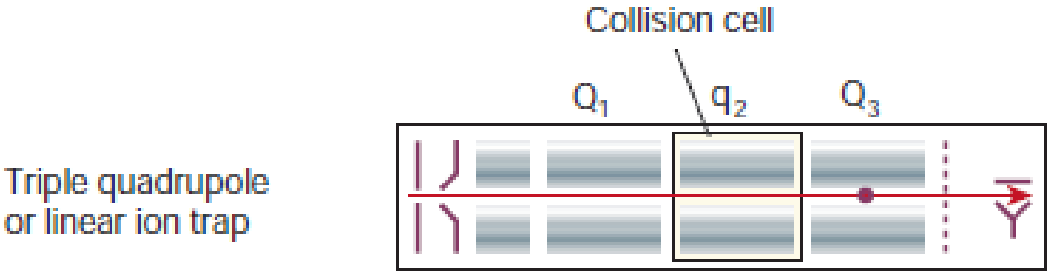


Figure 1.16: Triple quadrupole analyser[16]

The quadrupole TOF (Qq-TOF) analysers combine the front part of a triple quadrupole with the reflector TOF for measuring the m/z.[16] Illustration in Figure 1.17.

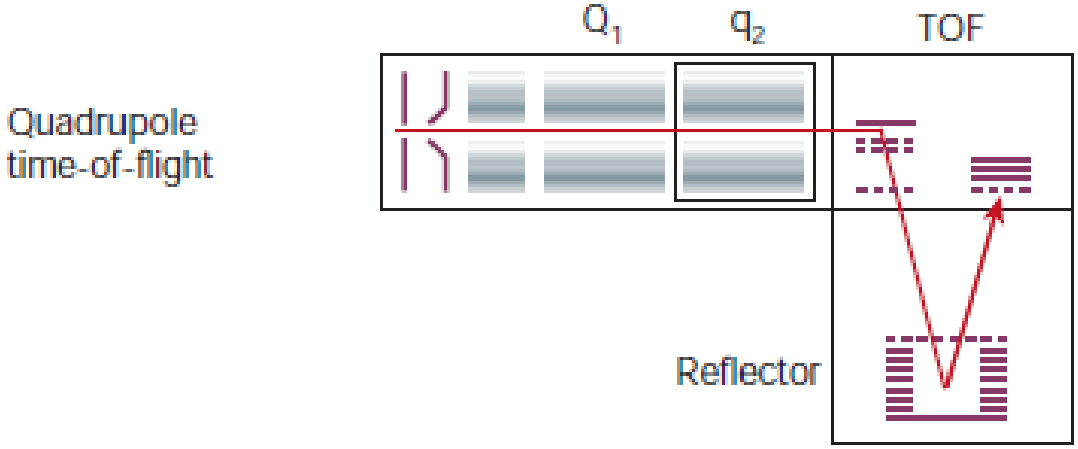


Figure 1.17: Qq-TOF analyser [16]

2 Objectives:

The work presented in this dissertation aims the development of a new nano-micro materials for mass spectrometry-based proteomics applications covering the key steps of protein digestion for identification of proteins by tandem MS as well as phosphopeptide enrichment for phosphoproteomics analysis. The topics covered by this research work include:

- Optimization of the conditions for robust and reproducible protein digestion using immobilized nano-trypsin.
- Synthesis of polystyrene nanoparticles using a 2^3 experimental design.
- Synthesis of a nano-micro immobilized lanthanide metal ion affinity material for phosphopeptide enrichment.
- Optimizing the experimental conditions for unbiased and reproducible phosphopeptide enrichment using new nano-micro immobilized lanthanide metal affinity chromatography.

3 Methods:

3.1 Preparation of a simple protein stock for tryptic digestion:

Reagents: Ammonium Bicarbonate (AmBic) 1M, Acetonitrile (ACN) (Carlo Erba Reagents), Milli-Q Water (MQH_2O), Bovine Serum Albumin (BSA) (Sigma-Aldrich), Ditiotreitol (DTT) (Nzytech) and Iodoacetamide (IAA) (Sigma-Aldrich)

Equipment: Vortex and Incubator

Procedure: In an Eppendorf dissolve 5mg of BSA in 1mL of MQH_2O. From this prepared stock transfer to another Eppendorf 100µg of Protein (20µL of solution) then add 2µL of DTT 110mM, agitate and incubate for 45 minutes at 37˚C. Add 2µL of IAA 400mM, agitate and incubate for 35 minutes at RT in the dark. To finish, add 476µL of AmBic 12.5mM 2% ACN and vortex. At this point, samples can be freeze for future use.

3.2 Preparation of E.Coli lysates stock for tryptic digestion:

Reagents: Urea, AmBic 1M, ACN, DTT, IAA and MQH_2O

Equipment: Centricon, Vortex, Centrifuge

Procedure: Transfer the lysates samples to a centricon to remove the buffer by centrifuge for 15 minutes at 6000rpm. Add 300µL of urea 3M dissolved in AmBic 12.5mM/2% ACN (pH ≈ 8.5) in the centricon and centrifuge for 15 minutes at 6000rpm. Next, add 100µL of urea 3M dissolved in AmBic 12.5mM/2% ACN and centrifuge for 15 minutes at 6000rpm and repeat. Remove the supernatant. This supernatant can be freeze for future analysis. Perform the Bradford technique to quantify the amount of protein in the samples. Transfer the desired amount of protein and add DTT 110mM to have 10mM of DTT in solution, vortex and incubate for 45 minutes at 37˚C. Add IAA 400mM to have in the final solution 33.33mM of IAA, vortex and incubate for 35 minutes at RT in a dark place. Add AmBic 12.5mM/2% ACN to reach a final volume of 500µL. Before tryptic digestion, urea must be removed from the solution using a Zip-Tip technique. Samples can be freeze for future use.

3.3 Bradford assay:

Reagents: BSA, MQH_2O and Bradford Reagent (Sigma-Aldrich).

Equipment: Vortex, ClarioStar (Spectrophotometer), 96-well plate

Procedure: Prepare a working solution containing 2µg/µL of BSA. Using a 96-well plate prepare the calibration curve in duplicates by loading 5µL of the solution prepared according to Table 3.1 and 250µL of Bradford reagent. To each well used, add 5µL of a sample with

appropriate dilution and 250µL of the Bradford reagent and mix. Measure the absorbance at 595 nm. The protein-dye complex is stable up to 60 minutes, assure your measure are before the time limit. Plot the net absorbance vs the protein concentration of each standard. Calculate the protein concentration of unknown samples by comparing the absorbance values against the standard curve (Table 3.1)[37,38].

Table 3.1: Calibration curve

Concentration (µg/µL)	Volume of BSA 2µg/µL (µL)	Volume of H_2O (µL)
0	0	200
0.2	20	180
0.4	40	160
0.6	60	140
0.8	80	120
1	100	100
1.2	120	80
1.4	140	60

3.4 Zip-Tip Technique:

Reagents: ACN, MQH_2O, Trifluoracetic acid (TFA) (Sigma-Aldrich), Formic Acid (FA) (Fluka Analytical) and Zip Tips (Thermo Scientific).

Equipment: Vortex

Procedure: First, aspirate 100µL of ACN and discard, repeat once. Aspirate 100µL of 0.1% TFA and discard, repeat. Aspirate 100µL of the sample and do up and down ten times and discard. Aspirate 100µL of 0.1% TFA/2% ACN and discard, repeat. Elute the proteins with 5µL of 0.1% FA/50% ACN. Elute again now using 5µL of 0.1% FA/90% ACN.[39]

3.5 Tryptic digestion of simple protein using immobilized trypsin nanoparticles:

Reagents: AmBic 1M, ACN, MQH_2O, BSA and Immobilized Trypsin nanoparticles.

Equipment: Vortex, Ultra Sonic Bath (US) and Incubator

Procedure: Sonicate the stock of immobilized trypsin for 10 minutes. Prepare the target immobilized trypsin concentration. In an Eppendorf mix 1, 5 and 10µg (5, 25 and 50µL) of BSA with 20µL of immobilized trypsin and complete the volume using AmBic 12.5mM/2% ACN for a final volume of 120µL. Incubate the samples overnight with gentle stirring at 37˚C. Remove the samples from the incubator, shake and spin down them. With the help of a magnet separate the

supernatant from the pellet transferring the supernatant to a new Eppendorf. Samples can be freeze for future analysis. Figure 3.1 shows a quick walk-through for this technique.

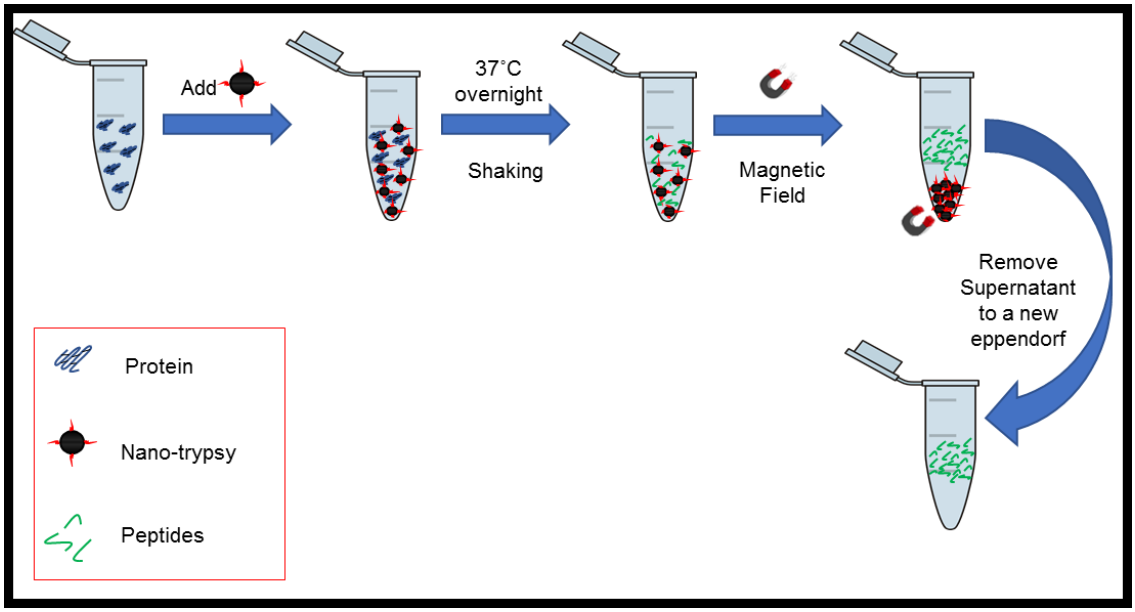


Figure 3.1: Quick walk-through for tryptic digestion using immobilized trypsin magnetic nanoparticles.

3.6 Tryptic digestion of simple protein using commercial immobilized trypsin microparticles:

Reagents: AmBic 1M, ACN, MQH_2O, BSA and Commercial Immobilized Trypsin (ClonTech Laboratories).

Equipment: Vortex, Ultrasonic Bath (US) and Incubator

Procedure: Sonicate the immobilized trypsin for 10 minutes. Remove the amount of trypsin desired and wash with MQH2O twice. Prepare the target immobilized trypsin concentration. In an Eppendorf mix 1, 5 and 10µg (5, 25 and 50µL) of BSA with 20µL of immobilized trypsin and complete the volume using AmBic 12.5mM/2% ACN for a final volume of 120µL. Incubate the samples overnight with gentle stirring at 37˚C. Remove the samples from the incubator, shake and spin down them. With the help of a magnet separate the supernatant from the pellet transferring the supernatant to a new Eppendorf. Samples can be freeze for future analysis.

3.7 MALDI analysis of simple protein digestion for protein identification:

Reagents: Ammonium dihydrogen phosphate (ADP) (Sigma-Aldrich) α-cyano-4-hydroxycinnamic acid (CHCA) (Fluka), ACN, MQH_2O, TFA, FA.

Equipment: Speed Vacuum, Incubator, Vortex

Procedure: Dry the peptidic samples in the speed vacuum. Resuspend the samples in 10 µL of 0.3% FA, vortex the samples, incubate for 15 minutes at 37˚C and vortex again. Prepare a stock solution by dissolving 10mg od ADP in 1 mL of MQH_2O Prepare the MALDI matrix by dissolving 7mg of CHCA in a mixture of 100µL of Stock solution, 500µ of ACN, 400µL of MQH_2O and 1µL of 0.1% TFA. In the MALDI target, placate 0.5µL of the sample and on the sample, placate 1µL of MALDI matrix. Samples in the MALDI target can be read after dried. What remains of samples can be freeze for future analysis.

3.8 ESI LC analysis proteins:

Reagents: FA, ACN,

Equipment: Easy-nLC II

Procedure: Before ESI MS/MS analysis all samples were diluted with 100 µL of 0.1% (v/v) aqueous FA before loading onto an EASY-nLC II equipped with an EASY-Column, 2cm, ID100µm, 5µm, C18-A1 (Thermo Fisher Scientific) and an EASY-Column, 10cm, ID75µm, 3µm, C18-A2 (Thermo Fisher Scientific). Chromatographic separation was carried out using a multistep linear gradient at 300 nL/min (mobile phase A: aqueous formic acid 0.1% (v/v); mobile phase B 90% (v/v) ACN and 0.1% (v/v) FA) 0-90 min linear gradient from 0% to 35% of mobile phase B, 90-100 min linear gradient from 35% to 95% of mobile phase B. For each sample two replicate injections were performed.

MS acquisition was set to cycles of MS (2 Hz), followed by MS/MS (8–32Hz), cycle time 3.0 seconds, active exclusion, exclude after one spectrum, release after 0.5 min. Reconsider precursor if current intensity, previous intensity 3.0 an intensity threshold for fragmentation of 2000 counts. All spectra were acquired in the range 150–2200 Da. LC-MS/MS data were analysed using Data Analysis 4.2 software (Bruker).

Proteins were identified using Mascot (Matrix Science, UK). MS/MS spectra were searched against the SwissProt database 57.15 (515,203 sequences; 181,334,896 residues), setting the taxonomy to *E.coli* (22,646 sequences). Tandem MS data were searched with MASCOT search engine with the following parameters: precursor mass tolerance of 20 ppm, fragment tolerance of 0.05 Da, trypsin specificity with a maximum of 2 missed cleavages, cysteine carbamidomethylation set as fixed modification and methionine oxidation, as variable modification. Significance threshold for the identification was set to $p < 0.05$ and false discovery rate (FDR) was estimated by running the searches against a randomized decoy database. Results of the identification step were filtered to proteins with a FDR below 1%.[40]

Proteins were identified using Mascot (Matrix Science, UK). MS/MS spectra were searched against the SwissProt database 57.15 (515,203 sequences; 181,334,896 residues), setting the taxonomy to Other Mammalia (12,633 sequences). Tandem MS data were searched with MASCOT search engine with the following parameters: precursor mass tolerance of 20 ppm,

fragment tolerance of 0.05 Da, trypsin specificity with a maximum of 1 missed cleavage, cysteine carbamidomethylation set as fixed modification and methionine oxidation, serine, threonine and tyrosine phosphorylation as variable modification. Significance threshold for the identification was set to $p < 0.05$ and false discovery rate (FDR) was estimated by running the searches against a randomized decoy database. Results of the identification step were filtered to proteins with a FDR below 1%.[40]

3.9 Polyacrylamide gel

Reagents: Solution I, Solution II, Solution III, SDS 10%, Butanol 50% (Sigma-Aldrich), MQH_2O, APS and TMED.

Equipment: Gel supporter

Procedure:

Table 3.2: Polyacrylamide Gel Protocol.[41]

Stock Solution	Stacking Gel	Running Gel	
% acrylamide	4	12	
Solution I*	-	2.5	mL
Solution II*	1	-	
Solution III (acrylamide/bisacrylamide) (37.5:1)	0.52	4	
SDS 10%	0.04	0.1	
H2O Milli-Q	2.48	3.4	
APS 10 %	30	70	µL
TMED	2	5	

*Solution I: Tris-Base 27.23g, add HCl until pH=8.8 and MQH_2O until 150mL; Solution II: Tris-Base 6.06g, add HCl until pH=6.8 and MQH_2O until 100mL

To produce the stacking gel and running gel mix the above quantities in a centrifugal tube. First, produce the running gel, place it in the gel supporter, add a 50% butanol solution to create a plane surface avoiding air entrance and wait for polymerization to occur. Next prepare the stacking gel and place it over the polymerized running gel, add the well comb and wait for polymerization. Keep in mind that polymerization time is about 15 to 30 minutes and it starts after mixing APS and TMED together. Remove the comb, and the gel is prepared.[41]

3.10 1D Gel Electrophoresis:

Reagents: 12.5% polyacrylamide gel

Equipment: Gel supporters and electrodes.

Procedure: After sample clean up, protein samples were re-suspended in 10µL of 1x Laemmli sample buffer and then heated in a dry bath at 100ºC for 5 minutes. The denatured proteins were loaded on 12.5% polyacrylamide gels with 1mm thickness. Proteins were

separated at 200 V(constant voltage) until the tracking dye front reaches the bottom of the gel.[42,43]

3.11 Gel staining and image analysis

Reagents: Ethanol (Carlo Erba Reagents), acetic acid (Panreac), Coomassie blue G-250m distilled H_2O and NaCl 0.5M

Equipment: No specific equipment

Procedure: Finished the gel electrophoresis, the gel was fixed for 30 minutes with 40% (v/v) ethanol and 7.5% (v/v) acetic acid and then stained overnight with colloidal Coomassie blueG-250. Gels were rinsed 4×20 min with 100mL of distilled water and further washed twice with 100 mL of 0.5 M sodium chloride until a clear background was observed. Gel imaging was carried out with a ProPicII–robot using 16ms of exposure time and a resolution of 70µm.[44]

3.12 Synthesis of polystyrene nanoparticles:

Reagents: Ammonium Persulfate (APS) (Sigma-Aldrich), Sodium Dodecyl Sulphate (SDS) (Panreac), MQH_2O, Styrene and 1-penthanol.

Equipment: Heating mantle, thermometer, mini pump, vortex, US bath, Round-bottom flasks, centrifuge tubes and speed vacuum.

Procedure: Dissolve APS and SDS in MQH_2O. Sonicate the mixture to assure a good and fast dissolution. Transfer the solution to a round-bottom flask and increase the temperature to 80˚C while stirring at 1000rpm. After reach 80˚C start adding, at constant rate, a mix of styrene and 1-penthanol for 30 minutes. After all styrene been added, maintain the temperature between 80 and 85˚C for 1 hour. Filtrate the obtained solution through a 220nm pore filter to a centrifugal tube. To quantify the amount of seeds, present in solution, assure a good homogenization and remove and dry 1mL of the solution, weight the precipitate keeping in mind that the precipitate has Seeds and SDS. As soon as possible clean the flask with xylene, dried styrene can be a hard to remove.

Quantities of APS, SDS, MQH_2O and Styrene vary according to the experience performed, Table 3.3 provides quantity of reagents for each experience.[45]

Table 3.3: Amount of reagents for each experience.

Experiment	APS (g)	SDS (g)	Styrene (mL)	1-Penthanol (mL)	MQH2O (mL)
1	0.004	0.2	0.7	0.01	48.3
2	0.04	0.2	0.7	0.01	48.3
3	0.004	2	0.7	0.01	48.3
4	0.04	2	0.7	0.01	48.3
5	0.004	0.2	7	0.1	42
6	0.04	0.2	7	0.1	42
7	0.004	2	7	0.1	42
8	0.04	2	7	0.1	42

3.13 Synthesis of monodisperse nano spheres-based immobilized lanthanides ion affinity chromatography:

Reagents: Polystyrene nano spherical seeds (Seeds), Polyvinyl Alcohol (PVA) (Sigma-Aldrich), Glycidyl Methacrylate (GMA) (Sigma-Aldrich), Trimethylolpropane Trimethacrylate (TMPTMA) (Sigma-Aldrich), 2,2'-Azobis(2-methyl-propionitrile) (AIBN) (Sigma-Aldrich), Toluene (Panreac), Tetrahydrofuran (THF) (Carlo Erba Reagents), Acetone (Sigma-Aldrich), Ethylenediamine (Scharlau), Ethanol, MQH_2O, ACN, TFA, Phosphoric acid (Alfa Aesar), HCl (), Formaldehyde (Panreac) Titanium Chloride (Sigma-Aldrich) and Lanthanum Chloride Heptahydrate (Sigma-Aldrich).

Equipment: Round-bottom flasks, heating mantle, condensation column, thermometer, centrifugal tubes, rotavapor and centrifuge.

Procedure: Prepare a 15mL solution containing 450mg of seeds and 1% (w/w) PVA and 0.25% (w/w) SDS, sonicate the solution to homogenise. Prepare a 150mL solution containing 1% (w/w) PVA and 0.25% (w/w) SDS. Prepare an oil-phase solution adding 6.7mL of GMA, 6.7mL of TMPTMA, 140mg of AIBN and 16.6mL of toluene. Add the oil-phase solution with 150mL of the seedless solution and sonicate for 10 minutes or more assuring that an emulsion is created. Add this new solution to the seeds in a round-bottom flask. Start stirring at 1200rpm. Increase the temperature to 30 ˚C. Maintain this condition for 20 hours to perform seed swelling. Increase the temperature now to 70˚C to start the polymerization, keep the stirring. Reaction time is 24 hours. Transfer the obtained solution to centrifugal tubes and wash with 15mL of THF and 15mL of acetone for each tube. Repeat the washes five times. Dry the solution in the rotavapor. Add 7g of this new dried compound with 150mL of ethylenediamine in a round-bottom flask. The reaction is conducted at 80˚C, 1200 rpm for 3 hours. Transfer the obtained solution to centrifugal tubes and wash with 30mL of 50% Ethanol, repeat the washes five times. Dry the solution in the rotavapor.

Add 7g of the last compound obtained with 5.1mL of phosphoric acid, 10mL of 37%HCl and 8mL of formaldehyde successively. Increase the temperature to 100 ˚C and stir at 1200rpm. Maintain these conditions for 24 hours. Transfer the solution to centrifugal tubes and wash each one with 30mL of 50% Ethanol, wash five times. Dry the solution in the rotavapor. The compound produced can be stored at RT for several months avoiding the light. Incubate 100mg of this compound with 20mL of a solution of the chosen metal (0.09M in 20% HCl). Stir at 1200rpm for 8 hours at RT. Transfer the solution to centrifugal tubes and wash with 10mL of 30% ACN and 0.1% TFA, wash five times.[13] Reaction scheme in Figure 3.2.

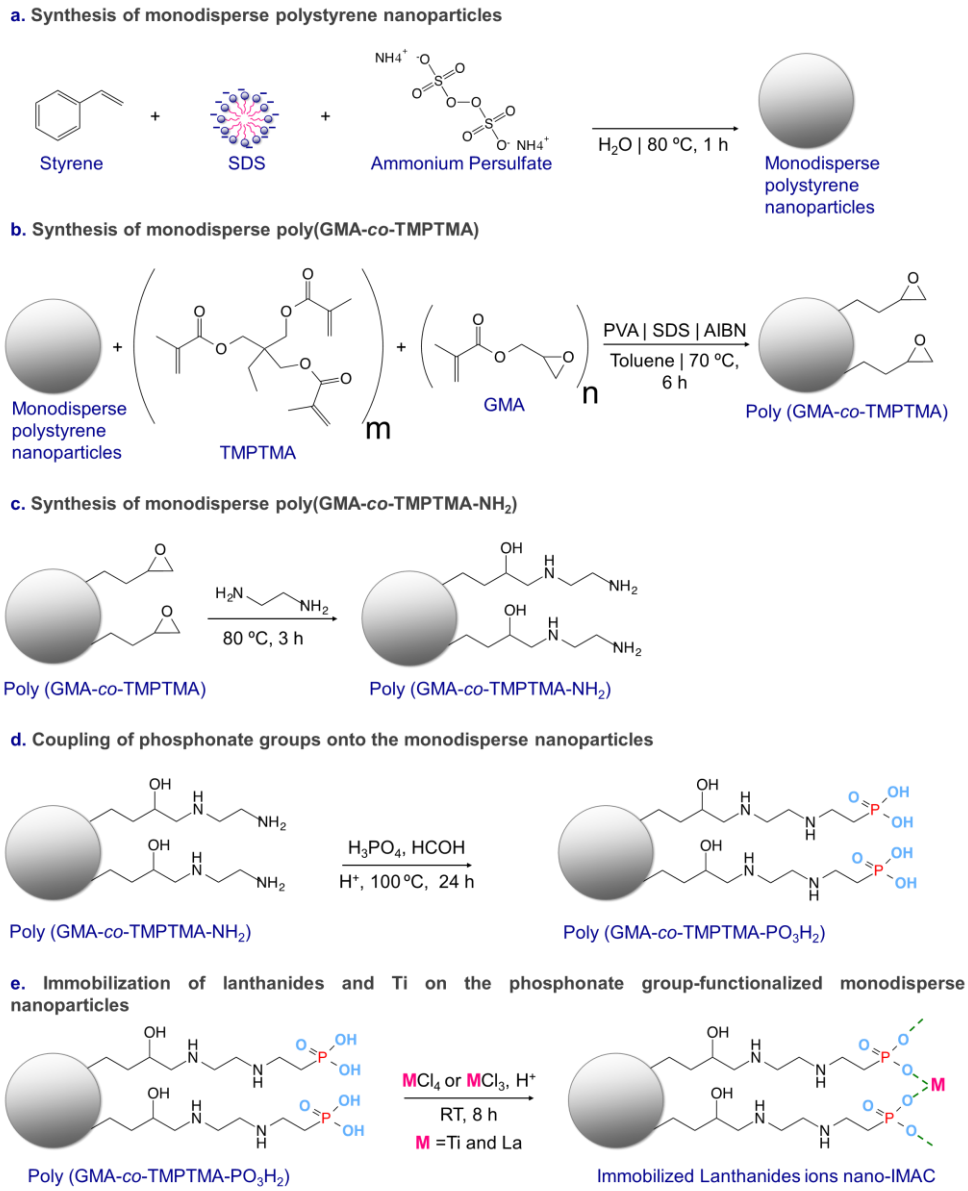


Figure 3.2: Scheme of IMAC synthesis

3.14 Enrichment of phosphopeptides using an IMAC column:

Reagents: Methanol (Carlo Erba Reagents), FA, Loading Buffer 1 (80% ACN/6%TFA) (pH ≈ 1), Washing Buffer 1 (50% ACN/0.5% TFA/200mM NaCl) (pH ≈ 2), Washing Buffer 2 (50% ACN/0.1% TFA) (pH ≈ 2), Elution Buffer 1 (10% Ammonium) (pH ≈ 12) and Elution Buffer 2 (80% ACN/2% FA) (pH ≈ 3).

Equipment: C8 (Supelco), vortex and centrifuge.

Procedure: Place a C8 membrane in a 10µL tip and wash with 20µL of methanol. Add multiples of 50µL of IMAC with a concentration of 10mg/mL until reach the target amount. For each time of IMAC added centrifuge at 200G for 2.5 minutes. Equilibrate the IMAC with 50µL of Loading Buffer, centrifuge at 200G for 2.5 minutes and repeat this step. Add 100µL of sample and centrifuge at 20G for 6 minutes. Wash the IMAC column with 50µL of Washing Buffer 1, centrifuge at 100G for 4 minutes. Wash the IMAC a second time using 50µL of Washing Buffer 2, centrifuge at 100G for 4 minutes. In a new Eppendorf add 35µL of 10% FA and elute the phosphopeptides with 20µL of Elution Buffer 1, centrifuge at 100G for 3 minutes. Preform the second elution using 20µL of Elution Buffer 2, centrifuge at 100G for 3 minutes. Add 3µL of FA to acidify the sample, perform a Zip-Tip to remove salts and impurities from the sample. Samples can be freeze for further use.[13]

4 Results and Discussion:

4.1 Protein digestion with immobilized trypsin:

To perform protein digestion, trypsin, was the chosen enzyme. As already stated trypsin is the desired enzyme due to its specificity. Trypsin was immobilized in a nanoparticle (80nm) with an iron core that confers magnetic properties to this particle. Since this immobilized trypsin device is brand new and has never been tested, it is imperative to evaluate its performance. In the first phase, it was important to discover how the device responded in general conditions since the amount of trypsin per mg of nanoparticle was unknown. Therefore, a large array of protein and immobilized trypsin concentrations were tested. BSA was used as a reference because it is a simple protein with a molecular weight of 66 kDa[46], which is easy to digest. Digestion success will be evaluated by sequence coverage percentage, a higher sequence coverage implies a good digestion. To assess the efficiency of this immobilized trypsin, BSA was digested in the following conditions: 4 different concentrations of trypsin (0.005, 0.01, 0.1 and 0.5 mg/mL) and three different amounts of protein (1, 5 and ten µg). The ratios of particles of immobilized enzyme to protein are between 1:0.02 and 1:17. The ratio of enzyme-protein for in-solution digestion with trypsin is usually 1:20. According to the provided procedure, commercial immobilized trypsin microparticles should be used at a concentration of 5% (w/w). Results in Figure 4.1 and important MS spectra in Figures 4.2 and 4.3.

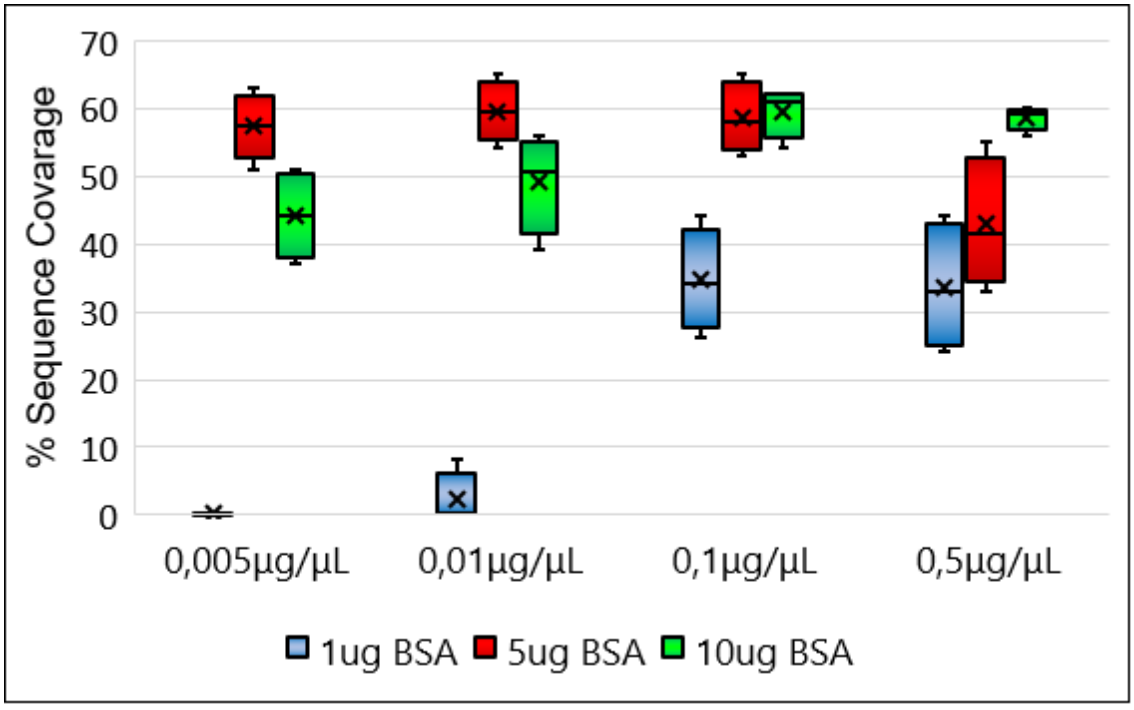


Figure 4.1: Sequence Coverage of the digestion of 1, 5 and 10 µg of BSA using 0.005, 0.01, 0.1 and 0.5µg/µL of Immobilized trypsin nanoparticles.

Protein identification was successfully achieved in these range of protein and enzyme concentrations, except when the lowest concentrations of protein and enzyme are combined (1µg of protein and 0.005µg/µL of nanoparticle), extreme conditions are rarely used, therefore, this will not be considered as a failure, but as the limitation of the system. Also, sequence coverage above 60% proves that digestion performance is quite good. Results also highlight the concentration of nanoparticles which assure better digestion and, therefore, a higher percentage of sequence coverages and possibly better identifications, ratios of immobilized trypsin particles to protein of 1:4 or 1:1 have higher percentages of sequence coverage. When 0.005µg/µL and 1µg of BSA were used, no identification was made, with 0.01µg/µL and 1µg of BSA only one identification was made. In all the remain conditions four identifications were possible to make; four replicates were make.

Other mass spectra of this experiment can be found in Annex I

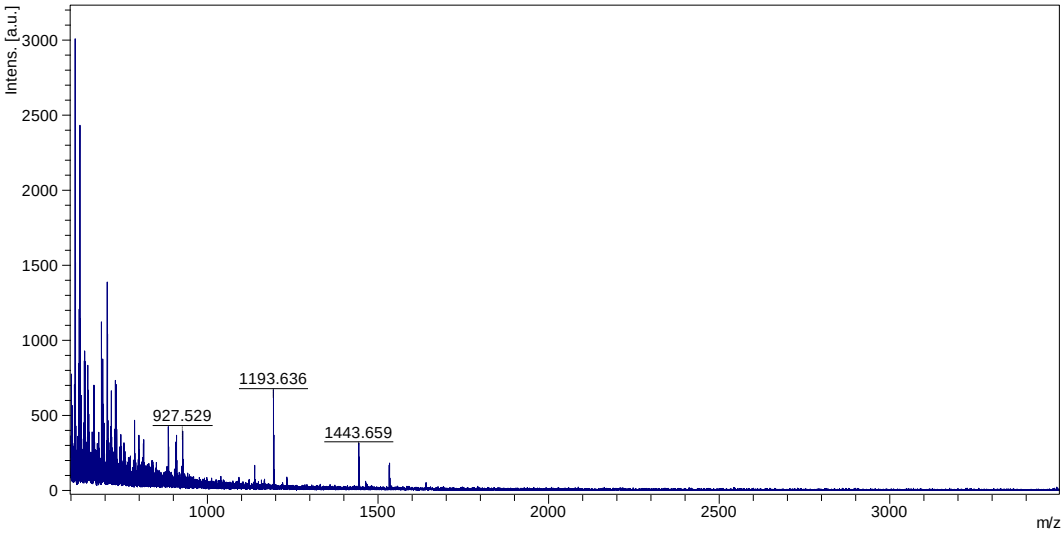


Figure 4.2: MS spectrum of the digestion of 1µg of BSA with 0.005µg/µL of immobilized trypsin nanoparticles.

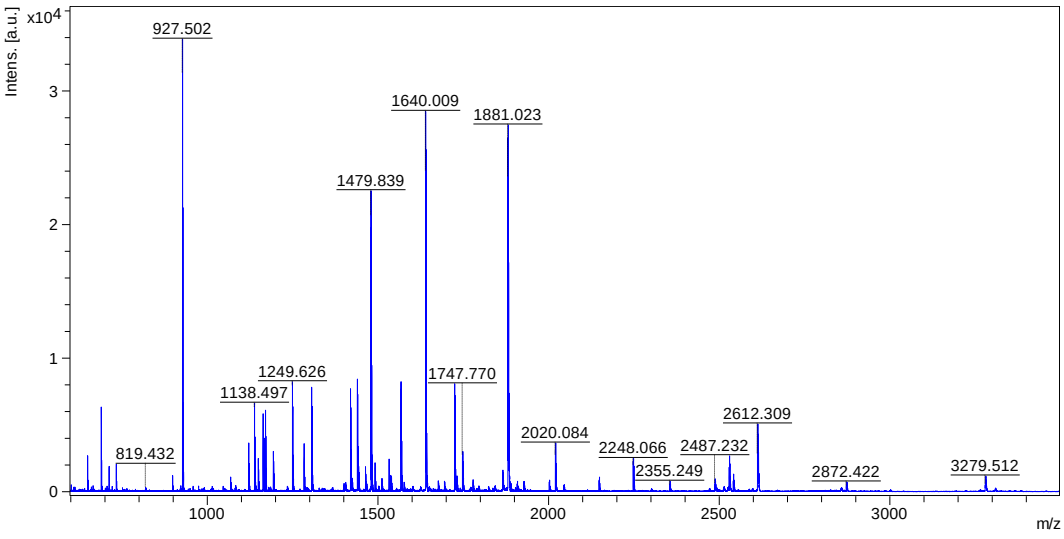


Figure 4.3: MS spectrum of the digestion of 5µg of BSA with 0.1µg/µL of immobilized trypsin nanoparticles.

Although this system works, it was still needed to assure its reproducibility.

In the following experimentation, four batches of immobilized trypsin were prepared. To test the reproducibility of the device, two different concentrations of immobilized trypsin nanoparticles and two separate amounts of protein were used and, therefore, four different enzyme-protein ratios were tested. For each condition, four replicates were made. Results in Figures 4.4, 4.9, 4.10 and 4.11 and significant MS spectra in Figures 4.5, 4.6, 4.7 and 4.8.

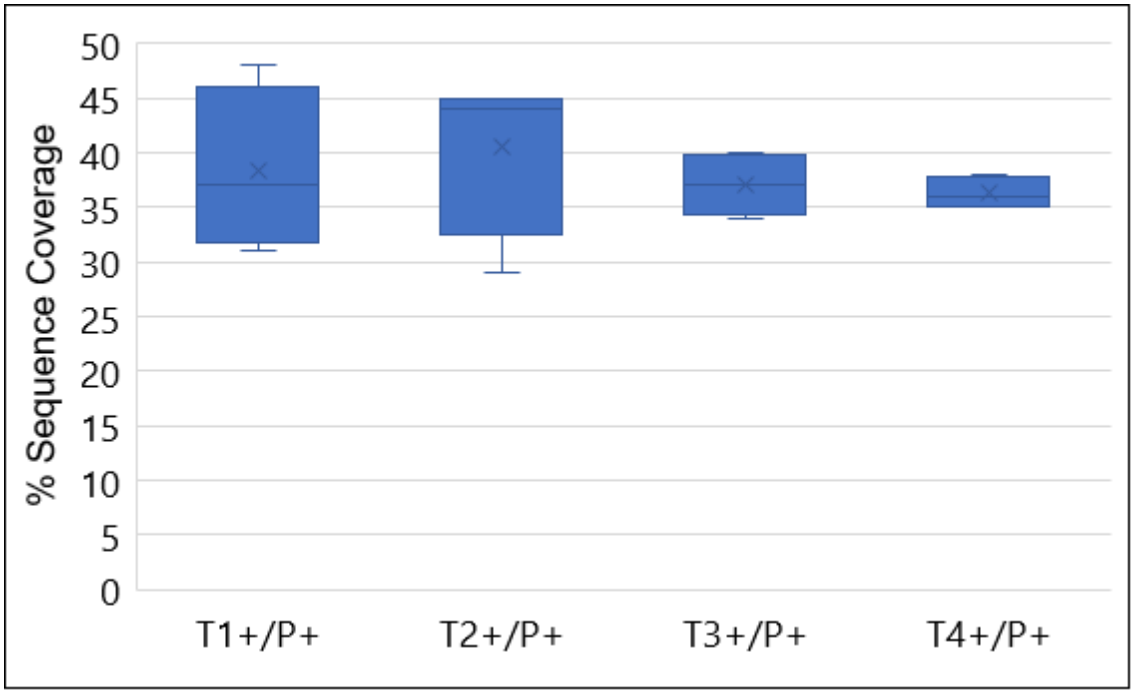


Figure 4.4: % Sequence Coverage of the digestion of 10µg of BSA with 0.5µg/µL of 4 different batches of immobilized trypsin nanoparticles.

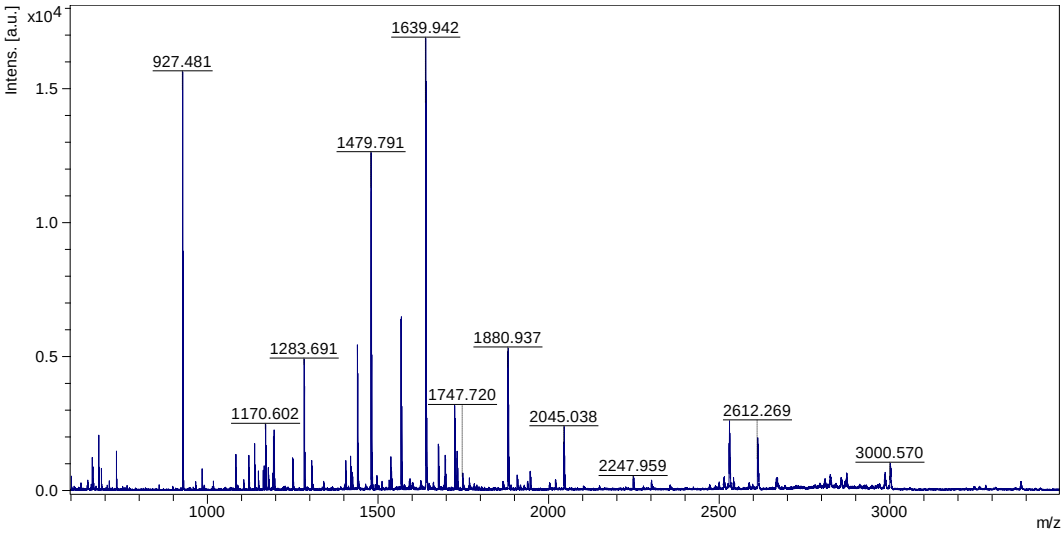


Figure 4.5: MS spectrum of the digestion of 10µg of BSA with 0.5µg/µL of immobilized trypsin nanoparticles – batch 1.

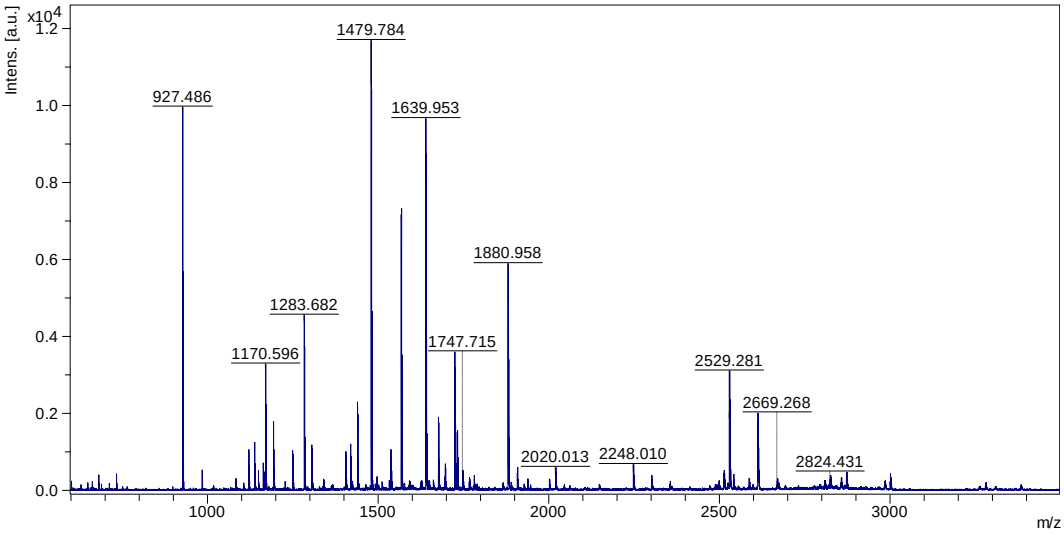


Figure 4.6: MS spectrum of the digestion of 10µg of BSA with 0.5µg/µL of immobilized trypsin nanoparticles – batch 2.

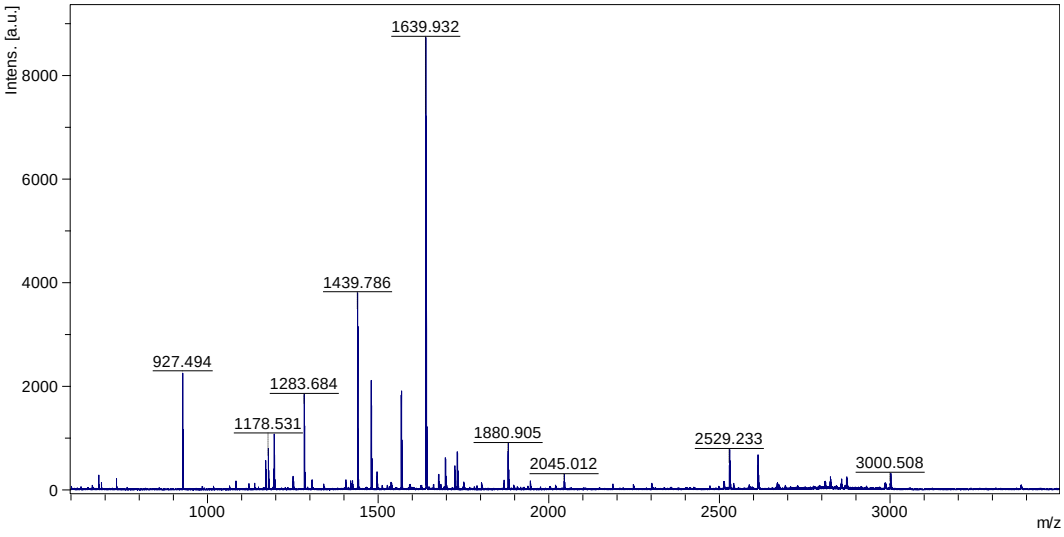


Figure 4.7: MS spectrum of the digestion of 10µg of BSA with 0.5µg/µL of immobilized trypsin nanoparticles – batch 3.

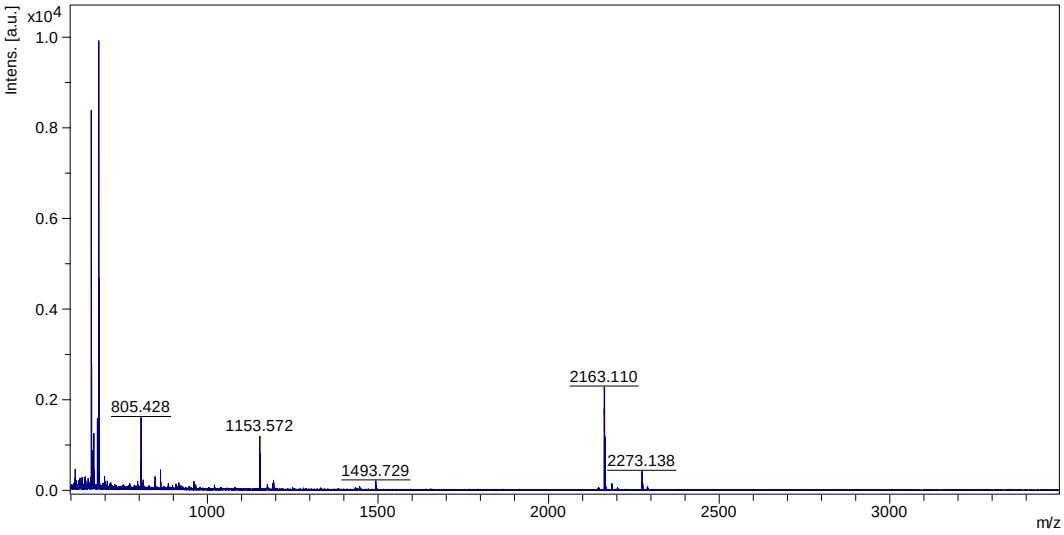


Figure 4.8: MS spectrum of the digestion of 10µg of BSA with 0.5µg/µL of immobilized trypsin nanoparticles – batch 4.

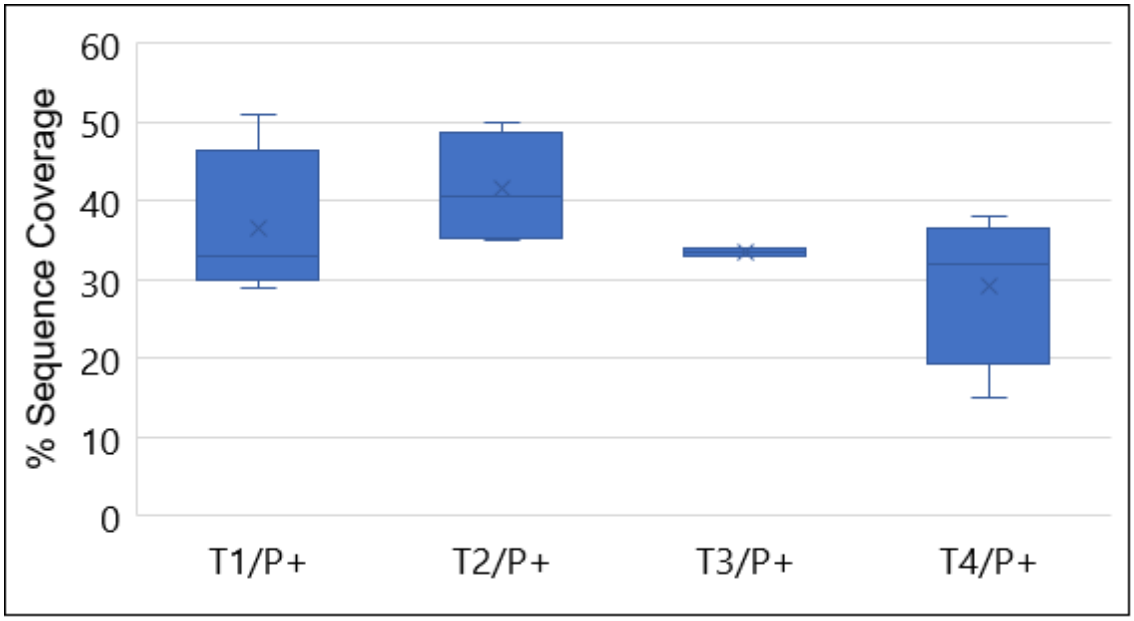


Figure 4.9: % Sequence Coverage of the digestion of 10µg of BSA with 0.1µg/µL of 4 different batches of immobilized trypsin nanoparticles.

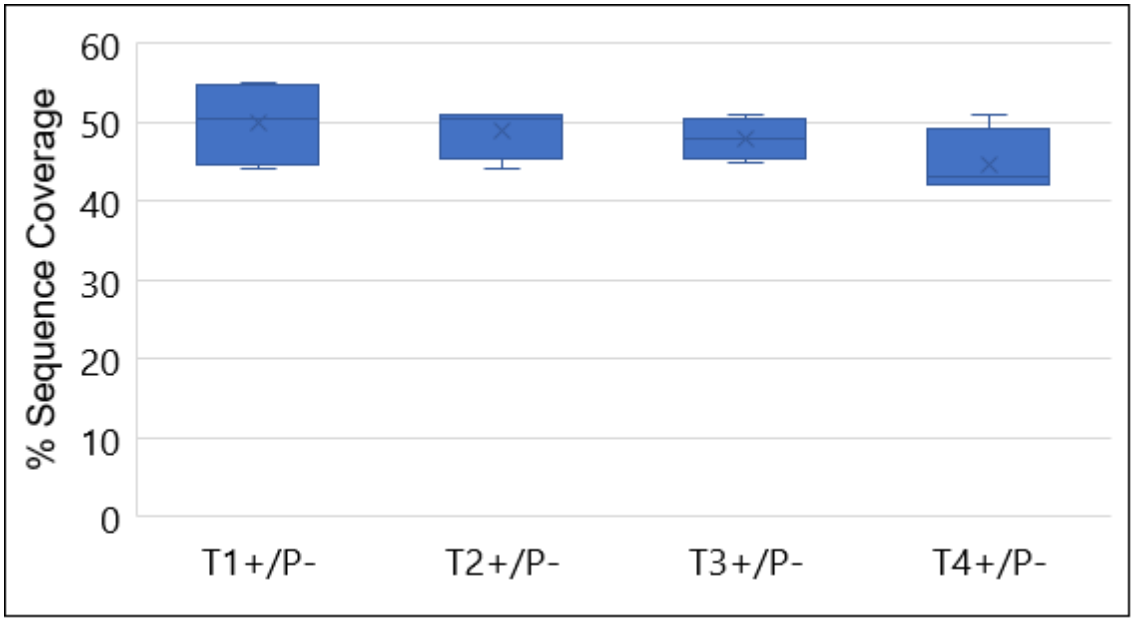


Figure 4.10: % Sequence Coverage of the digestion of 1µg of BSA with 0.5µg/µL of 4 different batches of immobilized trypsin nanoparticles.

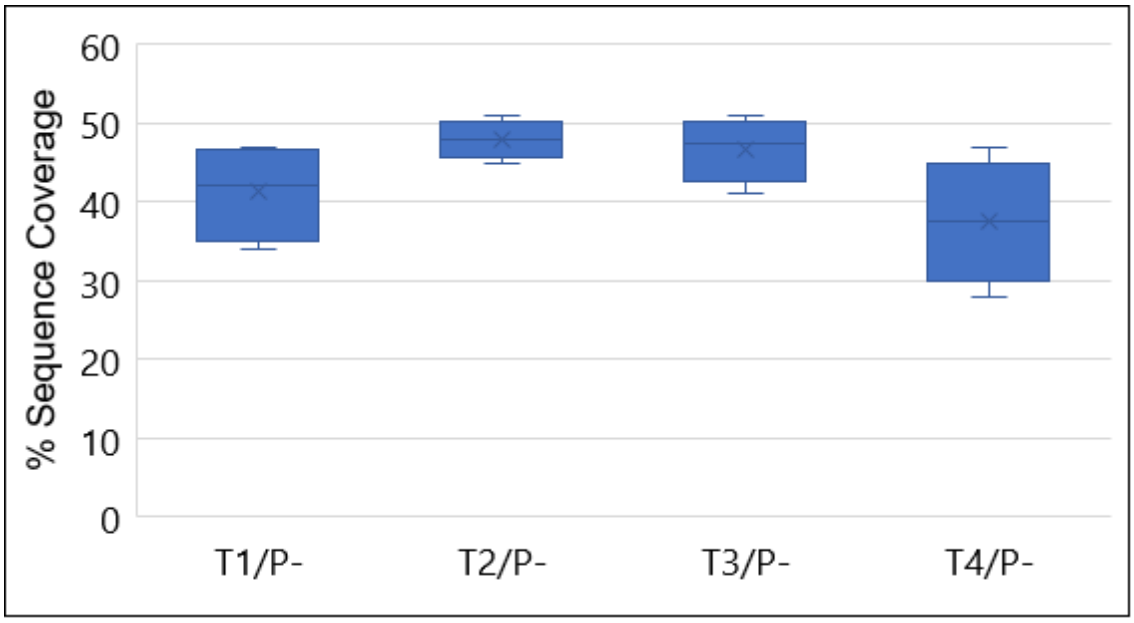


Figure 4.11: % Sequence Coverage of the digestion of 1µg of BSA with 0.1µg/µL of 4 different batches of immobilized trypsin nanoparticles.

Other mass spectra of this experiment can be found in Annex II

To ensure similarity between all the different batches, t-tests were performed for all batches and all enzyme-protein ratios. To decide whether the difference between two means, X1 and X2, is significant, to test the null hypothesis where H0: u1=u2 the statistic t is then calculated from Equation 1 (Annex III).

If the standard deviations, s1 and s2 are not significantly different, this assumption must be tested, s can be calculated using Equation 2 (Annex III).

To check if s1 and s2 are not significantly different, F test must be tested using Equation 3 (Annex III).

If F exceeds a critical value, s1 and s2 are significantly different, and therefore equation 1 cannot be used instead Equation 4 (Annex III) must be employed.

In Annex III results are displayed. All the batches are equal between them at a level of significance of 5%.[47]

Now it is possible to validate the efficiency and reproducibility of these immobilized trypsin nanoparticles.

Taking into consideration that the ratio of enzyme-protein has already been optimized, this ratio will be used for all the subsequent experiments.

However, in an attempt to increase sequence coverage and digestion speed, other variables were tested. In the following experience, stirring while samples are incubated overnight was analysed. Stirring improves the reaction giving motion to the immobilized trypsin nanoparticles. Otherwise, they would deposit, which would reduce the surface of contact and therefore the digestion efficiency as well. Results in Figures 4.12 and 4.13 and important MS spectra in Figures 4.14 and 4.15.

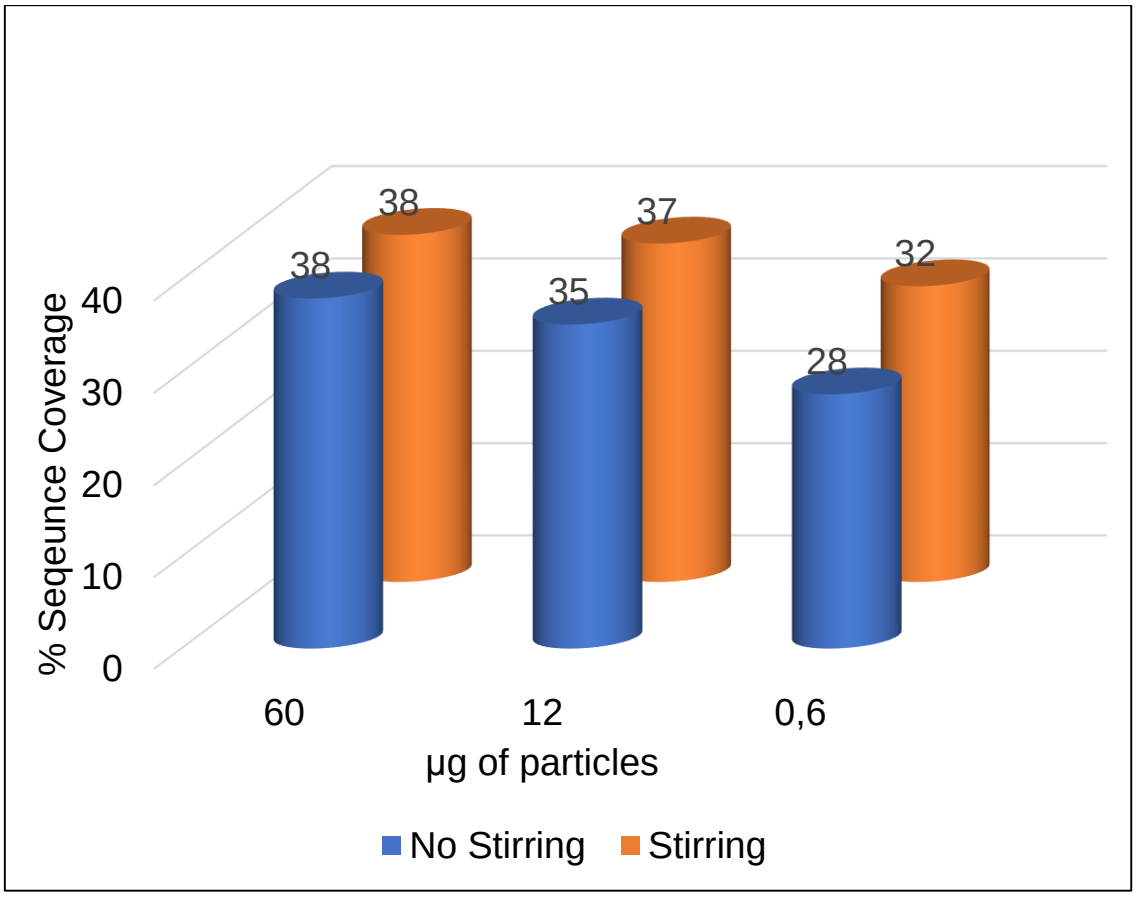


Figure 4.12: % Sequence Coverage of the digestion of 10 µg of BSA and 0.6, 12 and 60 µg of immobilized trypsin nanoparticles while stirring and non-stirring.

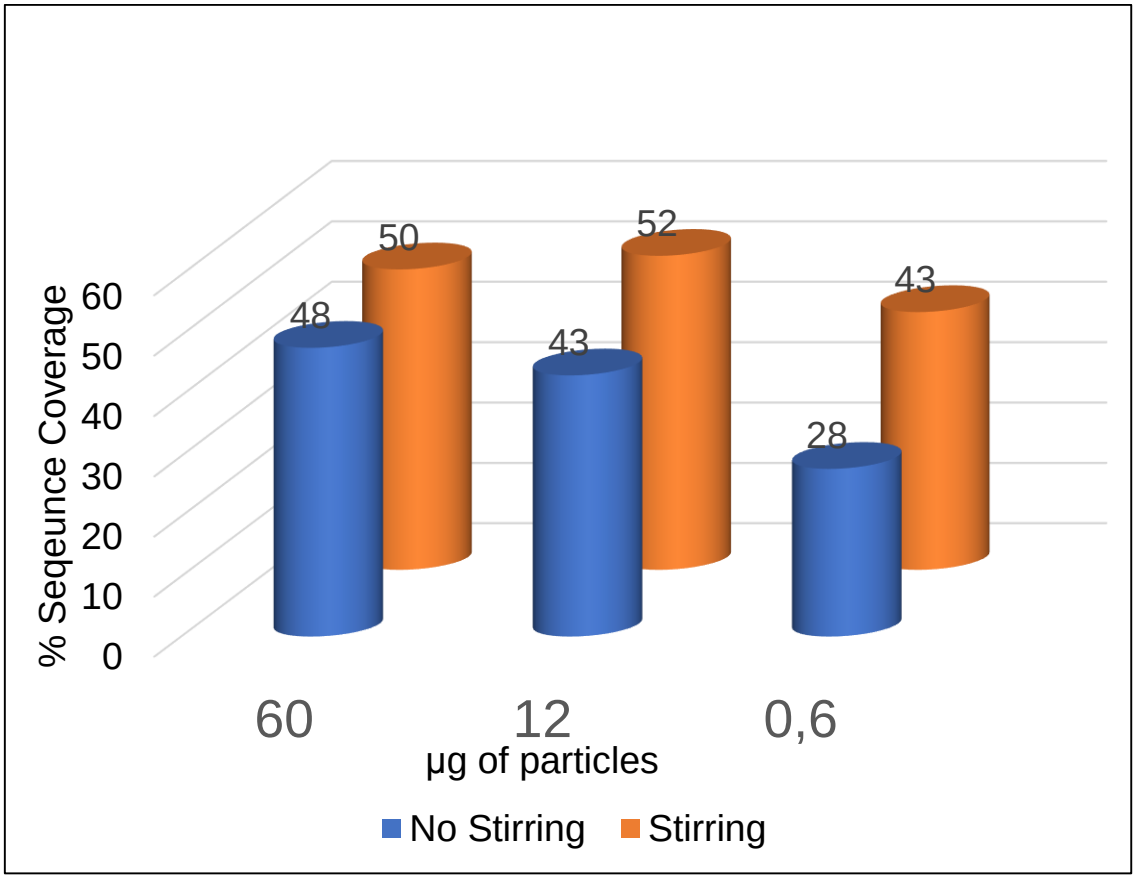


Figure 4.13: % Sequence Coverage of the digestion of 1 µg of BSA and 0.6, 12 and 60 µg of immobilized trypsin nanoparticles while stirring and non-stirring.

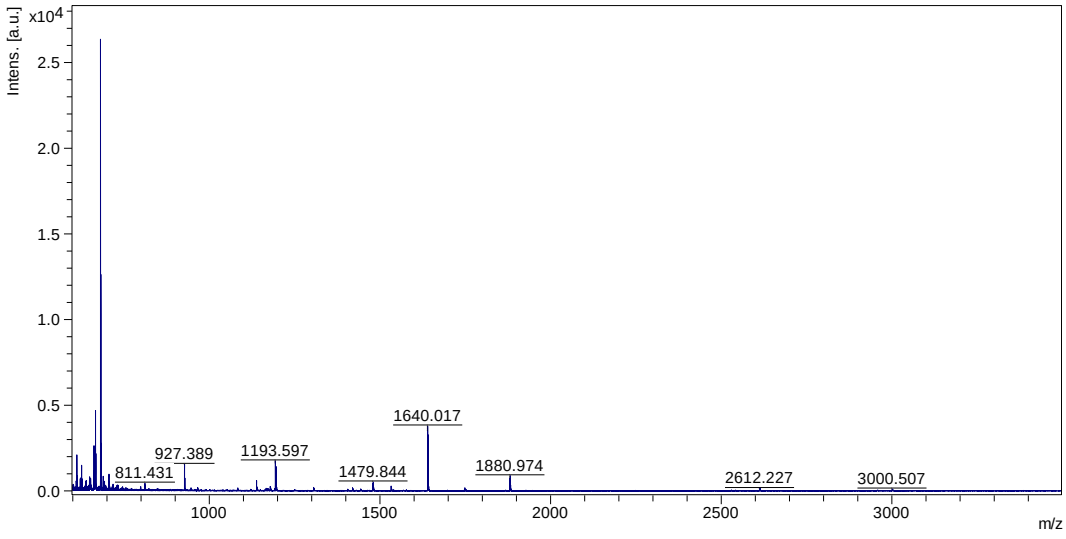


Figure 4.14: MS spectrum of the digestion of 1µg of BSA with 0.6µg of immobilized trypsin nanoparticles, non-stirring while digestion.

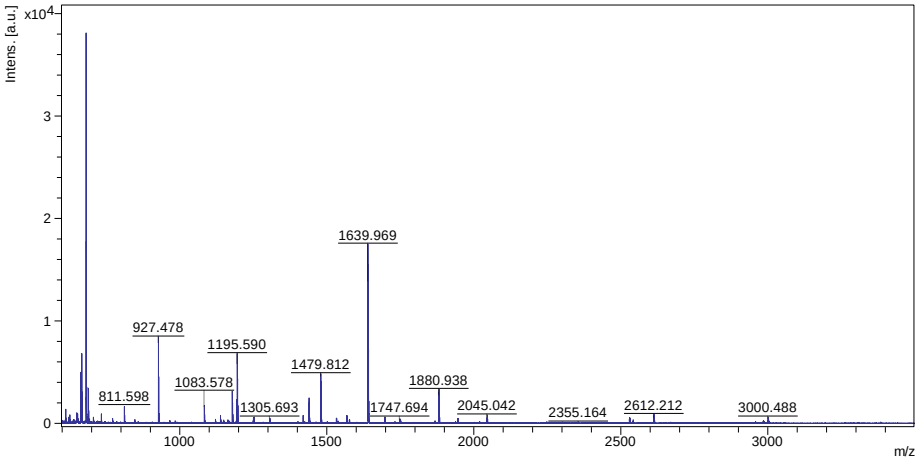


Figure 4.15: MS spectrum of the digestion of 1µg of BSA with 0.6µg of immobilized trypsin nanoparticles, stirring while digestion.

Other mass spectra of this experiment can be found in Annex IV

Sequence coverage increased when stirred, particularly when concentration of proteins and enzyme are very low. Therefore, stirring is now included in the process.

Many of immobilized trypsin systems try to not only avoid trypsin proteolysis, but also decrease the time of digestion.[24,25] Overnight digestion (around 12 hours) delay results for one day. The next set of experiments examined how sequence coverage responded to a decrease in digestion time. Commercial immobilized trypsin particles efficiency has begun to be compared to our immobilized trypsin nanoparticles. Results in Figures 4.16 and significant MS spectra in Figures 4.17 and 4.18.

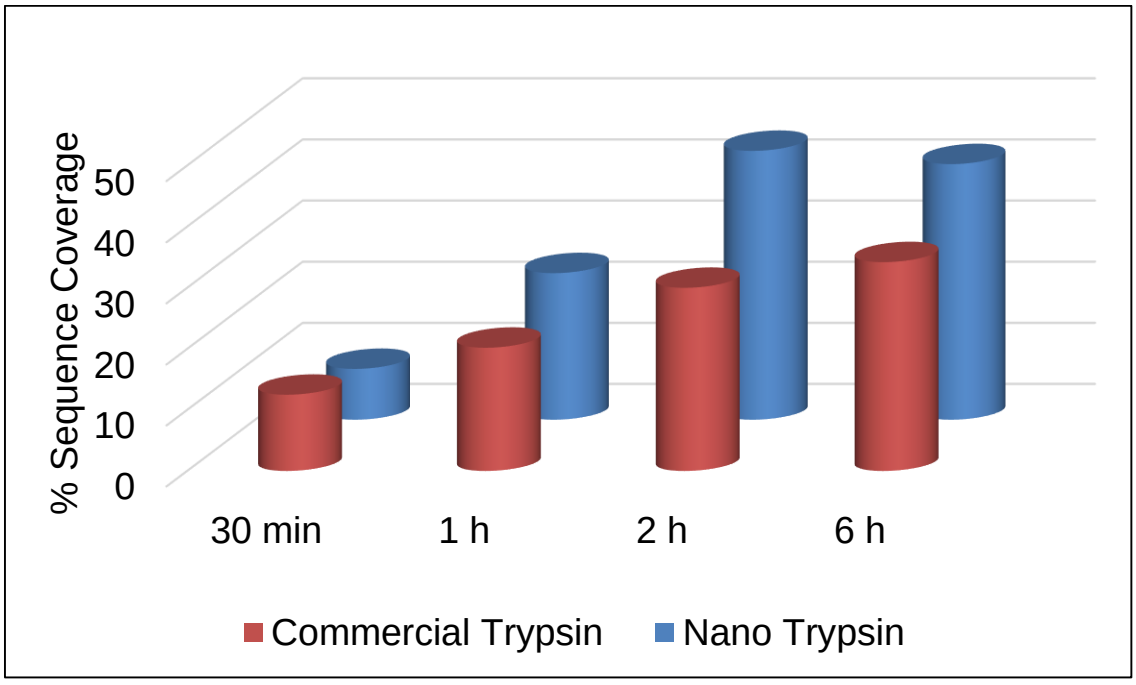


Figure 4.16: % Sequence Coverage of the digestion of 10µg of BSA and 0.5µg/µL immobilized trypsin nanoparticles, digestion time of 30 minutes, 1 hour, 2 hours and 6 hours.

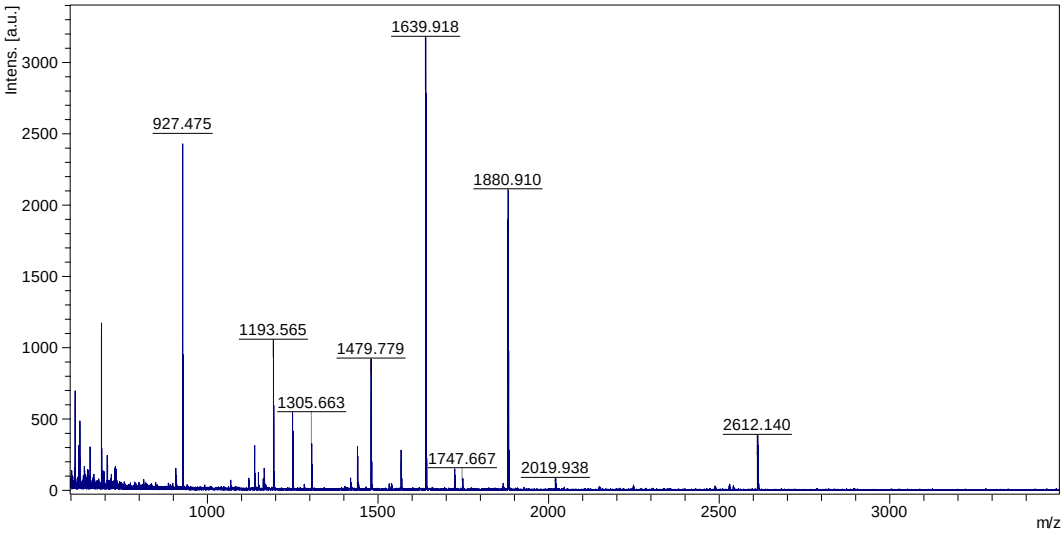


Figure 4.17: MS spectrum of the digestion of 10µg of BSA and 0.5µg/µL of commercial immobilized trypsin microparticles, 2 hours of digestion time.

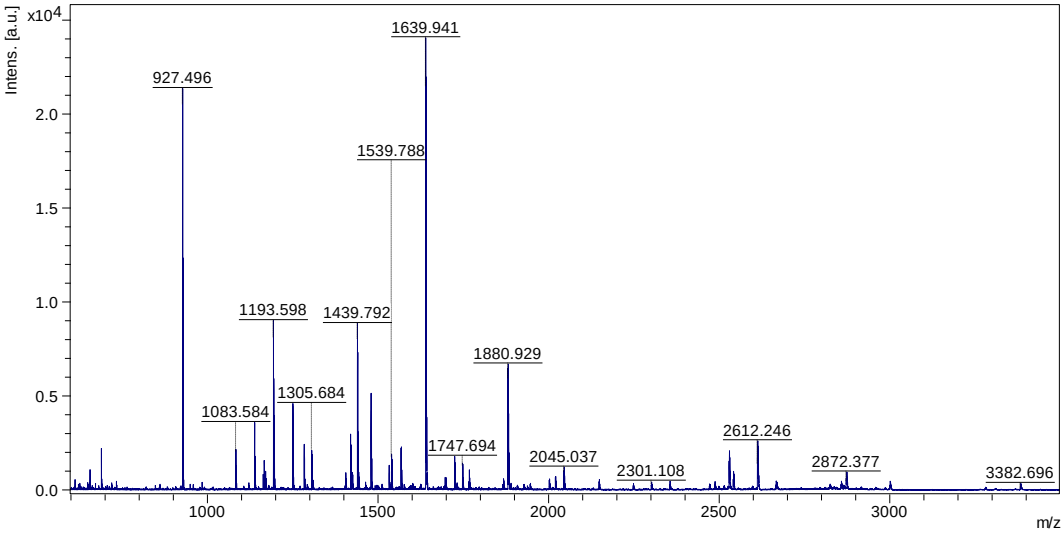


Figure 4.18: MS spectrum of the digestion of 10µg of BSA and 0.5µg/µL of immobilized trypsin nanoparticles, 2 hours of digestion time.

Other mass spectra of this experiment can be found in Annex V

Although commercial immobilized trypsin particles were able to digest and allow for protein identification. The digestion performed with immobilized trypsin magnetic nanoparticles proves to be more efficient, having a higher sequence coverage when digestion time was 1, 2 and 6 hours. After all these optimization tests a procedure was set to achieve the best results possible for protein digestion using immobilized trypsin nanoparticles. Results displayed in Table 4.1.

Table 4.1: Best conditions for protein digestion using immobilized trypsin nanoparticles

Variable	Best Condition
Ratio protein-nanoparticle	1:6
Stir	Yes
Digestion Time	Overnight (≈12 hours)

Although this system has achieved reliable results, so far, digestion was performed solely on a single protein. To ensure an indisputable place in the field of proteomics, this system must be able to digest complex samples with thousands of proteins. E. coli lysates were used in this crucial assay.

Before digestion, it was necessary to perform a Bradford assay to quantify the amount of protein present in the samples. The Bradford assay is a colorimetric protein assay based on an absorbance shift of the Bradford reagent dye used. In its acidic form, the brownish dye changes its colour to blue. This phenomenon is a consequence of the bond formed between proteins and the dye creating a noncovalent complex, which links the dye to the protein's carboxyl group by Van der Waals forces. Another feature of the Bradford assay is that the maximum absorbance occurs at a wavelength of approximately 595nm. The Bradford assay calibration curve (Table 4.2) is only linear over a certain range. Therefore, it is often necessary to dilute the samples before analysis. Another disadvantage of this assay is that this assay hinge on comparing the absorbance of the samples to the absorbance of protein standards, hence problems with these standards may lead to misleading quantifications.[37,38]

Table 4.2: Absorbance of the standard concentrations of BSA for Bradford assay.

Standard Concentration (mg/mL)	**Absorbance**
0.0	0.54 ± 0.03
0.2	0.59 ± 0.03
0.4	0.75 ± 0.01
0.6	0.85 ± 0.03
0.8	0.98 ± 0.12
1.0	1.09 ± 0.01
1.2	1.25 ± 0.00

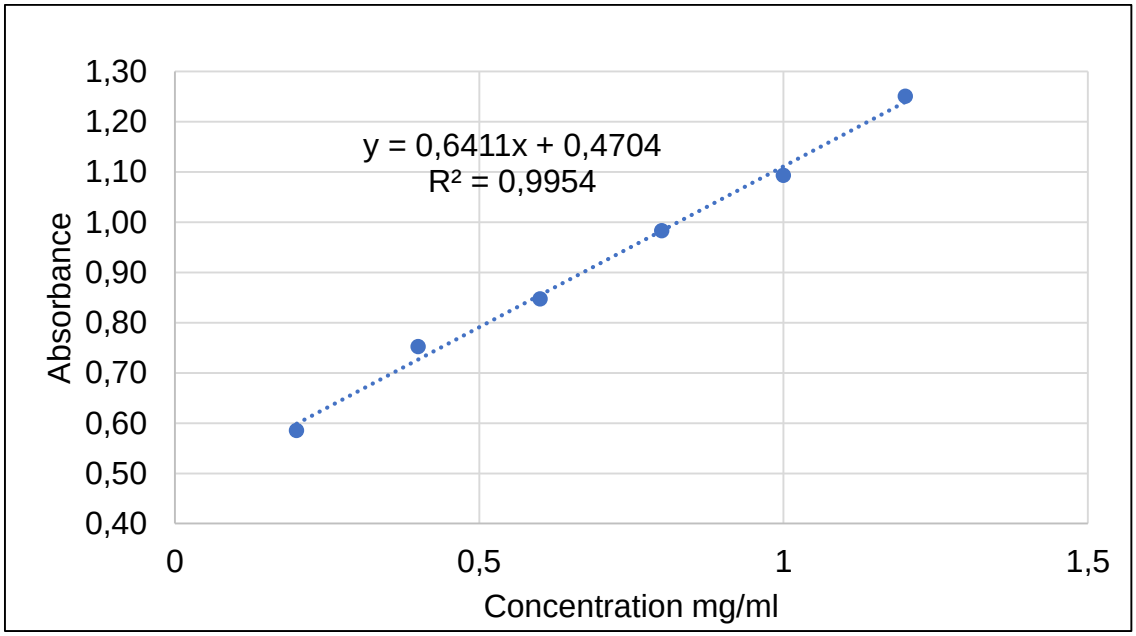


Figure 4.19: Graph absorbance vs concentration of the linear regression of the standard concentration of BSA for Bradford assay.

The curve obtained (Figure 4.19) is linear with almost any error. Therefore, it is suitable for quantifying the amount of protein present in these complex samples. Results of protein samples concentration in Table 4.3.

Table 4.3: Absorbance and calculated concentration for E.coli samples A and B.

Samples	*Absorbance*	*Concentration (mg/mL)*
A	0.99 ± 0.00	**4.0**
B	1.04 ± 0.01	**4.5**

Absorbance results delivered by sample analysis must be in the middle of the curve of linear regression, that region assures linearity. Closer to the end results may not be reliable (the first point of the standard was remove when calculated linear regression). To assure linearity and a good quantification samples had to be diluted (1:5 was the appropriate dilution).

Finally, the complex sample was digested using 10µg of protein and 0.5µg/µL of immobilized trypsin particles, comparing the efficiency of immobilized trypsin nanoparticles and commercial immobilized trypsin microparticles. The samples were analysed in the LC/MS. Results in Figure 4.20.

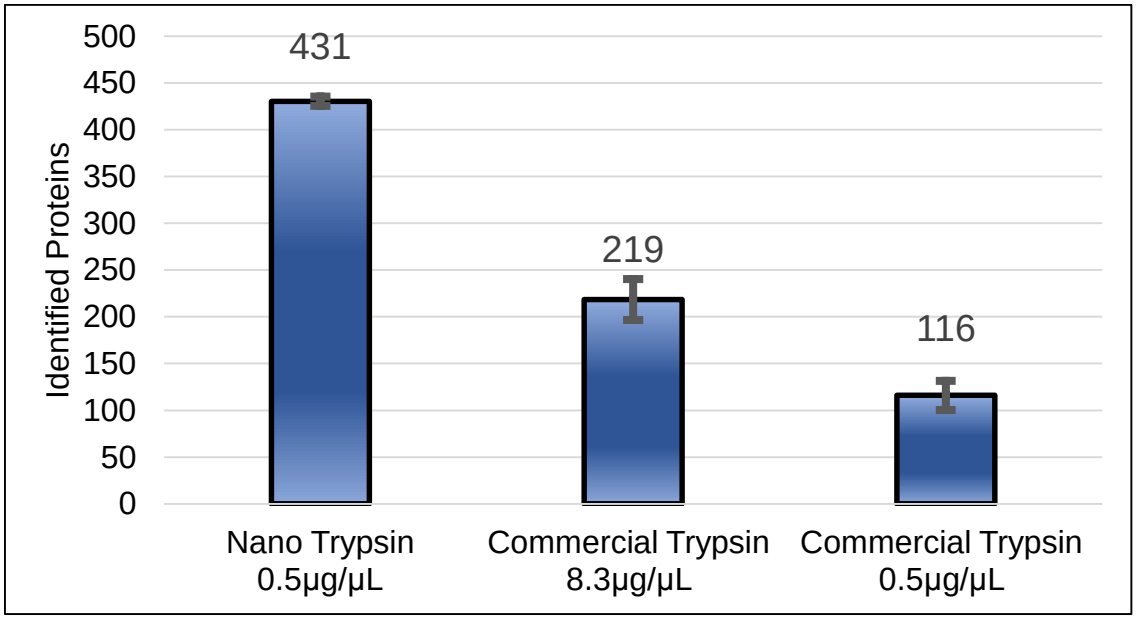


Figure 4.20: Number of protein identified using 10µg of E.coli lysates and three different immobilized trypsin particles: 0.5µg/µL of immobilized trypsin nanoparticles, 8.3µg/µL of commercial immobilized trypsin microparticles and 0.5µg/µL of commercial immobilized trypsin microparticles.

It is clear that immobilized trypsin nanoparticles perform better than commercial immobilized trypsin microparticles. Immobilized trypsin nanoparticles allow for the identification of nearly four times more proteins than commercial immobilized trypsin microparticles using the same concentration. Furthermore, commercial immobilized trypsin microparticles recommended concentrations are 16 times higher than the concentration of immobilized trypsin nanoparticles used, but still the number of protein identifications is almost two times higher with the immobilized trypsin nanoparticles. Therefore, immobilized trypsin nanoparticles are extremely efficient when compared to commercial immobilized trypsin microparticles.

To prove that these immobilized trypsin nanoparticles contain, in fact, trypsin we perform a 1D-GE by loading our nanoparticles with trypsin onto a 12.5% acrylamide SDS-PAGE as shown in Figure 4.21.

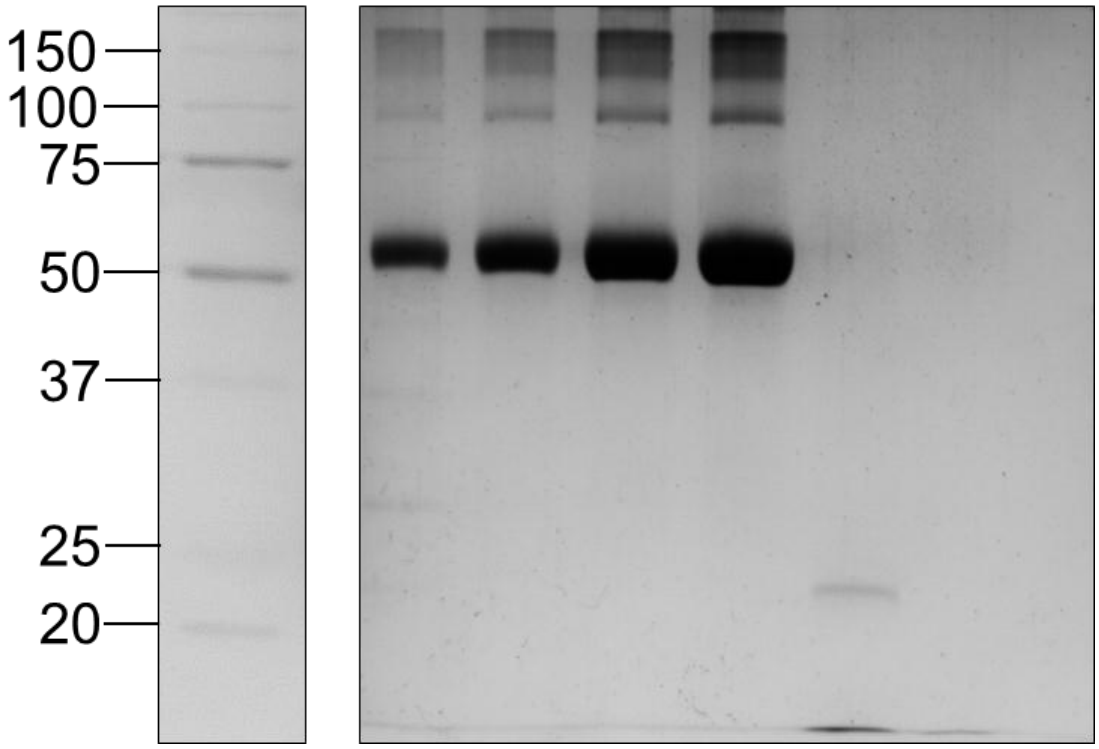


Figure 4.21: 1D-GE of Immobilized Trypsin nanoparticles

It is possible to observe that a unique protein band appearing at 23.8 kDa corresponding to the MW of trypsin.

A large number of immobilization methods and materials were created in this past years with success. Carboxyl functionalized magnetic microparticles, iron oxide microparticles with tannin layer and immobilized trypsin microparticles used in columns are a few examples of it.[24,25]

Unlike others, these particles differentiation is the size, real nanoparticles with 80nm, and the concentration needed to obtain results that are incredibly lower. Lowering the size of particles, increases the surface-area-to-volume ratio, having more surface in the same volume enhances the surface of reaction between trypsin and proteins. Thus, lower quantities of immobilized trypsin are required to reach the same efficiency. Since these nanoparticles are spherical surface-area-to-volume ratio can be calculated by dividing the area of a sphere by the volume of the sphere resulting in 3/r.

These proprieties make these nanoparticles a good system to perform tryptic digestion of various protein samples, whether they are complex or simple samples, in different scenarios.

4.2 Synthesis of Polystyrene nanoparticles:

The main goal of this work is to create a nano-sized IMAC for phosphopeptide enrichment. To do so, it is necessary to have a core from which to start functionalizing the particle. The smaller this core is, the smaller the final particle will be. To pursue that objective, polystyrene nanoparticles were produced to be the core of the IMAC.

Polymeric monodisperse microparticles are well studied and have an excellent performance in the field of separation sciences. Since chromatography is a separation technique, the use of polymeric particles is adequate. Taking into consideration the benefits of nanotechnology and nano-sized system, is to expect an improvement in the enrichment of phosphopeptides using nano-sized IMACs.

This synthesis of this polymer is a dispersion polymerization. Dispersion polymerization is a type of precipitation polymerization where the solvent solubilizes well the initiator and the monomer but does not solubilizes the polymer, therefore, when the polymer is formed it precipitates. Also, polystyrene synthesis is radical, this type of polymerization needs an initiator to start the reaction, in this case, the initiator suffers a homolithic fission through heat which forms the radical species where the monomers start to link forming the polymer, scheme in Figure 4.22.[48,49,50]

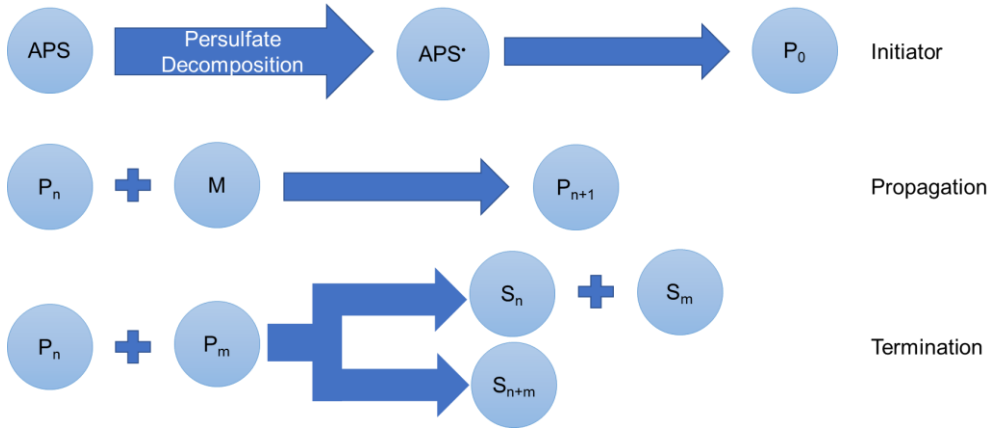


Figure 4.22: Scheme of PS seeds polymerization, radical polymerization.(adapted)[49] P is the polymer in formation, M is the monomer and S is the polymer in the final size.

Temperature, initiator type and concentration, stabilizer type and concentration, monomer concentration and solvency are key parameters in the polymerization that affect polymer size.[45,51,52]

A 2^3 experimental design was executed to optimize the condition of production of these PS nanoparticles. 2^3 means two levels with three factors and 2^3=8 experiences. The factors being optimized are the amount of 3 reagents: APS, SDS and styrene. Levels and factors in Table 4.4. Experimental design in Table 4.5.

Table 4.4: Level and factors of the experimental design and respectively amounts

Level/Factor	APS (g)	SDS (g)	Styrene (mL)
-	0.004	0.2	0.7
+	0.04	2	7

Table 4.5: Level of the factors for the different experiments

Experiment	APS (g)	SDS (g)	Styrene (mL)
1	-	-	-
2	+	-	-
3	-	+	-
4	+	+	-
5	-	-	+
6	+	-	+
7	-	+	+
8	+	+	+

For all experiences temperature and stirring was constant (between 80-90˚C and 1000 rpm). At the end of the reaction, a syringe filter with 220nm pores was used to discard bigger nanoparticles.

To verify the size of the produced nanoparticles Zetasizer (Malvern) equipment was used. This equipment uses a technique called dynamic light scattering (DLS). This technique is ideal for colloidal and nanoparticles size analysis. It consists of the random thermal motion of particles called the Brownian motion. Smaller particles diffuse light faster than larger particles. Brownian motion is also affected by temperature; therefore, a rigorous control of temperature is needed.

To measure the diffusion speed, the speckle pattern produced by illuminating the particles with a laser is observed. The scattering intensity at a specific angle will fluctuate with time, and this is detected using a sensitive avalanche photodiode detector. The intensity changes are analysed with a digital autocorrelator which generates a correlation function. This curve can be analysed to give the size and the size distribution.[53]

The target sized wanted for this PS nanoparticles are between 20-70nm. This size was determined based on the final size wanted for the IMAC (<500nm) and the fact that it is still needed to perform a functionalization of the nanoparticles that requires a swelling time. Furthermore, the concentration of nanoparticles is also necessary to take in consideration once a certain amount of them are needed to produce the IMAC, concentrations of above 40mg/mL are convenient because it allows producing three replicates of IMAC with a 20% security. Critical concentration was established at 12mg/mL allowing one replicate with 5% security, inferior to this concentration it is not possible to produce the IMAC. To determine the concentration of PS nanoparticles in solution, 1mL of solution was dried and weighted, it is important to withdraw from the amount weighted the weight of SDS that precipitates too. Other attributes that were taken into consideration is population size, that must be superior to 90%, and polydispersity (PDI) that must be inferior to 0.3. PDI value varies between 0 (totally monodisperse) and 1 (totally polydisperse).

To discard errors of analysis using Zetasizer, a blank sample of APS, SDS and Styrene was analysed for the different amounts. Blank samples values in Table 4.6.

Table 4.6: Values of the blank samples for each factor and level.

Level/Factor	APS (g)	Size (nm)	SDS (g)	Size (nm)	Styrene (mL)	Size (nm)
-	0.004	**0**	0.2	**2**	0.7	**0**
+	0.04	**0**	2	**2**	7	**0**

For each experience were made two batches except when results were conflicting, and an extra batch was made to confirm data. For each batch three replicates of size, population size and polydispersity. Results displayed in Table 4.7.

Table 4.7: Size, population and PDI of PS seeds of each experiment

Experiment	***Size (nm)***	***Population Size (%)***	***PDI***
1	132 ± 116	94 ± 5	0.3 ± 0.2
2	18.7 ± 0.4	100 ± 0	0.09 ± 0.00
3	1.8 ± 0.1	56 ± 6	0.4 ± 0.1
4	13.6 ± 0.9	96 ± 3	0.23 ± 0.01
5	1473 ± 1729	98 ± 4	0.31 ± 0.07
6	39 ± 1	99.3 ± 0.8	0.21 ± 0.01
7	49.6 ± 0.9	100 ± 0	0.20 ± 0.00
8	43 ± 5	92 ± 8	0.28 ± 0.03

Experiences 1, 3 and 5 were excluded because of size, population size and PDI, these parameters were outside the target area. This leaves a last result to analyse, concentration. Also, experience 3 might not have formed any polymer since the size obtained by SDS micelles is similar. Results displayed in Table 4.8.

Table 4.8: Concentration of PS seeds of each experiment

Experiment	***Concentration (mg/mL)***
1	Excluded
2	9 ± 2
3	Excluded
4	22 ± 4
5	Excluded
6	105 ± 33
7	69 ± 10
8	110 ± 3

Experiment 6, 7 and 8 have the conditions to be used in IMAC preparation.

From all the possible parameters, which might influence the size of the polymers in this work, the concentration of initiator (APS), concentration of stabilizer (SDS) and concentration of monomer (Styrene) were studied.

Keep in mind that our conclusion is produced by a combination of 3 factors, the parameters above described. Neither Shen[51] or Yun[52] have analysed the effects of 3 factors simultaneously.

To reach small size PS nano particles, the optimal conditions include low initiator concentration, high stabilizer concentration and low monomer concentration. Joining these three conditions together (experiment 4) we have the smallest acceptable size since experiment 3 was considered not to produce any PS nanoparticles.

In this work, we have concluded this information, except for the initiator concentration. It was not evident that a low concentration of initiator would decrease particle size, it appears to have a dependence on stabilizer and monomer concentration. Once more the fact of this 3-way interaction could explain the results. For our target size, initiator and monomer must be at the highest level regardless of the stabilizer level.

The main purpose of this experimental design was to seek which factor was the main responsible for size but, due to results, parametric statistics analysis was not possible, because data gather was non-parametric. Therefore, for a quick understanding of the experimental design and to try to withdraw some conclusion a cube illustration was created, Figures 4.23 and 4.24.

Regarding the concentration of PS nanoparticles, it is clear that a high concentration of monomer will produce more PS particles.

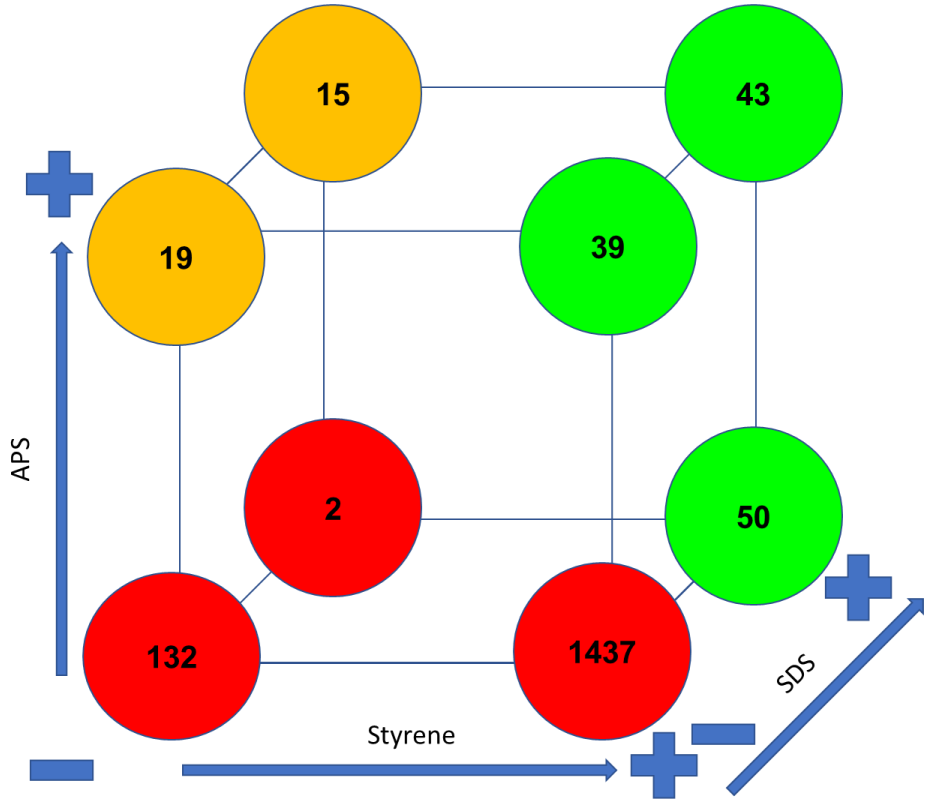


Figure 4.23: Synthesized PS seeds size according to the experimental conditions used displayed in 3D cube. Values in nm.

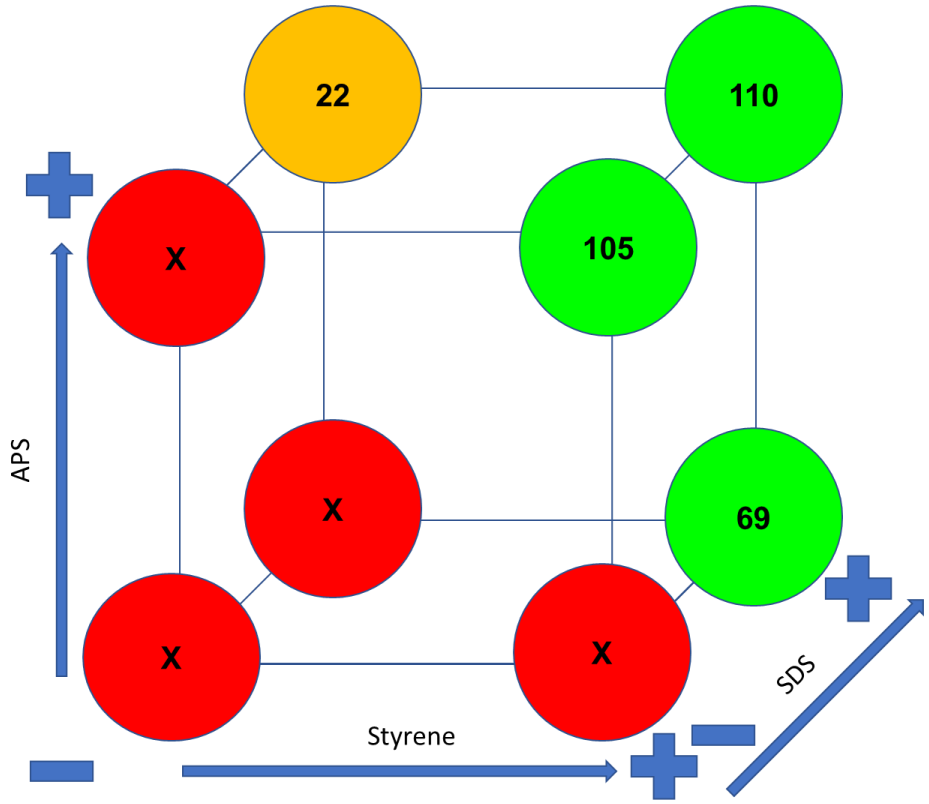


Figure 4.24: Synthesized PS seeds concentration according to the experimental conditions used displayed in 3D cube. X means that the concentration was not analyse because they were excluded by the size test. Values in mg/mL.

Regarding the temperature of reaction, increasing temperature also increases the solubility of styrene (30mg/100g of water at 20˚C) and since this is a dispersion polymerization, high solubility increases size particle, the reaction temperature cannot be above 140˚C because styrene boiling point is at 145˚C. PS is insoluble in water therefore, changes in the temperature will not be critical to the size of the particle.[51]

Immediately after styrene is added, the solution changes colour from transparent to white, that evidences the formation of polystyrene particles, this quick reaction possibly means the formation of small PS particles since they have a low quantity of styrene it also evidences the importance of a constant flux of styrene to achieve monodisperse PS. Due to technical difficulties on adding small amounts of styrene at constant flux was not possible. Therefore, experiences 1 to 4 should have high PDI. This is faithful to all except for experience 2 and 4. This exception can be explained by the fact of having a high concentration of inhibitor creating the perfect conditions for tiny nanoparticles (<20nm). Figure 4.25 shows a SEM image of PS seeds (Experiment 8) and as stated before, the size of the nanoparticles is inferior on SEM analysis than on DLS analysis.

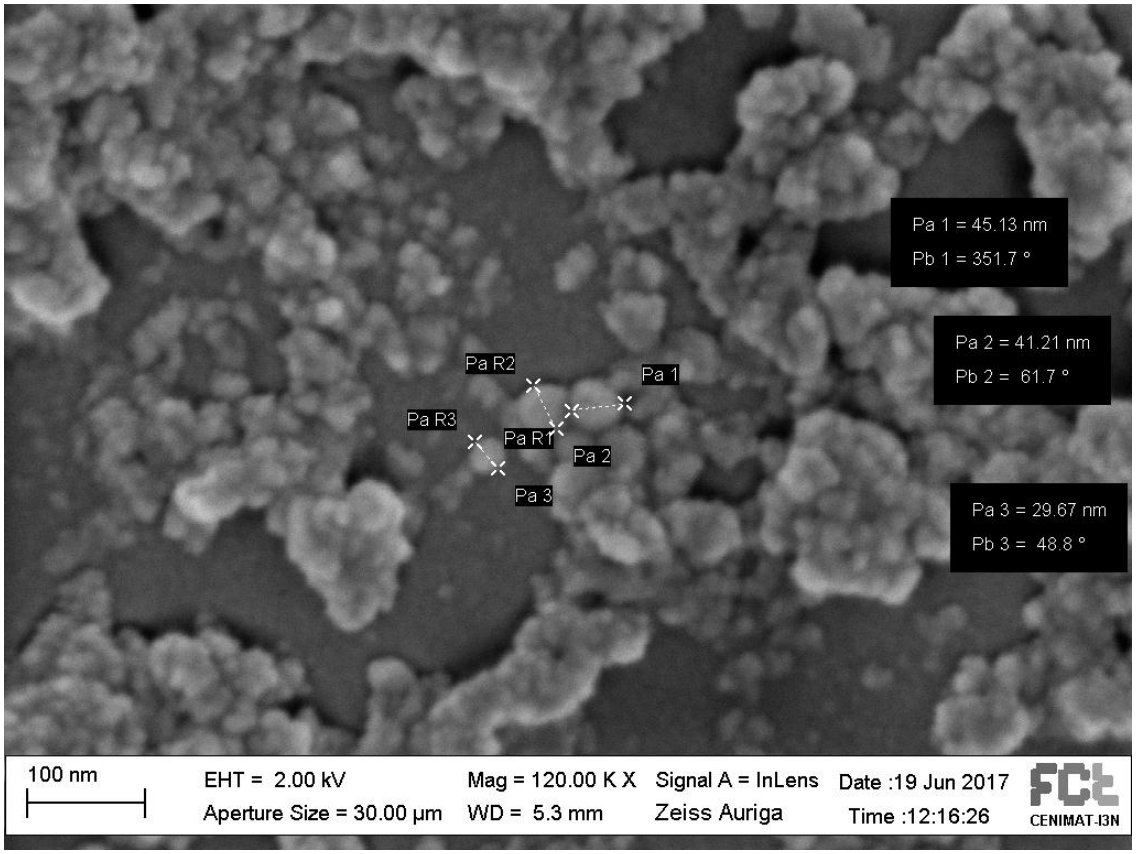


Figure 4.25: SEM image of PS seeds (Experiment 8)

4.3 Synthesis of IMAC:

This synthesis consists of 5 steps. The first step is the synthesis of PS seeds already commented in the chapter above. Monodisperse nanospheres have a uniform monodisperse size distribution, uniform column packing, low column pressure and high column efficiency. The prepared

nanospheres are stable whether strong acids or alkaline buffers are used. Hydrophilic surface of PS seeds minimizes nonspecific adsorption. The second step involves the creation of a flexible linker which increases the spatial distance between the metal and the matrix (seed). This flexible link provides a beneficial spatial orientation for the phosphopeptide binding. The third step is the addition of amino groups to the flexible linker increasing the hydrophilicity of the IMAC. Fourth, the addition of phosphonate groups for chelation and immobilization of the metal. This immobilization creates a beneficial structural orientation for phosphopeptides binding. This step is also called pre-IMAC. The fifth and last step is the addition of the metal. Usually, the metals used are titanium and zirconium. In this work, the metals used will be titanium, lanthanum, europium and gadolinium.[13]

To attach the flexible linker to the matrix PS seeds a preparation of the PS nanosphere surface is required. This preparation involves a swelling of the nanospheres. This step is a critical parameter for the final size of the IMAC. Zhou[13] turns a PS seed of 1.2µm into an IMAC of 12µm. In this work, the target is to maintain nano-size to all products. therefore, IMAC particle should be <500nm. Step 3 to 5 should not affect the size of the particle because they involve the addition of atoms to the particle, atoms size is very small when compared with the particle core, thus, it can be despised.

Three pre-IMAC were created, each one with different PS seeds. DLS analysis of pre-IMAC's were made. Results displayed in Table 4.9.

Table 4.9: Comparison of PS seeds size and Pre-IMAC size

Experiment	***Seed Size (nm)***	***Seed PDI***	***Pre-IMAC Size (nm)***	***Pre-IMAC PDI***
1	47.6 ± 0.4	0.25 ± 0.01	539 ± 136	0.9 ± 0.1
2	38 ± 1	0.21 ± 0.01	275 ± 20	0.4 ± 0.2
3	39.5 ± 0.7	0.21 ± 0.01	292 ± 32	0.75 ± 0.07

Results show that the quantity of PS seeds is critical parameter for size, quantity, and quality of the pre-IMAC, in experiment 1, the quantity of PS seeds used was 1603mg when the protocol recommends 450mg as a result of its size and PDI were much higher than in other experiments, the IMACs prepared in experiment 1 were discarded. Analysing the colour of pre-IMAC can provide a quick information about the purity of the pre-IMAC and how dry it is. Allegedly pre-IMAC should be light yellow. Produced IMAC's are orange/light-brown colour. I believe that this change of colour corresponds to a shift in size. Other nanoparticles such as gold nanoparticles also change colour depending on size. Changes in stirring speed apparently do not cause problems. I believe that the reaction should take place in a turbulent mode to assure the mix and reaction of all reagents. Prepared IMAC's exhibits the same colour as the pre-IMAC except for gadolinium IMAC that changes the colour to light yellow. This is probably a mark of metal binding and consequent formation of an IMAC. Figure 4.26 is a SEM image of the lanthanum IMAC.

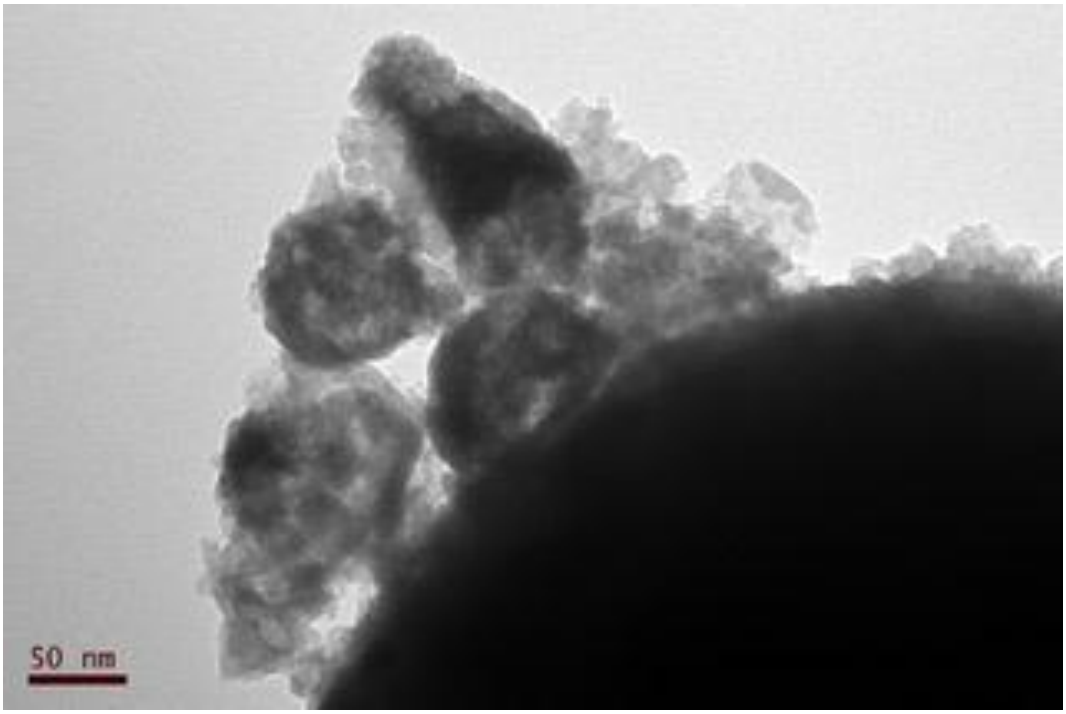


Figure 4.26: SEM image of lanthanum IMAC produced

4.4 Enrichment of phosphopeptides:

For enrichment of phosphopeptides, IMACs with two different metals were used: titanium^{4+} and lanthanum^{3+}. IMACs nanoparticles were displayed inside a pipette tip forming a compact column. A C8 membrane was used as a support for IMACs nanoparticles packaging. C18 membrane was also used but it resulted in no enrichment of phosphopeptides, therefore, de membrane choose as a support was C8 as indicated in the protocol. The difference between these two membranes is the number of carbon in each molecule, being the membrane with more carbons in each molecule (C18) more hydrophobic.

Critical variables in this process are the quantity of IMAC and the amount of sample. The use of high quantities of protein sample and low quantities of IMAC results in an incomplete adsorption of phosphopeptides. To assure this fact, a new enrichment of the flow-through must be done. This process was not done in the present work. The use of high quantities of IMAC and low quantities of protein sample results in nonspecific adsorption of nonphosphopeptides, in this situation too much IMAC was utilized for the amount of phosphopeptides. To assure that no nonphosphopeptides are observed in elution loading buffer should have pH<3. Other problems with nonphosphopeptides in the elution are due to nonphosphopeptides that are highly hydrophobic or highly acidic, in these scenarios increasing the amount of ACN in the loading buffer and increase the TFA concentration or add NaCl to the washing buffer 1, respectively. In every experiment realize, phosphopeptides were never observed in the flow-through.[13]

In the first experiments, 150µg of IMACs of titanium and lanthanum were used for enrichment passing through the column 7µg of protein. The protein used in every enrichment experiment was α-casein. The results of these first experiments were very negative, with no phosphopeptide observed. First, the pre-IMAC used for these first IMACs was the number 1 which was not in perfect condition however with the same conditions and using other pre-IMAC for producing the IMACs the same result was observed, no phosphopeptides. After these results first thoughts were targeted to protein amount, there was a clear deficit of protein. In further experiments, the quantity of protein was increased to 35µg. However, this quantity was not enough to identify the protein tested, but four phosphopeptides were identified, among these peptides also some nonphosphopeptides were identified thus the quantity of IMAC used is very high for the amount of protein. Regarding these last results, in the next enrichment, 100µg of IMAC and 105µg of protein were used. Results in Figure 4.27.

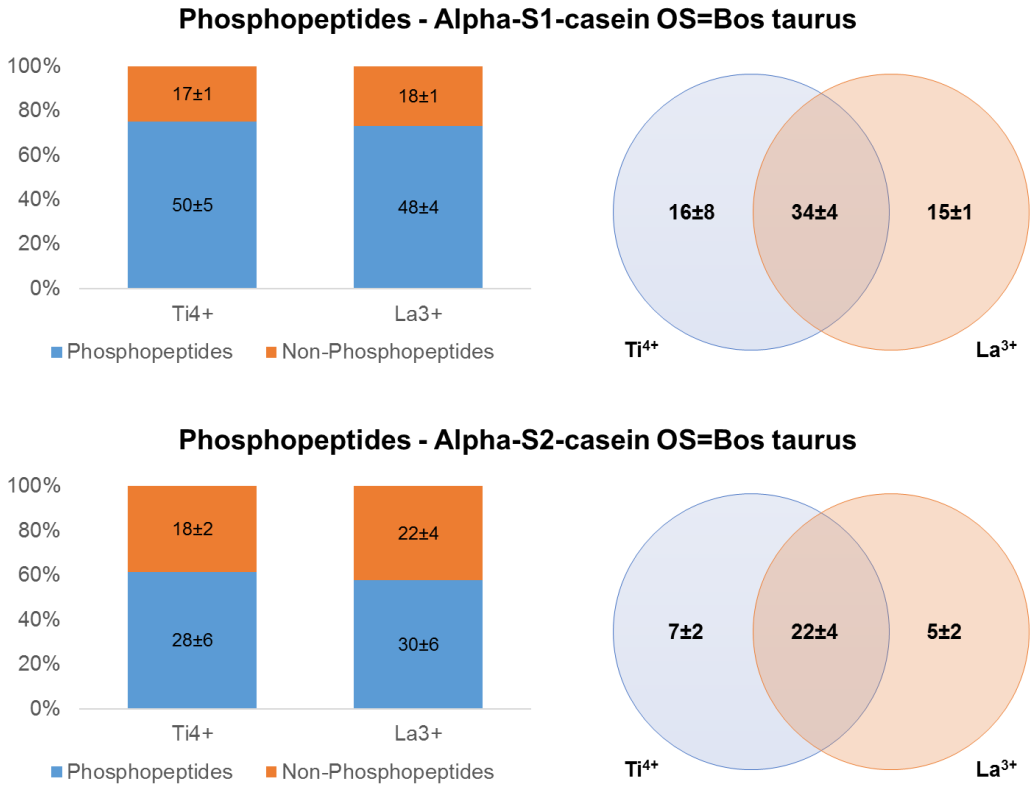


Figure 4.27: Analysis of the number of phosphopeptides and nonphosphopeptides identified using titanium and lanthanum IMACs nanoparticles for enrichment of α-casein protein sample.

With this experiment, it was discovered that the protein used is, in fact, two different protein, α-casein 1 (24.5kDa)[54] and α-casein 2 (26kDa).[55] According to Zhou[13], it was expected to find 18 different phosphopeptides or 20 according to Yu[56]. With this technique, we were able to identify 99 different phosphopeptides. Furthermore, results demonstrate that some phosphopeptides prefer to bind only with titanium IMAC and others prefer only to bind with lanthanum IMAC. Other authors have compared the enrichment of phosphopeptides using different IMACs and MOACs concluding that iron, titanium and zirconium IMACs have some peptides that only bonds with each type of metal. Also for the enrichment of phosphopeptides form α-casein samples it has been

proved that IMACs are more efficient (except iron IMAC) than MOACs[56], Figure 4.28 but still the number of phosphopeptides found whether with the titanium or the lanthanum IMAC nanoparticles produced are five times higher.

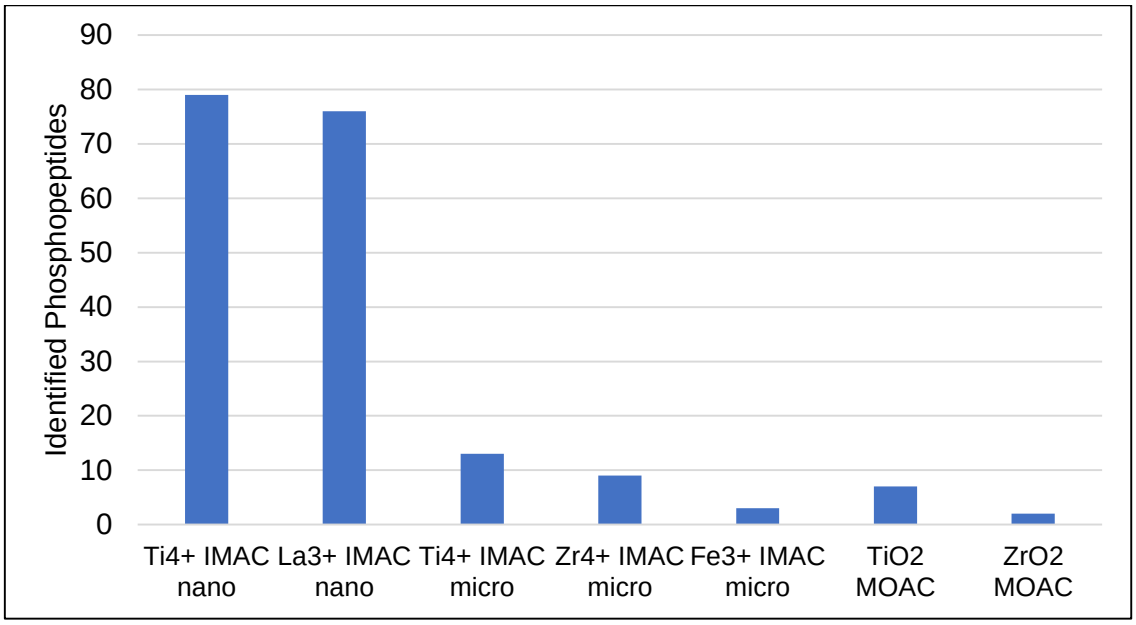


Figure 4.28: Number of Identified phosphopeptides with different techniques

Figure 4.29 shows the ratio of identified phosphopeptides versus the total number of identified proteins.

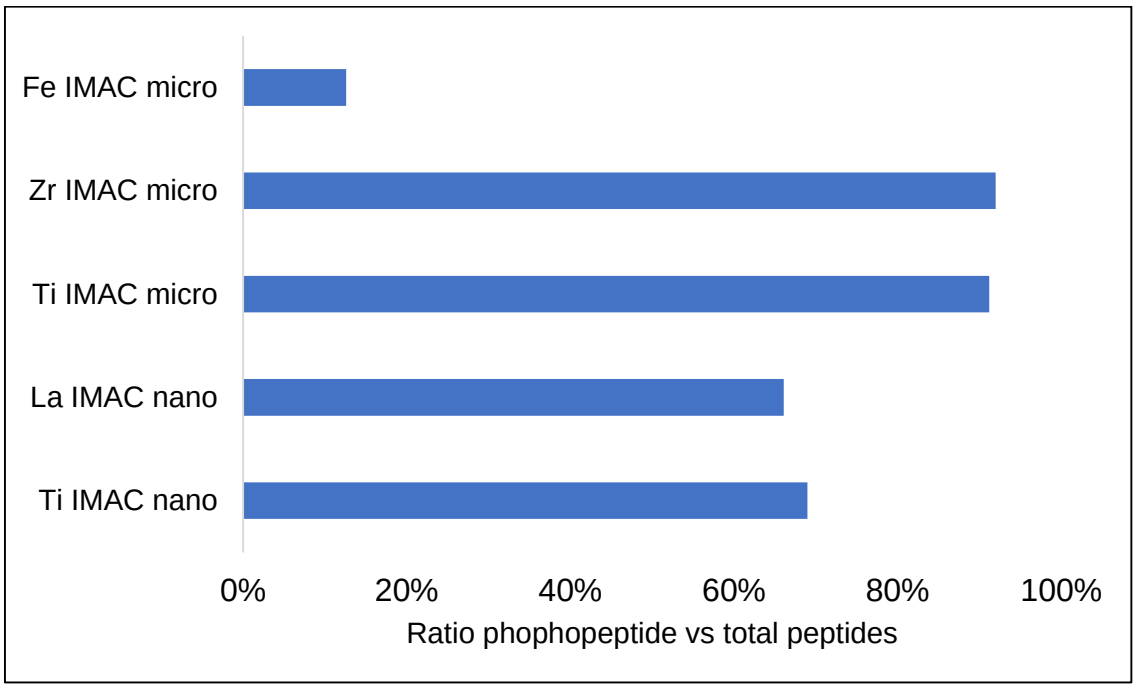


Figure 4.29: Ratio phosphopeptides identified vs total number of peptides identified.

Although the ratio of phosphopeptides versus total peptides is lower than other works this can be easily explained, the optimization of this technique was not performed, the optimal amount of

IMAC versus quantity of protein and buffers conditions, resulting in a nonspecific adsorption of nonphosphopeptides. Another possible explanation is the isoelectric point of the nonphosphopeptides. Normally nonphosphopeptides that bond to IMACs have an isoelectric point inferior or equal to five therefore, IMACs also enrich acidic peptides.[56]

Regarding the number of phosphopeptides in the flow-through the number of phosphopeptides found can be despised in both IMACs regardless the metal used, Figures 4.30 and 4.31.

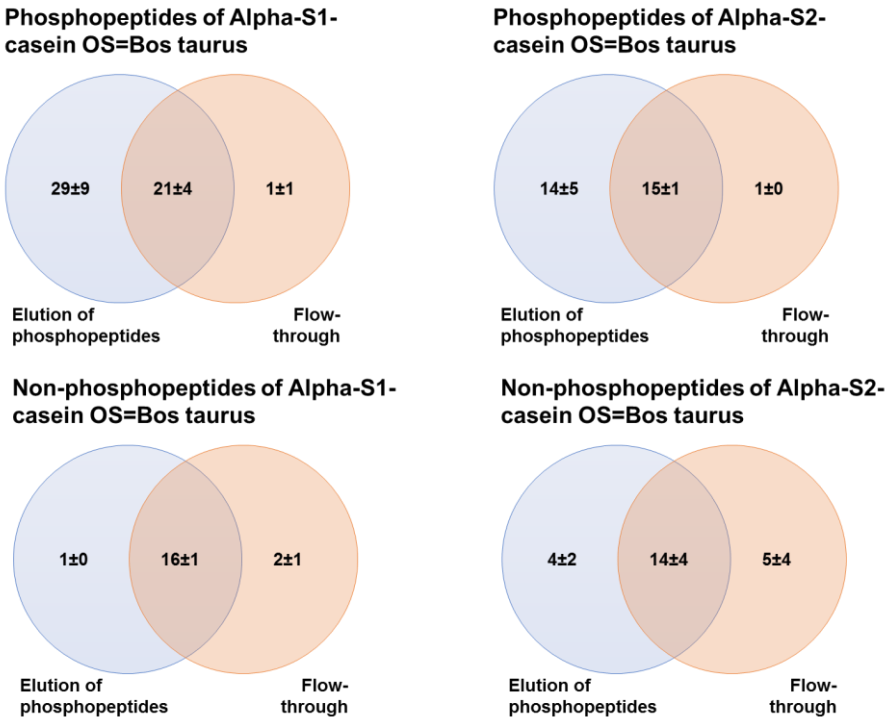


Figure 4.30: Comparing of the number of phosphopeptides and nonphosphopeptides identified in the elution and the flow-through using titanium IMACs nanoparticles for enrichment of α-casein protein sample.

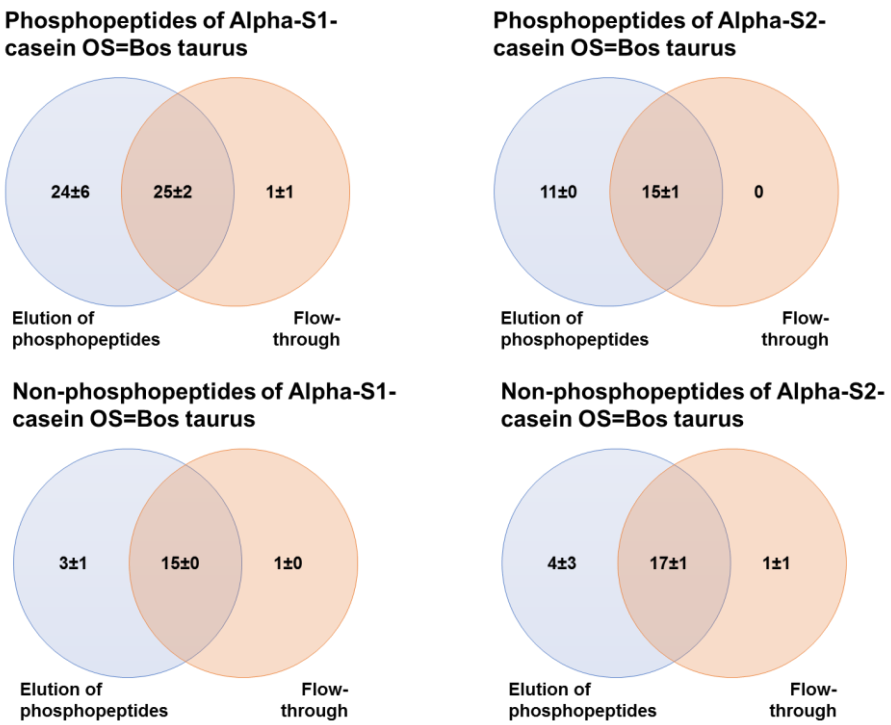


Figure 4.31: Comparing of the number of phosphopeptides and nonphosphopeptides identified in the elution and the flow-through using lanthanum IMACs nanoparticles for enrichment of α-casein protein sample.

Figure 4.32 is a representative tandem MS spectrum of a phosphopeptide.

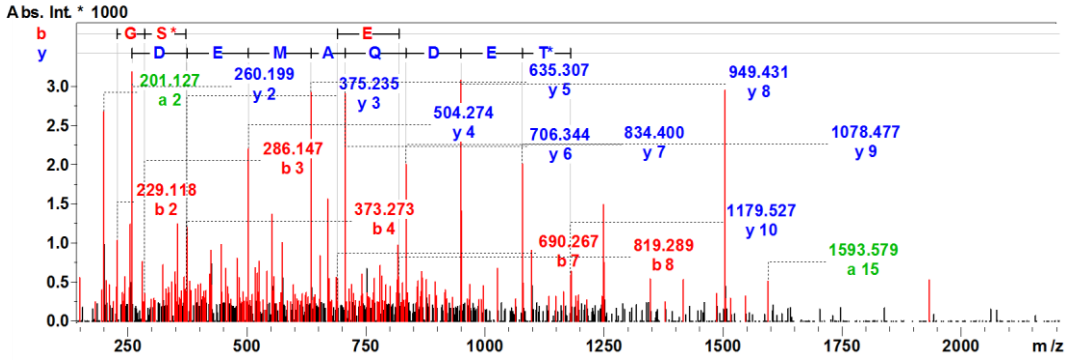


Figure 4.32: Representative tandem MS spectrum of a phosphopeptide.

Literature refers that iron IMACs have an affinity for peptides containing histidine. This occurs because the amino acid has strong chelating interaction with the metal.[56] To verify if this happens with the produced IMACs nanoparticles of titanium and lanthanum Figure 4.33 shows the percentage of phosphopeptides and nonphosphopeptides containing histidine in their structure.

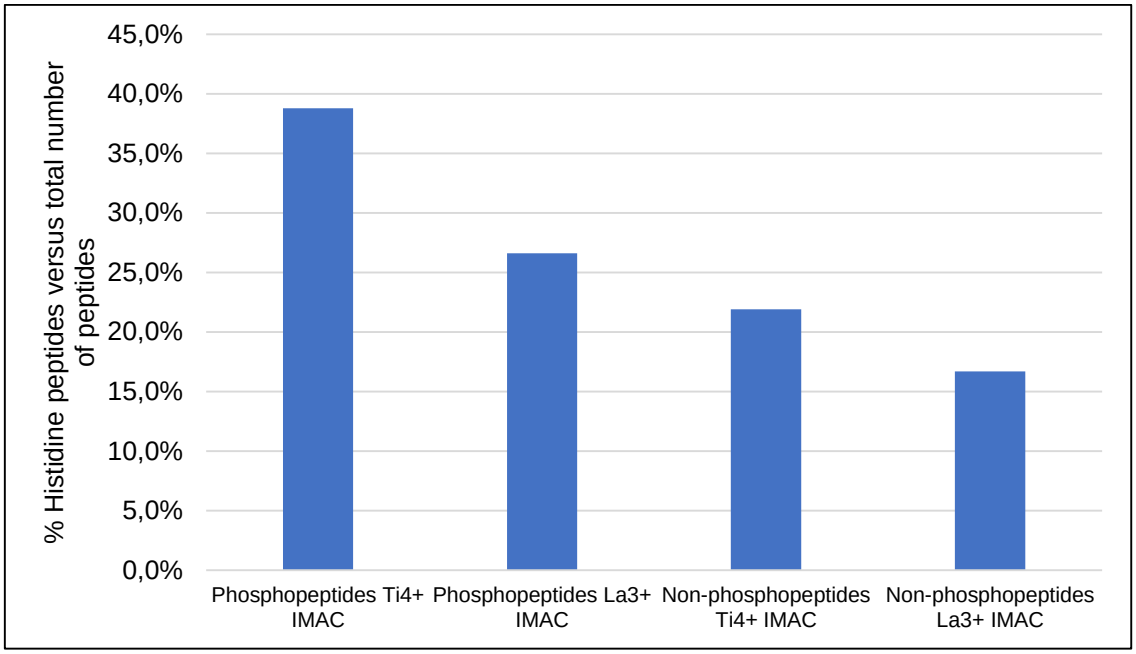


Figure 4.33: Percentage of histidine peptides versus total number of identified peptides.

5 Conclusions

The objective of obtaining a material with a higher performance than the current ones available in the market was achieved as the digestion performed with our immobilized trypsin magnetic nanoparticles was five times most effective that the digestion performed with commercial immobilized trypsin.

The objective of obtaining nano-material with a diameter range comprised between 20 and 50 nm was achieved. The parameter found as the most important is the level of styrene, which must be at the highest value studied. The levels of the other two variables studied, are directly linked to the styrene concentration. Thus, if styrene concentration is at the highest value, It is enough one of the other two variables to be at the highest value as well, in order to obtain the desired size. As for the amount of nanoparticles obtained it was clear established that the styrene and the APS must be at the highest concentrations studied. Therefore, we have concluded that for obtaining the nano-material at the desired size and in large quantities the APS and the styrene must be at the highest level.

Small particles of around 40 nm and concentrations higher than 40 mg/mL, combined with a critical assessment and some adjustments to the standard protocol for Pre-IMACs synthesis have allowed us to obtain a Pre-IMAC size of about 250nm.

The objective of developing a material to enrich phosphopeptides was achieved. The IMACs prepared, both with titanium or lanthanum, were used with a performance five times better than other similar materials described in literature for the protein used as proof of concept, alfa casein: 90 versus 20. It was verified that the phosphopeptide enrichment was function of the metal used to prepare the nanomaterials, Ti vs La.

6 Future Prospects

1. It is necessary to investigate if our immobilized trypsin nanoparticles digestion compares with the standard procedures for label-free protein quantification.
2. The enrichment of phosphopeptides from complex proteomes needs to be proven with the nano IMACs developed in this work.
3. The performance of nano IMAC prepared with both metals, Ti and La at the same time needs to be tested in order to avoid selective extraction.

7 References:

1. Twyman, R. M. *Principles of Proteomics*. (2014).
2. Mishra, N. *Introduction to Proteomics - Principles and Applications*. (2010).
3. What is Proteomics? Available at: https://www.ebi.ac.uk/training/online/course/proteomics-introduction-ebi-resources/what-proteomics. (Accessed: 15th September 2017)
4. Graves, P. R. & Haystead, T. a J. Molecular Biologist ' s Guide to Proteomics. *Microbiol. Mol. Biol. Rev. Microbiol. Mol. Biol. Rev.* **66,** 39–63 (2002).
5. Overview of Post-Translational Modifications - Thermo Fischer Scientific. Available at: https://www.thermofisher.com/pt/en/home/life-science/protein-biology/protein-biology-learning-center/protein-biology-resource-library/pierce-protein-methods/overview-post-translational-modification.html. (Accessed: 19th September 2017)
6. Mann, M. & Jensen, O. N. Proteomic analysis of post-translational modifications. *Nat. Biotechnol.* **21,** 255–261 (2003).
7. Khoury, G. A., Baliban, R. C. & Floudas, C. A. Proteome-wide post-translational modification statistics: frequency analysis and curation of the swiss-prot database. *Sci. Rep.* **1,** 90 (2011).
8. Liu, Y. & Chance, M. R. Integrating phosphoproteomics in systems biology. *Comput. Struct. Biotechnol. J.* **10,** 90–97 (2014).
9. Secko, D. Protein phosphorylation: A globalr regulator of cellular activity. (2006). Available at: https://www.scq.ubc.ca/protein-phosphorylation-a-global-regulator-of-cellular-activity/. (Accessed: 19th September 2017)
10. Engholm-Keller, K. & Larsen, M. R. Technologies and challenges in large-scale phosphoproteomics. *Proteomics* **13,** 910–931 (2013).
11. Klement, E., Raffai, T. & Medzihradszky, K. F. Immobilized metal affinity chromatography optimized for the analysis of extracellular phosphorylation. *Proteomics* **16,** 1858–1862 (2016).
12. Reinders, J. & Sickmann, A. State-of-the-art in phosphoproteomics. *Proteomics* **5,** 4052–4061 (2005).

13. Zhou, H. *et al.* Robust phosphoproteome enrichment using monodisperse microsphere-based immobilized titanium (IV) ion affinity chromatography. *Nat. Protoc.* **8,** 461–80 (2013).

14. Chandramouli, K. & Qian, P.-Y. Proteomics: challenges, techniques and possibilities to overcome biological sample complexity. *Hum. Genomics Proteomics* **2009,** 22 (2009).

15. Tian, R. Exploring intercellular signaling by proteomic approaches. *Proteomics* **14,** 498–512 (2014).

16. Aebersold, R. & Mann, M. Mass spectrometry-based proteomics. *Nature* **422,** 198–207 (2003).

17. Abdallah, C., Dumas-Gaudot, E., Renaut, J. & Sergeant, K. Gel-based and gel-free quantitative proteomics approaches at a glance. *Int. J. Plant Genomics* **2012,** (2012).

18. Brunelle, J. L. & Green, R. *One-dimensional SDS-polyacrylamide gel electrophoresis (1D SDS-PAGE)*. *Methods in Enzymology* **541,** (Elsevier Inc., 2014).

19. Rabilloud, T. & Lelong, C. Two-dimensional gel electrophoresis in proteomics: A tutorial. *J. Proteomics* **74,** 1829–1841 (2011).

20. Baggerman, G., Vierstraete, E., Loof, A. De & Schoofs, L. Gel-Based Versus Gel-Free Proteomics : A Review. 669–677 (2005). doi:10.2174/138620705774962490

21. Vale, G. *et al.* An assessment of the ultrasonic probe-based enhancement of protein cleavage with immobilized trypsin. *Proteomics* **11,** 3866–3876 (2011).

22. Hustoft, H. K. *et al.* A Critical Review of Trypsin Digestion for LC-MS Based Proteomics. *Integr. Proteomics* 73–92 (2010). doi:10.5772/2473

23. Capelo, J. L. *et al.* Overview on modern approaches to speed up protein identification workflows relying on enzymatic cleavage and mass spectrometry-based techniques. *Anal. Chim. Acta* **650,** 151–159 (2009).

24. Sun, L., Li, Y., Yang, P., Zhu, G. & Dovichi, N. J. High efficiency and quantitatively reproducible protein digestion by trypsin-immobilized magnetic microspheres. *J. Chromatogr. A* **1220,** 68–74 (2012).

25. Atacan, K., Çakıroğlu, B. & Özacar, M. Efficient protein digestion using

immobilized trypsin onto tannin modified Fe3O4 magnetic nanoparticles. *Colloids Surfaces B Biointerfaces* **156,** 9–18 (2017).

26. Dunn, J., Reid, G. & Bruening, M. TECHNIQUES FOR PHOSPHOPEPTIDE ENRICHMENT PRIOR TO ANALYSIS BY MASS SPECTROMETRY. *Mass Spectrom. Rev.* **26,** 223–257 (2007).

27. Fílla, J. & Honys, D. Enrichment techniques employed in phosphoproteomics. *Amino Acids* **43,** 1025–1047 (2012).

28. Burlingame, A. L. Mass Spectrometry-based Detection and Assignment of Protein Posttranslational Modifications. *ACS Chem. Biol.* (2014). doi:10.1021/cb500904b

29. Lombardi, B., Rendell, N., Edwards, M., Katan, M. & Godovac, J. Evaluation of phosphopeptide enrichment strategies for quantitative TMT analysis of complex network dynamics in cancer-associated cell signalling. *Eur. Proteomics Assoc.* **6,** 10–15 (2015).

30. Yates, J. R., Ruse, C. I. & Nakorchevsky, A. Proteomics by mass spectrometry: approaches, advances, and applications. *Annu. Rev. Biomed. Eng.* **11,** 49–79 (2009).

31. Han, X., Aslanian, A. & Yates, J. R. Mass spectrometry for proteomics. *Curr. Opin. Chem. Biol.* **12,** 483–490 (2008).

32. Watson, J. *Intorduction to Mass Spectrometry*. (Wiley, 2009).

33. El-Aneed, A., Cohen, A. & Banoub, J. Mass Spectrometry, Review of the Basics: Electrospray, MALDI, and Commonly Used Mass Analyzers. *Appl. Spectrosc. Rev.* **44,** 210–230 (2009).

34. Lin, J.-L., Bonnichsen, M. H., Nogeh, E. U., Raftery, M. J. & Thomas, P. S. Proteomics in detection and monitoring of asthma and smoking-related lung diseases. *Expert Rev. Proteomics* **7,** 361–72 (2010).

35. Water for MALDI-ToF - Application - Water purification - Merck. (2017). Available at: http://www.merckmillipore.com/PT/en/water-purification/learning-centers/applications/organic-analysis/maldi-tof/Co6b.qB.HZYAAAFAc2YBvVBo,nav?ReferrerURL=https%3A%2F%2Fwww.google.pt%2F&bd=1. (Accessed: 19th September 2017)

36. Kim, J. Sample preparation for matrix-assisted laser desorption/ionization mass

spectrometry. *Mass Spectrom. Lett.* **6,** 27–30 (2015).

37. Bradford, M. M. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* **72,** 248–254 (1976).

38. Santos, H. M. *Bioscope Research Team - Bradford Assay Using A 96 Well Plate Protocol*. (2015).

39. Pierce Biotechnology - Thermo Scientific. *Pierce C18 Tips*. (2011).

40. Jeong, K., Kim, S. & Bandeira, N. False discovery rates in spectral identification. *BMC Bioinformatics* **13,** S2 (2012).

41. Santos, H. M. *Bioscope Research Team - Polyacrylamide Gel*. (2015).

42. Santos, H. M. *et al.* A novel 18O inverse labeling-based workflow for accurate bottom-up mass spectrometry quantification of proteins separated by gel electrophoresis. *Electrophoresis* **31,** 3407–3419 (2010).

43. Oliveira, E. *et al.* Proteomics analysis of the peritoneal dialysate effluent reveals the presence of calcium-regulation proteins and acute inflammatory response. *Clin. Proteomics* **11,** 17 (2014).

44. Gauci, V. J., Padula, M. P. & Coorssen, J. R. Coomassie blue staining for high sensitivity gel-based proteomics. *J. Proteomics* **90,** 96–106 (2013).

45. He, G. & Pan, Q. Synthesis of polystyrene and polystyrene/poly(methyl methacrylate) nanoparticles. *Macromol. Rapid Commun.* **25,** 1545–1548 (2004).

46. Bovine Serum Albumins - Albumin Sigma Aldrich. Available at: http://www.sigmaaldrich.com/life-science/biochemicals/biochemical-products.html?TablePage=103994915. (Accessed: 19th September 2017)

47. Ziegel, E. R. *Statistics and Chemometrics for Analytical Chemistry*. *Technometrics* **46,** (2004).

48. Brandup, J. & Immergut, E. H. *Polymer Handbook*. (Wiley, 1989).

49. Flory, P. *Principles pf Polymer Chemistry*. (1972).

50. Mark, J. E. *Polymer data handbook*. *Journal of the American Chemical Society* **131,** (1999).

51. Shen, S., Sudol, E. D. & El-Aasser, M. S. Control of particle size in dispersion

polymerization of methyl methacrylate. *J. Polym. Sci. Part A Polym. Chem.* **31,** 1393–1402 (2003).

52. Yun, D. S., Lee, H. S., Jang, H. G. & Yoo, J. W. Controlling size and distribution for nano-sized polystyrene spheres. *Bull. Korean Chem. Soc.* **31,** 1345–1348 (2010).

53. MALVERN. Zetasizer Nano Series. *Malvern Instruments Ltd* 20 (2014).

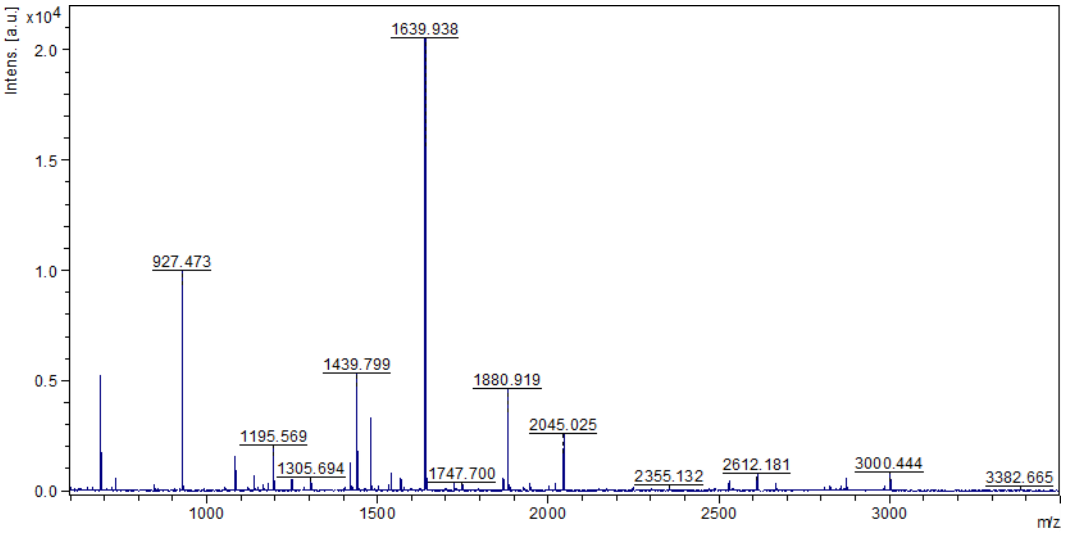
54. Uniprot. CSN1S1 - Alpha-S1-casein precursor - Bos taurus (Bovine) - CSN1S1 gene & protein. Available at: http://www.uniprot.org/uniprot/P02662. (Accessed: 19th September 2017)

55. Uniprot. CSN1S2 - Alpha-S2-casein precursor - Bos taurus (Bovine) - CSN1S2 gene & protein. Available at: http://www.uniprot.org/uniprot/P02662. (Accessed: 19th September 2017)

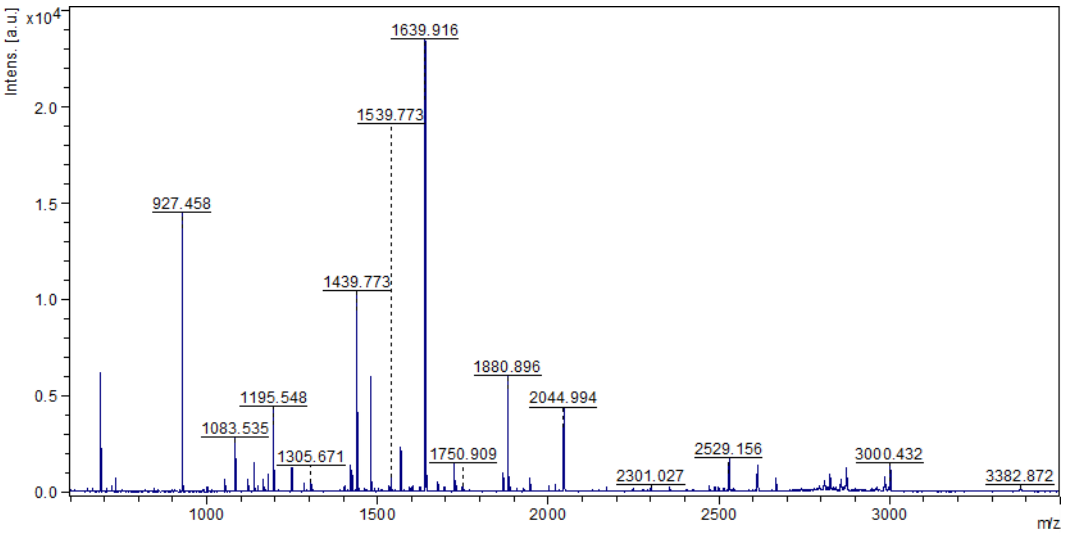
56. Yu, Z. *et al.* Preparation of monodisperse immobilized Ti4+ affinity chromatography microspheres for specific enrichment of phosphopeptides. *Anal. Chim. Acta* **636,** 34–41 (2009).

Annexes

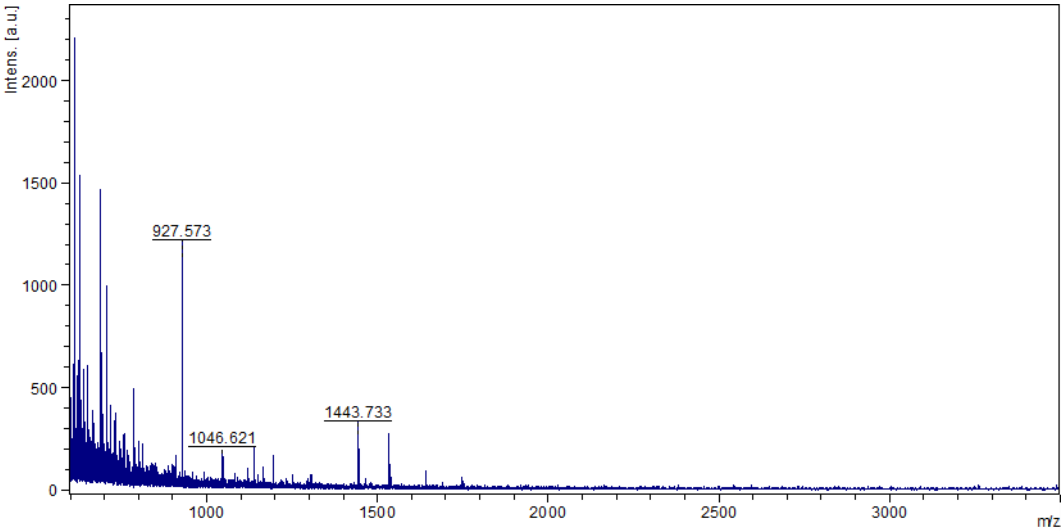
Annex I: MS spectrums



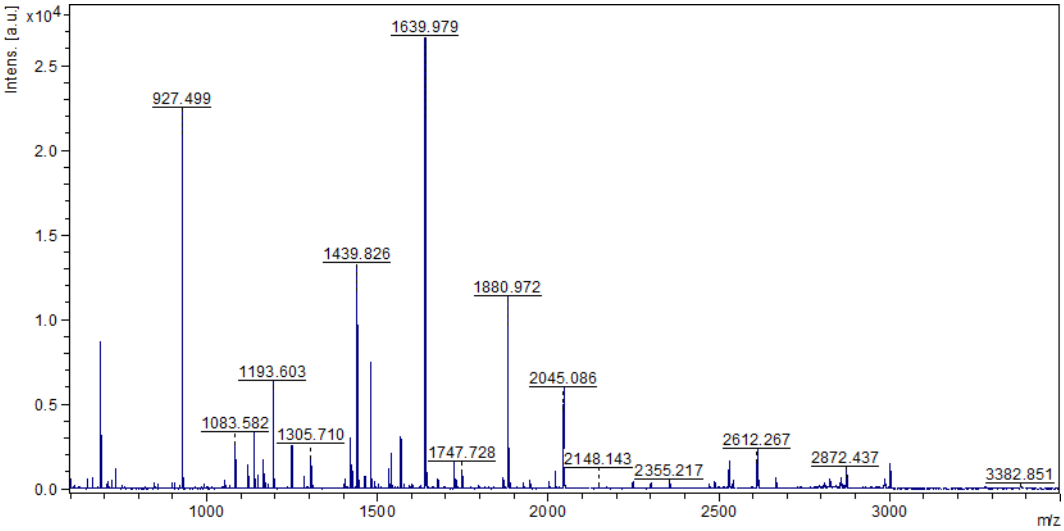
MS spectrum of the digestion of 5µg of BSA with 0.005µg/µL of immobilized trypsin nanoparticles.



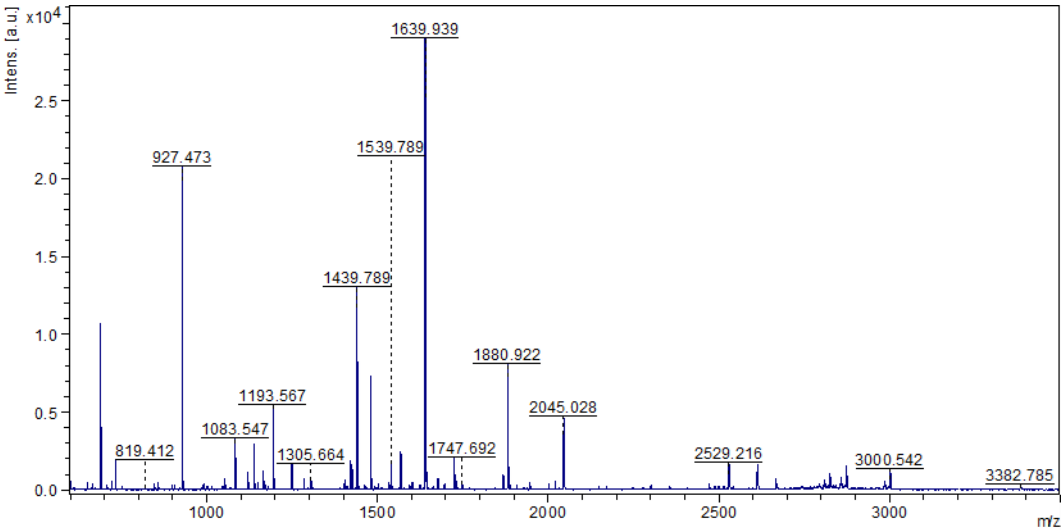
MS spectrum of the digestion of 10µg of BSA with 0.005µg/µL of immobilized trypsin nanoparticles.



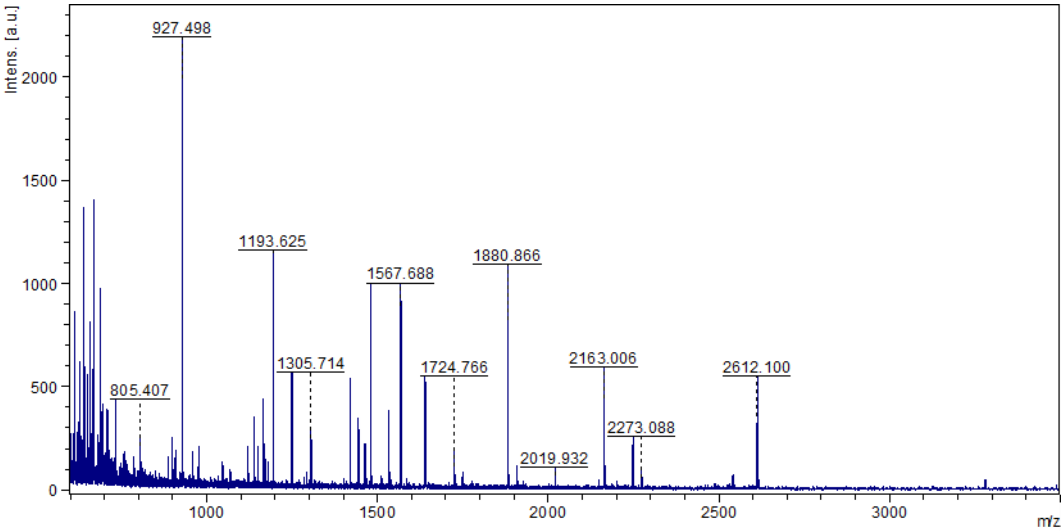
MS spectrum of the digestion of 1µg of BSA with 0.01µg/µL of immobilized trypsin nanoparticles.



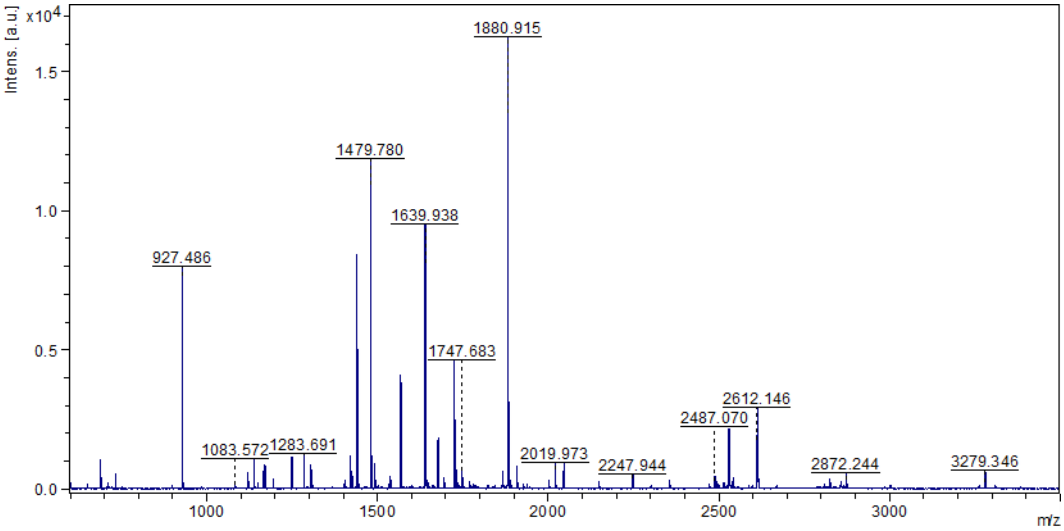
MS spectrum of the digestion of 5µg of BSA with 0.01µg/µL of immobilized trypsin nanoparticles.



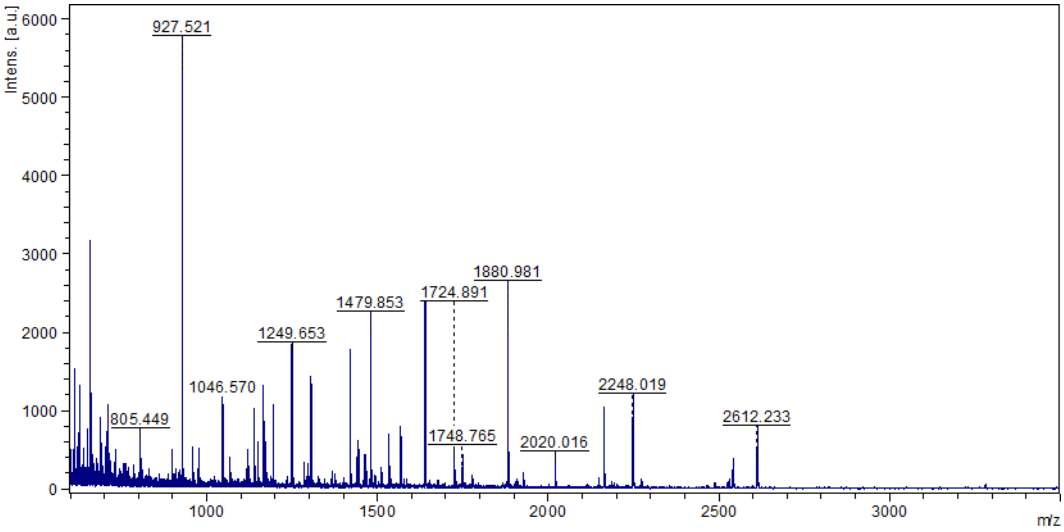
MS spectrum of the digestion of 10µg of BSA with 0.01µg/µL of immobilized trypsin nanoparticles.



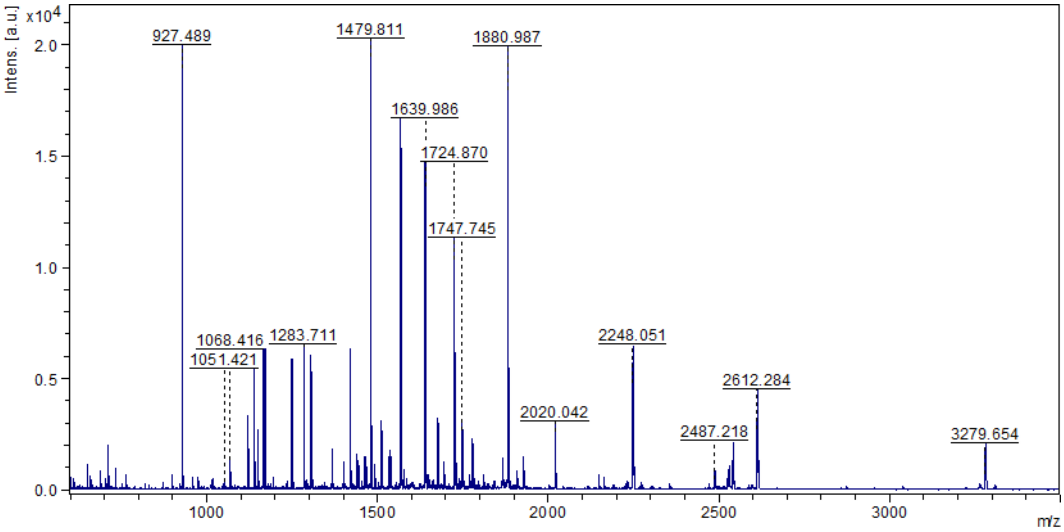
MS spectrum of the digestion of 1µg of BSA with 0.1µg/µL of immobilized trypsin nanoparticles.



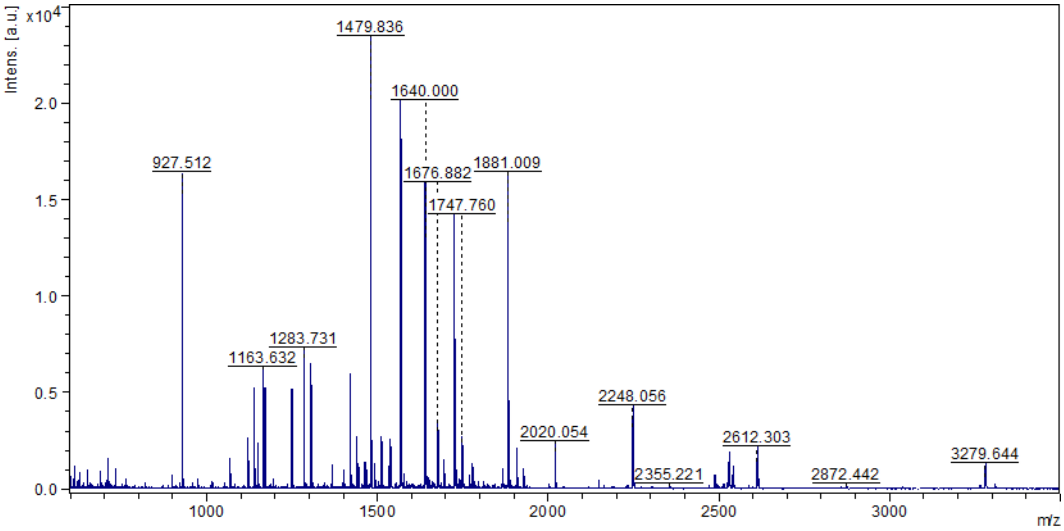
MS spectrum of the digestion of 10µg of BSA with 0.1µg/µL of immobilized trypsin nanoparticles.



MS spectrum of the digestion of 1µg of BSA with 0.5µg/µL of immobilized trypsin nanoparticles.

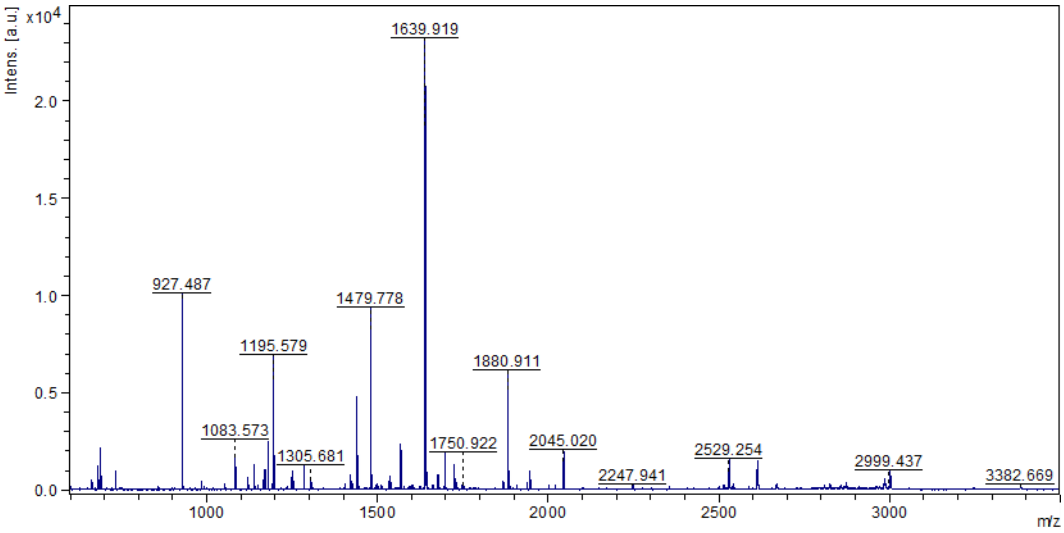


MS spectrum of the digestion of 5µg of BSA with 0.5µg/µL of immobilized trypsin nanoparticles.

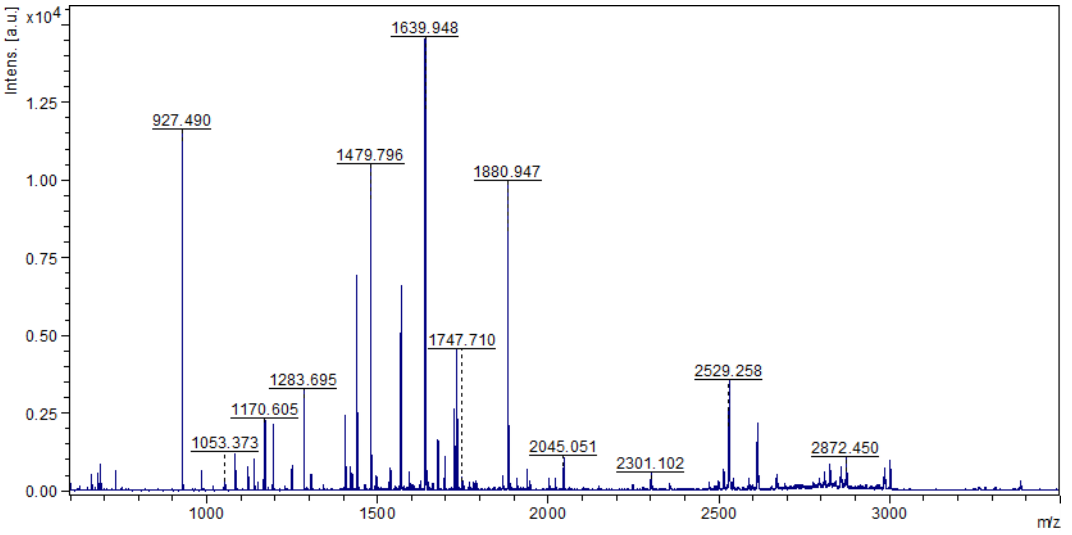


MS spectrum of the digestion of 10µg of BSA with 0.5µg/µL of immobilized trypsin nanoparticles.

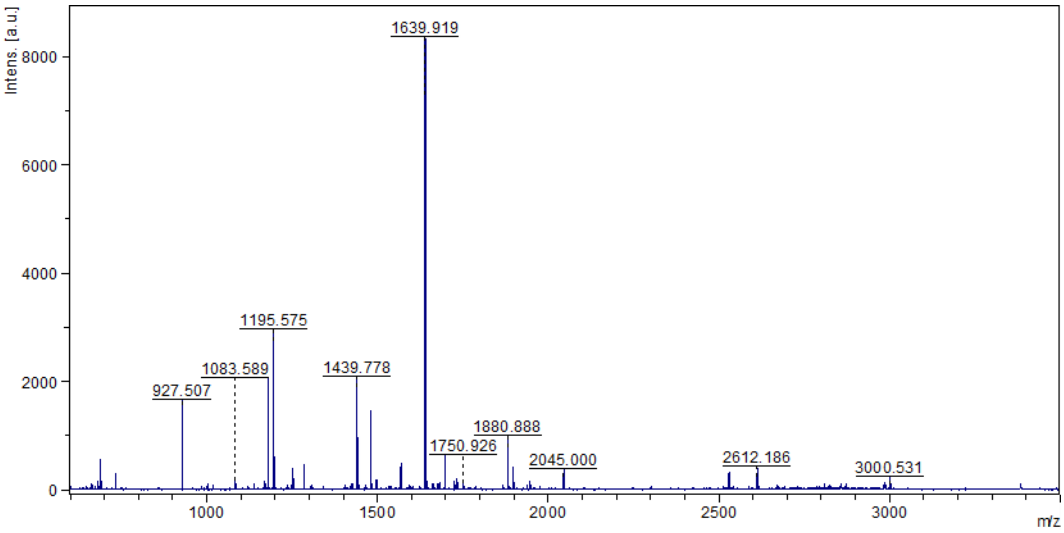
Annex II:



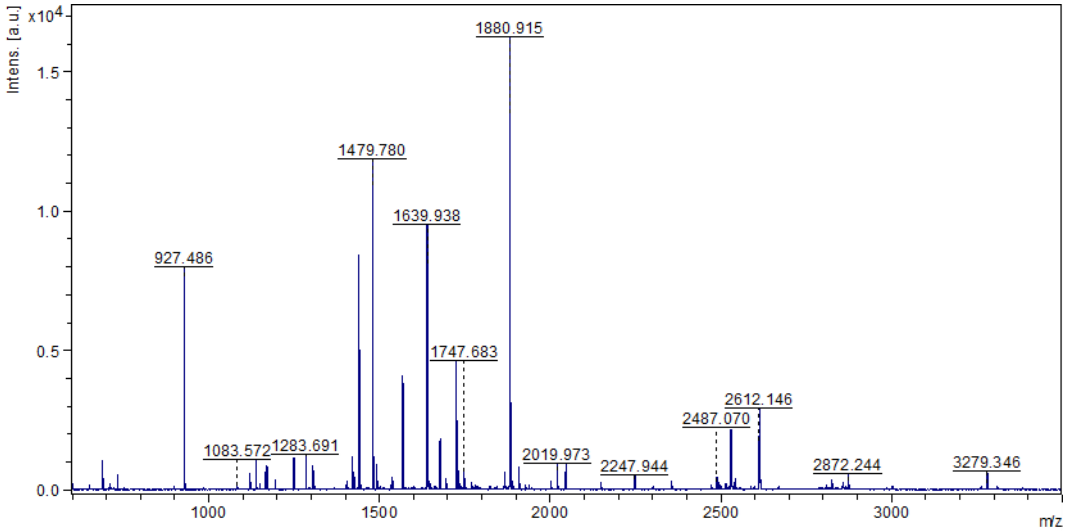
MS spectrum of the digestion of 10µg of BSA with 0.1µg/µL of immobilized trypsin nanoparticles – batch1.



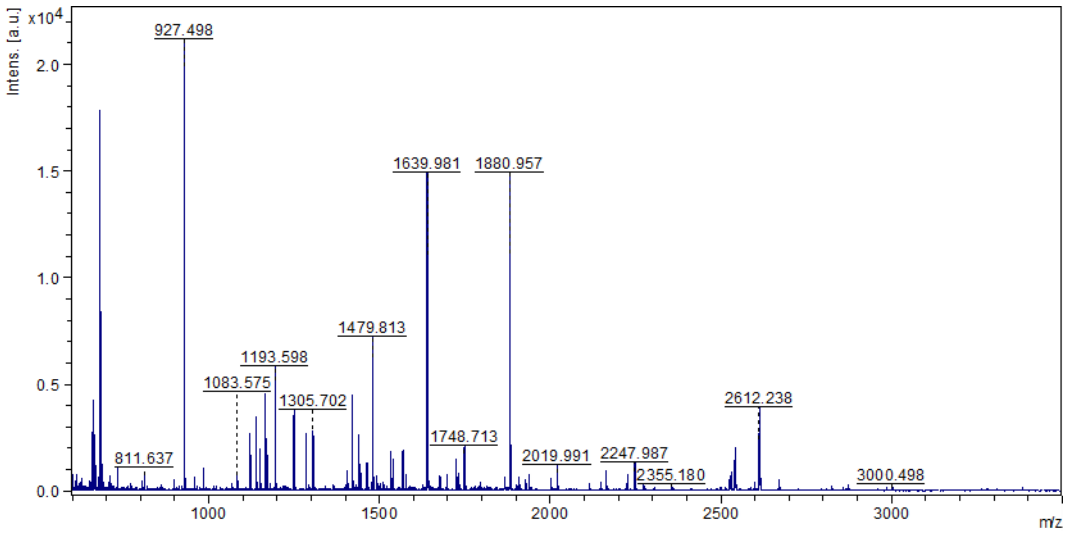
MS spectrum of the digestion of 10µg of BSA with 0.1µg/µL of immobilized trypsin nanoparticles – batch2.



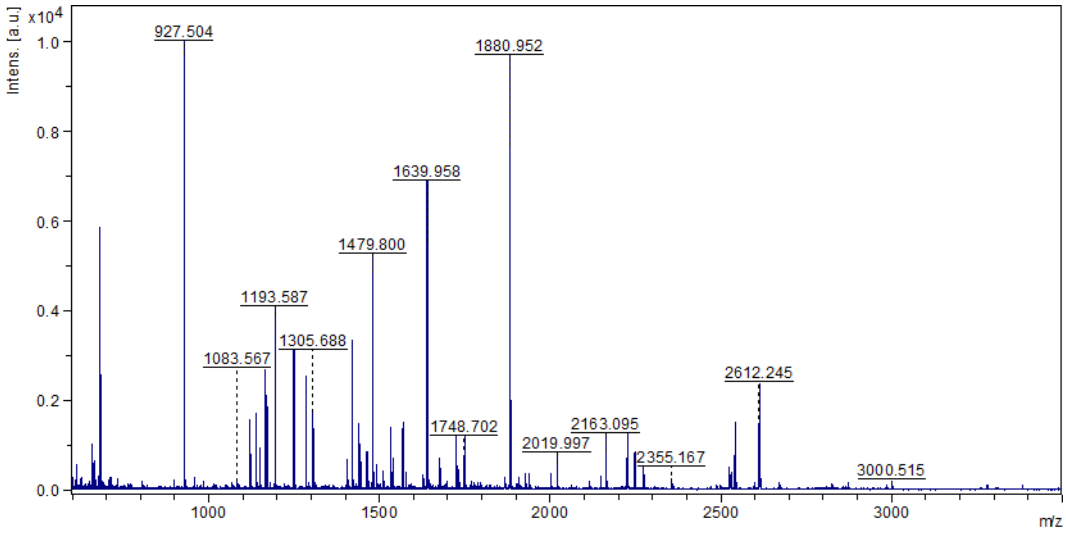
MS spectrum of the digestion of 10µg of BSA with 0.1µg/µL of immobilized trypsin nanoparticles – batch3.



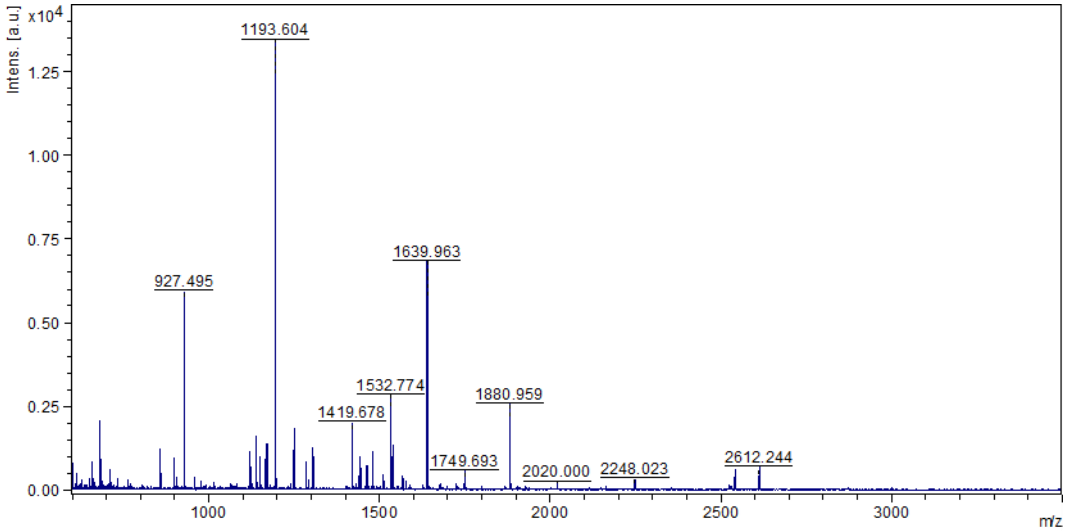
MS spectrum of the digestion of 10µg of BSA with 0.1µg/µL of immobilized trypsin nanoparticles – batch4.



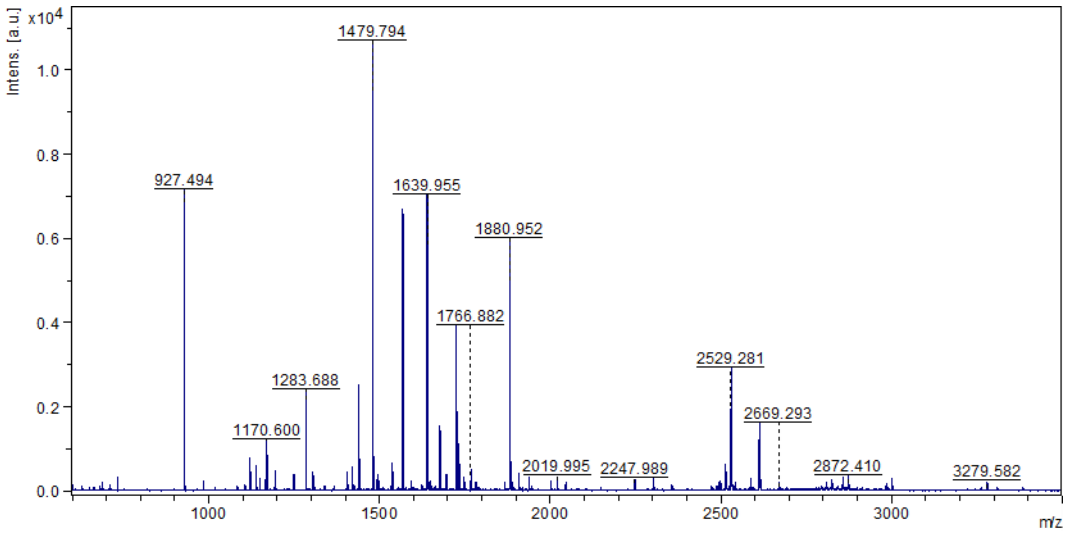
MS spectrum of the digestion of 1µg of BSA with 0.5µg/µL of immobilized trypsin nanoparticles – batch1.



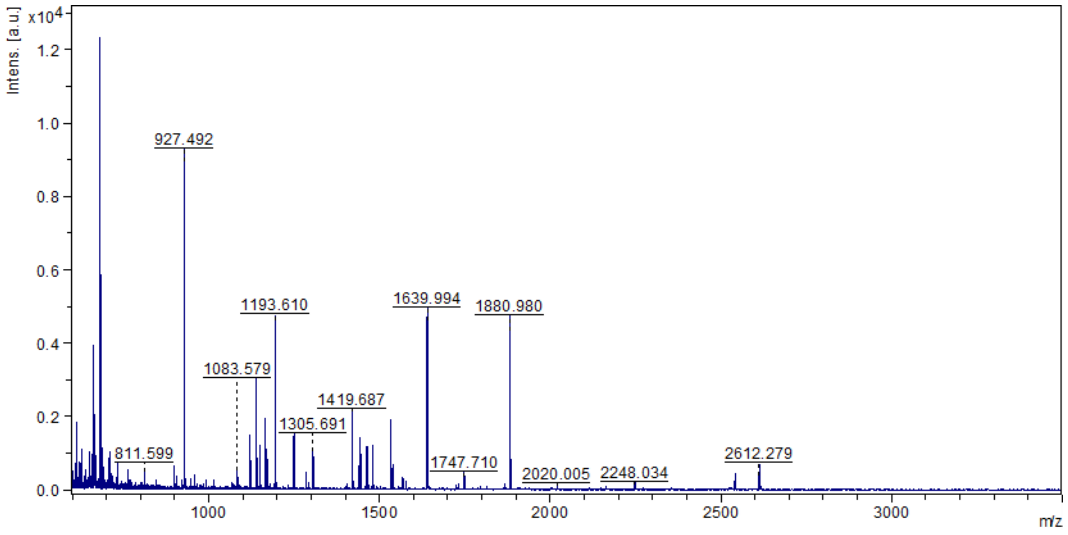
MS spectrum of the digestion of 1µg of BSA with 0.5µg/µL of immobilized trypsin nanoparticles – batch2.



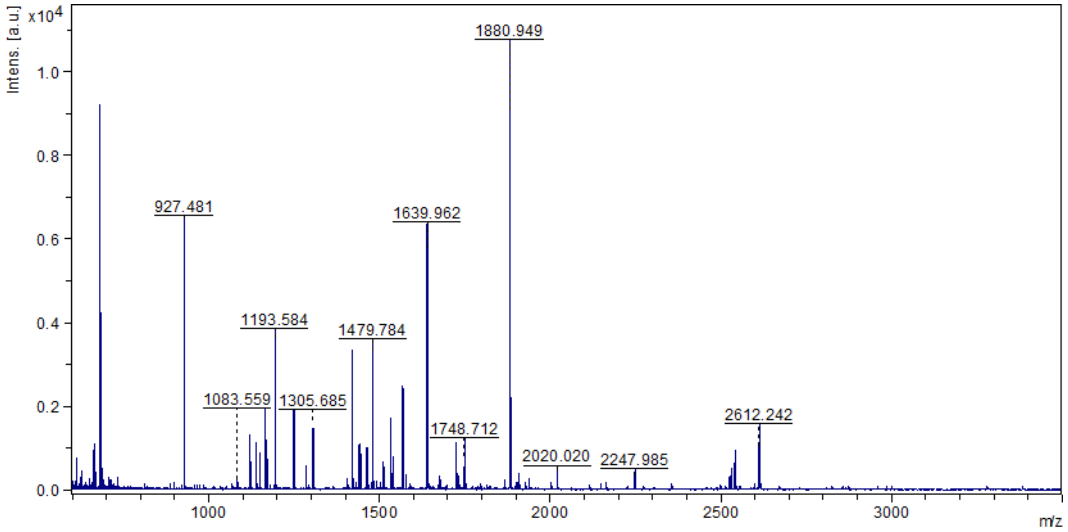
MS spectrum of the digestion of 1µg of BSA with 0.5µg/µL of immobilized trypsin nanoparticles – batch3.



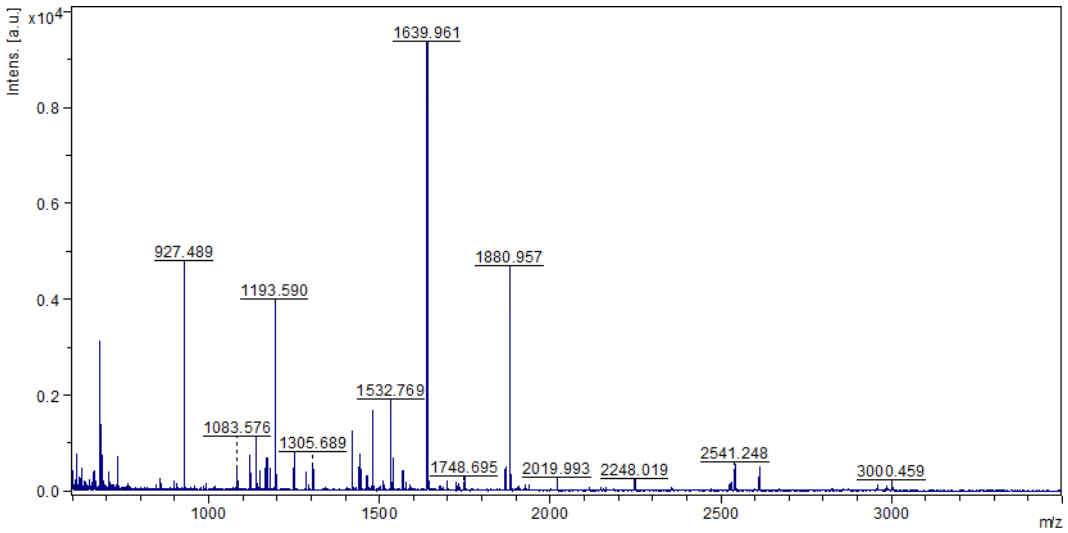
MS spectrum of the digestion of 1µg of BSA with 0.5µg/µL of immobilized trypsin nanoparticles – batch4.



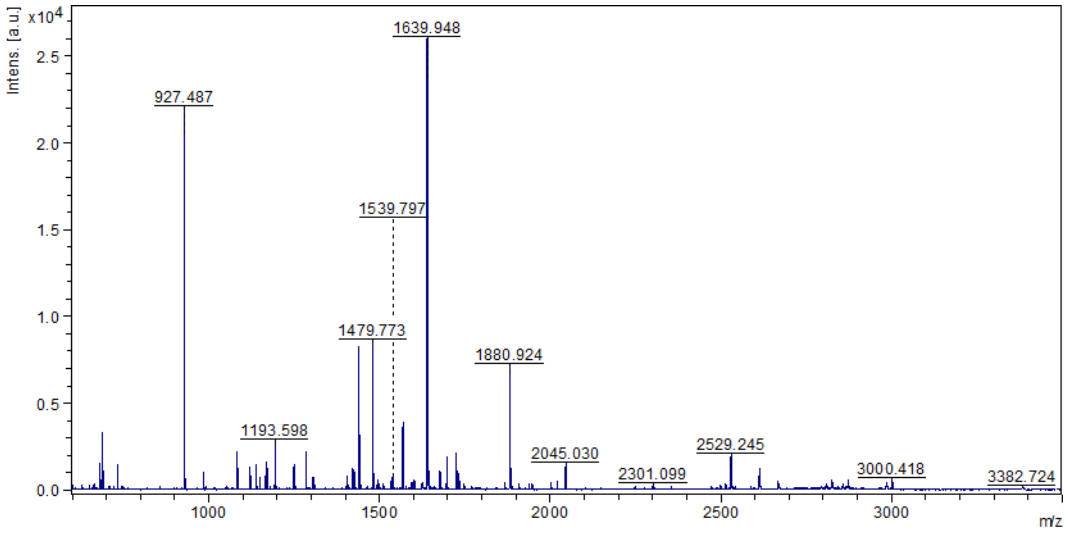
MS spectrum of the digestion of 1µg of BSA with 0.1µg/µL of immobilized trypsin nanoparticles – batch1.



MS spectrum of the digestion of 1µg of BSA with 0.1µg/µL of immobilized trypsin nanoparticles – batch2.



MS spectrum of the digestion of 1µg of BSA with 0.1µg/µL of immobilized trypsin nanoparticles – batch3.



MS spectrum of the digestion of 1µg of BSA with 0.1µg/µL of immobilized trypsin nanoparticles – batch4.

Annex III: T-Test equations and results

$$t = \frac{x_1 - x_2}{s * \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

Equation 1: T-test equation where t has n_1+n_2-2 degrees of freedom, x_1 and x_2 are the means of sample 1 and 2, s is the estimated standard deviation between the two means and n_1 and n_2 are the size of the samples 1 and 2.

$$s = \frac{(n_1 - 1) * s_1^2 + (n_2 - 1) * s_2^2}{(n_1 + n_2 - 2)}$$

Equation 2: Estimative of the standard deviation between 2 means, where, s is the estimated standard deviation between the two means, s_1 and s_2 are the standard deviations of sample 1 and 2 and n_1 and n_2 are the size of the samples 1 and 2.

$$F = \frac{s_1^2}{s_2^2}$$

Equation 3: Fischer's test, where s_1 and s_2 are the standard deviations of sample 1 and 2. subscripts 1 and 2 are assigned in the equation so that F is always ≥1. The degrees of freedom of the numerator and the denominator are n_1-1 and n_2-1 respectively.

$$t = \frac{x_1 - x_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

Equation 4: T-test equation where x_1 and x_2 are the means of sample 1 and 2, s_1 and s_2 are the standard deviations of sample 1 and 2 and n_1 and n_2 are the size of the samples 1 and 2. With the degrees of freedom calculated by Equation 5.

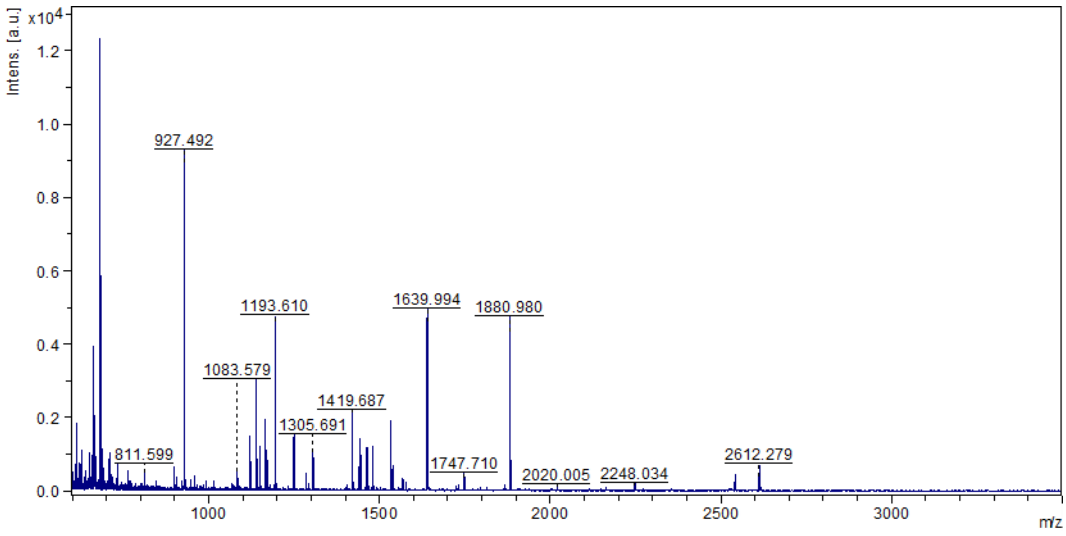
$$degrees\ of\ freedom = \frac{(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2})^2}{\frac{s_1^4}{n_1^2 * (n_1 - 1)} + \frac{s_2^4}{n_2^2 * (n_2 - 1)}}$$

Equation 5: Degrees of freedom where s_1 and s_2 are the standard deviations of sample 1 and 2 and n_1 and n_2 are the size of the samples 1 and 2.

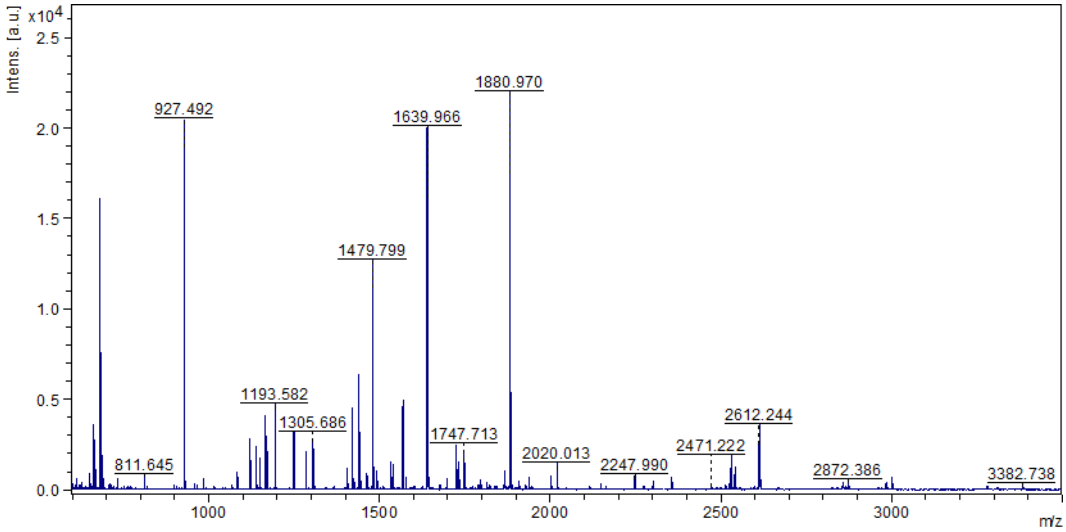
Test 1	t,6,5%	t obtained	Test 2	t,6,5%	t obtained
T1/T2	2.45	0.42	T1/T2	2.45	0.82
T1/T3	2.45	0.56	T1/T3	2.45	0.61
T1/T4	2.45	0.52	T1/T4	2.45	1.04
T2/T3	2.45	0.85	T2/T3	2.45	2.20
T2/T4	2.45	1.08	T2/T4	2.45	2.00
T3/T4	2.45	0.45	T3/T4	2.45	0.86

Test 3	t,6,5%	t obtained	Test 4	t,6,5%	t obtained
T1/T2	2.45	0.32	T1/T2	2.45	2.00
T1/T3	2.45	0.67	T1/T3	2.45	1.45
T1/T4	2.45	1.53	T1/T4	2.45	0.75
T2/T3	2.45	0.47	T2/T3	2.45	0.51
T2/T4	2.45	1.56	T2/T4	2.45	2.35
T3/T4	2.45	1.30	T3/T4	2.45	2.08

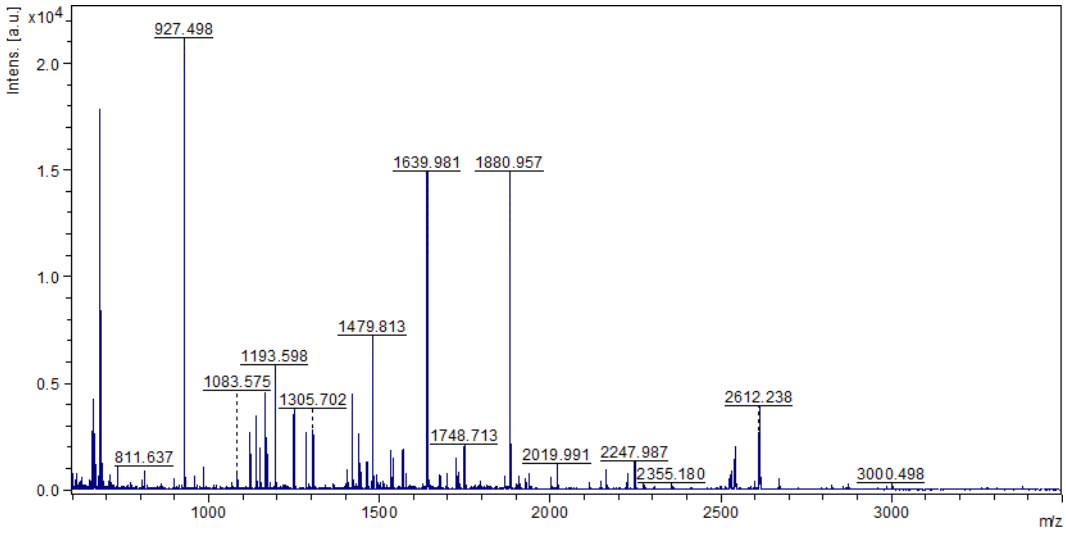
Annex IV:



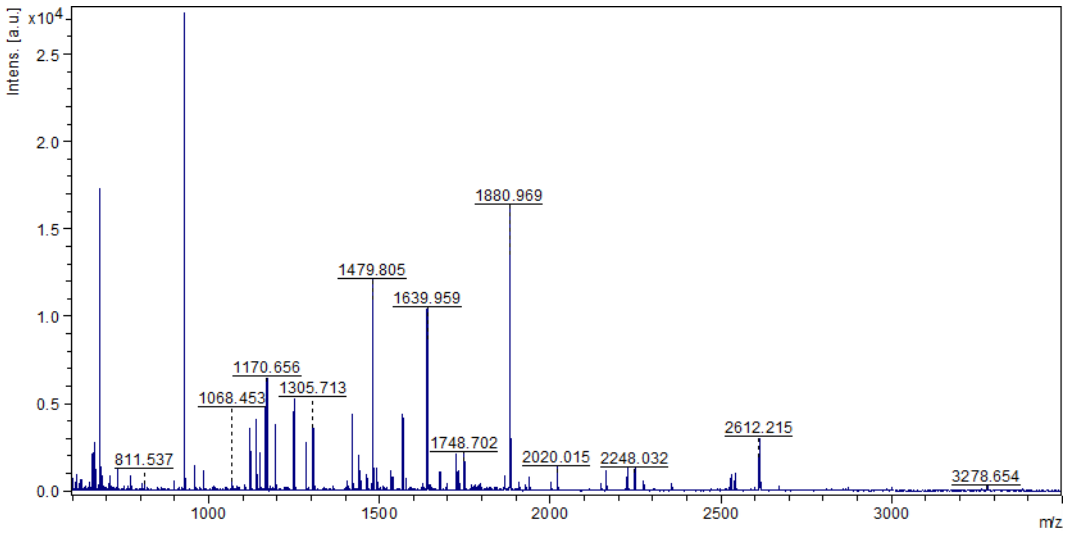
MS spectrum of the digestion of 1µg of BSA with 12µg of immobilized trypsin nanoparticles, non-stirring while digestion.



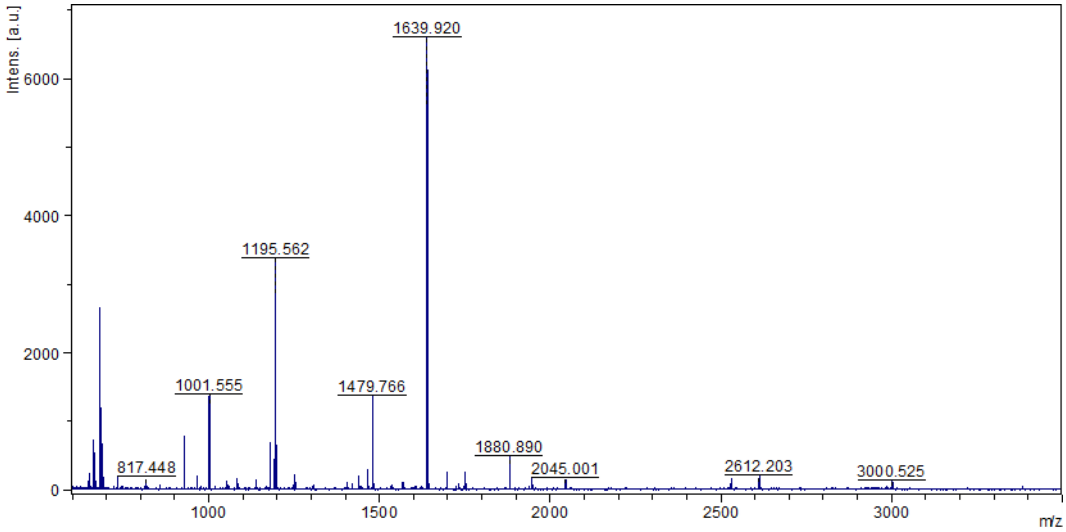
MS spectrum of the digestion of 1µg of BSA with 12µg of immobilized trypsin nanoparticles, stirring while digestion.



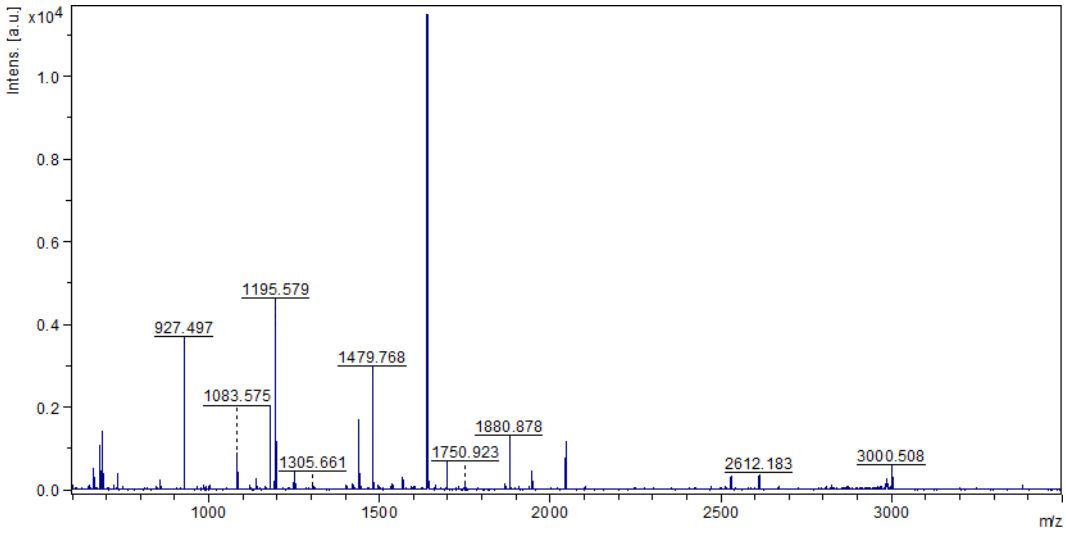
MS spectrum of the digestion of 1µg of BSA with 60µg of immobilized trypsin nanoparticles, non-stirring while digestion.



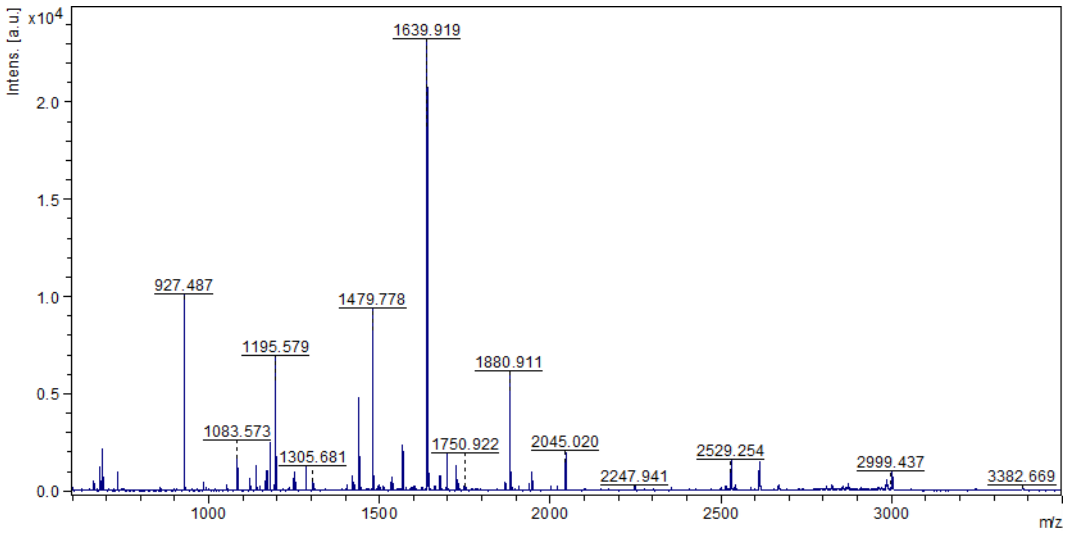
MS spectrum of the digestion of 1µg of BSA with 60µg of immobilized trypsin nanoparticles, stirring while digestion.



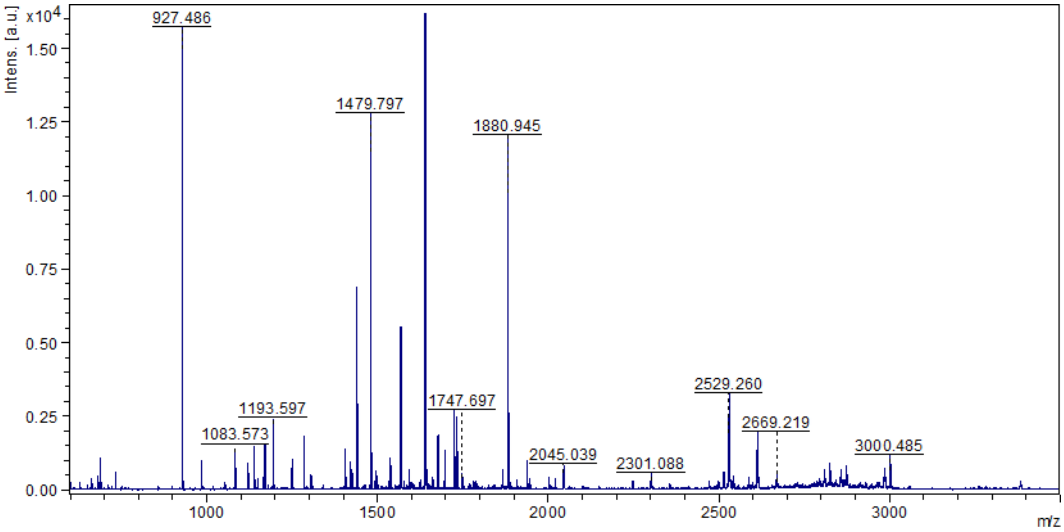
MS spectrum of the digestion of 10µg of BSA with 0.6µg of immobilized trypsin nanoparticles, non-stirring while digestion.



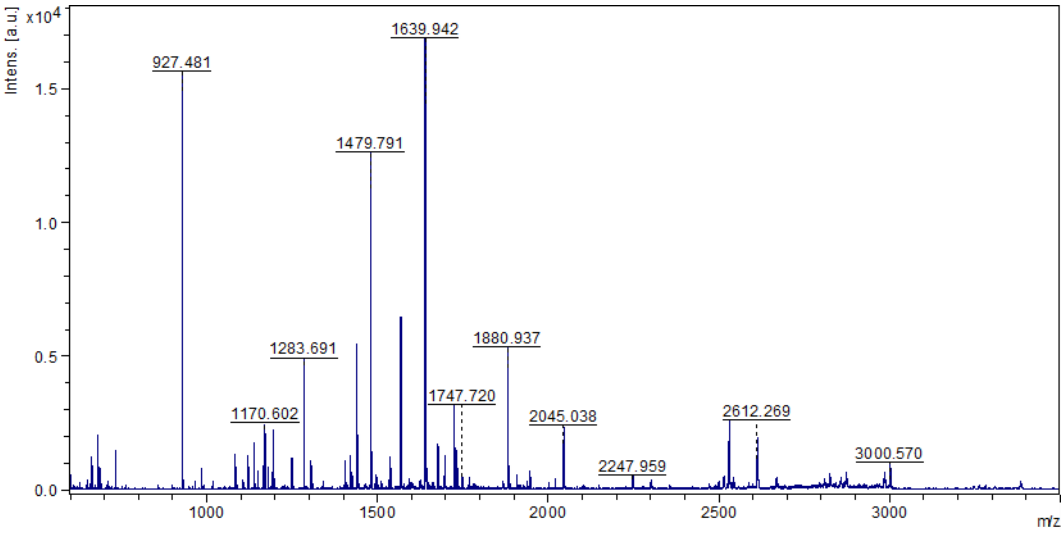
MS spectrum of the digestion of 10µg of BSA with 0.6µg of immobilized trypsin nanoparticles, stirring while digestion.



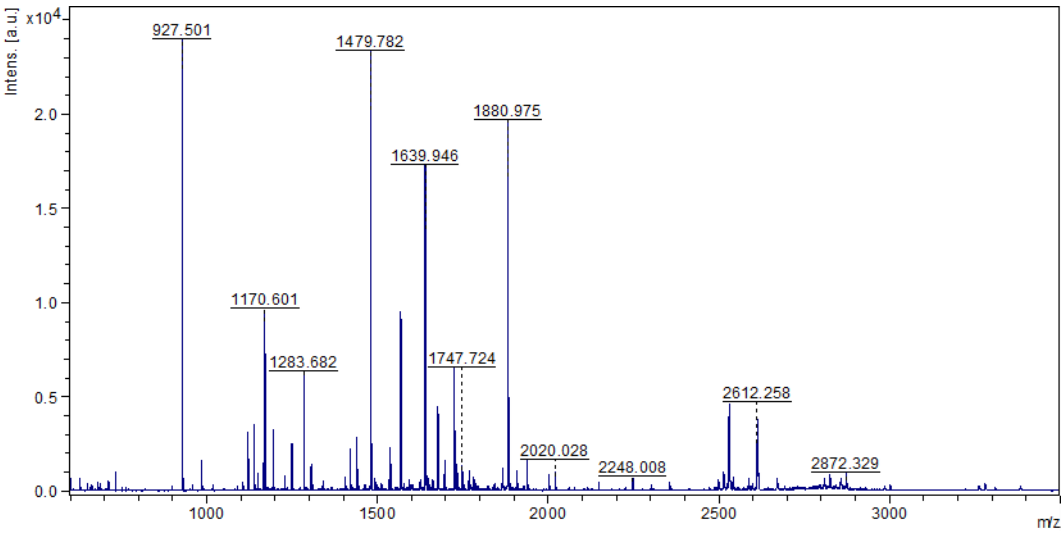
MS spectrum of the digestion of 10µg of BSA with 12µg of immobilized trypsin nanoparticles, non-stirring while digestion.



MS spectrum of the digestion of 10µg of BSA with 12µg of immobilized trypsin nanoparticles, stirring while digestion.

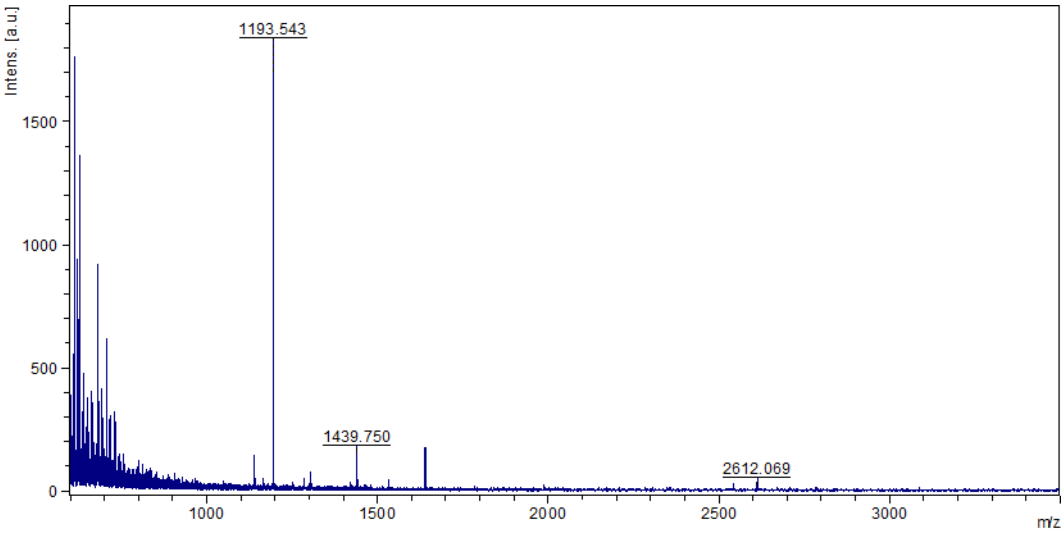


MS spectrum of the digestion of 10µg of BSA with 60µg of immobilized trypsin nanoparticles, non-stirring while digestion.

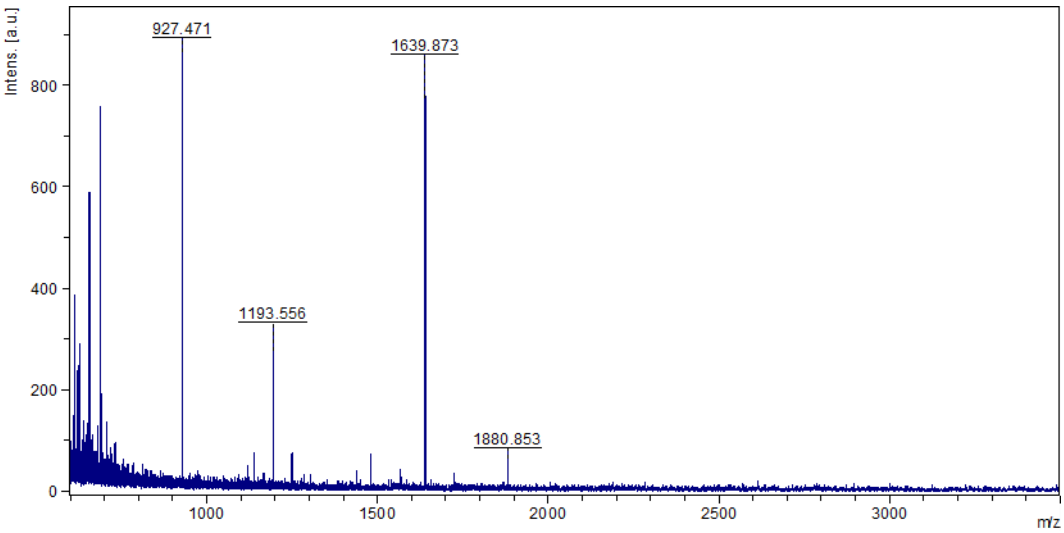


MS spectrum of the digestion of 10µg of BSA with 60µg of immobilized trypsin nanoparticles, stirring while digestion.

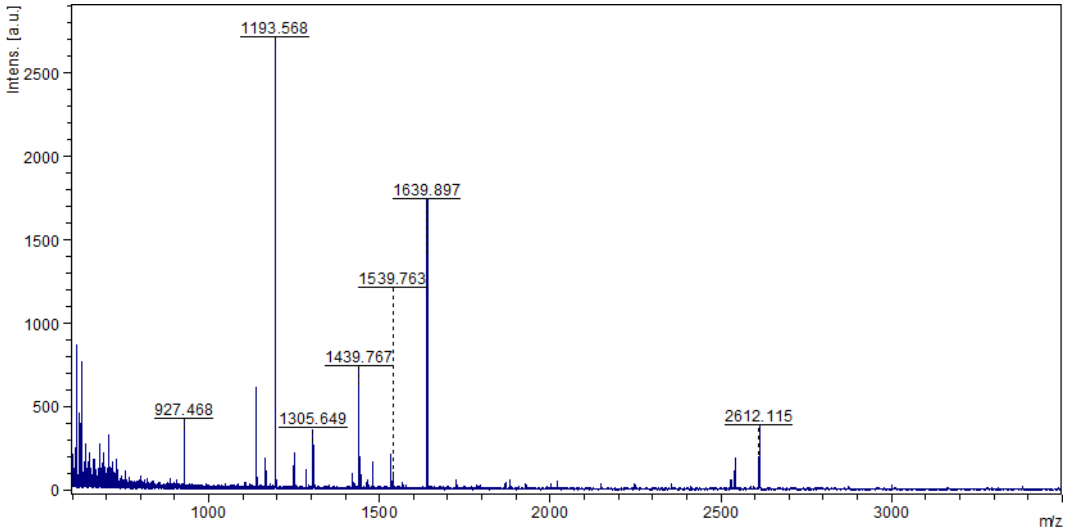
Annex V:



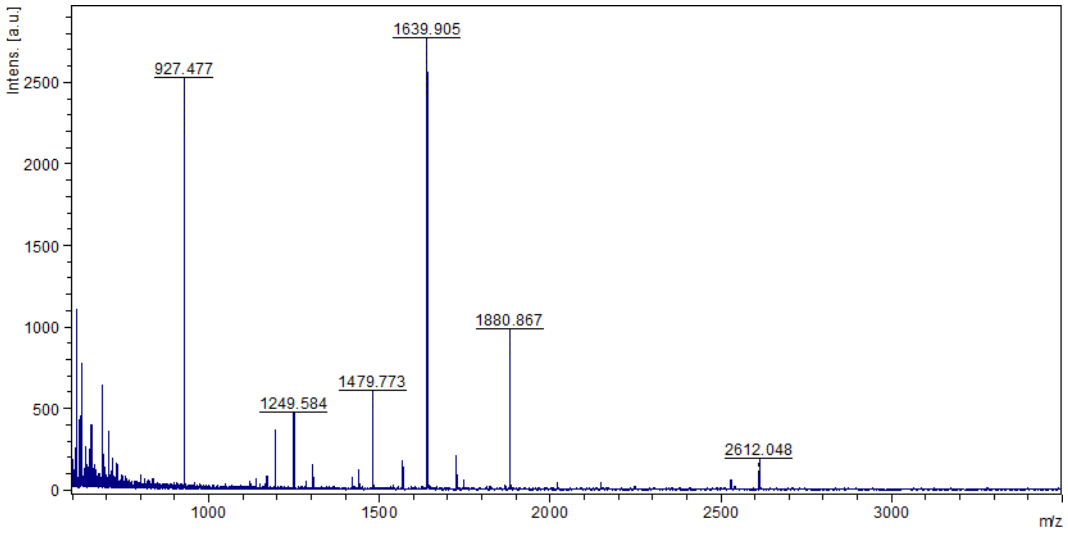
MS spectrum of the digestion of 10µg of BSA and 0.5µg/µL of immobilized trypsin nanoparticles, 30 minutes of digestion time.



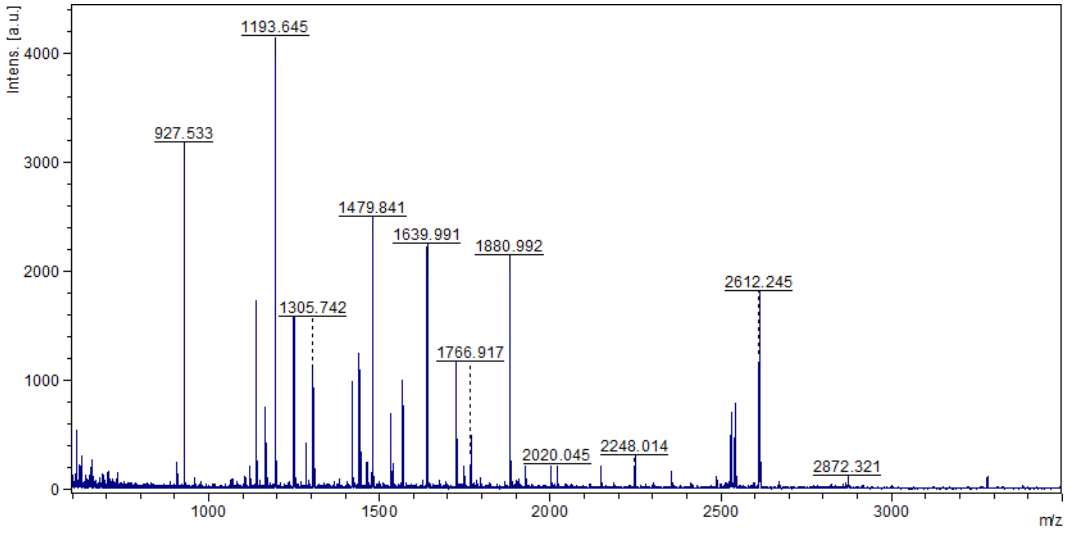
MS spectrum of the digestion of 10µg of BSA and 0.5µg/µL of commercial immobilized trypsin microparticles, 30 minutes of digestion time.



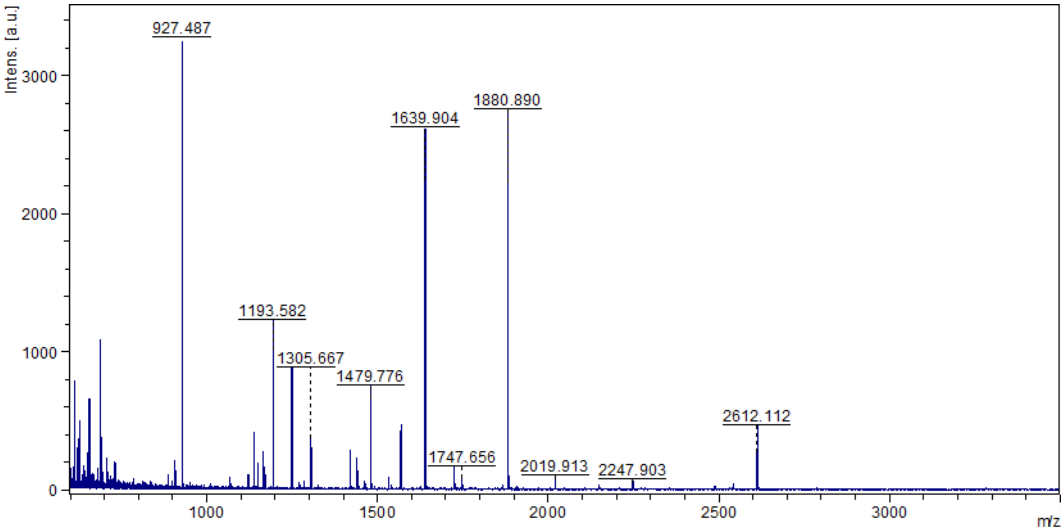
MS spectrum of the digestion of 10µg of BSA and 0.5µg/µL of immobilized trypsin nanoparticles, 1 hours of digestion time.



MS spectrum of the digestion of 10µg of BSA and 0.5µg/µL of commercial immobilized trypsin microparticles, 1 hours of digestion time.



MS spectrum of the digestion of 10µg of BSA and 0.5µg/µL of immobilized trypsin microparticles, 6 hours of digestion time.



MS spectrum of the digestion of 10µg of BSA and 0.5µg/µL of commercial immobilized trypsin microparticles, 6 hours of digestion time.