Università degli Studi di Napoli Federico II

Ph.D. Thesis in Chemical Sciences

XXXIII Cycle

# DEVELOPMENT OF A NEW METHOD FOR IDENTIFICATION OF FSH GLYCOFORMS INVOLVED IN HUMAN FERTILITY

Chiara Melchiorre

Tutor                                                    Supervisor

Prof. A. Amoresano                          Prof. A. Silipo
Dr. A. Palmese
Dr. I.C. Guerrera


Coordinator

Prof. A. Lombardi

# Acknowledgments

This PhD has been a great formative experience for me and it would not have been possible without the support and supervision that I received from many people.

First and foremost, I am extremely grateful to my supervisors:

Prof. **Angela Amoresano** from University of Naples Federico II, for her invaluable advice, her support and patience throughout my PhD. Her scientific suggestions had been fundamental for the realization of the entire experimental design during this PhD project.

Dr. **Ida Chiara Guerrera** Head of the Proteomics Platform at SFR Necker (PPN) in Paris for her help and advice during my period abroad. Thanks to her expertise I could learn the most advanced biomolecular mass spectrometry methodologies (PRM-MS). In her laboratory I had the chance to work in an international and stimulating team.

Dr. **Angelo Palmese** Head of Characterization & Innovative Analytics Unit from Healthcare Business of Merck for his supervision, support and tutelage during the internship in Merck group company. Thanks to his support I could design and characterize the different glycovariants described in this thesis.

I would like to acknowledge my examinator Prof. **Alba Silipo** from University of Federico II of Naples for her continuous advices and suggestions during my PhD study.

Additionally, I would like to express my gratitude to Prof. **Andrea Carpentieri** from University Federico II of Naples for his mentoring, his invaluable support and for encouraged me in day-life during all PhD period.

Least but not last, I would like to thank the PhD coordinator, Professor **Angelina Lombardi** for her guidance over all the time of the doctorate.

I would like to thank my friends, lab mates, colleagues and research team BIMA lab, SCHI & SCHII labs and PPN lab for the precious time spent together in the lab, and in social settings. My gratitude also goes to my family and friends for their encouragement and support throughout my studies.

# TABLE OF CONTENT

**Abstract**

Follicle Stimulating Hormone (FSH) is one of the most important glycohormones involved in human fertility and it is widely used to assess the ovarian reserve in women for *in vitro* fertilization. FSH is a heterodimeric glycohormone composed by two polypeptide subunits, the α-subunit (common to all gonadotropins) and a FSH specific β-subunit. The molecule has 4 glycosylation site, Asn52-Asn78 of α-subunit and Asn7-Asn24 of β-subunit. Each glycosite hosts a wide microheterogeneity of glycans, thus generating a heterogeneous population of different glycoforms.

The biological activity of human hFSH is highly regulated by its glycosylation, but the role of the different glycoforms in FSH mechanism of action is still poorly understood. The development of new methodologies aimed at characterizing and identifying the site-specific glycosidic profile of hFSH could contribute to understand how the glycosylation pattern modulates hFSH function.

Glycomic studies on serum hFSH are very challenging because of the huge heterogeneity of glycoforms and their very low concentration in serum. The current reference method for serum hFSH dosage is ELISA, but it fails in distinguishing proteoforms because of the lack of specific antibodies. A combination of mass spectrometry approaches was used here to characterise recombinant r-FSH and ultimately quantify endogenous hFSH proteoforms in serum.

Site-specific glycosylation profiles of different commercialised r-FSH molecules (GONAL-F ®, BEMFOLA, OVALEAP, FOSTIMON) were determined by untargeted proteomics and glycan analysis. The MS analysis led to the identification of at least 5 known glycoforms for each of 4 glycosylation sites of the molecule.

In order to find higher efficiency r-FSH isoforms of the molecule, artificial glycovariant were produced by stress conditions (temperature, pH, etc.) and

analysed in the same way. Glycosylation of the molecule subject to pH stress condition shows a slight decrease in the sialic acid content as well as in hyper O-Acetylated species. AEX-pH acidic fraction was characterised by highly sialylated end-capping species, while the basic fraction showed the presence of neutral galactosylated species and a lower content of sialylated moieties. Totally de-sialylated and partially de-galactosylated FSH glycovariants were obtained by enzymatic treatments. *In vivo* and *in vitro* experiments will be conducted by Merck (www.merckgroup.com) to determine the efficiency of these glycovariants in fertility therapy.

On the basis of mass spectral r-FSH characterisation, two innovative targeted MS methods were developed to identify and quantify the global amount of hFSH as well as different glycoforms in woman serum. A Multiple Reaction Monitoring (MRM) MS-based method for the absolute quantification of circulating hFSH molecule was optimised, and it was successfully applied to the analysis of 10 women sera samples of different ages showing a great variability of the amount of the hormone. Results show comparable sensitivity (>1 ng/ml) and higher selectivity and specificity than conventional ELISA assays, and therefore this MRM-MS method could be proposed as an alternative for the quantification of hFSH. As for glycoforms detection and dosage, a method based on high-resolution mass spectrometry in Parallel Reaction Monitoring (PRM-MS) mode was developed to monitor the glycosylation profile of hFSH in pooled sera sample. The sensitivity was boosted by an optimised sample preparation involving immunoaffinity purification of hFSH from serum. The 9 most abundant hFSH site-specific glycoforms were identified in a very complex matrix such as serum. Moreover, a relative quantification of the identified glycoforms was obtained by post acquisition data processing using Skyline software. In both recombinant and serum hFSH the same most abundant glycoforms were found, the biantennary fully sialylated (A2G2S2) on α-Asn52, the mono-sialylated (A2G2S1) on α-Asn78 and fucosylated biantennary (FA2G2S2 and FA2G2S1)

structures on β-Asn 24. Glycopeptide spectra containing β-Asn 7 glycosite didn't pass our quality criteria for quantitative analysis.

Although the importance of glycosylation in the biological activity of FSH is well documented, this is the first study reporting site-specific characterization of circulating h-FSH glycoforms by MS. The proposed strategy, if applied to a large cohort of women, can be useful to investigate the existing relations between the hFSH glycan moieties and the hormone function. Such knowledge will bring new insights on the complex mechanisms in which hFSH glycoforms are involved and, ultimately, will improve its usage in infertility treatments. Furthermore, these targeted MS methods can be easily implemented for a simultaneous analysis of several serum proteins, making the PRM/MRM-MS methodology advantageous in terms of time and cost.

# Chapter 1    Introduction

## 1.1    Glycomics and glycoproteomics

'Omics' studies are aimed at the complete understanding and profiling of genes (genomics), mRNA (transcriptomics), proteins (proteomics) and metabolites (metabolomics) in a specific biological system in a non-targeted and impartial way [1].

Although many scientific discoveries have emerged from genomic and transcriptomic approaches, this information still does not provide a complete picture of the physiology of a cell or organism. The proteins expressed by the cell, collectively termed "proteome," perform many of the cell's functions. Most eukaryotic proteins are post-translationally modified (e.g., by phosphorylation, oxidation, ubiquitination, lipidation, or glycosylation). These modifications, combined with alternative splicing in eukaryotes, make the proteome considerably more complex than the transcriptome. Direct characterisation of the proteome is required to understand both its complexity and its global functions. The analysis of all proteins expressed by cells, tissues, or organisms is referred to as "proteomics". Unlike the genome, which is constant (every cell of an organism has the same set of genes), the proteome is dynamic and varies. The set of proteins expressed by a cell is highly dependent on its tissue type, microenvironment, and stage within its life cycle. When cells receive cues in the form of growth factors, hormones, metabolites, or other agents, various genes are turned on or off; hence, proteomes vary during cell differentiation, activation, trafficking, and during malignant transformation. Moreover, many proteins are secreted from cells and circulate in the blood or lymphatic fluid or are excreted in the saliva, mucus, tear fluid, or urine [2].

As for the -omics mentioned above, glycomics refers to comprehensive studies that outline the "glycome", which describes the complete repertoire of glycans and glycoconjugates that cells produce under certain conditions of time, space

and environment. Glycomic analyses try to understand how a collection of glycans relates to a particular biological event. Glycans participate in almost all biological processes, from intracellular signalling to organ development to tumour growth. Understanding how the totality of glycans governs these processes is a central goal of glycobiology.

In the last decade, glycomics studies have contributed significantly to diverse areas of science, including early disease detection, vaccine development, and control of the reproductive and immune systems. The sugar components of glycoconjugates modulate or mediate a wide variety of functions in physiological and pathophysiological states [3]. For example, cellulose serves as a structural scaffold, starch and glycogen represent energy sources, glycoproteins and glycolipids, at cell surfaces, regulate and modulate complex biological processes. Glycobiology refers to the study of the glycoconjugates role in complex biological and pathological processes and it has become crucial both for developmental biology and immunology, as well as for the study of various cardiovascular, infectious diseases and cancer [3]. Glycoproteomic studies are focused on protein glycosylation. Glycoproteins and glycolipids are particularly abundant at cell surfaces and make up a layer called the glycocalyx [4], which plays a key role in cellular interactions. Moreover, many proteins secreted by the cell, are glycosylated and the nature of the glycans affects their function and half-life [5, 6]. Additionally, abnormal glycosylation is correlated with several disease states such as cancer [7, 8], inflammatory diseases, congenial disorders, etc [9, 10]. Recent studies explored the different role of envelope glycoproteins along the virus pathobiology, ranging from immune evasion by glycan mimicry/shielding toward the recognition of glycans on host cell receptors up to induction of innate immune cell response mediated by complement activation [11]. As an example, during last year many authors lingered on the spike (S) envelope protein of the currently emerging virus (CoV) inducing severe acute

respiratory syndrome (SARS) to explain the crucial role of glycoprotein in infection initiation [12-15].

While the importance of studying glycosylation has clearly been demonstrated, the glycosylation state of most proteins, along with the role of the glycans on those proteins, is still unknown. For instance, although 50% of the total number of proteins entered into the Swiss-Prot database are expected to be glycosylated, only around 10% of these glycans have been characterized,[16] indicating that significantly more advances in glycomics are necessary, in order to better understand the structural and functional roles of glycans.

### 1.1.1 Proteins glycosylation

Glycosylation is a very common post-translational modification (PTM) of protein. Although the transfer of initial sugar(s) to glycoproteins occurs in the ER or on the ER membrane, the subsequent addition of the many different sugars that make up a mature glycan is accomplished in the Golgi. Glycoproteins occur in nature as heterogeneous mixtures of structures, called glycoforms, which differ in the composition of glycans and their attachment site. The two most common forms of glycosylation are N- and O-linked glycosylation. (Figure 1.1).

**A)**                                          **B)**
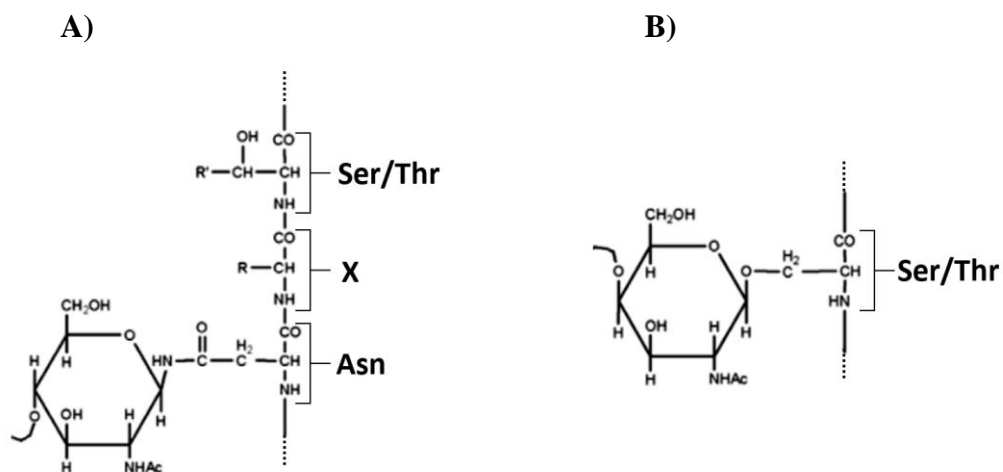


*Figure 1.1 A) N-linked and, B) O-linked glycosylation. Figure adapted from [17]*

Regarding N-glycosylation, oligosaccharides bind to the amide nitrogen of asparagine side chains within the Asn-X-Ser/Thr consensus sequence in which X is any amino acid except proline [17]. In O-glycosylation there is no consensus sequence and oligosaccharides can link to the hydroxyl oxygen of any serine (Ser) or Threonine (Thr) side chains [18]. All N-linked glycans are derived from the precursor Glc3Man9GlcNAc2, which is attached to the protein during translation. While many of the pendant residues on this initial building block are eventually removed, a well-defined trimannosyl pentasaccharide core (Man3GlcNAc2) always remains, even after glycan modification. The precursor is subjected to various modifications in the ER and Golgi apparatus, which results in three types of glycan structures: high mannose, complex and hybrid. (Figure 1.2)
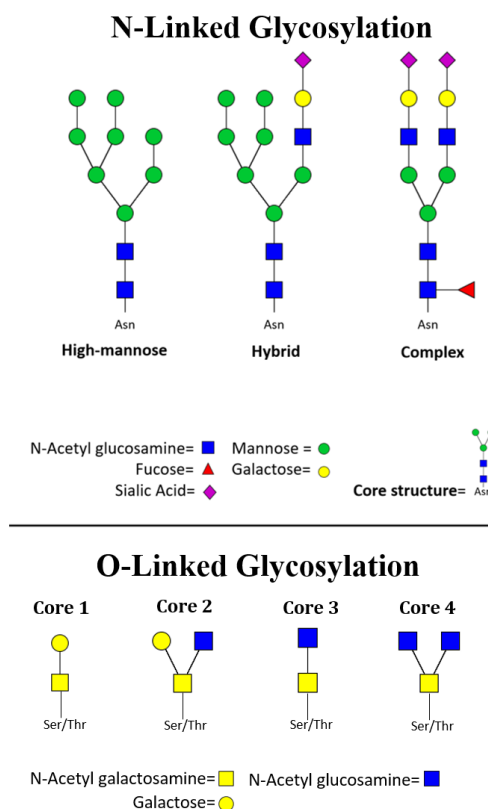


*Figure 1.2 Types of N-glycan structures (High-mannose, Hybrid and Complex) and O-glycan structures (core1, core2, core3 and core 4). Original figure created by GlycoWorkbench software.*

Depending on the number of branches attached, complex-type glycans could be subdivided to bi-, tri and tetra-antennary structures. N-Glycans play important roles in the quality control of glycoprotein folding in the ER lumen; glycoproteins will not pass ER quality control if they fail to reach the correct conformation [19]. Although N-glycosylation is the most common, O-glycosylation of proteins also play an essential role in cell biology. O-glycosylation is crucial in the mucin biosynthesis, high-molecular weight and heavily O-glycosylated proteins that form mucus secretions. Furthermore, this type of glycosylation is also critical for the formation of proteoglycan core proteins which are used to create extracellular matrix components. Additionally, antibodies can be heavily O-glycosylated [20].

N-glycosylation does not preclude the other from occurring, as O-glycosylation commonly occurs on glycoproteins that were N-glycosylated in the ER. Besides the different linkage, O-glycosylation also differs in the O-glycosidic mechanism, which is not as complex as N-glycosidic one; while a precursor glycan is transferred en bloc to Asn via N-glycosylation, sugars are added one-at-a-time to serine or threonine residues in O-glycosylation.

Glycan structural heterogeneity is the result of the non-template driven synthesis in the Golgi where monosaccharide units can be coupled in many different ways (adopting different conformations and binding positions) without following a specific pattern, unlike proteins or DNA. Mammalian glycans are synthesized with complex biosynthetic pathways that involved the assembly of the following monosaccharide units: fucose (Fuc), glucose (Glc), galactose (Gal), N-acetylglucosamine (GlcNAc), N-acetylgalactosamine (GalNAc), iduronic acid (IdoA), glucuronic acid (GLCA),mannose (Man), xylose (Xyl) and sialic acid (SA). In this way a complex mixture of glycosylated variants, called glycoforms, were generated. It has been assessed that around 700 proteins and more than 7000 different structures are required to generate the complete variety of glycans in mammals [21].

## 1.1.2 Glycoproteome characterisation

Glycome mapping shows a bigger level of complexity than genomics or proteomics. Monosaccharide units may be coupled to each other in many different ways (Figure 1.3), as opposed to the amino acids in proteins or the nucleotides in DNA, which are always coupled in a standard fashion [21]. Glycans are generated by the coordinated action of many enzymes in the subcellular compartments of a cell and therefore they are secondary gene products. Since the glycan's structure are subject to the expression, activity and accessibility of the several biosynthetic enzymes, it is not possible to produce large amounts of glycans for structural and functional studies by using recombinant DNA technology as it is for proteins. Moreover, the reactions involved in glycosylation and de-glycosylation at specific sites are frequently incomplete, thus increasing the heterogeneity of glycoproteins [22, 23]. Furthermore, glycans can adopt complex branched structures of monomeric units linked together, and these structures have different conformations (glycoforms). They can be covalently attached to other compounds forming glycoconjugates.
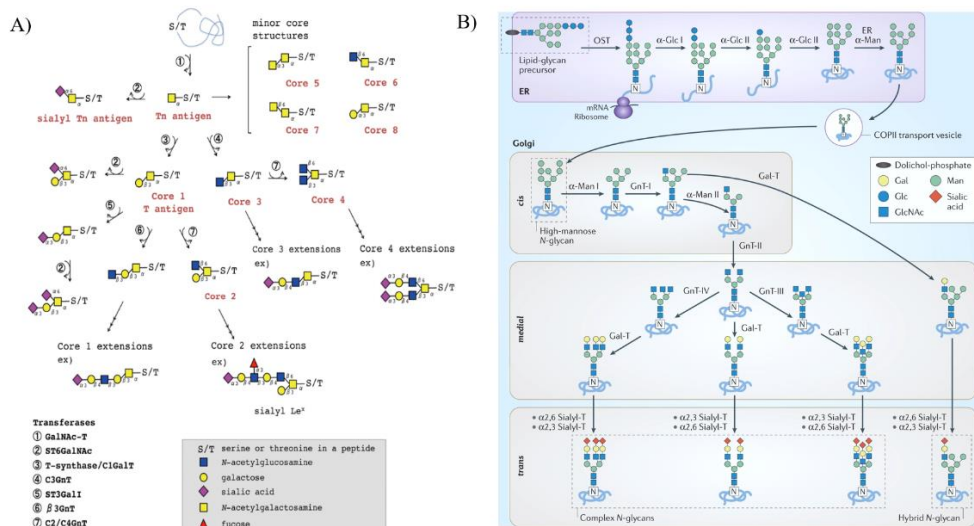


*Figure 1.3 A) O-glycan biosynthesis generated by the coordinated action of many enzymes B) N-glycan synthesis is initiated in the endoplasmic reticulum (ER) by the en bloc transfer of a lipid-glycan to Asp by the multi-subunit oligosaccharyl-transferase (OST) (www.nature.com, www.intechopen.com )[22].*

The combination of accurate technologies and advanced software programs can solve the mystery of glycan structures.

Mass spectrometry, thanks to its high sensitivity and selectivity, is a powerful tool in glycomic and glycoproteomic [24-28], allowing characterisation of glycoproteins either at the glycopeptides level or after de-glycosylation, characterizing separately the peptides containing the glycosylation sites and the glycans [29, 30]. Mass spectrometry analytical power lies on its capability to measure mass to charge (m/z) ratio; moreover MS / MS data acquisition is the key to structural characterisation, providing information with high sensitivity and accuracy. The presence of the glycan moiety tends to increase the acidic characteristics of glycoproteins and glycopeptides, simultaneously increasing hydrophilicity and surface activity [31]. This is the reason why glycopeptides are very difficult to detect in complex mixtures, generally requiring enrichment procedures prior to MS analyses. Thus, derivatization procedures or separation methodologies are commonly used in order to equalize analytes with quite similar ionization efficiencies in the mass spectrometer. The complex heterogeneity of carbohydrates in biological systems complicates the analysis, as a single peptide can be modified by many glycans, each having a different mass and structure. In the frame of glycoproteomic, there are new and innovative techniques that aim to determine the sequence of the protein backbone and glycan component in addition to site occupancy.

### 1.1.3    Analysis and profiling of glycan

The general glycoproteomic approach for the analysis of glycans involves the release of the latter from the glycoprotein. Two methods are available for releasing carbohydrates: enzymatic methods and chemical methods. Enzymatic methods allow the conservation of the information of the glycoprotein. Specific enzymes release N-linked glycans from glycoproteins. Two styles of enzymes are available, endoglycosidases and exoglycosidases; the previous releases all or

most of the glycan whereas the latter removes only a specific portion of the non-reducing end of the glycan [32]. PNGase (protein N-glycosidase) are a common endoglycosidases family, that catalyse the release of most N-glycans (e.g. PNGase F cleaves the bond between the innermost GlcNAc and asparagine residues of glycoproteins (Figure 1.4)), releasing the carbohydrate and preserving the intact glycosylamine at the reducing end, although the amino is quickly hydrolysed and replaced by a hydroxyl) [23].
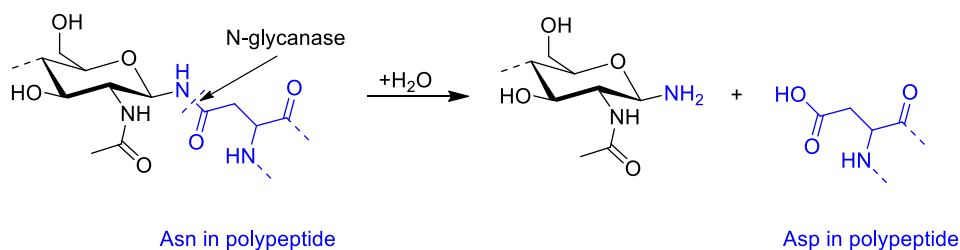


*Figure 1.4 Reaction mechanisms of N-glycanase (PNGase F). Original figure created by ChemDraw software.*

Methods for the release, and analysis of N-glycans are well established, however, the release and recovery of O-glycans, particularly intact O-glycans, remains very challenging [33]. One major reason for this is that the enzymatic release of O-glycans is limited to endo-α-N-acetylgalactosamidase, an O-glycanase with a high specificity that only enables the removal of core 1 disaccharides (Galβ1-3GalNAc) from serine or threonine. Therefore, the existing best methods for universal removal of O-glycans from glycoproteins rely on chemical release. Over the last 20 years, several methods for the chemical release of O-glycans have been developed, including techniques that use reductive β-elimination. The most common conditions of reductive β-elimination have been developed by Carlson in 1968. The method classically involves the use of sodium hydroxide and sodium borohydride. Under alkaline sodium hydroxide conditions, the O-glycans are cleaved from the glycoprotein and the sodium borohydride immediately reduces the terminal monosaccharide of the free sugars to an alditol. Reduction of the terminal sugar by converting it to an alditol reduces or completely stops a side reaction called 'peeling', which may be a side reaction

which will be defined as stepwise degradation of the polysaccharide starting at the reducing end and removing one sugar residue at a time (Figure 1.5) [34].
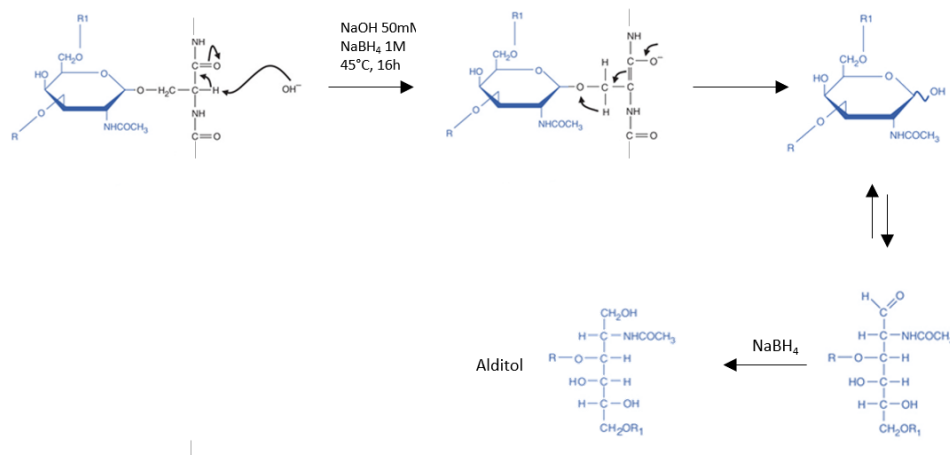


*Figure 1.5 Reductive β-elimination. Original figure created by ChemDraw software.*

To improve the detection of glycans, it is possible to derivatize the hydroxyl groups by substituting the proton by a non-polar group like a methyl such as in the permethylation [35] (Figure 1.6). Carbohydrates are not ionized as efficiently as proteins that can be easily protonated on several basic sites, neither do they seem to be vaporizable as efficiently. Permethylation decreases the intermolecular hydrogen bonding, thus increasing the sample volatility and therefore the intensity of the ion signals. It also improves the formation of diagnostic fragment ions increasing the yield of useful structural information. It avoids gas-phase rearrangements of glycan residues that can introduce high-level of ambiguity and complexity into structural analyses. Furthermore, it can remove the negative charge associated with acidic residues, resulting in a stabilization of acidic residues in the positive mode; this feature is particularly important for the analysis of sialylated glycans in which the carboxyl group is located on the same carbon as the glycosidic bond.
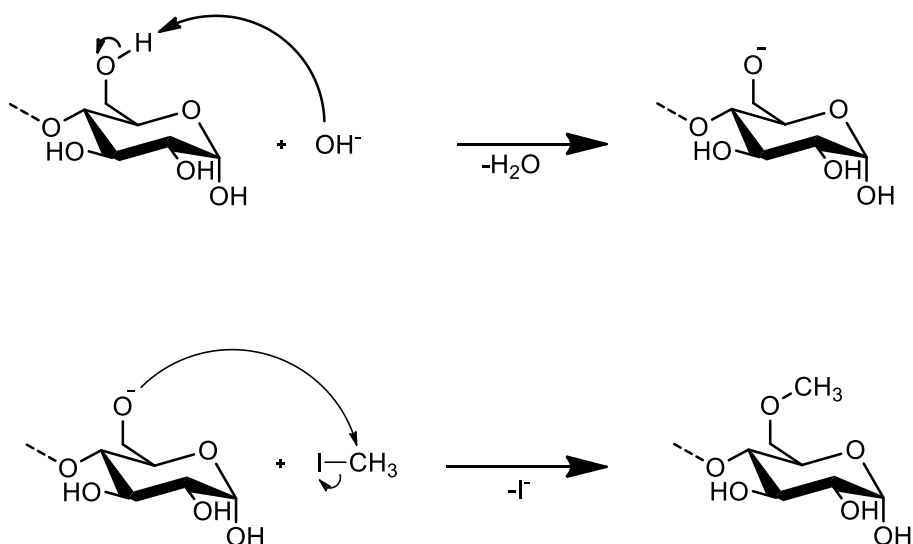
*Figure 1.6 Permethylation mechanism. Adapted from [35]. Original figure created by ChemDraw software.*

### 1.1.4 Glycopeptide analysis

Glycopeptide analysis allows an identification of site-specific glycosylation properties. To this aim, several approaches and experimental protocols have been developed [30, 36]. In most cases, the glycoprotein is hydrolysed with specific endoproteinases, resulting in a mixture of peptides and glycopeptides. The most common enzymes used to obtain glycopeptides carrying individual glycosites are: trypsin, chymotrypsin, Asp-N, Glu-C and Lys-C. Mass spectrometric identification of glycopeptides within a complex protein digest is still challenging due to the following reasons: i) glycopeptides often represent only a minor percentage of the total peptide mixture; ii) the glycopeptides signal intensities are generally very low compared to those of non-modified peptides, mainly because the signal is distributed between a population of peptide species carrying different glycan moiety and to their lower ionization efficiency; iii) glycopeptide signals are often suppressed in the presence of other non-glycosylated peptides [31], especially if the glycans are terminated with a sialic acid negatively charged

moiety. Therefore, it is frequently necessary to use different analytical techniques, either in parallel or sequentially, to investigate on glycopeptides.
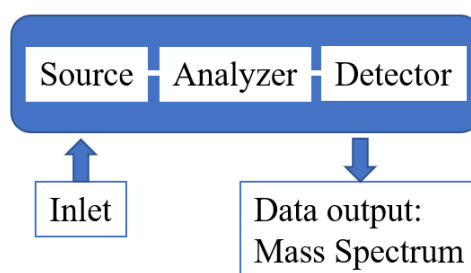
## 1.2    Mass spectrometry

Mass spectrometry, in its most simple definition, is the production and detection of ions separated according to their mass to charge (m/z) ratios [37-39]. The detection of such ions results in a mass spectrum, which is a plot of the relative abundance of the ions as a function of their m/z ratio. Most important features of mass spectrometry are reproducibility, sensitivity, accuracy. Today a wide variety of mass spectrometers is available [40, 41], all sharing the capability to assign mass-to-charge values to ions, although the principles of operation and the types of experiments that can be done on these instruments differ greatly.

Mass spectrometers have four essential parts:

a) system for sample introduction

b) source that produces gas phase ions from the sample

c) one or more mass analyser to separate ions

d) ion detector.



*Figure 1.7 Working process of a mass spectrometer, from sample introduction system to the mass spectrum display.*

Different type of mass spectrometers can be distinguished by their ionization system and by the type of analyser that is an essential component to define the

accessible mass range, sensitivity and resolution [41]. All mass spectrometers are operated at very low pressure to prevent collisions of ions with residual gas molecules in the analyser during the flight from the ion source to the detector.

The most widespread ionization methods in biochemical analyses are Electrospray Ionization (ESI) and Matrix Assisted Laser Desorption Ionization (MALDI) [42, 43]. Mass analysers as quadrupoles (Q), ion traps (IT), time-of-flight (TOF), Orbitrap, or combination of these in "hybrid instruments", are commonly used for their good resolution and sensitivity. The coupling of Liquid Chromatography (LC) and tandem Mass Spectrometry (MS/MS) is a widely used analytical technique for quantitative and qualitative analysis [44].

- *MALDI ion source:*

Laser light (generally a pulsed nitrogen laser at 337 nm) irradiation is used in MALDI source to induce sample ionisation [45-47].
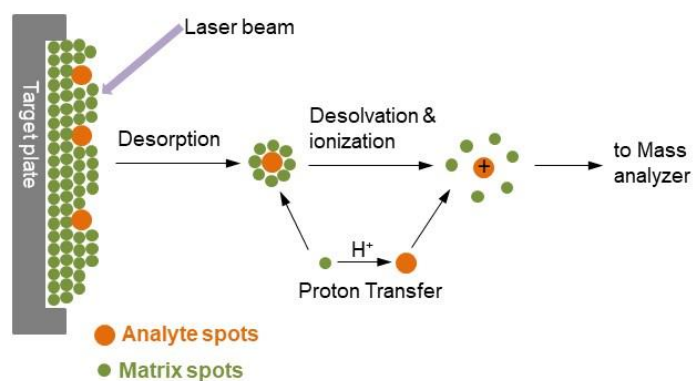


*Figure 1.8 Ionization of analytes by MALDI (www.creative-proteomics.com)*

The sample is pre-mixed on a metal target with a highly absorbing matrix, a small aromatic molecule, that upon drying, co-crystallises with the sample [48]. This process takes place in the ion source under a high vacuum and a strong electrical field between the target and an extraction plate. Energy deposition into the matrix molecules induces sublimation of the matrix crystals and subsequent expansion into the gas phase. In positive ion mode, protonated molecular ions $(M+H)^+$ are

usually the dominant species, although they can be accompanied by salt adducts [43]. In negative ionisation mode the deprotonated molecular ions (M-H)⁻ are usually the most abundant species [49, 50]. Carbohydrates generally ionize as $MNa^+$ species but this ionization is not as sensitive as ionization of peptides, so a higher laser power is generally required and sometimes it is difficult to maintain a stable signal. Signal intensity even depends on the nature of the matrix and the method of sample preparation. Numerous matrices have been studied for the analysis of glycans and glycoconjugates, and a matrix very effective with a class of compounds can be ineffective with others. In table below are shown most common MALDI matrix with their application.

*Table 1.1 UV-MALDI matrices[38]*

| Compound | Acronym | Application to |
|---|---|---|
| Picolinic acid | PA | Oligonucleotides, DNA |
| 3-Hydroxypicolinic acid | HPA, 3-HPA | Oligonucleotides, DNA |
| 2,5-Dihydroxybenzoic acid | DHB | Proteins, oligosaccharides |
| α-Cyano-4-hydroxycinnamic acid | α-CHC, α-CHCA, 4-HCCA, CHCA | Peptides, smaller proteins, triacylglycerols, numerous other compounds |
| 4-Chloro-α-cyano-cinnamic acid | ClCCA | Peptides |
| 3,5-Dimethoxy-4 hydroxycinnamic acid | SA | Proteins |
| 2-(4-Hydroxyphenylazo) benzoic acid | HABA | Peptides, proteins, glycoproteins, polystyrene |
| 2,6-Dihydroxyacetophenone | DHAP | Glycopeptides, phosphopeptides, proteins |
| 2,4,6-Trihydroxyacetophenone | THAP | Solid-supported oligonucleotides |

One of the drawbacks of MALDI-MS for oligosaccharides analysis is the high frequency of in-source decay fragmentation in the delayed extraction process with long delay times, often leading to extensive losses of acidic residues, such as sialic acid [51]. However, MALDI-MS is rather sensitive and tolerant to contaminants, allowing rapid analyses without much effort.

- *Electrospray ion source (ESI)*

ESI ionization is obtained on the application of a high voltage to the tip of a needle, from which the sample emerges forming a dispersed spray of highly charged droplets, with a co-axially introduced as nebulising gas [52, 53].
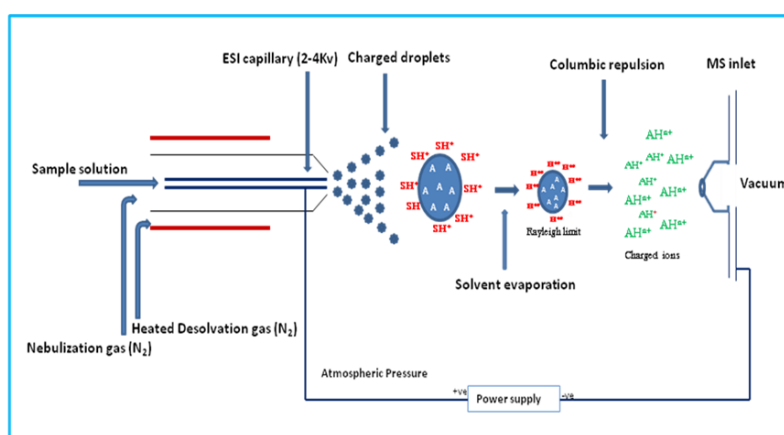


*Figure 1.9 Schematic representation of ESI source (www.chem.pitt.edu)*

This gas is usually nitrogen, and its function is to aid the spray emerging from the capillary tip to direct towards the mass spectrometer. Solvent evaporation causes charged droplets to diminish in size, with the help of a drying. Multiply charged sample ions are extracted into the analyser. An important feature of ESI-MS is that it confers very little extra energy to the sample (it is known as a "soft" ionization method) [54] generally preventing in source fragmentation of the sample ions. Regarding the analysis of glycans and glycoconjugates, the ESI source generates multicharged ions, producing very complex and difficult to interpret spectra. To overcome this problem, especially in the analysis of complex mixtures, a pre-fractionation step is usually performed. Most of ESI-MS

instruments are equipped with an upstream liquid chromatographic system (LC), hence the name of the LC-MS technique [55, 56]. A great enhancement of ESI ion sources has come from the reduction of the flow rate of the liquid used to create the spray to a nano-scale level [57]. This device leads to a higher efficiency in creating ions because the charge density at the Rayleigh limit increases significantly with decreasing droplet size ($zR$ $\alpha$ $8R3/2$). Other advantages related to the use on μLC and nLC are the low consumption of sample and the higher sensitivity of the methods because of the increase in the concentration of the analyte as it elutes off the column.

After ionization, analytes enter the second region of the mass spectrometer, the analyser, whose main function is the separation of the ions according to their m/z values. There are several type of mass analysers currently used for analysis of biomolecules. The main types are quadrupole (Q), time-of-flight (TOF) and Orbitrap. They can be self-consistent or they can be combined together in tandem.

- *Quadrupole mass analyser*

Quadrupole mass analyser is generally coupled to ESI ion source and consists of four parallel rods that have fixed DC and alternating RF potentials applied to them (Figure 1.10).The application of these voltages forces in a plane normal to the direction of ions drift allows ions oscillations which can be stable or unstable, according the m/z value [58]. Only ions with stable trajectories will reach the detector, otherwise they will collide with the quadrupole rods.
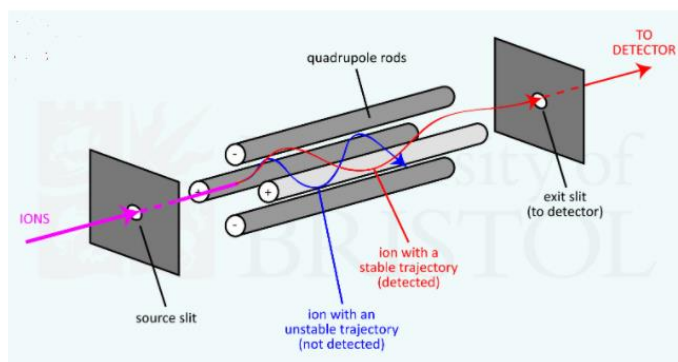
*Figure 1.10 Schematization of a Quadrupole mass analyser (www.chm.bris.ac.uk/ © Paul. J gates 2014)*

- *Time Of Flight mass analyser (TOF)*

TOF analyser is generally coupled to MALDI ion source and consisting in a tube under a high vacuum. All ions entering the TOF tube are endowed with a fixed kinetic energy, proportional to the applied voltage and the charge [59]. Separation is based on the principle that the higher is the mass of the ion, the lower its velocity. Thus, ions can be separated measuring the time each ion takes to travel through the field free region of the tube. With delayed extraction (DE), the extraction voltage is applied after a defined interval of time (expressed in ns) from the laser pulse. During this "delay time", the extraction voltage, applied as a potential gradient, compensates for the distribution of initial kinetic energies, so that ions will not be influence by differences in initial kinetic energies in their path [59]. Another point for kinetic energy correction concerns the introduction of an electrostatic mirror. Thereby the flight times of ions with identical m/z values, but different kinetic energy values will be corrected when the ions arrive to the detector [60].
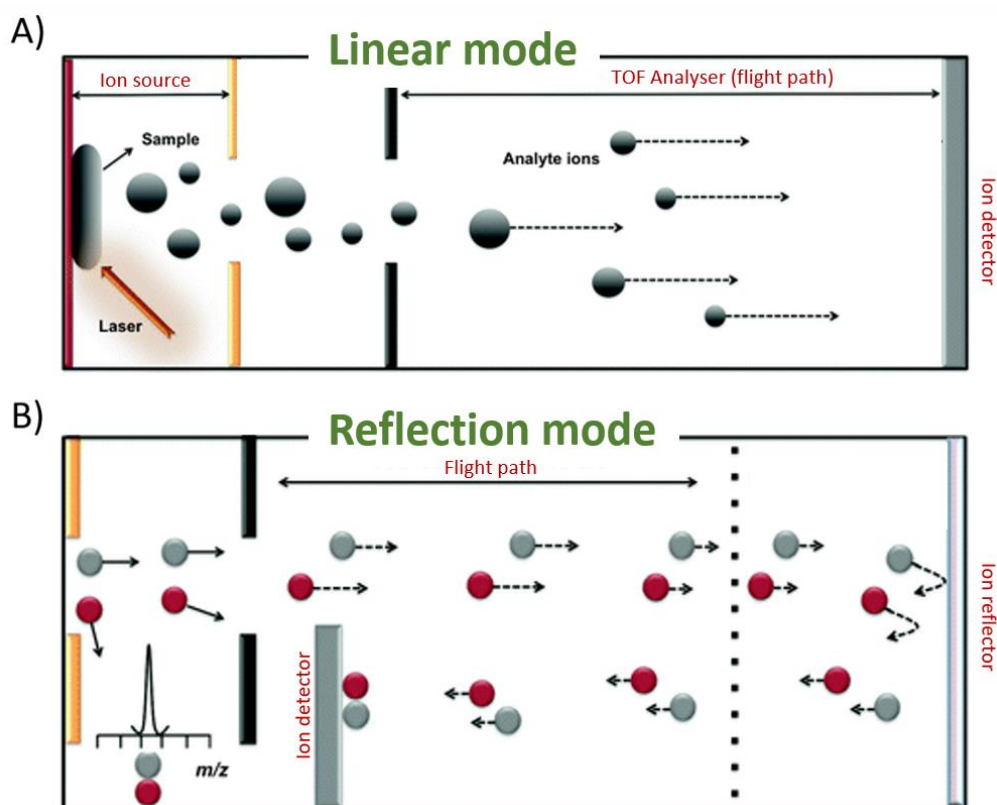
*Figure 1.11 Basic components of a linear (A) and reflecting (B) TOF mass spectrometer. Adapted from [60].*

- ***Orbitrap mass analyser***

The Orbitrap mass analyser is the first high-performance mass analyser that employs trapping of ions in electrostatic fields [61]. The electrostatic field is produced by two electrodes: a central spindle-shaped and an outer barrel-like electrode (Figure 1.12). Ions are moving in harmonic, complex spiral-like movements around the central electrode while shuttling back and forth over its long axis in harmonic motion with frequencies dependent only on their *m/z* values. The oscillating ions create electric current in the outer electrode, and the mass spectrum is obtained by Fourier transform of the recorded current [62]. The longer the current recording period the better is the resolution [63, 64].
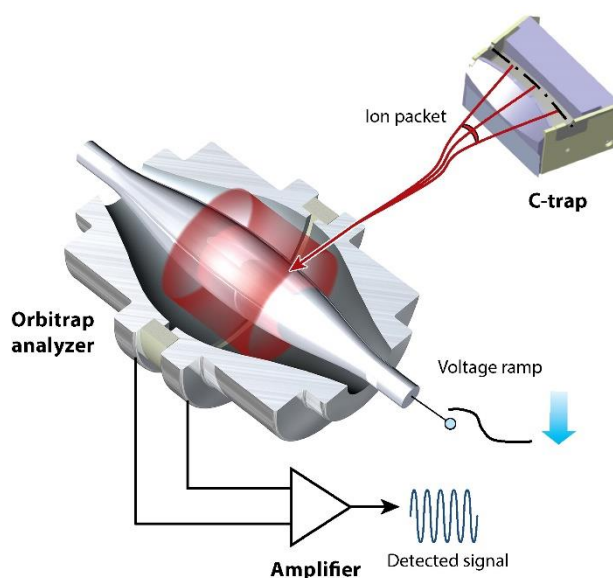
*Figure 1.12 Cross section of the C-trap ion accumulation device and the Orbitrap mass analyser with an example of an ion trajectory. During the voltage ramp, the ion packets enter the Orbitrap mass analyser forming rings that induce current, which is detected by the amplifier (www.mass-spec.chem.ufl.edu) [64].*

Tight integration with the ion injection process allows the high-resolution, mass accuracy, and sensitivity that have become crucial for addressing analytical needs in numerous areas of research, as well as in routine analysis [65]. Nowadays three are three major families of Orbitrap instruments: Linear trap quadrupole (LTQ) Orbitrap, Q Exactive, and Orbitrap Fusion mass spectrometers [64]. LTQ Orbitrap mass spectrometer was the first commercial instrument to incorporate an Orbitrap mass analyser [66]. The design of the LTQ Orbitrap instrument, an ion trap followed by the Orbitrap mass analyser (Figure 1.13), allowed the flexibility required by researchers. Both MS and MS$^n$ spectra could be recorded by using either the Orbitrap analyser for highest resolution and mass accuracy or the ion trap analyser for highest speed and sensitivity [65]. The most commonly employed operation mode for the instrument became acquisition of full scans in the Orbitrap analyser and data-dependent MS/MS scans in the ion trap analyser. This mode allowed full utilization of the resolution and mass accuracy for the detection of precursors in complex mixed spectra and the speed and sensitivity for MS/MS spectra on an LC timescale in discovery-based experiments [65].
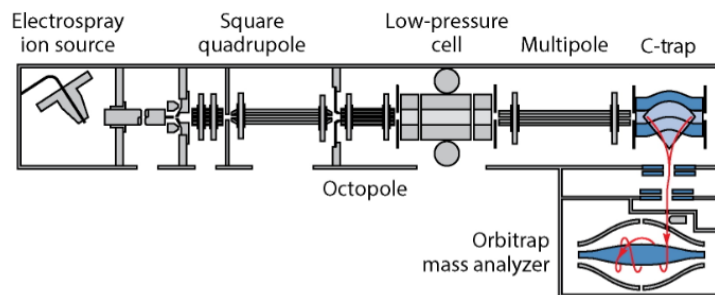
*Figure 1.13 Schematic of the LTQ Orbitrap mass spectrometer with traditional ion trap followed by Orbitrap mass spectrometer architecture (www.mass-spec.chem.ufl.edu)*

Despite the great analytical power of high mass resolution and accuracy, LTQ Orbitrap has some limitations. First, the only fragmentation method available was ion trap–based CID, a method that proved powerful for peptide identification but was limited for modified peptides with important post-translational modifications (PTMs) such as phosphorylation and glycosylation [67]. In addition, the practical accurate mass MS/MS scan rate was slow, impeding its utility in experiments requiring online chromatography [67].

To obtain structural information not afforded by low-energy CID fragmentation, additional fragmentation techniques were implemented that required adding a new multipole for higher energy collision-induced dissociation (HCD) as well as a newly developed electron transfer dissociation (ETD) reagent ion source, which will be described in detail in the following paragraph.

The subsequently developed Q Exactive instrument permitted precursor ion isolation on an Exactive-type mass spectrometer [68] (Figure 1.14). A mass filtering quadrupole was utilized for isolation of precursors and an Orbitrap analyser employed for detection of full scans and MS/MS spectra, using an HCD cell for fragmentation.
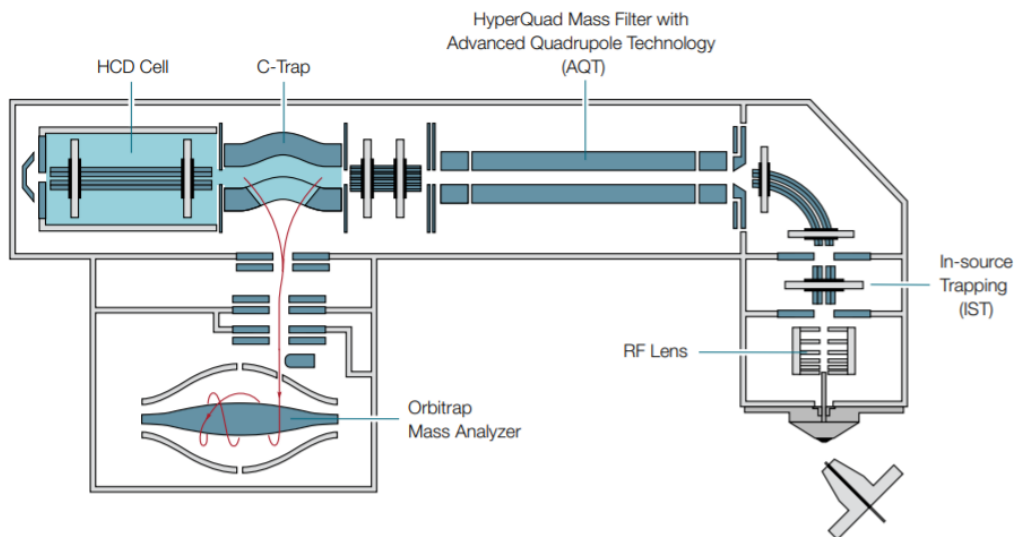
*Figure 1.14 Quadrupole/Orbitrap hybrid mass spectrometer Q Exactive architecture.*
*Abbreviation: HCD, higher energy collision-induced dissociation (www.thermofisher.com)*

The Orbitrap Fusion instrument, incorporating a quadrupole mass filter and Orbitrap and linear ion trap mass analysers, combined advanced ion trap Orbitrap hybrid technology with the quadrupole Orbitrap hybrid systems [69] (Figure 1.15). This architecture enabled significant performance improvements. With three mass analysers, operation can be fully parallelized, maximizing the use of the ion current. The instrument architecture facilitates the realization and rapid execution of complex modes of analysis, due to its characteristic ability to concurrently isolate ions with one analyser and separately detect ions in the two remaining analysers.
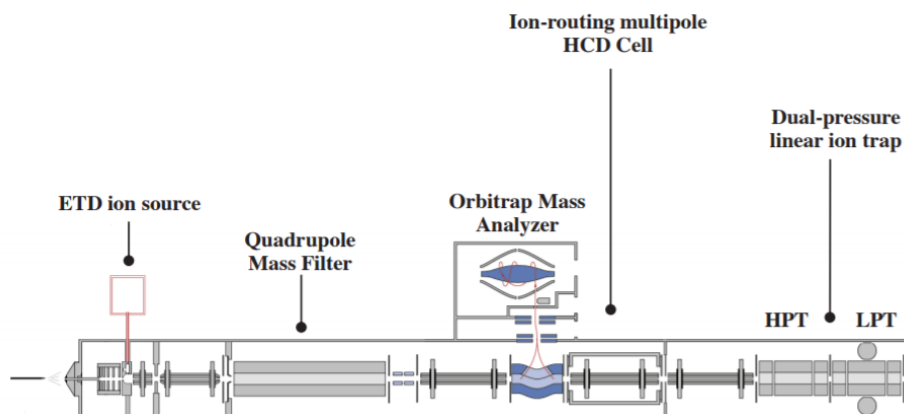
*Figure 1.15 Quadrupole/Orbitrap/ion trap hybrid mass spectrometer architecture; Schematic of the Orbitrap Fusion mass spectrometer. Abbreviations: ETD, electron transfer dissociation (www.thermofisher.com).*

- **Quadrupole ion trap (IT)**

The quadrupole ion trap is based on the same principle as the quadrupole mass filter, except that the quadrupole field is generated within a three-dimensional trap [70]. The trap consists of three electrodes, a ring electrode and two hemispherical end caps electrodes. These electrodes allow ion trapping in a small volume (Figure 1.16). In IT, ions are dynamically stored in a three-dimensional quadrupole ion storage device. The RF and DC potentials can be scanned to eject successive m/z ratios from the trap into the detector [71].
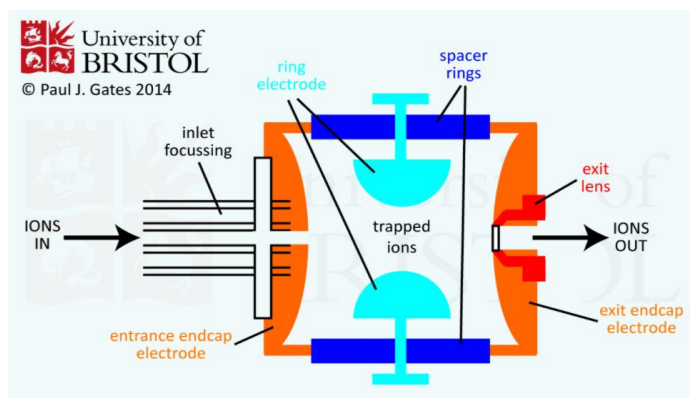


*Figure 1.16 Graphical representation of a Quadrupole Ion Trap geometry (www.chm.bris.ac.uk © Paul. J gates 2014)*

The advantages of the ion-trap mass spectrometer include compact size, and the ability to trap and accumulate ions to increase the signal-to-noise ratio of a measurement. For these reasons, ion traps are used in fragmentation experiments. Conventional ion trap mass spectrometers operate with a three-dimensional (3D) quadrupole field, which confers to the analysis very high efficiency as regard the time to fill the ion trap and to generate a complete mass spectrum, but presents some problems as regard the trapping efficiencies, primarily due to their small volume. Thanks to the introduction of linear ion traps, characterized by a greater ion accumulation capacity and greater trapping efficiency, these problems have been overcome.

### 1.2.1 Tandem mass spectrometry

Tandem mass spectrometry (MS/MS) is used to produce structural information about a compound by fragmenting specific sample ions inside the mass spectrometer and identifying the fragment ions. Tandem mass spectrometry also enables specific compounds to be detected in complex mixtures because of their characteristic fragmentation patterns. MS/MS is based on two stages of mass analysis, one to select a precursor (or parent ion) and the second to analyse fragment (or daughter ions). Generally, in a tandem mass spectrometer the two analysers are separated by a collision cell where an inert gas (e.g. argon, xenon) collides with the selected sample ions and brings about their fragmentation.
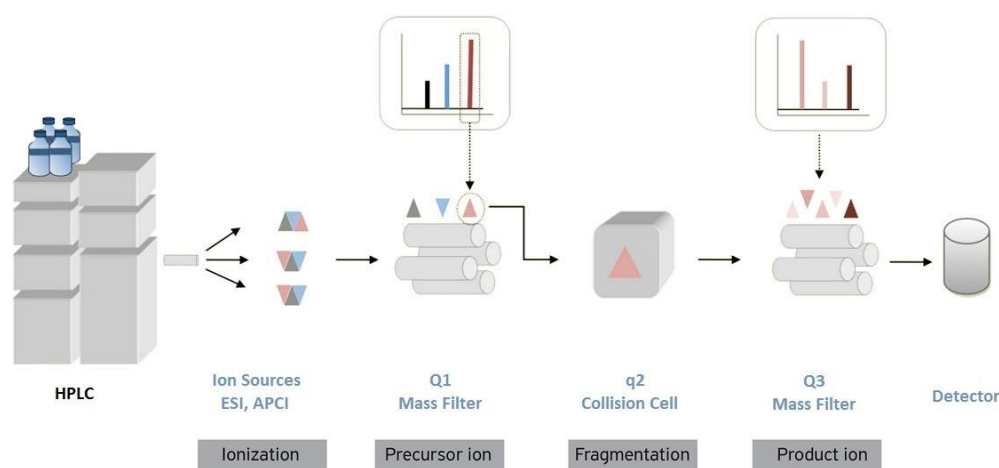
*Figure 1.17  Example of LC-MS/MS instrumentation. Triple quadrupole mass spectrometer (QqQ), is a tandem mass spectrometer, which has two quadrupole mass analysers, with a (non-mass-resolving) radio frequency–only quadrupole between them, acting as a collision cell for collision-induced dissociation (CID) to fragment the selected precursors/parent ions, and to generate fragment/daughter ions (www.creative-proteomics.com).*

Peptide precursor ions, dissociated by the most usual low-energy collision conditions, fragment along the backbone at the amide bonds, forming structurally informative sequence ions and less useful non-sequence ions by losing small neutrals like water, ammonia, etc. The amino acid backbone has three different types of bonds (NH-CaH, CaH-CO and CO-NH) and each of them can be fragmented originating different fragment ions. Each bond splintering gives rise to two species, one neutral and the other one charged, and only the charged one could be monitored by the mass spectrometer. Hence there are six possible fragment ions for each amino acid residue and these are labeled with the a, b, and c ions having the charge retained on the N-terminal fragment, and the x, y and z ions having the charge on the C-terminal fragment (Figure 1.18) [72]. The most common cleavage sites are at the CONH bonds, which give rise to the b and/or the y ions.
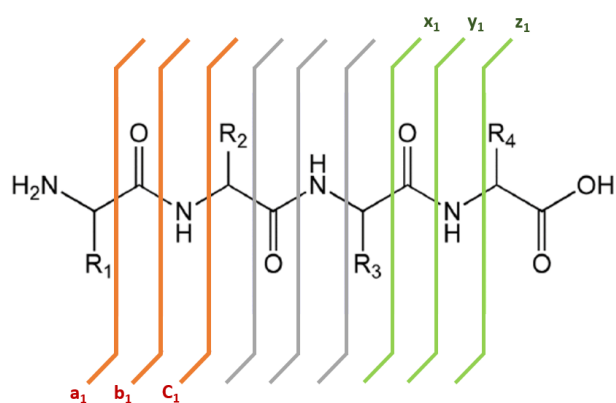
*Figure 1.18 Peptides fragmentation scheme. Adapted from [73].*

Investigation of glycoconjugates (glycolipids and glycopeptides) by MS and MS/MS has shown that the fragmentation patterns within the carbohydrate portion can be rather complex, especially for CID-MS/MS spectra [30].

A systematic nomenclature [74] for labelling the fragment-ions observed both in FAB-MS and FAB-MS/MS spectra has been devised by analogy to that in use for peptides, although the structures of carbohydrates necessitate the use of somewhat more complex symbols. When lower-case letters are used to designate peptide cleavages, the system proposed here serves to describe cleavages within the carbohydrate portion of glycopeptides. The Figure 1.19 shows the nomenclature system can be applied to assign product ions.
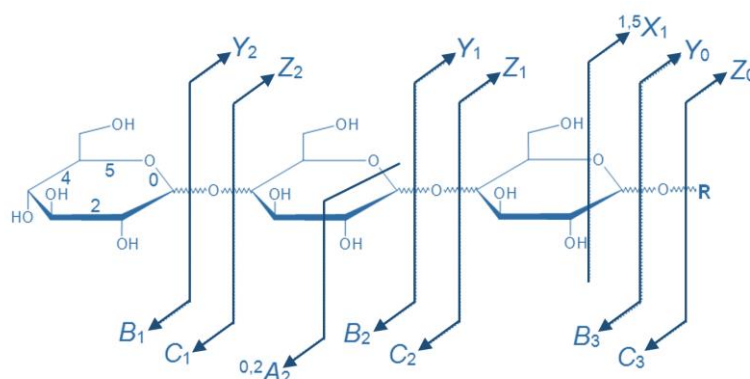


*Figure 1.19 Domon and Costello systematic nomenclature. Adapted from [46].*

The simplest fragmentation of the carbohydrate moiety of glycoconjugates and glycosides occurring during FAB-MS and FAB-MS/MS results from the cleavage of the glycosidic bonds and thus yields information about the sugar sequences. More complex processes involving the fragmentation of the sugar ring have been observed, particularly in CID-MS/MS spectra [75]. As illustrated in Figure 1.19, when the charge is retained on the carbohydrate portion, fragments are designated as $A_i$, $B_i$ and $C_i$, where i represents the number of the glycosidic bond cleaved, counted from the non-reducing end.

On the other hand, the reducing sugar unit are labeled as $X_j$, $Y_j$ and $Z_j$, where j is the number of the inter-glycosidic bond counted from the reducing end.

### 1.2.2 Instruments for MS/MS analysis

Instruments for tandem mass spectrometry can be classified as tandem in space or tandem in time (Figure 1.20). Tandem in space means that ion selection, ion fragmentation and fragments analysis, are events that occur in three different regions of the spectrometer; instruments of this type are, for examples, the triple quadrupole (QqQ) showed in Figure 1.17 and hybrid instruments such as QqTOF or TOF/TOF [26]. In Tandem in time instruments, those three steps of analysis occur in the same region of the spectrometer but in different times; QIT and LIT are instruments classified as tandem in time.
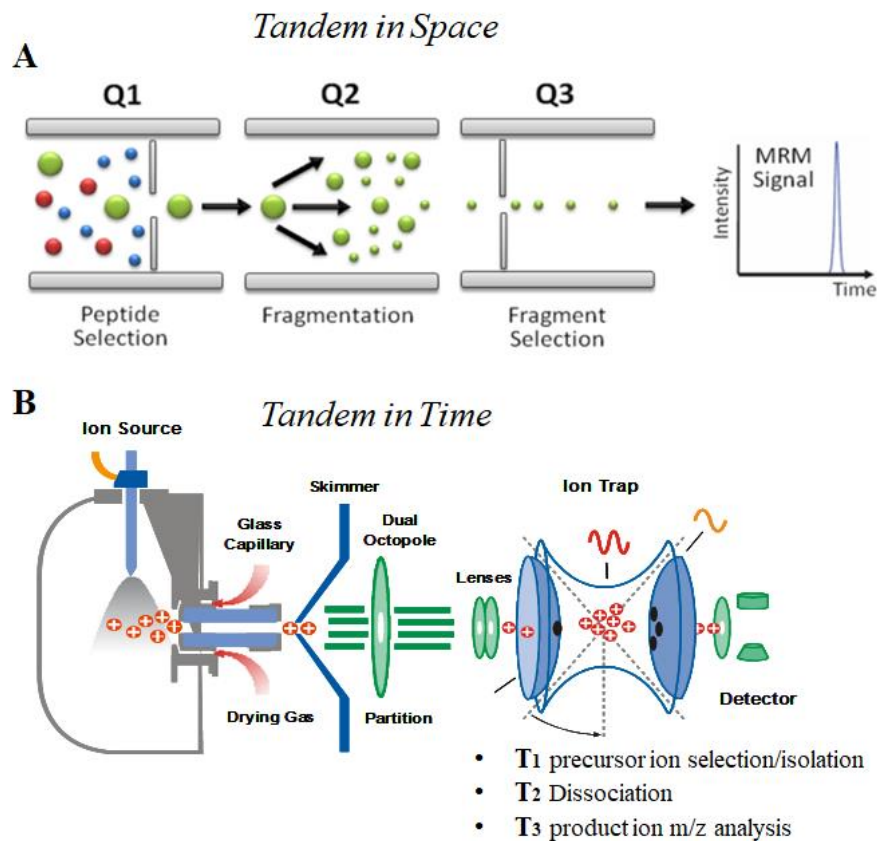
*Figure 1.20 Scheme of tandem in space A) and in time B) mass spectrometer (www.mrmatlas.org, www.slideserve.com).*

A tandem mass spectrometer like QqQ can operate in different modality:

- <u>Product ion scan</u>: the first quadrupole Q1 selects a specific precursor ion having a specific ratio m/z; this is fragmented in the collision cell Q2 and Q3 (the second analyser) is scanning the resulting fragments or "product ions" (MS2). In this mode, it is possible to obtain important structural information of the analyte under examination [71].

- <u>Precursor ion scan</u>: the first analyser (Q1) operates a scan to detect all precursor ions that generate a specific fragment, the second analyser (Q3) is set to output a single specific product ion (m/z) to the detector. Typically, this method is used to detect a subset of molecules that contain a specific functional group [71].

- <u>Neutral loss scan</u>: In the constant neutral loss scanning both the Q1 and Q3 scan and collect data across the whole m/z range. The two analysers are off set, hence, the Q3 allows the acquisition of only those ions which differ by a certain number of mass units (equivalent to a neutral fragment) from the ions transmitted through Q1. This scan mode is used, for example, to identify phosphorylated peptides in a peptide mixture due to the typical fragmentation of these peptides consisting in loss of a phosphoric acid residue that correspond a mass shift of -98 Da [71].

- <u>Multiple reaction monitoring (MRM)</u>: the first and the third quadrupoles act as filters to specifically select predefined m/z values corresponding to the peptide ion and a specific fragment ion or a set of fragment ions of the peptide, whereas the second quadrupole serves as collision cell [71].

### 1.2.3 Targeted mass spectrometry

### 1.2.3.1 Multiple Reaction Monitoring (MRM)

In a targeted proteomics experimental workflow, selected peptides that are surrogates exclusively of proteins of interest are measured in a predefined m/z ranges and retention time window. Targeted experiments are primarily performed on a triple-quadrupole (QQQ) and hybrid quadrupole-linear ion trap (QTrap) mass spectrometers using an MRM (or SRM: Selected Reaction Monitoring) data acquisition method [6].

MRM is the most common mode of using a triple quadrupole MS/MS for quantitative analysis [71], allowing enhanced sensitivity and selectivity. In MRM mode, two stages of mass filtering are employed on a triple quadrupole mass spectrometer. In the first step, an ion of interest (the precursor) is preselected in Q1 and induced to fragment by collisional excitation with a neutral gas in a pressurized collision cell (Q2). In the second step, instead of obtaining full scan MS/MS where all the possible fragment ions derived from the precursor are mass

analysed in Q3, only a small number of sequence-specific fragment ions (transition ions) are mass analysed in Q3 (Figure 1.21).
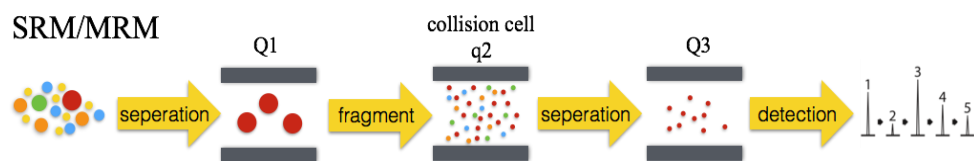


*Figure 1.21 Typical diagram showing the selected reaction monitoring experiment. In this triple quad, Q1 and Q3 act like a mass filter whereas Q2 acts as a collision cell for selected peptide ion (www.thermofisher.com).*

Thus, MRM mode works like a double mass filter that reduces drastically the noise and increases sensitivity and selectivity. Triple quadrupole systems allow the detection of many MRM transitions [76]. This enables quantitation of many targeted analytes in a single experiment. Typically, additional MRM transitions must be detected to perform identification of quantified compounds. Herewith, the most intense ion is known as 'quantifier' and all additional ions are called 'qualifiers' [6, 76].

MRM methods offer both absolute structural specificity for the analyte and relative or absolute quantification of analyte when stable, isotopically labelled standards are added to a sample in known amount. When a synthetic, stable isotope labelled peptide is used as an internal standard, the concentration can be measured by comparing the signals from the exogenous labelled and endogenous unlabelled species. This can be done because they have the same physicochemical properties and differ only by mass.

In contrast to conventional shotgun proteomic studies, SRM or MRM measurements are quantitative analyses strictly targeting a predetermined set of peptides and depend on specific SRM/MRM transitions for each targeted peptide. Previous information is required to define and optimize these transitions [76]:

-Proteins: i) The proteins that constitute the targeted protein set have to be selected. ii) For each targeted protein, those peptides that present good MS

responses and uniquely identify the targeted protein, or a specific isoform thereof, should be identified. Such peptides were called proteotypic peptides (PTPs). iii) For each PTP, those fragment ions that provide optimal signal intensity and discriminate the targeted peptide from other species present in the sample have to be identified.

-Peptides: i) The peptides that constitute the targeted analytes set have to be selected. ii) For each peptides those fragment ions that provide optimal signal intensity and discriminate the targeted peptide from other species present in the sample have to be identified.

-Optimization of transition's parameters: To perform qualitative analysis, high sensitivity is desired, thus each MRM transition is maximized by tuning acquisition parameters of the mass spectrometer. The signal intensity is determined by the combination of peptide ionization efficiency, its transfer into the analyser and its dissociation into, ideally, a few intense fragments. The setting of parameters for the ionization process and ion optics are critical; for example, at too low Extraction Voltage: peptides are not efficiently transferred, while at high interface voltage, peptides may undergo fragmentation in the ion source. Typically, optimal conditions are determined by using a set of reference compounds spanning the m/z range of the instrument. Fragmentation conditions can further tuned to increase the signal response for each target peptide. The acquisition of fragmentation patterns under different collision conditions on a QqQ instrument and the derived pseudo-breakdown curves provide to meticulously determine the optimal collision energies [77]. Typically, optimal collision conditions for doubly and triply charged peptides ranges between 20 and 40 V; and most of the singly charged high mass y-ions appear to have very similar behaviour [78]. The fragmentation patterns are also affected by the nature of collision gas and its pressure, in fact, different nature of collision gas may also affect the optimal collision b-ions are generated energy as the energy resulting of collisions between neutral gas molecules and peptide ions that can be converted

into internal energy, is directly related to the mass of the collision gas. Consequently, these instrumental parameters need to be carefully monitored and controlled to ensure the inter- and intra-laboratory reproducibility of fragmentation patterns.

The optimized transitions are the essence of this kind of assay. The time and effort required to establish these conditions is the price to pay for the excellent quantitative performance experiments. However, once validate, such assays can be used in any study that involves the targeted protein. In QQQ-based MRM analysis no full mass spectra are recorded unlike in other MS-based proteomic techniques [79]. The non-scanning nature of this methodologies translates into an increased sensitivity by one or two orders of magnitude compared with conventional 'full scan' techniques. Additionally, it results in a linear response over a wide dynamic range up to five orders of magnitude. This enables the detection of low abundance proteins in highly complex mixtures, which is crucial for systematic quantitative studies [80].

However, the development of a glycopeptide quantification method is a very complex task mainly due to the absence of exogenous glycopeptide standards and to the incomplete proteolytic digestion caused by steric hindrance due to glycan moiety.

### 1.2.3.2 Parallel Reaction Monitoring (PRM)

Recent studies report that targeted experiments known as Parallel Reaction Monitoring (PRM) can also be performed on hybrid quadrupole-Orbitrap (q-OT) and quadrupole time-of-flight (q-TOF) mass spectrometers [81].
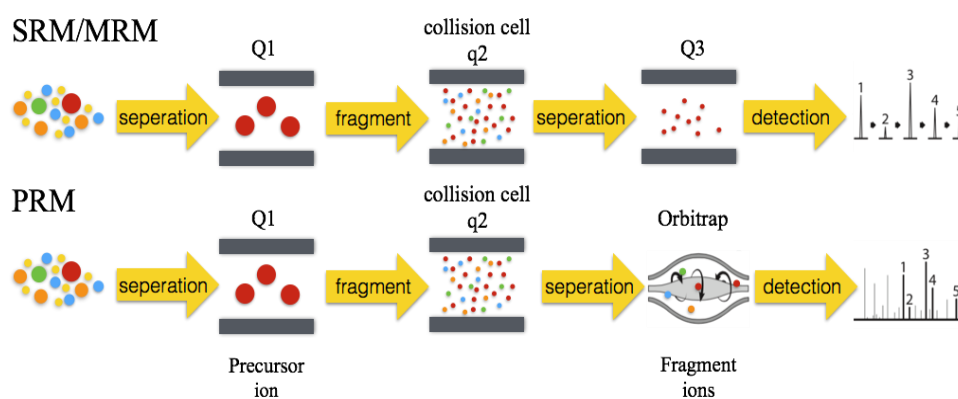
*Figure 1.22 Diagram comparison between MRM and PRM methodologies (www.thermofisher.com).*

The Orbitrap in a q-OT replaces the third quadrupole (Q3) mass analyser of a QQQ. Studies using targeted MRM and PRM methodologies have demonstrated that both have comparable sensitivity with similar linearity, dynamic range, precision, and repeatability for proteins quantification [81, 82]. However, PRM has certain advantages over MRM, such as it is relatively easier to build the data acquisition method because a priori selection of target transitions is not required [83]. Furthermore, PRM provides high specificity as the MS/MS data is acquired in high resolution modality that can separate co-isolated background ions from the target ions. In MRM, only three to five transitions are monitored, whereas in PRM a full MS/MS spectrum is acquired that contains all the potential product ions and confirms identity of the target peptide.

Targeted method based on PRM-MS has been successfully applied in relative quantification of proteins and their posttranslational modifications (PTMs). PRM is also convenient to study PTMs that are very low in percentage and are challenging to identify and quantify by untargeted proteomics methods.

In PRM, as in MRM, endogenous peptides that are quantifiable surrogates of protein of interest are first selected. The selected peptides should be specific and stoichiometric to the protein of interest [84]. Multiple MS/MS data points of the peptides are then acquired across the elution profile. The data is then useful to quantify peptides, and thus their corresponding proteins, in the samples.

## 1.3 Mass spectrometry applied to protein glycosylation studies

With the developed method, which couple electrospray ionization (ESI)-MS whit liquid chromatography (LC-MS), it is even possible to separate isomers to give a comprehensive view on all structures present in a complex sample. Mass spectrometry, thanks to its high sensitivity and selectivity, is a powerful tool in glycomics and glycoproteomics [30, 85], allowing characterisation of glycoproteins either at the glycopeptides level or after de-glycosylation, characterizing separately the peptides containing the glycosylation sites and the glycans [86].
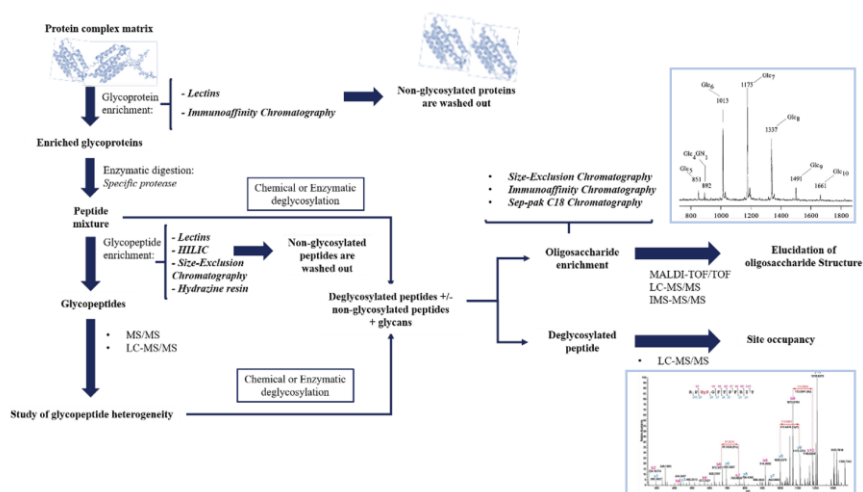


*Figure 1.23 Mass spectrometry (MS)-based glycoproteomics approach workflow. Adapted from [29].*

The bottom-up proteomic studies provide the use of simple protocols of in solution or in situ digestion to release peptides from proteins to be submitted to the MS analysis. Unlike canonical peptides, glycosylated peptides have a significantly higher negative charge and acidity, which affects their ionization efficiency. This event clearly speculates that the MS signals of glycosylated peptides undergo an effect of ion suppression due to the competition with a much higher number of non-modified peptide counterpart displaying a more intense ion current.

Many methods have been developed during the last decades to overcome the challenges associated to the glycoproteome analysis by mass spectrometry. Several purification methods are combined to overcome ionization difficulties and to avoid the suppression of the glycopeptide signal from the non-modified peptides in complex matrix. Commonly used methods for purifying glycopeptides before mass spectrometry analysis is reverse phase purification (RP) on the high-pressure liquid chromatography (HPLC) system. This method ensures a separation of the different glycopeptides based on their different amino acid sequence since the retention mechanism is governed by the hydrophobicity of peptide portion [87]. The techniques that could be hyphenated with RP-HPLC include purification approaches based on hydrazine resins, lectin affinity chromatography and gel filtration, or size exclusion chromatography (Figure 1.23).

Hydrophilic interaction chromatography (HILIC) is a purification technique widely used in the last few years to purify glycopeptides. HILIC separation exploits the polar interactions between the hydroxyl groups of glycans and the stationary phase. The efficient removal of the non-glycosylated counterpart takes place by using organic solvent washing followed by glycopeptide elution with an aqueous buffer [88]. This method consents the glycopeptides separations based on their oligosaccharide moiety, and its results are extremely useful when there is more than one glycosylation on the same peptide. Among the main enrichment techniques, the use of immunoaffinity columns plays an important role in the characterisation of site-specific occupancy due to the natural affinity of glycan epitopes to the specific antibodies on the functional regions [89, 90].

### 1.3.1 Glycans and glycopeptides characterisation

Glycans or glycopeptides analysis is very challenging due to the limited quantities that are released from glycoproteins. Since the structures of a glycans may are connected to the expression, activity, and accessibility of the several

biosynthetic enzymes, it is not possible to use recombinant DNA technology in order to produce large quantities of glycans for structural and functional studies as it is for proteins. Although today technologies such as ESI-MS/MS are commonly preferred for proteomics studies, MALDI-MS is still an effective technique for N-glycan analysis of simple or complex matrices [91, 92].

MALDI-TOF analysis of the intact glycan mixture and the attribution of the different structures is carried out by checking the molecular weight and the knowledge of molecular pathways for the biosynthesis of oligosaccharides. However, this approach is convenient in glycoforms profiling, but nevertheless it does not provide structural information such as sugar anomericity, neither on glycans site-specificity. To obtain this type of information, the combination of a profile by MALDI-TOF, with experiments of tandem mass spectrometry by post-source decay (PSD) or collision-induced dissociation (CID), is generally required. The LC-MS/MS of whole glycopeptides provide, instead, more information about the site-specificity of glycans.

Usually the CID fragmentation of the glycopeptides produces a wide fragmentation on the oligosaccharide portion (such as typical oxonium ion fragment showed in Table 1.2), and y-and b-type ions from the peptide moiety, therefore these MS/MS data are useful for assigning the glycan compositions. In an analogous way, neutral losses of saccharides such as hexose (Hex: 162 Da), N-acetylhexosamine (HexNac: 203 Da), fucose (Fuc: 146 Da), N-acetylneuraminic acid (NeuAc: 291 Da) could be used to indicate the presence of glycopeptides in the mass spectra. $MS^n$ experiments, on glycans moiety or directly on glycopeptides, are useful to characterize glycosidic structures present in glycoproteins as well as the type of branching, the sequence of the antennas, and the possible presence of modifying groups (e.g., sulphate, phosphate, acetyl groups, etc.).

*Table 1.2 Typical oxonium ion and relative cross-ting fragments.*

| Glycan fragment + H$^+$ | Oxonium ions (m/z) | Oxonium-derived ions (m/z) |
|---|---|---|
| Fuc | 147.07 | 129.05 |
| Hex | 163.06 | 109.03 |
| | | 115.04 |
| | | 127.04 |
| | | 145.05 |
| HexNac | 204.09 | 138.06 |
| | | 126.06 |
| | | 144.07 |
| | | 168.07 |
| | | 186.08 |
| NeuAc | 292.10 | 274.09 |
| HexNac+ Hex | 366.14 | 204.09 |

Structural characterisation of complex carbohydrates is labor-intensive and time-consuming and requires orthogonal methods to identify: (i) the constituent monomers, (ii) their sequence including branching points, (iii) configuration of the glycosidic bond, (iv) position of the glycosidic bond, (iv) anomeric configurations [93]. Many software applications (open access and commercial) are currently supportive of the above objectives, and although these tools differ in computational strategies they share the same fundamentals: (i) determining the accurate mass in the precursor MS spectra; (ii) supposition of glycan structural composition and the peptide sequence from the analysis of the LC-MS/MS spectra; (iii) interrogation of theoretical spectra contained in databases of proteins and glycans that matches the marker fragment ions for the peptide backbone and glycan, or attempt to de-novo sequencing the portion of the glycan; and (iv) data validity score [94]. Such predictive software are principally valuable for implementation of the LC-MS/MS method based on MRM; an MS modality widely used for the accurate quantification of molecules. The main applications of MRM were limited to the metabolomics and proteomics, but their usefulness is also expanding in the field of glycoscience. MRM methods are very suitable for robust, fast, sensitive, and specific quantitative analysis of multiple target

compounds, simultaneously detected also in the presence of other more abundant compounds (for instance, in complicated mixtures, such as biofluids and serum).

Since MRM is a targeted approach, both the knowledge of the mass and charge state of the analyte and its fragmentation behaviour in CID are essential for the choice of best transitions during the method development. Numerous studies have focused on the fragmentation model of glycans and glycopeptides [95], emphasizing that the typical fragmentation occurs on glycosidic bonds under low energy CID conditions normally used in triple quadrupole instruments. In these low energy conditions, the cross-ring cleavages are often low abundant, while intense fragments of glycan (oxonium ions) and those related to peptide fragmentation have been shown to be characteristic of each N-glycopeptides [96, 97]. The fragmentation pattern of high-energy collision dissociation (HCD) is an alternative type of fragmentation method characterized by higher activation energy and shorter activation time comparing to the traditional ion trap CID (Figure 1.24). HCD for peptides with post-translational modifications (PTMs) can provide both the sequence information (b- and y-type fragment ions) and the localization of the modification sites as it can identify CID-labile PTMs. Thus, high mass accuracy $MS^2$ spectra have been successfully applied for PTMs studies, as certain diagnostic ions specific for HCD could be recognized for PTMs identification [98].
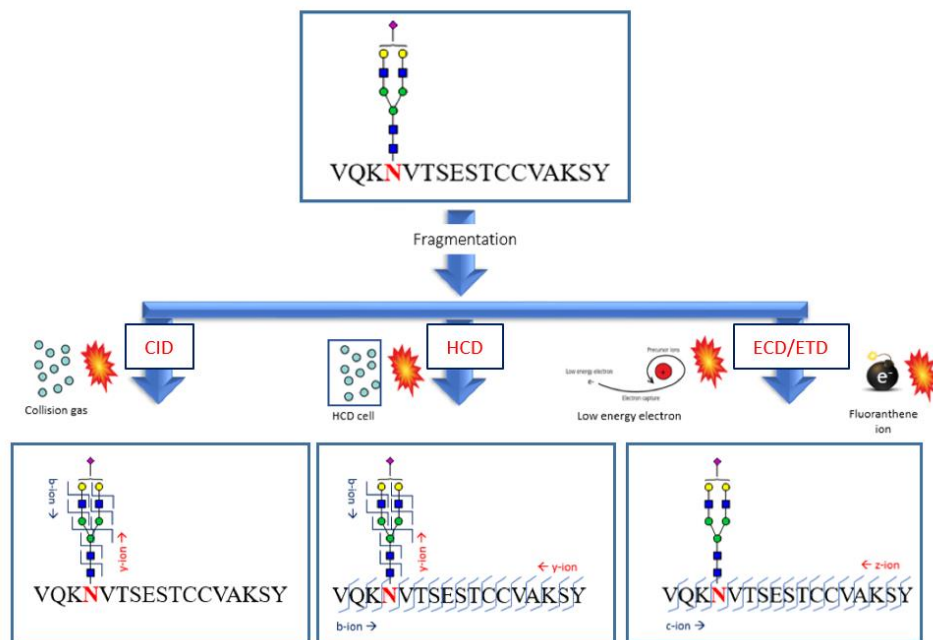
*Figure 1.24 Mechanism of collision induced dissociation (CID), high-energy collision dissociation (HCD), electron-capture dissociation (ECD), and electron transfer dissociation (ETD) fragmentation. Adapted from [29]*

Another fragmentation method used in the analysis of glycopeptides are electron-capture dissociation (ECD) and electron transfer dissociation (ETD). In both techniques, the glycan portion does not undergo fragmentation while the peptide fragments provide both the z and c ions (Figure 1.24). ECD experiments are typically performed on ionic resonance instruments of the Fourier transform cyclotron (FT-ICR) while the ETD can be performed in an ion trap and orbitrap mass spectrometers.

## 1.4 Aim of the study

Follicle Stimulating Hormone (FSH) is one of the most important hormones involved in human fertility and it is widely used for *in vitro* fertilization. While its role and its function are nowadays well known, the structural and functional contribution of glycosylation in hormonal action are not completely understood [99, 100]. In clinical settings, ELISA is the official method for quantifying serum hFSH but, with this methodology, some clinically important targets remain undistinguishable and/or undetectable due to the lack of specific antibodies capable of recognizing the different proteoforms generated by glycosylation.

Mass spectrometry applied to glycomic studies of glycosylated hormones is a powerful tool to characterise and quantify specifically and precisely glycoforms.

The aim of this project is to develop an innovative targeted strategy based on MRM and PRM tandem mass spectrometry analysis to characterise FSH glyco-isoforms in serum in a site-specific manner. In order to achieve this aim, the PhD project was divided into two major phases:

1) Characterisation of glycosylation patterns of FSH. Through the analysis of different FSH molecules, including the pituitary and some existing commercial forms such as reference r-hFSH based drug (GONAL-F ®) (Chapter 2). To fully characterise FSH, different complementary approaches have been used, such as bottom up and glycomic analysis. Furthermore, different glycovariants were generated through stress conditions (temperature, pH, etc.) and analysed by mass spectrometry for a deeper understanding of FSH glycosylation (Chapter 3). This part of the thesis was performed at Merck KGaA company (www.merckgroup.com) (project supported by an innovative and industrial National Operational Program)

2) Development of two advanced biomolecular targeted mass spectrometry methodologies in human serum: i) an MRM-based mass spectrometry method for absolute FSH quantification in woman sera using unmodified peptides (Chapter

4); ii) a sensitive and robust PRM-based MS method for the identification and the relative quantification of specific FSH glycoforms (Chapter 5).

Ultimately, the overall ambition of the project is to contribute in understanding the complex mechanism of action of glycans in human fertility and provide a robust alternative to ELISA for FSH quantification.

## 1.5 References

1.      Horgan, R.P. and L.C. Kenny, *'Omic'technologies: genomics, transcriptomics, proteomics and metabolomics.* The Obstetrician & Gynaecologist, 2011. **13**(3): p. 189-195.
2.      Bertozzi, C.R. and R. Sasisekharan, *Glycomics*, in *Essentials of Glycobiology. 2nd edition*. 2009, Cold Spring Harbor Laboratory Press.
3.      Reily, C., et al., *Glycosylation in health and disease.* Nature Reviews Nephrology, 2019. **15**(6): p. 346-366.
4.      Li, Q., et al., *Comprehensive structural glycomic characterisation of the glycocalyxes of cells and tissues.* Nature Protocols, 2020: p. 1-37.
5.      Dwek, R.A., *Glycobiology: toward understanding the function of sugars.* Chemical reviews, 1996. **96**(2): p. 683-720.
6.      Rudd, P.M. and R.A. Dwek, *Glycosylation: heterogeneity and the 3D structure of proteins.* Critical reviews in biochemistry and molecular biology, 1997. **32**(1): p. 1-100.
7.      Slovin, S.F., S.J. Keding, and G. Ragupathi, *Carbohydrate vaccines as immunotherapy for cancer.* Immunology and cell biology, 2005. **83**(4): p. 418-428.
8.      Musselli, C., P.O. Livingston, and G. Ragupathi, *Keyhole limpet hemocyanin conjugate vaccines against cancer: the Memorial Sloan Kettering experience.* Journal of cancer research and clinical oncology, 2001. **127**(2): p. R20-R26.
9.      Daniels, M.A., K.A. Hogquist, and S.C. Jameson, *Sweet'n'sour: the impact of differential glycosylation on T cell responses.* Nature immunology, 2002. **3**(10): p. 903-910.
10.     Dwek, M.V., H.A. Ross, and A.J. Leathem, *Proteome and glycosylation mapping identifies post-translational modifications associated with aggressive breast cancer.* PROTEOMICS: International Edition, 2001. **1**(6): p. 756-762.
11.     Watanabe, Y., et al., *Exploitation of glycosylation in enveloped virus pathobiology.* Biochimica et Biophysica Acta (BBA)-General Subjects, 2019. **1863**(10): p. 1480-1497.
12.     Vankadari, N. and J.A. Wilce, *Emerging COVID-19 coronavirus: glycan shield and structure prediction of spike glycoprotein and its interaction with human CD26.* Emerging microbes & infections, 2020. **9**(1): p. 601-604.
13.     Padoan, A., et al., *IgA-Ab response to spike glycoprotein of SARS-CoV-2 in patients with COVID-19: a longitudinal study.* Clinica chimica acta, 2020.
14.     Ou, X., et al., *Characterisation of spike glycoprotein of SARS-CoV-2 on virus entry and its immune cross-reactivity with SARS-CoV.* Nature communications, 2020. **11**(1): p. 1-12.
15.     Walls, A.C., et al., *Structure, function, and antigenicity of the SARS-CoV-2 spike glycoprotein.* Cell, 2020.
16.     Apweiler, R., H. Hermjakob, and N. Sharon, *On the frequency of protein glycosylation, as deduced from analysis of the SWISS-PROT database.* Biochimica et Biophysica Acta (BBA)-General Subjects, 1999. **1473**(1): p. 4-8.
17.     Han, L. and C.E. Costello, *Mass spectrometry of glycans.* Biochemistry (Moscow), 2013. **78**(7): p. 710-720.
18.     Stanley, P., *Golgi glycosylation.* Cold Spring Harbor perspectives in biology, 2011. **3**(4): p. a005199.

19. Ferris, S.P., V.K. Kodali, and R.J. Kaufman, *Glycoprotein folding and quality-control mechanisms in protein-folding diseases.* Disease models & mechanisms, 2014. **7**(3): p. 331-341.

20. Gemmill, T.R. and R.B. Trimble, *Overview of N-and O-linked oligosaccharide structures found in various yeast species.* Biochimica et Biophysica Acta (BBA)-General Subjects, 1999. **1426**(2): p. 227-237.

21. Moremen, K.W., M. Tiemeyer, and A.V. Nairn, *Vertebrate protein glycosylation: diversity, synthesis and function.* Nature reviews Molecular cell biology, 2012. **13**(7): p. 448-462.

22. Reily, C., et al., *Glycosylation in health and disease.* Nature Reviews Nephrology, 2019: p. 1.

23. Shajahan, A., et al., *Glycomic and glycoproteomic analysis of glycoproteins—a tutorial.* Analytical and bioanalytical chemistry, 2017. **409**(19): p. 4483-4505.

24. Chen, Zhengwei, Junfeng Huang, and Lingjun Li. *"Recent advances in mass spectrometry (MS)-based glycoproteomics in complex biological samples."* TrAC Trends in Analytical Chemistry, 2019.**118**: 880-892.

25. Dong, Xue, et al. *"Advances in mass spectrometry-based glycomics."* Electrophoresis, 2018. **39** (24): 3063-3081.

26. Domon, Bruno, and Ruedi Aebersold. *"Mass spectrometry and protein analysis."* Science, 2006. **312**(5771): 212-217.

27. Zaia, Joseph. *"Mass spectrometry and glycomics."* Omics: a journal of integrative biology, 2010. **14** (4): 401-418.

28. Ruhaak LR, Xu G, Li Q, Goonatilleke E, Lebrilla CB. *Mass Spectrometry Approaches to Glycomic and Glycoproteomic Analyses*. Chem Rev. 2018;**118**(17):7886-7930. doi:10.1021/acs.chemrev.7b00732

29. Illiano, A., et al., *Protein Glycosylation Investigated by Mass Spectrometry: An Overview.* Cells, 2020. **9**(9).

30. Dalpathado, D.S. and H. Desaire, *Glycopeptide analysis by mass spectrometry.* Analyst, 2008. **133**(6): p. 731-738.

31. Annesley, T.M., *Ion suppression in mass spectrometry.* Clinical chemistry, 2003. **49**(7): p. 1041-1044.

32. Tarentino, A.L., C.M. Gomez, and T.H. Plummer Jr, *Deglycosylation of asparagine-linked glycans by peptide: N-glycosidase F.* Biochemistry, 1985. **24**(17): p. 4665-4671.

33. Zauner, G., et al., *Protein O-glycosylation analysis.* Biological chemistry, 2012. **393**(8): p. 687-708.

34. Geyer, H. and R. Geyer, *Strategies for analysis of glycoprotein glycosylation.* Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics, 2006. **1764**(12): p. 1853-1869.

35. Ciucanu, I. and F. Kerek, *A simple and rapid method for the permethylation of carbohydrates.* Carbohydrate research, 1984. **131**(2): p. 209-217.

36. Srikanth, J., R. Agalyadevi, and P. Babu, *Targeted, Site-specific quantitation of N-and O-glycopeptides using 18 O–labeling and product ion based mass spectrometry.* Glycoconjugate journal, 2017. **34**(1): p. 95-105.

37. De Hoffmann, Edmond. *"Mass spectrometry."* Kirk-Othmer Encyclopedia of Chemical Technology, 2000.

38. Gross, Jürgen H. *Mass spectrometry: a textbook*. Springer Science & Business Media, 2006.

39.	Yates, John R., Cristian I. Ruse, and Aleksey Nakorchevsky. *"Proteomics by mass spectrometry: approaches, advances, and applications."* Annual review of biomedical engineering, 2009.**11**: 49-79.

40.	Whitehouse, Craig M., et al. "*Electrospray interface for liquid chromatographs and mass spectrometers."* Analytical chemistry, 1985. **57** (3): 675-679.

41.	Hopley, Chris, et al. "*Towards a universal product ion mass spectral library–reproducibility of product ion spectra across eleven different mass spectrometers." Rapid Communications in Mass Spectrometry*: An International Journal Devoted to the Rapid Dissemination of Up-to-the-Minute Research in Mass Spectrometry, 2008. **22** (12): 1779-1786.

42.	Awad, Hanan, Mona M. Khamis, and Anas El-Aneed. "*Mass spectrometry, review of the basics: ionization."* Applied Spectroscopy Reviews, 2015. **50**.(2): 158-175.

43.	Wolk DM, Clark AE. *Matrix-Assisted Laser Desorption Time of Flight Mass Spectrometry.* Clin Lab Med. 2018 Sep; **38** (3): 471-486. doi: 10.1016/j.cll.2018.05.008. PMID: 30115392.

44.	Aceña, Jaume, et al. "*Advances in liquid chromatography–high-resolution mass spectrometry for quantitative and qualitative environmental analysis."* Analytical and bioanalytical chemistry, 2015. **407** (21): 6289-6299.

45.	Zenobi, Renato, and Richard Knochenmuss. "*Ion formation in MALDI mass spectrometry*." Mass spectrometry reviews, 1998. **17** (5): 337-366.

46.	Debois, Delphine, et al. *"MALDI in-source decay, from sequencing to imaging".* Top Curr Chem. 2013; 331:117-41. doi: 10.1007/128_2012_363. PMID: 22976457.

47.	Brown, Robert S., and John J. Lennon. *"Mass resolution improvement by incorporation of pulsed ion extraction in a matrix-assisted laser desorption/ionization linear time-of-flight mass spectrometer."* Analytical chemistry, 1995. **67** (13): 1998-2003.

48.	Kaufmann, Raimund. "*Matrix-assisted laser desorption ionization (MALDI) mass spectrometry: a novel analytical tool in molecular biology and biotechnology.*" Journal of biotechnology, 1995. **41** (2-3): 155-175.

49.	Yamagaki T, Suzuki H, and Tachibana K. *A comparative study of the fragmentation of neutral lactooligosaccharides in negative-ion mode by UV-MALDI-TOF and UV-MALDI ion-trap/TOF mass spectrometry. J Am Soc Mass Spectrom*. 2006 Jan;**17**(1):67-74. doi: 10.1016/j.jasms.2005.09.004. Epub 2005 Dec 15. PMID: 16352446.

50.	Karas, M. and R. Krüger, *Ion formation in MALDI: the cluster ionization mechanism.* Chemical reviews, 2003. **103**(2): p. 427-440.

51.	Harvey, D.J., *Matrix-assisted laser desorption/ionization mass spectrometry of carbohydrates.* Mass spectrometry reviews, 1999. **18**(6): p. 349-450.

52.	Fenn, John B., et al. "*Electrospray ionization–principles and practice.*" Mass Spectrometry Reviews, 1990. **9** (1): 37-70.

53.	Marginean, Ioan, et al. "*Analytical characterization of the electrospray ion source in the nanoflow regime*." Analytical chemistry, 2008. **80** (17): 6573-6579.

54.	Wang, Yan, et al. "*A "soft" and "hard" ionization method for comprehensive studies of molecules.*" Analytical chemistry, 2018. **90** (24): 14095-14099.

55. Holčapek, Michal, Robert Jirásko, and Miroslav Lísa. "*Recent developments in liquid chromatography–mass spectrometry and related techniques*." Journal of Chromatography A, 2012. **1259**: 3-15.

56. Bonvin, Grégoire, Julie Schappler, and Serge Rudaz. "*Capillary electrophoresis–electrospray ionization-mass spectrometry interfaces: Fundamental concepts and technical developments*." Journal of Chromatography A, 2012, **1267**: 17-31.

57. Amoresano, A., et al., *Technical advances in proteomics mass spectrometry: identification of post-translational modifications.* Clinical Chemistry and Laboratory Medicine (CCLM), 2009. **47**(6): p. 647-665.

58. Dawson, Peter H., ed. *Quadrupole mass spectrometry and its applications*. Elsevier, 2013.

59. Jurinke, Christian, Paul Oeth, and Dirk van den Boom. "*MALDI-TOF mass spectrometry.*" Molecular biotechnology, 2004. **26** (2): 147-163.

60. Hosseini, Samira, and Sergio O. Martinez-Chapa. "*Principles and mechanism of MALDI-ToF-MS analysis.*" Fundamentals of MALDI-ToF-MS Analysis 2017: 1-19.

61. Makarov, Alexander. "*Electrostatic axially harmonic orbital trapping: a high-performance technique of mass analysis*." Analytical chemistry, 2000. **72** (6): 1156-1162.

62. Nikolaev EN, Kostyukevich YI, Vladimirov GN. *Fourier transform ion cyclotron resonance (FT ICR) mass spectrometry: Theory and simulations. Mass* Spectrom Rev. 2016 Mar-Apr;**35**(2):219-58. doi: 10.1002/mas.21422. Epub 2014 Feb 10. PMID: 24515872.

63. Perry, Richard H., R. Graham Cooks, and Robert J. Noll. "*Orbitrap mass spectrometry: instrumentation, ion motion and applications."* Mass spectrometry reviews, 2008. **27** (6): 661-699.

64. Eliuk, S. and A. Makarov, *Evolution of orbitrap mass spectrometry instrumentation.* 2015.

65. Hecht, Elizabeth S., et al. "*Fundamentals and Advances of Orbitrap Mass Spectrometry*." Encyclopedia of Analytical Chemistry: Applications, Theory and Instrumentation, 2006: 1-40.

66. Makarov, A., et al., *Performance evaluation of a hybrid linear ion trap/orbitrap mass spectrometer.* Analytical chemistry, 2006. **78**(7): p. 2113-2120.

67. Zubarev RA, Makarov A. *Orbitrap mass spectrometry*. Anal Chem. 2013 Jun 4;**85**(11):5288-96. doi: 10.1021/ac4001223. Epub 2013 May 13. PMID: 23590404.

68. Michalski, A., et al., *Mass spectrometry-based proteomics using Q Exactive, a high-performance benchtop quadrupole Orbitrap mass spectrometer.* Molecular & Cellular Proteomics, 2011. **10**(9).

69. Senko, M.W., et al., *Novel parallelized quadrupole/linear ion trap/Orbitrap tribrid mass spectrometer improving proteome coverage and peptide identification rates.* Analytical chemistry, 2013. **85**(24): p. 11710-11714.

70. March RE. *Quadrupole ion traps*. Mass Spectrom Rev. 2009 Nov-Dec;**28**(6):961-89. doi: 10.1002/mas.20250. PMID: 19492348.

71. Haag AM. *Mass Analyzers and Mass Spectrometers*. Adv Exp Med Biol. 2016;919:157-169. doi: 10.1007/978-3-319-41448-5_7. PMID: 27975216.72. ROEPSTORFE, P., *Proposal for a common nomenclature for sequence*

*ions in mass spectra of peptides.* Biomed. Mass Spectrom., 1984. **11**: p. 601-605.

73.     Roepstorff, Peter, and Jan Fohlman. "*Proposal for a common nomenclature for sequence ions in mass spectra of peptides*." Biomedical mass spectrometry 1984. 11 (11): 601-601.

74.     Domon, B. and C.E. Costello, *A systematic nomenclature for carbohydrate fragmentations in FAB-MS/MS spectra of glycoconjugates.* Glycoconjugate journal, 1988. **5**(4): p. 397-409.

75.     Reinhold, V.N., B.B. Reinhold, and C.E. Costello, *Carbohydrate molecular weight profiling, sequence, linkage, and branching data: ES-MS and CID.* Analytical chemistry, 1995. **67**(11): p. 1772-1784.

76.     Vidova V, Spacil Z. *A review on mass spectrometry-based quantitative proteomics: Targeted and data independent acquisition.* Anal Chim Acta. 2017 Apr **29**;964:7-23. doi: 10.1016/j.aca.2017.01.059. Epub 2017 Feb 2. PMID: 28351641.

77.     Sleno, L. and D.A. Volmer, *Ion activation methods for tandem mass spectrometry.* Journal of mass spectrometry, 2004. **39**(10): p. 1091-1112.78.        Paizs, B. and S. Suhai, *Fragmentation pathways of protonated peptides.* Mass spectrometry reviews, 2005. **24**(4): p. 508-548.

79.     Liebler DC, Zimmerman LJ. *Targeted quantitation of proteins by mass spectrometry*. Biochemistry. 2013 Jun 4;52(22):3797-806. doi: 10.1021/bi400110b. Epub 2013 Mar 27. PMID: 23517332; PMCID: PMC3674507.

80.     Gallien, S., E. Duriez, and B. Domon, *Selected reaction monitoring applied to proteomics.* Journal of Mass Spectrometry, 2011. **46**(3): p. 298-312.81.

81.     Rauniyar, N., *Parallel reaction monitoring: a targeted experiment performed using high resolution and high mass accuracy mass spectrometry.* International journal of molecular sciences, 2015. **16**(12): p. 28566-28581.

82.     Gallien, S. and B. Domon, *Detection and quantification of proteins in clinical samples using high resolution mass spectrometry.* Methods, 2015. **81**: p. 15-23.

83.     Rauniyar, N., *Parallel reaction monitoring: a targeted experiment performed using high resolution and high mass accuracy mass spectrometry.* International journal of molecular sciences, 2015. **16**(12): p. 28566-28581.

84.     Worboys, J.D., et al., *Systematic evaluation of quantotypic peptides for targeted analysis of the human kinome.* Nature methods, 2014. **11**(10): p. 1041-1044.

85.     Girolamo, F.D., et al., *The role of mass spectrometry in the "omics" era.* Current organic chemistry, 2013. **17**(23): p. 2891-2905.

86.     Yu, Y.Q., et al., *A rapid sample preparation method for mass spectrometric characterisation of N-linked glycans.* Rapid Communications in Mass Spectrometry: An International Journal Devoted to the Rapid Dissemination of Up-to-the-Minute Research in Mass Spectrometry, 2005. **19**(16): p. 2331-2336.

87.     Vreeker, G.C. and M. Wuhrer, *Reversed-phase separation methods for glycan analysis.* Analytical and bioanalytical chemistry, 2017. **409**(2): p. 359-378.

88.     Zhu, R., et al., *Glycoproteins enrichment and LC-MS/MS glycoproteomics in central nervous system applications*, in *Neuroproteomics*. 2017, Springer. p. 213-227.

89. PEKONEN, F., D.M. WILLIAMS, and B.D. WEINTRAUB, *Purification of thyrotropin and other glycoprotein hormones by immunoaffinity chromatography.* Endocrinology, 1980. **106**(5): p. 1327-1332.

90. Liu, T., et al., *Human plasma N-glycoproteome analysis by immunoaffinity subtraction, hydrazide chemistry, and mass spectrometry.* Journal of proteome research, 2005. **4**(6): p. 2070-2080.

91. Kamiyama, T., et al., *Identification of novel serum biomarkers of hepatocellular carcinoma using glycomic analysis.* Hepatology, 2013. **57**(6): p. 2314-2325.

92. de Oliveira, R.M., C.A. Ornelas Ricart, and A.M. Araujo Martins, *Use of mass spectrometry to screen glycan early markers in hepatocellular carcinoma.* Frontiers in Oncology, 2018. **7**: p. 328.

93. Alley Jr, W.R. and M.V. Novotny, *Structural glycomic analyses at high sensitivity: a decade of progress.* Annual review of analytical chemistry, 2013. **6**: p. 237-265.

94. Abrahams, J.L., et al., *Recent advances in glycoinformatic platforms for glycomics and glycoproteomics.* Current Opinion in Structural Biology, 2020. **62**: p. 56-69.

95. Kolli, V. and E.D. Dodds, *Energy-resolved collision-induced dissociation pathways of model N-linked glycopeptides: implications for capturing glycan connectivity and peptide sequence in a single experiment.* Analyst, 2014. **139**(9): p. 2144-2153.

96. Huang, J., et al., *Quantitation of human milk proteins and their glycoforms using multiple reaction monitoring (MRM).* Analytical and Bioanalytical Chemistry, 2017. **409**(2): p. 589-606.

97. Carpentieri, A., et al., *Glycoproteome study in myocardial lesions serum by integrated mass spectrometry approach: preliminary insights.* European Journal of Mass Spectrometry, 2010. **16**(1): p. 123-149.

98. Quan, L. and M. Liu, *CID, ETD and HCD fragmentation to study protein post-translational modifications.* Modern Chemistry & Applications, 2012.

99. Willey, K.P., *An elusive role for glycosylation in the structure and function of reproductive hormones.* Human reproduction update, 1999. **5**(4): p. 330-355.

100. Varki, A., *Biological roles of oligosaccharides: all of the theories are correct.* Glycobiology, 1993. **3**(2): p. 97-130.

# Chapter 2    FSH glycosidic profile characterisation

## 2.1    Follicle Stimulating Hormone (FSH)

FSH is a member of the glycoprotein hormone family, which also includes LH, CG and thyroid stimulating hormone (TSH), belonging to the superfamily of cystine-knot growth factors (CKGF). FSH is secreted by the anterior pituitary gland as heterodimers (two dissimilar subunits) of ~30 kDa, consisting of non-covalently associated α- and β-subunits [1]. Glycoprotein hormones share a common α chain with the same amino acid sequence but derive functional specificity from different β-chains (hormone-specific) [2].

- The α subunit, with relative molecular mass of ~14 kDa, is composed of 92 amino acids carrying 2 carbohydrate moieties linked to Asn-52 and Asn-78 [1, 3].

- The β subunit, with relative molecular mass of ~17 kDa, is composed of 111 amino acids carrying 2 carbohydrate moieties linked to Asn-7 and Asn-24 [1, 3].

FSH plays a key role in the proliferation and development of ovarian granulosa cells in women and spermatogenesis in men [4].

Gonadotropin-releasing hormone (GnRH) from the hypothalamus regulates the synthesis and secretion of gonadotropins by the anterior pituitary. Interestingly, FSH is synthesized and secreted in different isoforms throughout the menstrual cycle [5]. Four naturally occurring FSH glycoforms (tetra-glycosylated FSH[24], tri-glycosylated FSH[21] and FSH[18], di-glycosylated FSH[15]) have been identified according to their migration in SDS-PAGE [6]. These isoforms differ in their glycosylation content (occupancy) and antennarity of the β subunit. In all four glycoforms the α subunit is fully glycosylated at both sites (Asn-52 and -78), while β subunit can be glycosylated at both (FSH[24]), only one glycosite (FSH[21] and FSH[18]) or not glycosylated at all (FSH[15]). The difference in migration

properties between FSH[21] and FSH[18] reflects the greater number of antennary structures on the β subunit at Asn-7 compared with at Asn-24, as tetra-antennary glycan structures have only been reported at Asn-7 [7].

The glycoforms composition of FSH in circulation fluctuates in a characteristic manner throughout the follicular phase; during the early follicular phase the more acidic (highly sialylated) fully glycosylated FSH glycoforms prevail, while in the mid-cycle (i.e. during the peri-ovulatory phase) there is a shift to less-acidic FSH glycoforms (hypo-glycosylated isoform) [8, 9]. Importantly, these more acidic isoforms survive longer in circulation but may exhibit partial differences in target organ bioactivity (i.e. FSH receptor, FSHR, binds hypo-glycosylated FSH at a 3-fold higher capacity than highly glycosylated FSH) [10-12].

Furthermore, di-glycosylated FSH[15] is more abundant in younger women, whilst tetra-glycosylated and highly sialylated forms are more predominant in peri/postmenopausal women [5, 13]; Hypo-glycosylated FSH dominates in 21-24 year old women, fully- and hypo-glycosylated FSH are roughly equivalent in 39-41 year-old women and fully-glycosylated FSH is dominant in 55-81 year-old women and men [14]. This suggests that the glycoforms composition of circulating FSH is functionally relevant in fertile women, reflecting their endocrine status. Nowadays, exogenous FSH (recombinant r-hFSH or urinary purified u-hFSH) is used clinically to induce ovulation in women as part of ART treatments (i.e. ovulation induction and controlled ovarian stimulation) and to treat infertility due to gonadotropin deficiency in men [15]. About the mechanism of action, r-hFSH binds to and activates the follicle stimulating hormone receptor, a G-coupled transmembrane receptor regulating cellular metabolism and related survival/maturation. The r-hFSH mimics the actions of endogenous FSH, which is required for normal follicular growth, maturation, and gonadal steroid production. The glycosylation of FSH plays a fundamental role in its mechanism of action; Glycans linked to the α-subunit are critical for signal transduction, and glycans of the β-subunit play a crucial role in clearance of the dimer from the

circulation [16, 17]. Alternative glycosylation on FSH not only affects the metabolic clearance and net *in vivo* biopotency of the hormone, but also some glycosylation variants provoke differential or even unique effects at the target cell level. Therefore, carbohydrate residues play a decisive role in the biological activity of FSH but to date their mechanism of action has not been fully deciphered [18]. Therefore, today the development of new methodologies aimed at characterizing and identifying the glycosidic profile of biomolecules is a current topic of great scientific interest. In the first year, the study was focused on the characterisation of the FSH glycosylation using different mass spectrometry methodology such as MALDI-MS (MALDI-TOF) and LC-MS/MS (LC-Q/TOF).

The first part is carried out on the recombinant FSH (r-hFSH) drug, GONAL-F ®, to study the molecule and its glycosylation by classical mass spectrometry methodologies. The strategy consists in a first liquid chromatography (LC) to separate the two subunits of the molecule; then, the glycan moiety was removed from the aminoacidic portion and analysed separately by MALDI-TOF mass spectrometry in order to obtain the overall glycan profile of the two subunits (α and β-FSH). The site-specific glycan profile was obtained by performing a further off-line liquid chromatography prior to glycan removing and mass spectrometry analysis. Subsequently, a faster and more sensitive methodology was implemented to perform a glycopeptides direct analysis; with this approach, we obtained and compared the site-specific glycosidic profile of various forms of FSH, including pituitary (standard), urinary (FOSTIMON drug) and other two recombinant molecules (BEMFOLA and OVALEAP drugs), without any off-line chromatography.

## 2.2 Materials and Methods

### 2.2.1 Chemicals and reagents

Commercial GONAL-F ® 900UI/1,5 ml (66μg/1.5ml), BEMFOLA 450UI/0.75ml, OVALEAP 450UI/0.75ml, FOSTIMON 75 EI/UI; Standard pituitary follitropin, Trifluoro- acetic acid (TFA), acetonitrile (ACN), urea, Tris, EDTA, dithiothreitol (DTT), iodoacetamide (IAM), dimethyl sulfoxide (DMSO), sodium hydroxide NaOH, methyl iodide and α-Cyano-4-hydroxycinnamic acid were purchase from sigma; Chloroform and Methanol were from Fluka; PNG-ase F, trypsin and chymotrypsin from Roche.

### 2.2.2 Characterisation of α and β-FSH glycans by classical MS approach

#### 2.2.2.1 FSH α and β subunits purification

About 400 ug r-hFSH (GONAL-F ®) were purified by HPLC on a Vydac C4 column (25X0.46 cm, 5 pm) using 0.1 % TFA (solvent A) and 0.07% TFA in 95% ACN (solvent B). The glycoproteins were eluted by means of a linear gradient of 18-43% solvent B over 25 min. Fractions containing α and β chains were collected, dried in a Speed-Vac centrifuge (Savant), lyophilized twice and stored at -20°C.

#### 2.2.2.2 Tryptic enzymatic hydrolyses

The two fractions containing α-FSH and β-FSH subunits (Figure 2.2) were collected and submitted to reduction, carbamidomethylation and tryptic digestion.

The two fractions were dissolved in denaturing buffer (6M urea, 300 mM Tris pH 8.0, 10mM EDTA) containing DTT (10-fold molar excess on the Cys residues) at 37 °C for 2 h and then iodoacetamide (IAM) was added to perform carbamidomethylation using an excess of alkylating agent (5-fold molar excess on thiol residues). The mixture was then incubated in the dark at room temperature for 30 minutes. The product was purified by

Chloroform/Methanol/H2O precipitation. Supernatant was removed and the pellet was dried down. Digestion of protein mixture was carried out in AMBIC 10mM using trypsin at a 50:1 protein: enzyme mass ratio. The sample was incubated at 37°C for 16 h.

### 2.2.2.3 Enzymatic de-glycosylation

An aliquot of tryptic digestion mixtures was de-glycosylated by PNG-ase F in 50mM AMBIC, pH=8.5, overnight at 37 °C. N-linked oligosaccharide chains released by PNGase F were separated from peptides by reverse phase chromatography on prepacked Sep-pak cartridges.

### 2.2.2.4 Peptides MALDI-TOF/TOF analysis

Peptides mixtures were resuspended in formic acid 0.1%. Subsequently samples were mixed to the matrix consisting of a solution of 10 mg/mL α-Cyano-4-hydroxycinnamic acid (α-Cyano) on a MALDI plate. The matrix was prepared dissolving α-Cyano in ACN and 50mM citric acid 7:3 (v/v). MALDI-profiling in reflectron ion mode was performed on a AB SCIEX 5800 MALDI-TOF/TOF mass spectrometer equipped with Voyager software and ms spectra were acquired in the positive mode.

### 2.2.2.5 Oligosaccharides MALDI-TOF/TOF analysis

The mixture of N-linked oligosaccharides was submitted to per-methylation by treatment with dimethyl sulfoxide (DMSO) containing powdered sodium hydroxide (NaOH) for 2 h. Successively methyl iodide was added to the sample and the reaction conducted for 1 h at room temperature. The extraction of oligosaccharides was performed with chloroform and water with the oligosaccharides extracted in chloroform fraction.

Permethylated N-linked glycans were resuspended in water and methanol 50% 1:1 (v/v). Subsequently samples were mixed on MALDI plate to the matrix consisting of a solution of 10 mg/mL Dihydroxybenzoic acids (DHB). Mass

spectrometry analysis was performed on an AB SCIEX 5800 MALDI-TOF/TOF mass spectrometer equipped with Voyager software.

### 2.2.2.6 Site-specific N-glycopeptides analysis

α and β-FSH glycopeptides were separate from peptides mixture by HPLC on a Vydac C18 column (25X0.46 cm, 5 pm) using 0.1 % TFA (solvent A) and 0.07% trifluoro- acetic acid in 95% acetonitrile (solvent B) and a linear gradient of 10-55% solvent B in 50 min and 55-95 % solvent B in 5min. Each fraction was manually collected, dried in a Speed-Vac centrifuge and lyophilized twice prior to mass spectrometric analysis. The fractions containing the glycopeptides were subjected to enzymatic de-glycosylation and per-methylation before MS and MS analysis as previously described.

### 2.2.3 Glycopeptides direct analysis by LC-MS/MS

### 2.2.3.1 Enzymatic digestions:

About 35µg of different FSH molecule, r-hFSH (GONAL-F ®, BEMBFOLA, OVALEAP), u-hFSH (FOSTIMON) and p-hFSH (pituitary hormone standard) were treated with 100 µl of 8 M GuCl, 1mM EDTA, 130mM Tris-HCl, pH 7.6 and then reacted with 6 µl of 500 mM DTT at 37 ° C for 1 hr. The glycoprotein was then alkylated with 12 µl of 500 mM IAM and the reaction was kept at ambient temperature for 30 min in dark.

The product was purified by Chloroform/Methanol/H2O precipitation. The supernatant was removed, and the pellet dried down. Digestion of glycoprotein was carried out in 2 M Urea, 50 mM Tris-HCl buffer, pH 8.0 with Chymotrypsin (Roche) in the ratio 1:20 for 4h at 37°C. The peptide and glycopeptides mixture was desalted on zip-tip C18 chromatography before mass spectrometry analysis.

### 2.2.3.2 Glycopeptides LC-MS/MS analyses

Chymotryptic mixture was resuspended in 20µl of 0,1% HCOOH, and an aliquot of 2 µl (10%) was analysed by LC-MS/MS using LTQ Orbitrap XL system

(Thermo Fisher) equipped with a nano-HPLC. After loading, the peptide mixture was first concentrated and desalted on the pre-column (C18 EasyColumn L=2 cm, ID=100mm, Thermo Fisher Scientific). The sample was then fractionated on a C18 reverse-phase capillary column (C18 EasyColumn L=10 cm, ID=7,5µm, 3 µm, Thermo Fisher Scientific) at a flow rate of 300nl/min with the gradient of eluent B (0,2% formic acid, 95% acetonitrile LC-MS Grade) and eluent A (0,2% formic acid, 2% acetonitrile LC-MS Grade) reported below.
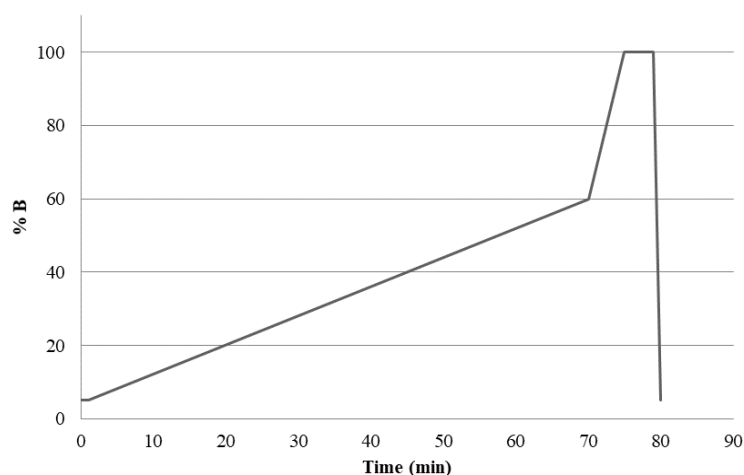


*Figure 2.1 Chromatographic Gradient used in peptide mixtures analysis*

Peptide analysis was performed using data-dependent acquisition (DDA) of one MS scan (mass range from 400 to 1800 m/z) followed by MS/MS scans of the five most abundant ions in each MS scan.

Raw data obtained from nano-LC-MS/MS analyses were analysed with MASCOT software (Matrix Science Boston, USA). Both peptide mass data and the data obtained by fragmentation spectra were included in the peak list for protein identification.

The research was done by setting the following parameters:

- Database: *NCBI*
- Taxonomy: *Homo Sapiens*
- Enzyme: *Chymotrypsin*

- Fixed modification: *Carbamidomethyl (C)*

- Variable modification: *Oxidation (M), Pyro-Glu (N-term Q)*

- Peptide tolerance: *10 ppm*

- MS/MS tolerance: *± 0,6 Da*

- Ions charge: *+2, +3, +4*

- Instrument: *ESI-FTICR*

## 2.3 Results and Discussions

### 2.3.1 Characterisation of α and β-FSH glycans by MALDI-TOF/TOF

#### 2.3.1.1 Overall profile of α and β-FSH glycans

Figure 2.2 shows the HPLC chromatographic profile of r-hFSH (GONAL-F ®) sample.
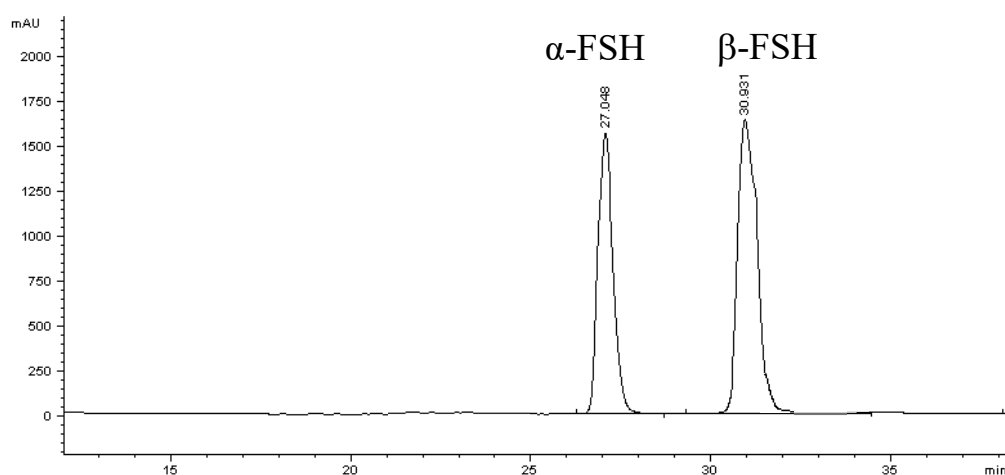


*Figure 2.2 Purified α (RT= 27.04) and β (RT=30.93) chain of r-hFSH.*

Since FSH molecule consist of two non-covalently linked subunits, liquid phase chromatography allowed us to efficiently separate them. MALDI-TOF analysis of the peptides mixtures from two collected fractions (27.04 and 30.93 min peaks), allowed the identification of the two polypeptide chains corresponding respectively to the α and β chain of the FSH heterodimer; the spectra analysis (Figure 2.3, Figure 2.4) showed 72% of sequence coverage of α subunit and

100% of sequence coverage of β subunit (identified peptides are marked in red end detail are showed in Table 2.1 and (Table 2.2 ):
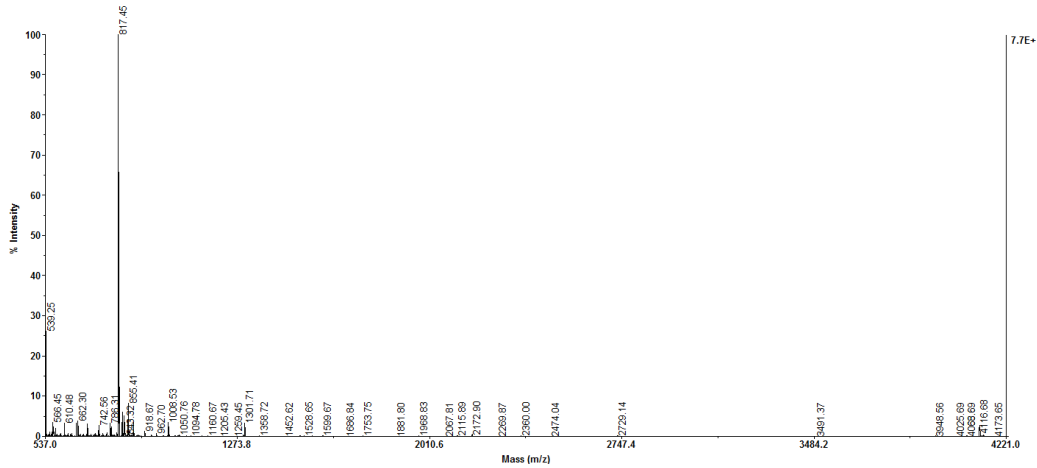


*Figure 2.3* MALDI-TOF spectrum of FSH **α**-chain tryptic peptides

## α-FSH sequence coverage (72%):

<span style="color:red">APDVQDCPEC</span>    <span style="color:red">TLQENPFFSQ</span>    <span style="color:red">PGAPILQCMG</span>    <span style="color:red">CCFSRAYPTP</span>    <span style="color:red">LRSKKTMLVQ</span>

<span style="color:red">K</span>N$_{52}$VTSESTCC    VAK<span style="color:red">SYNRVTV</span>    <span style="color:red">MGGFK</span>VEN$_{78}$HT    ACHCSTCYYH    KS

*Table 2.1 MH$^+$(m/z) recorded in MALDI-TOF spectrum (Figure 2.3) and corresponding peptides sequence.*

| Theoretical MH$^+$ (m/z) | Experimental MH$^+$ (m/z) | Peptide sequence | Start-end |
|---|---|---|---|
| 539.25 | 539.25 | SYNR | 64-67 |
| 719.41 | 719.40 | TMLVQK | 46-51 |
| 817.45 | 817.45 | AYPTPLR | 36-42 |
| 838.44 | 838.45 | VTVMGGFK | 68-75 |
| 1032.58 | 1032.57 | AYPTPLRSK | 36-44 |
| 1160.67 | 1160.67 | AYPTPLRSKK | 36-45 |
| 1358.69 | 1358.72 | SYNRVTVMGGFK | 64-75 |
| 4116.74 | 4116.68 | APDVQDCPECTLQENPFFSQPGAPILQCMGCCFSR | 1-35 |

The N$_{52}$VTSESTCCVAK and VEN$_{78}$HTACHCSTCYYHKS peptides were not identified as they both contain a N-glycosylation site (Asn-52 and Asn-78); this data confirms that the 100% site-occupancy of α-subunit glycosylation.
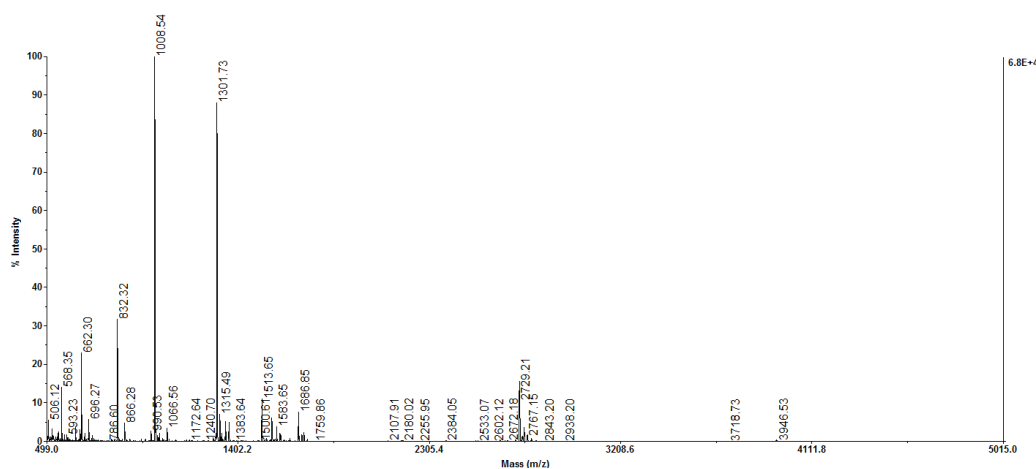


*Figure 2.4* MALDI-TOF spectrum of FSH **β**-chain tryptic peptides

### β-FSH sequence coverage (100%):

<span style="color:red">
<pre>
            10              20              30              40              50

NSCELTNITI    AIEKEECRFC    ISINTTWCAG   YCYTRDLVYK  DPARPKIQKT

            60              70              80              90             100

CTFKELVYET    VRVPGCAHHA    DSLYTYPVAT  QCHCGKCDSD  STDCTVRGLG

           110

PSYCSFGEMK    E
</pre>
</span>

*Table 2.2 MH+(m/z) recorded in MALDI-TOF spectrum (Figure 2.4) and corresponding peptides sequence.*

| Th. MH$^+$ (m/z) | Exp. MH$^+$ (m/z) | Peptide sequence | Start-end |
|---|---|---|---|
| 593.23 | 593.23 | EECR | 15-18 |
| 637.35 | 637.27 | DLVYK | 36-40 |
| 656.30 | 656.29 | TCTFK | 50-54 |
| 1008.53 | 1008.54 | ELVYETVR | 55-62 |
| 1301.72 | 1301.73 | DLVYKDPARPK | 36-46 |
| 1315.49 | 1315.49 | CDSDSTDCTVR | 87-97 |
| 1432.62 | 1432.60 | GLGPSYCSFGEMK | 98-110 |
| 1561.66 | 1561.64 | GLGPSYCSFGEMK | 98-111 |
| 1645.82 | 1645.82 | CTFKELVYETVR | 50-62 |

| | | | |
|---|---|---|---|
| 2172.93 | 2172.92 | FCISIN$_{24}$TTWCAGYCYTR | 19-35 |
| 2180.03 | 2180.02 | NSCELTN$_7$ITIAIEKEECR | 1-18 |
| 2729.09 | 2729.21 | CDSDSTDCTVRGLGPSYCSFGEMK | 87-110 |
| 2729.20 | 2729.21 | VPGCAHHADLSYTYPVATQCHCGK | 63-86 |
| 3718.72 | 3718.73 | ELVYETVRVPGCAHHADLSYTYPVATQCHCGK | 55-86 |

In its hypo-glycosylated form (hFSH[15]) the β-subunit of FSH is not glycosylated; in fact, in the spectrum of FSH β-chain tryptic peptides we found signals at m/z 2172.93 and 2180.03 attributed to the peptides 19-35 and 1-18 respectively, containing the unmodified N-glycosite Asn-7 and Asn-24.

The fractions containing α and β-FSH were then enzymatically de-glycosylated.

a)

b)

*Figure 2.5 MALDI-TOF spectrum of FSH α-chain (a) and β-chain (b) de-glycosylated tryptic peptides*
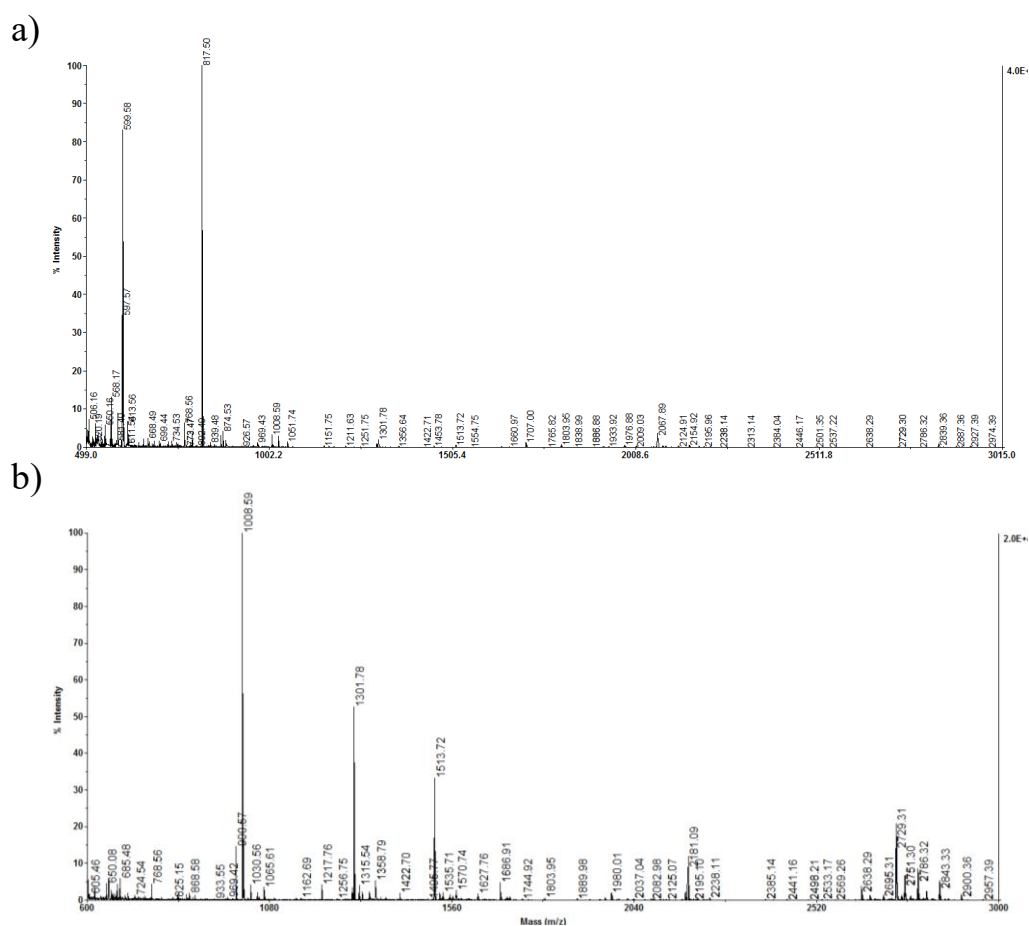
The mass spectrum analysis of peptides fraction (Figure 2.5) shows signals at m/z 1356.64 and 2067.89 attributable to de-glycosylated peptides of α-subunit, while mass signals at m/z 2173.99 and 2181.09 were attributed to de-glycosylated peptides of β-subunit (Table 2.3). The increase of 1 Da on the experimental MW of these peptides is due to the enzymatic de-glycosylation (hydrolysis mechanism Figure 1.4) of PNG-ase F, which converts the asparagine (Asn: 114.04 Da) involved in the N-glycosidic bond into aspartic acid (Asp: 115.03 Da).

*Table 2.3 MH+(m/z) recorded in MALDI spectrum (Figure 2.5) and corresponding peptides sequence.*

| **FSH: α-subunit** | | | |
|---|---|---|---|
| **Theoretical MH+ (m/z)** | **Experimental MH+ (m/z)** | **Peptide sequence** | **Start-end** |
| 539.25 | 539.25 | SYNR | 64-67 |
| 719.41 | 719.40 | TMLVQK | 46-51 |
| 817.45 | 817.45 | AYPTPLR | 36-42 |
| 838.44 | 838.45 | VTVMGGFK | 68-75 |
| 1032.58 | 1032.57 | AYPTPLRSK | 36-44 |
| 1160.67 | 1160.67 | AYPTPLRSKK | 36-45 |
| 1355.59 | 1356.64 | **N52**VTSESTCCVAK | 52-63 |
| 1358.69 | 1358.72 | SYNRVTVMGGFK | 64-75 |
| 2066.82 | 2067.89 | VE**N24**HTACHCSTCYYHK | 76-91 |
| 4116.74 | 4116.68 | APDVQDCPECTLQENPFFSQPGAPILQCMGCCFSR | 1-35 |
| **FSH: β-subunit** | | | |
| **Theoretical MH+ (m/z)** | **Experimental MH+ (m/z)** | **Peptide sequence** | **Start-end** |
| 656.30 | 656.29 | TCTFK | 50-54 |
| 1008.53 | 1008.54 | ELVYETVR | 55-62 |
| 1301.72 | 1301.73 | DLVYKDPARPK | 36-46 |
| 1315.49 | 1315.49 | CDSDSTDCTVR | 87-97 |
| 1561.66 | 1561.64 | GLGPSYCSFGEMK | 98-111 |
| 1645.82 | 1645.82 | CTFKELVYETVR | 50-62 |
| 2172.93 | 2173.99 | FCISI**N24**TTWCAGYCYTR | 19-35 |
| 2180.03 | 2181.09 | NSCELT**N7**ITIAIEKEECR | 1-18 |
| 2729.09 | 2729.21 | CDSDSTDCTVRGLGPSYCSFGEMK | 87-110 |
| 2729.20 | 2729.21 | VPGCAHHADLSYTYPVATQCHCGK | 63-86 |

The oligosaccharide moiety was derivatized as described before and analysed by MALDI-TOF. Table 2.4 and Table 2.6 summarize the glycans identified in α and β-FSH chains oligosaccharides spectra (Figure 2.6).

a)



b)



*Figure 2.6 MALDI-TOF spectrum of α-FSH (a) and β-FSH chain (b) oligosaccharides released by enzymatic de-glycosylation.*

Each signal was assigned to hypothetical glycan structures based on the theoretical molecular weight and on the known bio-synthetical pathway of N-linked oligosaccharides. The tables 2.4 - 2.7 report in detail, for each signal, the theoretical and experimental molecular weight, the attributed glycans monosaccharides compositions, the hypothetical glycan structure and MS/MS validation [19].

*Table 2.4 Overall glycan profile of α-FSH*

| FSH: α-subunit | | | | |
|---|---|---|---|---|
| **Theoretical Mw** | **Experimental Mw** | **Monosaccharide Composition** | **Hypothetical Glycan Structure** | **MS/MS validation** |
| 1497.77(H+) | 1497.26(H+) | Fuc2HexNac2Hex3 | | |
| 2070.03(Na+) | 2069.36(Na+) | HexNac4Hex5 | | X |
| 2186.08(Na+) | 2185.37(Na+) | HexNac3Hex5NeuAc | | |
| 2227.10(Na+) | 2226.37(Na+) | HexNac4Hex4NeuAc | | |
| 2431.20(Na+) | 2430.40(Na+) | HexNac4Hex5NeuAc | | X |
| 2792.40(Na+) | 2791.43(Na+) | HexNac4Hex5NeuAc2 | | X |
| 2822.39(Na+) | 2822.43(Na+) | HexNac4Hex5NeuGly | | |
| 2880.43(Na+) | 2879.44(Na+) | HexNac5Hex6NeuAc | | |

In order to confirm the N-glycan structures hypnotized, the most intense m/z signals were fragmented by MSMS experiments, as showed for example in Figure 2.7.

Precursor ion: 2791.43 m/z
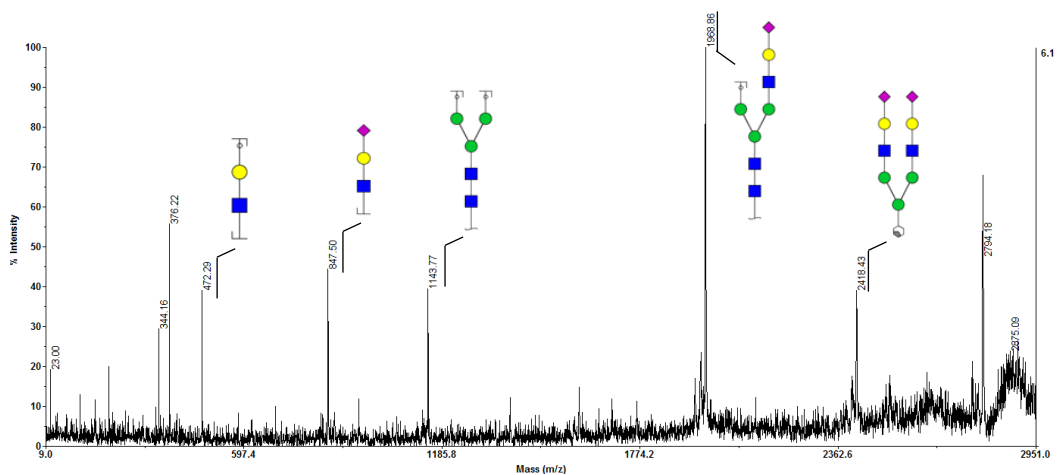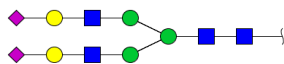


*Figure 2.7 MALDI-TOF/TOF spectrum of 2791.43 m/z precursor ion. Found fragment are resumed in table below*

*Table 2.5  MS/MS validation of glycan structure assigned to 2792.17 precursor ion.*

| Precursor Structure | Precursor Mw (Na+) | Experimental Fragments Mw (Na+) | Theoretical Fragments Mw (Na+) | Fragments Structure |
|---|---|---|---|---|
| | 2792.17 | 472.29 | 472.21 |  |
|  | | 847.50 | 847.4 |  |
| | | 1143.77 | 1143.55 |  |
| | | 1968.86 | 1967.97 |  |
| | | 2418.43 | 2418.17 |  |

*Table 2.6 Overall glycan profile of β-FSH*

| FSH: β-subunit | | | | |
|---|---|---|---|---|
| **Theoretical Mw** | **Experimental Mw** | **Monosaccharide Composition** | **Hypothetical Glycan Structure** | **MS/MS validation** |
| 2040.02 | 2039.36 | FucHexNac4Hex4 |  | |
| 2244.45 | 2243.37 | FucHexNac4Hex5 |  | |
| 2401.01 | 2400.39 | FucHexNac4HexNeuAc |  | |
| 2431.20(Na+) | 2430.39(Na+) | HexNac4Hex5NeuAc |  | |
| 2519.26 | 2518.39 | HexNac5Hex6 |  | |
| 2605.30 | 2604.41 | FucHexNac4Hex5NeuAc |  | X |
| 2693.35 | 2692.43 | FucHexNac5Hex6 |  | |
| 2792.40(Na+) | 2791.43(Na+) | HexNac4Hex5NeuAc2 |  | |
| 2850.42 | 2849.44 | FucHexNac5Hex5NeuAc |  | |
| 2880.43(Na+) | 2879.44(Na+) | HexNac5Hex6NeuAc |  | |
| 2966.40 | 2965.45 | FucHexNac4Hex5NeuAc2 |  | X |
| 3241.60 | 3240.47 | HexNac5Hex6NeuAc2 |  | |

| | | | | |
|---|---|---|---|---|
| 3415.69 | 3414.50 | FucHexNac5Hex6NeuAc2 |  | X |
| 3602.68 | 3601.51 | HexNac5Hex6NeuAc3 |  | |
| 3776.86 | 3775.52 | FucHexNac5Hex6NeuAc3 |  | |
| 4052.00 | 4052.98 | HexNac6Hex7NeuAc3 |  | |
| 4226.09 | 4224.58 | FucHexNac6Hex7NeuAc3 |  | |

As for α-subunit, the β-subunit N-glycan structures hypnotized were confirmed by fragmentation of the most intense m/z signals, as showed for example in Figure 2.8.

Precursor ion: 2966.54 m/z
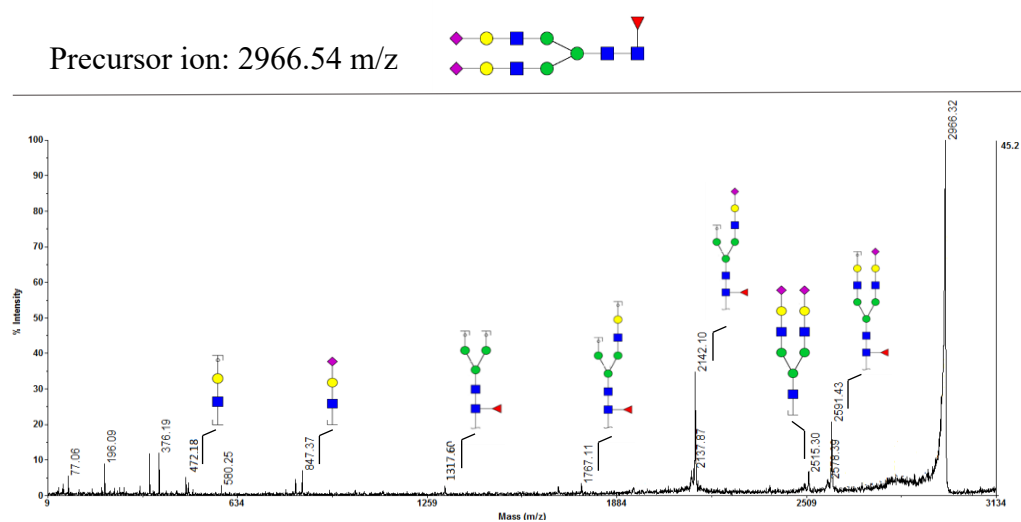


*Figure 2.8 MALDI-TOF/TOF spectrum of 2966.54 m/z precursor ion. Found fragment are resumed in table below*

*Table 2.7 MS/MS validation of glycan structure assigned to 2966.54 precursor ion*

| Precursor Structure | Mw (Na+) Precursor | Experimental Fragments Mw (Na+) | Theoretical Fragments Mw (Na+) | Fragments Structure |
|---|---|---|---|---|
| | 2966.54 | 472.18 | 472.21 | |
| | | 847.37 | 847.4 | |
| | | 1317.60 | 1317.64 | |
| | | 1767.11 | 1766.87 | |
| | | 2142.03 | 2142.05 | |
| | | 2515.30 | 2515.23 | |
| | | 2591.43 | 2591.28 | |

## 2.3.1.2 Site-specific glycans profile of α and β-FSH

Glycopeptides of the α and β enriched fractions were separated from peptides mixture by HPLC as described before. As an example, the chromatographic trace of α-subunit tryptic peptides mixture separation is showed in Figure 2.9.
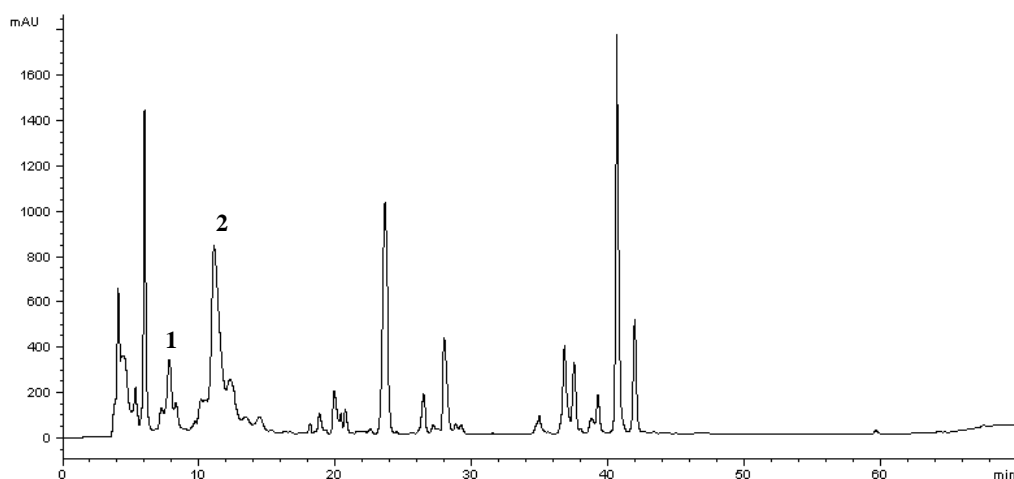


*Figure 2.9 HPLC chromatogram of α-chain tryptic peptides*

75

MALDI-TOF mass spectrometric analysis of individual fractions revealed that the peaks 1 and 2 (respectively at RT 8 and 11 min), contain glycopeptides. These two fractions were subjected to enzymatic de-glycosylation and oligosaccharide per-methylation before MS and MSMS analysis as previously described.

Peptides spectrum of first glycopeptide fraction shows the signal at m/z: 1356.57 attributed to the peptide (52-63) with the increase of 1 Da due to the conversion of Asn52 to Asp52 produced by PNG-ase F enzymatic de-glycosylation. Analogously, the mass signal at m/z 2067.70 recorded in the second glycopeptide fraction contain the peptide (76-91) with the same 1 Da increasing (Figure 2.10).
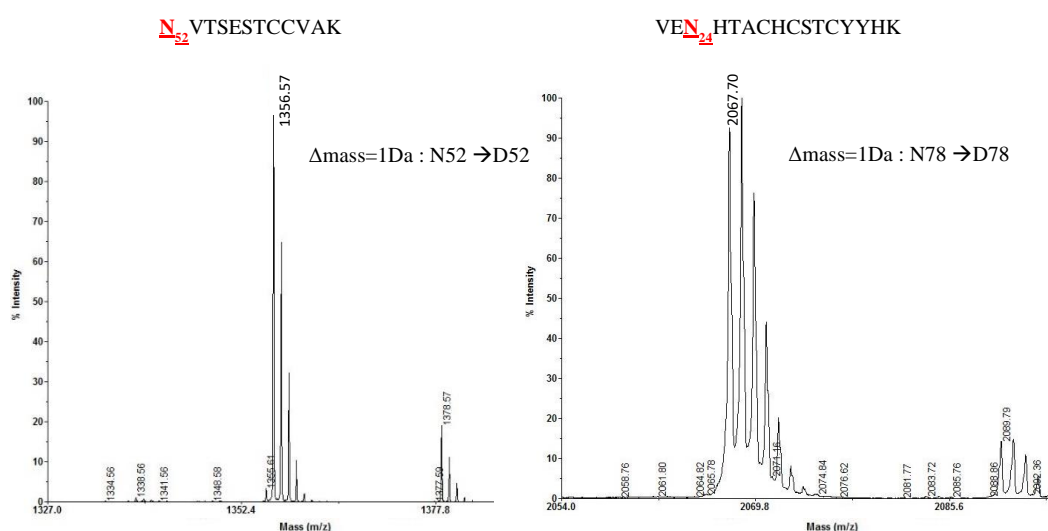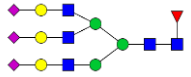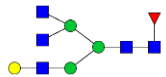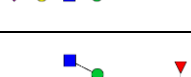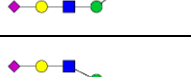


*Figure 2.10 MALDI-TOF spectra of FSH α-chain tryptic de-glycosylated peptides. Peak 1 contains glycosite N52; peak 2 contains N78 glycosylation site.*

MALDI-TOF/TOF oligosaccharides analysis of the two glycopeptide-enriched fractions allowed us to draw up the site-specific glycosylation profile of α-FSH (Table 2.8).

*Table 2.8 Site-specific N-glycan profile of α-FSH*

| FSH: α-subunit | | | | | |
|---|---|---|---|---|---|
| **Glycosite** | **Th. Mw (m/z)** | **Exp. Mw (m/z)** | **Monosaccharide Composition** | **Hypothetical Glycan Structure** | **MS/MS validation** |
| **N52** peptide [52-63]: 1356.57 m/z | 2070.03 | 2070.41 | HexNac4Hex5 | | |
| | 2285.15 | 2286.12 | FucHexNac5Hex4 | | |
| | 2363.34 | 2363.46 | Fuc2HexNac5Hex5 | | |
| | 2431.12 | 2431.54 | HexNAc4Hex5NeuAc | | X |
| | 2530.27 | 2531.63 | FucHexNAc6Hex4 | | |
| | 2605.29 | 2605,61 | FucHexNAc4Hex5Sia | | X |
| | 2792.38 | 2792,70 | HexNAc4Hex5NeuAc2 | | X |
| | 2966.47 | 2966,78 | FucHexNAc4Hex5NeuAc2 | | X |
| | 3211.59 | 3211.01 | FucHexNAc5Hex5NeuAc2 | | |
| | 3241.60 | 3241.93 | HexNAc5Hex6NeuAc2 | | X |
| | 3415.69 | 3415,62 | FucHexNAc5Hex6 NeuAc2 | | X |
| | 3603.78 | 3603.02 | HexNAc5Hex6NeuAc3 | | X |

| | Observed | Calculated | Composition | Structure | |
|---|---|---|---|---|---|
| | 3776.87 | 3777.10 | FucHexNAc5Hex6 NeuAc3 | | |
| **N78** peptide [76-91]: 2067.70 m/z | 2285.15 | 2286.16 | FucHexNac5Hex4 | | |
| | 2363.34 | 2363.34 | Fuc2HexNac5Hex5 | | |
| | 2431.12 | 2431.32 | HexNAc4Hex5NeuAc | | X |
| | 2530.27 | 2531.33 | FucHexNAc6Hex4 | | |
| | 2792.38 | 2792,49 | HexNAc4Hex5NeuAc2 | | X |
| | 2966.47 | 2966.56 | FucHexNAc4Hex5NeuAc2 | | X |
| | 3211.59 | 3209.65 | FucHexNAc5Hex5NeuAc2 | | X |
| | 3241.60 | 3241.67 | HexNAc5Hex6NeuAc2 | | X |
| | 3603.78 | 3602.81 | HexNAc5Hex6NeuAc3 | | X |

In β- subunit were identified two fractions rich in glycopeptides which were recovered and analysed as described before. Figure 2.11 shows a zoom of peptides spectra of glycopeptides rich fractions. Mass signal at m/z: 2180.91 was attributed to the peptide (1-18) with the increase of 1 Da due to the conversion of Asn7 to Asp7 produced by PNG-ase F enzymatic de-glycosylation. Analogously, the mass signal at m/z 2173.75 recorded in the second glycopeptide fraction contain the peptide (19-35) with the same 1 Da increasing (Asn24→Asp24) (Figure 2.11).

NSCELT**N**₇ITIAIEKEECR   Δmass=1Da : N7 →D7

FCISI**N**₂₄TTWCAGYCYTR   Δmass=1Da : N24 →D24
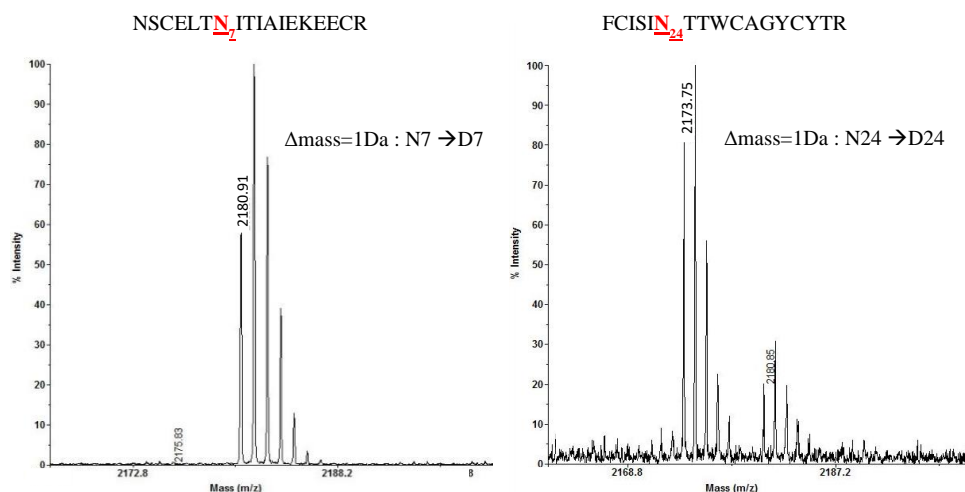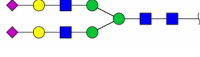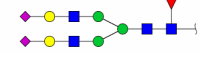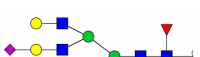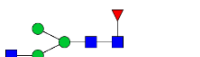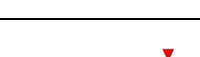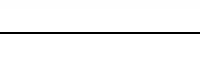
*Figure 2.11 MALDI-TOF spectra of FSH β-chain tryptic de-glycosylated peptides recovered in the two glycopeptides fractions.*

The site-specific glycosylation profile of β-FSH was obtained by MALDI-TOF/TOF oligosaccharides analysis. of the two glycopeptide-enriched fractions (Table 2.9).

*Table 2.9 Site-specific N-glycan profile of β-FSH*

| FSH: β-subunit | | | | | |
|---|---|---|---|---|---|
| **Glycosite** | **Th. Mw (m/z)** | **Exp. Mw (m/z)** | **Monosaccharide Composition** | **Hypothetical Glycan Structure** | **MS/MS validation** |
| **N7** peptide [1-18]: 2180.91 m/z | 1590.79 | 1590.61 | FucHexNAc3Hex3 |  | |
| | 1981.98 | 1981.76 | HexNAc3Hex4NeauAc |  | |
| | 2401.19 | 2400.91 | FucHexNAc4Hex4NeuAc |  | X |
| | 2431.12 | 2430.92 | HexNAc4Hex5NeuAc |  | X |
| | 2605.29 | 2604.95 | FucHexNAc4Hex5NeuAc |  | X |

79

| | | | | | |
|---|---|---|---|---|---|
| | 2792.03 | 2792.03 | HexNAc4Hex5NeuAc2 |  | X |
| | 2966.47 | 2966.09 | FucHexNAc4Hex5NeuAc2 |  | X |
| | 3054.52 | 3054.13 | FucHexNAc5Hex6NeuAc |  | X |
| | 3415.69 | 3415.24 | FucHexNAc5Hex6NeuAc2 |  | X |
| | 3776.87 | 3777.37 | FucHexNAc5Hex6NeuAc3 |  | X |
| **N24** peptide [19-35]: 2173.75 m/z | 1590.79 | 1590.65 | FucHexNAc3Hex3 |  | |
| | 2401.19 | 2400.94 | FucHexNAc4Hex4NeuAc |  | X |
| | 2431.20 | 2430.94 | HexNAc4Hex5NeuAc |  | X |
| | 2605.29 | 2605.01 | FucHexNAc4Hex5NeuAc |  | X |
| | 2966.47 | 2966.14 | FucHexNAc4Hex5NeuAc2 |  | X |
| | 3054.52 | 3054.17 | FucHexNAc5Hex6NeuAc |  | X |
| | 3415.69 | 3415.31 | FucHexNAc5Hex6NeuAc2 |  | X |
| | 3776.87 | 3776.44 | FucHexNAc5Hex6NeuAc3 |  | X |

The analysis of the oligosaccharides from α and β-FSH showed a heterogeneous profile of glycoforms. In particular, on α-chain, completely sialylated and mono-

sialylated biantennary and tri-antennary glycan structures were mainly observed at Asn52 glycosylation site, while di-sialylated and mono-sialylated bi-antennary glycan structures were mainly detected at Asn78. On β-chain Asn24 glycosite were mainly observed fucosylated fully sialylated and mono-sialylated biantennary and fucosylated tri-antennary glycan structures, while di-sialylated and mono-sialylated bi-antennary and fucosylated bi-antennary and tri-antennary glycan structures were mainly detected at Asn7.

For an easier representation of the glycans structures, in the following paragraphs the glycans Oxford notation will be used. The Oxford notation is based on building up N-glycan structures. Therefore, it can be used to denote very complex glycans. All N-glycans have two core GlcNAcs; F at the start of the abbreviation indicates a core fucose; Mx, number (x) of mannose on core GlcNAcs; Ax, number of antenna (GlcNAc) on trimannosyl core; Gx, number (x) of linked galactose on antenna; Sx, number (x) of sialic acids linked to galactose. The table below shows an explication of this kind of notation.

*Table 2.10 Symbol and Text Nomenclature for Representation of Glycan Structure: Consortium for Functional Glycomics(CFG) and Oxford Notation.*

| Text linear nomenclature | CFG | Oxford notation |
|---|---|---|
| HexNac4Hex5 | | A2G2 |
| HexNAc4Hex5NeuAc | | A2G2S1 |
| HexNAc4Hex5NeuAc2 | | A2G2S2 |
| FucHexNAc4Hex5NeuAc2 | | FA2G2S2 |
| FucHexNAc5Hex6NeuAc2 | | FA3G3S2 |
| FucHexNAc5Hex6NeuAc3 | | FA3G3S3 |

## 2.3.2 Glycopeptides direct analysis by nano LC-Q/TOF

The direct analysis of FSH glycopeptides was carried out on the chymotryptic hydrolysate of the molecule. Despite the low ionization efficiency of chymotryptic peptides (because of the lack of C-terminal charge), this enzyme was selected over trypsin because of the size of FSH glycopeptides that it would generate. In fact, due to the absence of basic residues (K and R) near the 4 glycosylation sites, tryptic hydrolysis would have originated glycopeptides too large and therefore too difficult to analyse by mass spectrometry. Figure 2.12 shows the total ion current (TIC) obtained from the nano LC-Q/TOF analysis of different FSH molecule, r-hFSH (GONAL-F ®, BEMBFOLA, OVALEAP), u-hFSH (FOSTIMON) and p-hFSH (pituitary hormone standard).
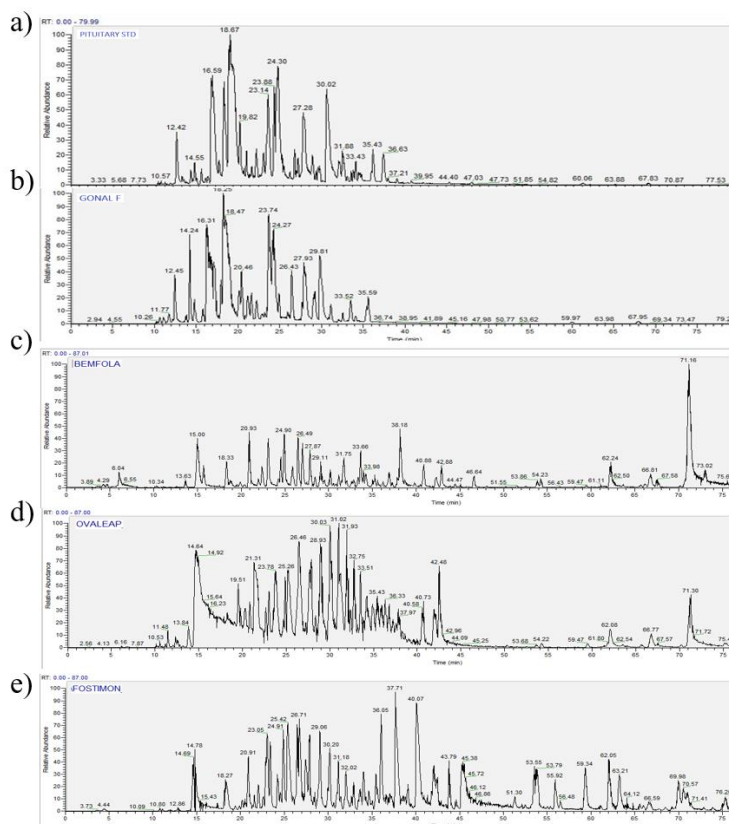


*Figure 2.12 Comparison between TIC obtained from different FSH molecule: a) Pituitary Std; b) Gonal F; c) Bemfola; d) Ovaleap; e) Fostimon.*

For each sample, data obtained from nano-LC-MS/MS analysis were processed by MASCOT software for protein identification [20]. As an example, the Table 2.11 shows the results obtained from the analysis of the GONAL-F ®.

*Table 2.11 Protein identification by MASCOT software against NCBI database.*

| Gonal F sample | | | | | | |
|---|---|---|---|---|---|---|
| **Protein name** | **NCBI Accession Number** | **Start-End** | **Th. Mw** | **Exp. Mw** | **Peptides** | **Score** |
| Chain A, Structure Of Human Chorionic Gonadotropin | gi\|640445 | 1-12 | 1403.57 | 1403.57 | APDVQDCPECTL | 56 |
| | | 1-17 | 2018.83 | 2018.83 | APDVQDCPECTLQENPF | 75 |
| | | 1-18 | 2165.90 | 2165.90 | APDVQDCPECTLQENPFF | 92 |
| | | 18-26 | 928.51 | 928.50 | FSQPGAPIL | 24 |
| | | 18-33 | 1871.79 | 1871.78 | FSQPGAPILQCMGCCF | 58 |
| | | 19-33 | 1724.73 | 1724.71 | SQPGAPILQCMGCCF | 66 |
| | | 34-41 | 903.48 | 903.48 | SRAYPTPL | 34 |
| | | 66-74 | 979.49 | 979.49 | NRVTVMGGF | 83 |
| Chain B, Crystal Structure Of Human Follicle Stimulating Hormone Complexed With Its Receptor | gi\|60594077 | 1-19 | 2326.10 | 2326.10 | NSCELTNITIAIEKEECRF | 41 |
| | | 6-19 | 1722.87 | 1722.86 | TNITIAIEKEECRF | 64 |
| | | 20-27 | 993.46 | 993.46 | CISINTTW | 28 |
| | | 32-39 | 1088.50 | 1088.50 | CYTRDLVY | 28 |
| | | 32-53 | 2759.38 | 2759.39 | CYTRDLVYKDPARPKIQKTCTF | 22 |
| | | 38-53 | 1951.03 | 1951.08 | VYKDPARPKIQKTCTF | 30 |
| | | 40-53 | 1688.90 | 1688.90 | KDPARPKIQKTCTF | 34 |
| | | 57-74 | 2072.98 | 2072.97 | VYETVRVPGCAHHADSLY | 41 |
| | | 59-74 | 1810.84 | 1810.87 | ETVRVPGCAHHADSLY | 40 |
| | | 75-103 | 3291.37 | 3291.37 | TYPVATQCHCGKCDSDSTDCTVRGLGPSY | 44 |
| | | 75-106 | 3685.50 | 3685.50 | TYPVATQCHCGKCDSDSTDCTVRGLGPSYCSF | 33 |
| | | 104-111 | 986.39 | 986.38 | CSFGEMKE | 56 |

Unfortunately, MASCOT software was not used for the automated attribution of glycopeptides because the version of the software used did not support the integration of glycan libraries as search parameter in "peptide variable modifications" section. For this reason, subsequently Byonic software was chosen for the identification of glycopeptides and skyline for XIC extraction and quantification of targeted ions [20, 21]. However, manual research of glycopeptides was performed for each sample. The knowledge of the FSH glycan

profile, obtained in the previous analyses, was a prerequisite for predicting the exact mass of each glycopeptides. The extrapolation of the known glycopeptide mass from the chromatographic trace allowed us to identify the glycan moiety on each glycosite as shown in Table 2.12.

*Table 2.12 Site specific glycan profile obtained by manual inspection of the samples raw data files.*

| FSH: α-subunit | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Glycan Notation | Glycopeptide Mw (m/z) | z | Peptide | Glycation Site | Gonal_F | Pituitary | Bemfola | Ovaleap | Fostimon |
| A2G2S2 | 1785.73 | 2 | 48-59 | ASN 52 | x | x | x | x | x |
| A3G3S1 | 1215.53 | 2 | 48-59 | ASN 52 | | | x | | |
| A3G3S3 | 1645.66 | 3 | 48-59 | ASN 52 | x | x | x | x | x |
| FA2G2S2 | 1239.52 | 3 | 48-59 | ASN 52 | x | x | x | x | |
| A3G3S2 | 1312.56 | 3 | 48-59 | ASN 52 | x | | x | | x |
| A4G4S3 | 1531.27 | 3 | 48-59 | ASN 52 | | | | | |
| A4G4S4 | 1628.29 | 3 | 48-59 | ASN 52 | | | x | x | x |
| A2G2S1 | 1324.49 | 3 | 75-88 | ASN 78 | x | x | x | x | |
| A2G2S2 | 1650.67 | 2 | 75-83 | ASN 78 | x | | x | x | x |
| A3G3S1 | 1687.68 | 2 | 75-83 | ASN7 8 | | | x | | |
| A3G3S2 | 1446.21 | 2 | 75-88 | ASN 78 | x | x | x | x | x |
| A3G3S3 | 1543.24 | 3 | 75-88 | ASN 78 | x | x | x | | x |
| A4G4S2 | 1344.18 | 3 | 75-83 | ASN 78 | | x | | x | x |
| A4G4S3 | 1441.24 | 3 | 75-83 | ASN 78 | | x | | x | x |
| FSH: β-subunit | | | | | | | | | |
| Glycan Notation | Glycopeptide Mw (m/z) | z | Peptide | Glycation Site | Gonal_F | Pituitary | Bemfola | Ovaleap | Fostimon |
| A3G1 | 1694.17 | 2 | 6_19 | ASN 7 | x | | x | | |
| FA2G2S1 | 1462.90 | 3 | 1_19 | ASN 7 | x | x | x | | |
| A3G4 | 1492.96 | 3 | 1_19 | ASN 7 | x | x | | | x |
| FA2G2S2 | 1559.98 | 3 | 1_19 | ASN 7 | x | x | x | x | x |
| F2A5G1 | 1362.55 | 3 | 1_19 | ASN 7 | | x | | | |
| F2A5G4 | 1725.43 | 3 | 1_19 | ASN 7 | | x | | x | |
| F2A2G3S1 | 1565.64 | 3 | 1_19 | ASN 7 | | | | x | |
| F3A4G2S2 | 1792.70 | 3 | 1_19 | ASN 7 | x | | | x | |
| A4G5 | 1614.68 | 3 | 1_19 | ASN 7 | x | x | x | | |
| A3G3S2 | 1431.92 | 3 | 6_19 | ASN 7 | | x | | x | |
| A3G3S2 | 1633.03 | 3 | 1_19 | ASN 7 | x | | x | x | x |
| FA3G3S2 | 1681.68 | 3 | 1_19 | ASN 7 | x | x | x | | |
| A4G5S1 | 1711.72 | 3 | 1_19 | ASN 7 | x | x | x | x | x |
| A4G2S1 | 1384.53 | 3 | 1_19 | ASN 7 | | | | | |
| A3G3S3 | 1730.06 | 3 | 1_19 | ASN 7 | x | | x | | x |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| FA3G3S3 | 1778.76 | 3 | 1_19 | ASN 7 | x | x | x | x | x |
| A5G6S1 | 1375.33 | 4 | 1_19 | ASN 7 | x | x | x | | |
| A4G4S3 | 1389.10 | 4 | 1_19 | ASN 7 | | | x | x | x |
| FA5G6S1 | 1411.59 | 4 | 1_19 | ASN 7 | | | x | | |
| FA4G4S3 | 1425.58 | 4 | 1_19 | ASN 7 | x | x | x | | x |
| A5G6S2 | 1448.1 | 4 | 1_19 | ASN 7 | x | x | x | | |
| A4G4S4 | 1461.91 | 4 | 1_19 | ASN 7 | | | | | |
| F2A3G3S2 | 1730.33 | 3 | 1_19 | ASN 7 | | | | x | x |
| FA4G4S4 | 1498.38 | 4 | 1_19 | ASN 7 | x | x | x | x | |
| FA5G5S3 | 1516.92 | 4 | 1_19 | ASN 7 | | | x | x | |
| A5G5S4 | 1553.14 | 4 | 1_19 | ASN 7 | | | x | | |
| A6G7S2 | 1539.36 | 4 | 1_19 | ASN 7 | x | x | x | | |
| FA5G5S4 | 1589.73 | 4 | 1_19 | ASN 7 | | | x | | x |
| FA2G1S1 | 1462.91 | 3 | 1_19 | ASN 7 | x | x | x | x | |
| A6G2 | 1715.20 | 2 | 20-27 | ASN 24 | x | | | | |
| FA2G2 | 1382.08 | 2 | 20-27 | ASN 24 | x | | | | |
| FA1G1S1 | 1345.06 | 2 | 20-27 | ASN 24 | | x | x | | x |
| FA2G2S1 | 1527.63 | 2 | 20-27 | ASN 24 | x | x | x | x | |
| FA2G2S2 | 1673.15 | 2 | 20-27 | ASN 24 | x | x | x | x | x |
| F2A2G2S2 | 1746.14 | 3 | 20-27 | ASN 24 | | | | | x |
| FA3G3S2 | 1855.75 | 2 | 20-27 | ASN 24 | x | x | x | x | x |
| FA4G2 | 1585.15 | 2 | 20-27 | ASN 24 | | | | | |
| FA3G3S3 | 1334.51 | 3 | 20-27 | ASN 24 | x | x | x | | x |
| FA4G2S1 | 1730.70 | 2 | 20-27 | ASN 24 | | | x | x | |
| FA4G4S3 | 1456.23 | 3 | 20-27 | ASN 24 | x | | x | | x |
| FA4G4S4 | 1553.28 | 3 | 20-27 | ASN 24 | | | | | x |
| FA5G5S4 | 1675.01 | 3 | 20-27 | ASN 24 | | | | | x |
| FA2G1S1 | 1527.60 | 2 | 20-27 | ASN 24 | x | x | x | x | |

The direct LC-MS/MS analyses of the chymotryptic digests of these glycohormones comprising the mixture of peptides and glycopeptides showed a heterogeneous profile of glycoforms occurring in the different samples with subtle differences in term of glycoforms. In particular, the biosimilars BEMFOLA and OVALEAP are characterized by bulkier glycan structures and greater sialylation extent than GONAL-F ® [22-25]. The urinary commercial FSH FOSTIMON showed a glycosylation pattern with an even higher level of antennary and higher extent of sialylation compared to GONAL-F ® and the other two biosimilars BEMFOLA and OVALEAP. These results are in

agreement with existing literature data [26, 27] on different r-hFSH preparations compared to GONAL-F ® displaying apparently identical polypeptide chains but a somewhat different glycosylation pattern [17, 18].

## 2.4    Conclusions

Biosimilars are gaining an increasing foothold in the global health panorama. These are biological medicines highly similar to another biological medicine already approved in the EU (called 'reference medicine') in terms of structure, biological activity and efficacy, safety and immunogenicity profile [27]. Biosimilars have been marketed in Europe since 2006 under the regulation of the EMA. In 2015, the FDA authorized the commercialization of the first biosimilar in the United States (US). Due to their intrinsic structural features, glycohormone biosimilars may differ in purity and composition of isoforms showing various glycosylation profiles, with possible consequent alterations in clinical efficacy or safety [28-33].

In this study an analytical method was used for the direct investigation of the glycosidic profile and the glycoforms structure of different FSH-based drugs (BEMFOLA, OVALEAP, FOSTIMON) using GONAL-F ® as reference molecule. BEMFOLA and OVALEAP showed the presence of bulkier glycan structures and greater sialylation extent than GONAL-F ®. The urinary commercial FSH (FOSTIMON) showed a glycosylation pattern with an even higher level of antennary and higher extent of sialylation compared to GONAL-F ® and the other two biosimilars BEMFOLA and OVALEAP. These results are in agreement with existing literature data [34, 35] on different r-hFSH preparations compared to GONAL-F ® displaying nominally identical polypeptide chains but a somewhat different glycosylation pattern.

## 2.5 References

1. Bousfield GR, Harvey DJ. *Follicle-Stimulating Hormone Glycobiology*. Endocrinology. 2019 Jun 1;**160**(6):1515-1535. doi: 10.1210/en.2019-00001. PMID: 31127275; PMCID: PMC6534497

2. Lapthorn, AJ1, et al. "*Crystal structure of human chorionic gonadotropin.*" Nature, 1994. **369** (6480): 455-461.

3. Das N, Kumar TR. *Molecular regulation of follicle-stimulating hormone synthesis, secretion and action*. J Mol Endocrinol. 2018 Apr;**60**(3):R131-R155. doi: 10.1530/JME-17-0308. Epub 2018 Feb 7. PMID: 29437880; PMCID: PMC5851872

4. Emperaire, J.C., *Review of physiology, Ovulation stimulation with gonadotropins*. 2015, Geneva: Springer International Publishing.

5. Wide, L. and K. Eriksson, *Dynamic changes in glycosylation and glycan composition of serum FSH and LH during natural ovarian stimulation*. Upsala journal of medical sciences, 2013. **118**(3): p. 153-164.

6. Davis, J.S., et al., *Naturally occurring follicle-stimulating hormone glycosylation variants*. Journal of glycomics & lipidomics, 2014. **4**(1): p. e117.

7. Dalpathado, D.S., et al., *Comparative glycomics of the glycoprotein follicle stimulating hormone: glycopeptide analysis of isolates from two mammalian species*. Biochemistry, 2006. **45**(28): p. 8665-8673.

8. Anobile, C., et al., *Glycoform composition of serum gonadotrophins through the normal menstrual cycle and in the post-menopausal state*. Molecular human reproduction, 1998. **4**(7): p. 631-639.

9. Elsaid, R., et al., *Hypo-glycosylated Human Follicle–Stimulating Hormone (hFSH21) Exhibits Higher Endocytic Rate for Recombinant hFSH Receptor than Fully-glycosylated hFSH (hFSH24)*. The FASEB Journal, 2015. **29**(1_supplement): p. 890.8.

10. Wide, L., et al., *Serum half-life of pituitary gonadotropins is decreased by sulfonation and increased by sialylation in women*. The Journal of Clinical Endocrinology & Metabolism, 2009. **94**(3): p. 958-964.

11. Jiang, X., et al., *Evidence for Follicle-stimulating Hormone Receptor as a Functional Trimer*. J Biol Chem, 2014. **289**(20): p. 14273-82.

12. Jiang, C., et al., *Hypoglycosylated hFSH has greater bioactivity than fully glycosylated recombinant hFSH in human granulosa cells*. The Journal of Clinical Endocrinology & Metabolism, 2015. **100**(6): p. E852-E860.

13. PADMANABHAN, V., et al., *Modulation of Serum Follicle-Stimulating Hormone Bioactivity and Isoform Distribution by Estrogenic Steroids in Normal Women and in Gonadal Dysgenesis\**. The Journal of Clinical Endocrinology & Metabolism, 1988. **67**(3): p. 465-473.

14. Bousfield, G.R., et al., *Macro-and micro-heterogeneity in pituitary and urinary follicle-stimulating hormone glycosylation*. Journal of glycomics & lipidomics, 2014. **4**.

15. Sam T, et al., *Follicle-stimulating hormone*. Pharmaceutical Biotechnology. 2013: New York: Springer. pp. 277–283.

16. Sairam, M. and G. Bhargavi, *A role for glycosylation of the alpha subunit in transduction of biological signal in glycoprotein hormones*. Science, 1985. **229**(4708): p. 65-67.

17. Ulloa-Aguirre, A., J.A. Dias, and G.R. Bousfield, *Gonadotropins and the Importance of Glycosylation.* Post-translational Modification of Protein Biopharmaceuticals, 2009: p. 109-147.

18. Bishop, L.A., et al., *Specific roles for the asparagine-linked carbohydrate residues of recombinant human follicle stimulating hormone in receptor binding and signal transduction.* Molecular Endocrinology, 1994. **8**(6): p. 722-731.

19. Varki, A., et al., *Symbol nomenclature for glycan representation.* Proteomics, 2009. **9**(24): p. 5398-5399.

20. Helsens, Kenny, et al. "*MascotDatfile: an open-source library to fully parse and analyse MASCOT MS/MS search results.*" Proteomics, 2007. **7** (3): 364-366.

21. Bollineni, Ravi Chand, et al. "*Large-scale intact glycopeptide identification by Mascot database search.*" Scientific reports, 2018. **8** (1): 1-13.

22. Grass, J., et al., *Analysis of recombinant human follicle-stimulating hormone (FSH) by mass spectrometric approaches.* Analytical and bioanalytical chemistry, 2011. **400**(8): p. 2427.

23 Orvieto, R. and D.B. Seifer, *Biosimilar FSH preparations-are they identical twins or just siblings?* 2016, Springer.

24. Mastrangeli, R., et al., *In-vivo biological activity and glycosylation analysis of a biosimilar recombinant human follicle-stimulating hormone product (Bemfola) compared with its reference medicinal product (GONAL-f).* PloS one, 2017. **12**(9).

25. Riccetti, Laura, et al. "*Glycosylation pattern and in vitro bioactivity of reference follitropin alfa and biosimilars.*" Frontiers in endocrinology, 2019. **10**: 503.

26. Lombardi A., Andreozzi C., Pavone V., Triglione V., Angiolini L., Caccia P. *Evaluation of the oligosaccharide composition of commercial follicle stimulating hormone preparations.* Electrophoresis. 2013; **34**:2394–2406. doi: 10.1002/elps.201300045

27. Bergandi, Loredana, et al. "*Human Recombinant FSH and Its Biosimilars: Clinical Efficacy, Safety, and Cost-Effectiveness in Controlled Ovarian Stimulation for In Vitro Fertilization.*" Pharmaceuticals, 2020. **13** (7): 136.

28. European Medicines Agency. "*Biosimilars in the EU: information guide for healthcare professionals.*" European Medicines Agency, 2017.

29. *GONAL-f.*; Available online:
https://www.ema.europa.eu/en/medicines/human/EPAR/gonal-f.

30. *Bemfola.*; Available online:
https://www.ema.europa.eu/en/medicines/human/EPAR/bemfola.

31. *Announcements Richter acquires all Bemfola® rights in respect of US.*; Available online:
https://www.richter.hu/en-
US/investors/announcements/Pages/extraord180710.aspx.

32. *Ovaleap*; Available online:
https://www.ema.europa.eu/en/medicines/human/EPAR/ovaleap.

33. Strowitzki T., Kuczynski W., Mueller A., Bias P. *Randomized, active-controlled, comparative phase 3 efficacy and safety equivalence trial of Ovaleap® (recombinant human follicle-stimulating hormone) in infertile women using assisted reproduction technology (ART)* Reprod. Biol. Endocrinol. RBE. 2016;**14**:1. doi: 10.1186/s12958-015-0135-8.

34.	Lombardi, A., et al., *Evaluation of the oligosaccharide composition of commercial follicle stimulating hormone preparations.* Electrophoresis, 2013. **34**(16): p. 2394-2406.

35.	Bergandi, L., et al., *Human Recombinant FSH and Its Biosimilars: Clinical Efficacy, Safety, and Cost-Effectiveness in Controlled Ovarian Stimulation for In Vitro Fertilization.* Pharmaceuticals, 2020. **13**(7): p. 136.

# Chapter 3    Recombinant r-FSH variants generation and characterisation of their glycan profile by LC-MS/MS

## 3.1    Introduction

Serum FSH governing the natural ovarian stimulation process exhibited dynamic changes of glycosylation and glycan carbohydrates composition. To date, the biological activity of the different FSH glycoforms is not yet clearly known, representing a fundamental requirement for understanding its structure-function relationships. This is demonstrated by the various forms of FSH present on the market, each having different glycoforms with a spectrum of biological activities that are not always comparable. There is ample evidence that several variants of gonadotropins and their receptors influence female reproductive health and ovarian response to exogenous gonadotropins [1-5]. Given its biological importance in folliculogenesis, pharmaceutical FSH products are currently adapted for multiple follicular growth in assisted reproductive technology (ART) [6]. The ovarian activity, as well as the ovarian response to exogenous gonadotropins appear to be influenced by the different isoforms of these gonadotropins. As aforementioned in hFSH heterodimer, 4 Asn residues, αN52, αN78, βN7, and βN24 are targets for N-linked glycosylation. The Asn-linked oligosaccharides in naturally occurring h-FSH typically include from 1 to 4 terminal sialic acid residues, which results in an overall negative charge for the glycoprotein. The degree of sialylation affects both the stability and biologic activity of the FSH isoform. During the course of the natural menstrual cycle, the acidity of FSH isoforms changes in response to changing estradiol serum concentrations Figure 3.1.
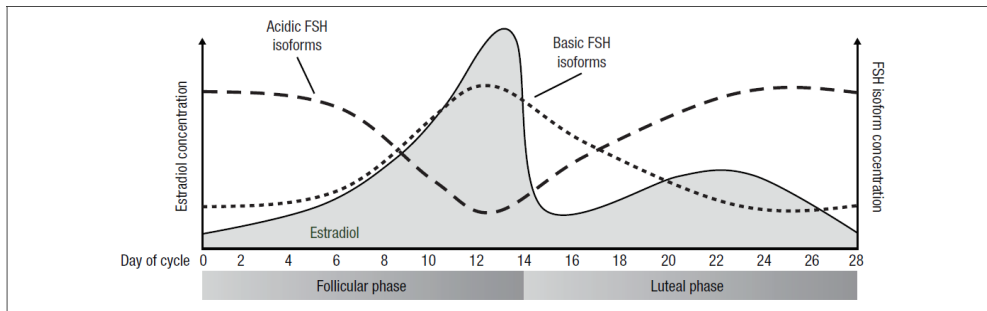
*Figure 3.1* Glycosylation patterns over the course of the menstrual cycle. FSH indicates follicle-stimulating hormone [1]

More acidic isoforms (highly sialylated isoforms) dominate during the transition from the luteal to the follicular phase and during the early follicular phase, and more basic isoform dominate during the late follicular and pre-ovulatory phases [7, 8]. Acidic isoforms seem to be involved in follicular recruitment at the end of menstrual cycles, while follicle selection and rupture seem to be promoted by basic FSH isoforms [9]. In addition to sialylation patterns, which determine the acidity or basicity of FSH isoforms, the complexity of oligosaccharide branching may also be relevant to the biologic activity of these isoforms; however, these effects are poorly understood. It is quite difficult to reconcile FSH glycosylation macro-heterogeneity, represented by the four naturally occurring glycoforms (hFSH[24], hFSH[21], hFSH[18], and hFSH[15]) [10] with micro heterogeneity resulting from the 30 to over 100 glycans attached to as many as 4 Asn residues on the $\alpha$ and $\beta$ FSH subunits [11-13]. Studies in transgenic mice indicate that the hypo glycosylated FSH[15] variant, has a very low level of secretion [14], hence only 3 most abundant isoforms are physiologically relevant: FSH[24], FSH[21], and FSH[18] [15]. A study conducted in adult women aged 21 to 81 revealed a progressive loss of hFSH[21] between 24 and 55 years, suggesting that the ratio of FSH[21] to FSH[24] decreases as a function of aging [10]. Unfortunately, to date the mechanisms responsible for the increasing of the fully glycosylated FSH[24] isoform during reproductive aging are not yet clear. What is known by isoform studies is that less acidic FSH isoforms are more active in receptor binding and

*in vitro* steroidogenesis assays [16-18]. In contrast, acidic forms of FSH are more active *in vivo*, presumably because of longer survival in the circulation [5, 19, 20]. In general, oligosaccharides on the β-subunit play a major role in determining the circulatory half-life and *in vivo* bioactivity of the gonadotropin [21], whereas the oligosaccharide in position αAsn52 is primarily involved in activation of the receptor/signal transducer (G protein) system and the ensuing biological response, an effect probably mediated through stabilizing the structure and/or conformation of the hFSH dimer particularly upon binding to its cognate receptor [22-24]. Nevertheless, more recent studies have shown that glycans at αAsn78 and oligosaccharides in the β-subunit also play an important role in hFSH-mediated signal transduction [15].

In this context, the aim of this study is presented. The work carried out at the Merck KGaA laboratory was focused on the generation of FSH variants panel and on the characterisation of their glycosylation. The data obtained will be of great interest for functional studies on each FSH variant to improve the knowledge we have today on the mechanism by which glycosylation influences the biological activity of the hormone. For our purposes, we generated the recombinant FSH variants summarized hereafter:

- **Acidic/basic:** Acidic/basic enriched fractions.

- **Temperature:** Temperature stressed sample.

- **Acidic/Basic pH:** Acidic and basic pH stressed sample.

- **De-sialylated:** Sialic acid residues removed by enzymatic hydrolysis.

- **De-sialylated and de-galactosylated:** Sialic acid and galactose residues removed by enzymatic hydrolysis.

For each variant of FSH, LC-MS/MS analysis was carried out to monitor the variation of the glycosidic profile and characterize it in comparison to that of the reference one.

## 3.2    Materials and methods

### 3.2.1    Chemicals and reagents

3 different batches of Drug Substance FSH (DS_ FSH), sodium citrate and sodium bicarbonate were provided from Merck. Tris(idrossimetil)amminometane chloridrate (TrisHCl), dithiothreitol (DTT); ethylendiaminetetraacetate (EDTA), Galactosidase, chymotrypsin, iodoacetamide (IAM), ammonium bicarbonate (AMBIC), trichloroacetic acid (TCA) were purchased from Sigma-Aldrich. Sialidase A was provided from Agilent (prozyme). Centrifugal filter Amicon 3k were purchased from Merck Millipore.

### 3.2.2    FSH variants generation

#### 3.2.2.1    Temperature -Temperature stressed sample

The thermal stressed variant was obtained by incubation of DS_FSH at 40° C for 4 weeks. Before subjecting the sample to thermal stress, an aliquot (DS_FSH_T0) of each of the 3 batches was taken to be used as reference material for the glycosylation profile. The experiment sample table is reported below.

| Sample | | Stress Conditions |
|---|---|---|
| DS_FSH_ T0 | Batch 1 | Untreated |
| | Batch 2 | |
| | Batch 3 | |
| DS_FSH_ Thermic Stressed variant | Batch 1 | 40° C for 4 weeks |
| | Batch 2 | |
| | Batch 3 | |

### 3.2.2.2 Acidic/basic variants - Acidic/basic enriched fractions

For each FSH batch acidic and basic fractions were provided by Merck chromatography laboratories. Basic and acidic fractions were enriched by preparative AEX-salt gradient chromatography and subsequent analytical AEX-pH gradient chromatography.

The experiment table is reported below.

| Sample | |
|---|---|
| DS_FSH_ Acidic enriched variant | Batch 1 |
| | Batch 2 |
| | Batch 3 |
| DS_FSH_ Basic enriched variant | Batch 1 |
| | Batch 2 |
| | Batch 3 |

### 3.2.2.3 Acidic/Basic pH - Acidic and basic pH stressed sample

To obtain the stressed variants at acidic and basic pH, 18.5 ml of DS_FSH were treated respectively with the same volume of sodium citrate 0.5 M pH 3.0 and sodium bicarbonate 0.5 M pH 9.0 for 3 days at 25 ° C. For both conditions, after incubation, buffer exchange was performed by Amicon 3K with milliQ water until a pH between 5-6 is obtained (DS_FSH pH at T0). The experiment table is reported below.

| Sample | | Stress Conditions |
|---|---|---|
| DS_FSH_ Acidic Stressed variant | Batch 1 | sodium citrate 0.5 M pH 3.0 for 3 days at 25°C |
| | Batch 2 | |
| | Batch 3 | |
| DS_FSH_ Basic Stressed variant | Batch 1 | sodium bicarbonate 0.5 M pH 9.0 for 3 days at 25°C |
| | Batch 2 | |
| | Batch 3 | |

### 3.2.2.4   De-sialylated - Sialic acid enzymatic hydrolysis

To obtain the de-sialylated variant the hydrolysis was performed by adding 0.1U
sialidase per mg of glycoprotein. Then the reaction was incubated at 37° C for
18h. The experiment table is reported below.

| Sample | | Stress Conditions |
|---|---|---|
| DS_FSH_ De-sialylated variant | Batch 1 | 0.1U sialidase per mg of glycoprotein. for 18h at 37° C |
| | Batch 2 | |
| | Batch 3 | |

### 3.2.2.5   De-sialylated/de-galactosylated- enzymatic hydrolysis

To obtain the de-sialylated and de-galactosylated variant, 0.1U sialidase and 0.1U
galactosidase per mg of glycoprotein were added. Then the reaction was
incubated at 37° C for 18h. The experiment table is reported below.

| Sample | | Stress Conditions |
|---|---|---|
| DS_FSH_ De-sialo/De-gal variant | Batch 1 | 0.1U sialidase and 0.1U galactosidase per mg of glycoprotein. for 18h at 37° C |
| | Batch 2 | |
| | Batch 3 | |

### 3.2.3   FSH variants characterisation

### 3.2.3.1   Glycopeptides Sample preparation

For each sample, 200 µg of glycoprotein were concentrated to 1 µg/µL by using
Amicon Ultra 3K cartridges. Then an equal volume of 8 M Guanidine, 130 mM
Tris, 1 mM EDTA pH 7.6 was added. Samples have been reduced by adding 20
µL of DTT 500 mM (1h, 37 °C) and then alkylated with 40 µL of IAM 500 mM
(30 minutes in the dark). Then each sample were washed five times with 200 µL
of 2 M Urea, 50 mM Tris pH 8.0 by using Amicon Ultra 3K cartridges. Washed
samples were subjected to enzymatic hydrolysis with 10 µL of Chymotrypsin
(E/S 1:20, 4h, 37 °C).

### 3.2.3.2 LC-MS/MS analysis

Samples have been analysed by UPLC-ESI-MS/MS on a Waters Xevo G2-XS Q-TOF mass spectrometer in MS$^E$ mode, injecting 25 µL using the auto dilution mode starting from 10 µL of each sample in vial. A hydrophilic interaction liquid chromatography (HILIC) based column was chosen to improve the separation of glycopeptides.

MS$^E$ is a technique that allows acquiring accurate peptide and fragment mass data for all peptides and glycopeptides of a sample within a single run. The MS$^E$ fragmentation data can be used to confirm the accurate mass assignment of a peptide and to localize the assigned modifications within a peptide sequence. Fragmentation information is collected in parallel for all modified or not modified peptide ions, avoiding the bias and analytical irreproducibility that occurs with data-dependent approaches that require serial pre-selection of peptide precursor ions. LC/MS$^E$ datasets are acquired with two alternating MS functions: one for MS of peptide precursor acquired at low collision cell energy and one (MS$^E$) for collecting peptide fragmentation data at elevated collision cell energies. Using this methodology, the resulting peptide information includes retention time, accurate mass and intensity and fragmentation profiles for all detected peptides and glycopeptides within a peptide map. The experimental conditions used for the UPLC and Q-TOF are reported below.

**UPLC Experimental conditions:**

| Column | Acquity UPLC BEH glycan, 1.7 µm, 2.1 µm x 150 mm (Waters) |
|---|---|
| Flow rate | 0.2 mL/min |
| Injection Volume | 25mL (autodilution mode) |
| Column Temperature | 50°C ± 2°C |
| Autosampler Temperature | 8°C ± 3°C |
| Eluents | A: 0.1% TFA in H$_2$O MilliQ<br>B: 0.1% TFA in ACN |
| UV detector | 214 nm |

**UPLC Gradient:**

| Time | % A | % B |
|------|-----|-----|
| 0.0 | 10 | 90 |
| 3.0 | 10 | 90 |
| 3.1 | 30 | 70 |
| 8.0 | 30 | 70 |
| 68.0 | 55 | 45 |
| 68.1 | 80 | 20 |
| 72.0 | 80 | 20 |
| 72.1 | 10 | 90 |
| 90.0 | 10 | 90 |

**Xevo-G2-XS Q-TOF Experimental conditions (MS$^E$):**

| | |
|---|---|
| **MS and MS/MS Scan range (m/z)** | 100-2500 |
| **Scan time (sec)** | 0.5 |
| **Capillary Voltage** | 3.0 kV |
| **Sampling Cone** | 30.0 V |
| **Source Temperature (°C)** | 120 |
| **Desolvatation Temperature (°C)** | 300 |
| **Cone Gas Flow (L/Hr)** | 50 |
| **Desolvatation Gas Flow (L/Hr)** | 800 |

### 3.2.3.3 Data interpretation

The mass spectrometry data were analysed by Expressionist v13.0 (Genedata) against custom glycopeptides and glycan libraries. The software allows a semi-automated processing of whole obtained dataset using dedicated libraries-MS for each FSH variants, focused both on the detection of MS glycopeptides precursors and the expected retention times.

### 3.3    Results and discussion

### 3.3.1    DS_FSH_T0 glycopeptides mapping and Library construction

Figure 3.2 shows the total ionic current (TIC) obtained from the LC-MS/MS analysis of DS_FSH T0 sample; the TIC shows the great glycopeptides separation from the unmodified peptides. This separation was crucial in simplifying  the attribution of each glycoform to each of the respective glycosite.
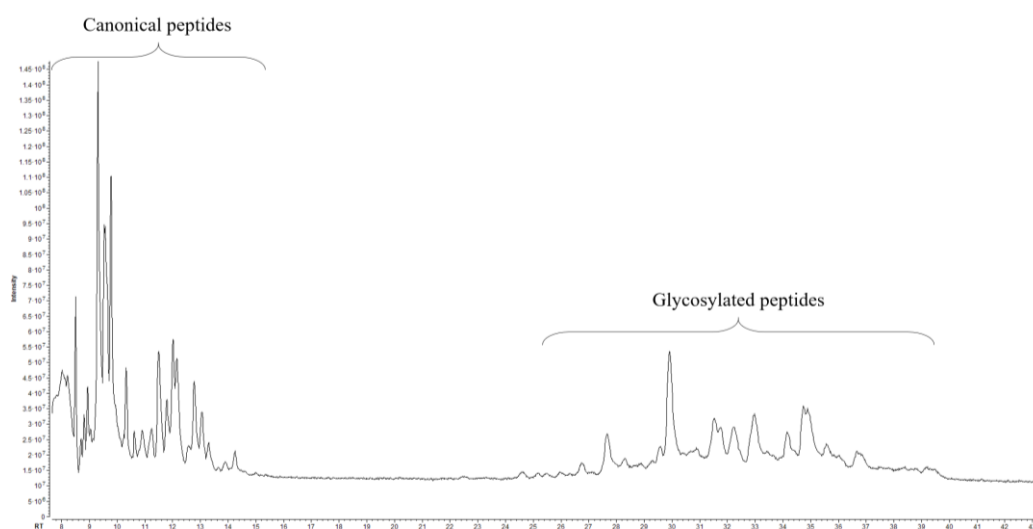


*Figure 3.2 Total Ion Current of DS_FSH T0 sample. Glycosylated peptides are separated from unmodified ones (canonical peptides).*

The LC-MS/MS data was then processed to obtain the distribution of glycoforms based on the N-glycosylation site of the molecule; Glycopeptide mapping profiles were obtained by Expressionist software carrying out a targeted search of the 4 glycosylated peptides (Table 3.1).

*Table 3.1 Glycopeptides identification by Expressionist software.*

| N-glycan site α- Asn 52 | | | | | |
|---|---|---|---|---|---|
| **Glycosylation** | **Calc. Mass** | **RT** | **z** | **Coverage** | **Integrated Max. Int.** |
| A2G2 | 2987.24 | 24.9 | 2 : 3 | MS | 6351 |
| A2G2S1 | 3278.34 | 26.9 | 2 : 3 | MS | 88392 |
| A2G2S2 | 3569.43 | 27.9 | 2 : 3 | MS/MS and MS | 169761 |
| A2G2S2_OAC1 | 3611.44 | 26.1 | 3 | MS | 164 |
| A3G3S1 | 3643.47 | 29.9 | 2 | MS | 201 |
| A3G3S1 | 3643.47 | 29.9 | 2 : 3 | MS | 500 |
| FA2G2S2 | 3715.49 | 30.4 | 3 | MS | 778 |
| A3G3S2 | 3934.56 | 31.4 | 2 : 3 | MS | 12445 |
| A3G3S3 | 4225.66 | 32.8 | 2 : 3 | MS | 14537 |
| A4G4S4 | 4881.89 | 36.6 | 3 | MS | 159 |
| N-glycan site α- Asn 78 | | | | | |
| **Glycosylation** | **Calc. MS** | **RT** | **z** | **Coverage** | **Integrated Max. Int.** |
| A2G2 | 2717.07 | 24.4 | 2 : 3 | MS | 14439 |
| A2G1S1 | 2846.12 | 23.9 | 2 : 3 | MS | 476 |
| A2G2S1 | 3008.17 | 26.3 | 2 : 3 | MS | 227546 |
| A2G2S2 | 3299.26 | 28.1 | 2 : 3 : 4 | MS/MS and MS | 658989 |
| A2G2S2_OAC1 | 3341.27 | 25.3 | 2 : 3 | MS | 10665 |
| A3G3S1 | 3373.30 | 29.7 | 2 : 3 | MS | 7559 |
| A2G2S2_OAC2 | 3383.29 | 22.5 | 2 : 3 | MS | 3133 |
| A2G2S2_OAC2 | 3383.29 | 24.3 | 2 : 3 | MS | 6530 |
| FA2G2S2 | 3445.32 | 29.8 | 2 : 3 | MS | 7088 |
| A3G3S2 | 3664.40 | 31.3 | 2 : 3 | MS | 28533 |
| A3G3S3 | 3955.49 | 32.3 | 2 : 3 | MS | 35150 |
| N-glycan site β- Asn 7 | | | | | |
| **Glycosylation** | **Calc. MS** | **RT** | **z** | **Coverage** | **Integrated Max. Int.** |
| FA2G2S1 | 4385.83 | 22.2 | 2 : 3 | MS | 2427709 |
| A3G3S1 | 4604.90 | 24.2 | 3 | MS | 511010 |
| FA2G2S2 | 4676.92 | 23.8 | 3 : 4 | MS/MS and MS | 11032671 |
| FA3G3S1 | 4750.96 | 25.1 | 3 | MS | 1676013 |
| A3G3S2 | 4896.00 | 25.9 | 3 : 4 | MS | 3718950 |
| FA3G3S2 | 5042.05 | 26.6 | 3 | MS/MS and MS | 13544092 |
| A3G3S3 | 5187.09 | 27.2 | 3 : 4 | MS/MS and MS | 12719087 |
| A4G4S2 | 5261.13 | 28.9 | 3 | MS | 475659 |
| FA3G3S3 | 5333.15 | 27.9 | 3 : 4 | MS/MS and MS | 45656072 |
| FA3G3S3_OAC1 | 5375.16 | 25.3 | 3 | MS | 349253 |

| Glycosylation | Calc. MS | RT | z | Coverage | Integrated Max. Int. |
|---|---|---|---|---|---|
| FA3G3S3_OAC1 | 5375.16 | 25.8 | 3 | MS | 313930 |
| FA3G3S3_OAC2 | 5417.17 | 22.8 | 3 | MS | 42507 |
| FA3G3S3_OAC2 | 5417.17 | 24.4 | 3 | MS | 175089 |
| FA3G3S3_OAC2 | 5417.17 | 25.2 | 3 | MS | 48542 |
| A4G4S3 | 5552.22 | 30.1 | 3 : 4 | MS/MS and MS | 2616686 |
| A4G4S3 | 5552.22 | 29.7 | 4 | MS | 195046 |
| FA4G4S3 | 5698.28 | 30.3 | 3 : 4 | MS/MS and MS | 8259333 |
| FA4G4S3 | 5698.28 | 30.8 | 3 : 4 | MS/MS and MS | 12914340 |
| A4G4S4 | 5843.32 | 30.8 | 3 : 4 | MS/MS and MS | 5006065 |
| FA4G4S4 | 5989.38 | 31.3 | 3 : 4 | MS/MS and MS | 25160582 |
| FA4G4S4_OAC1 | 6031.39 | 29.0 | 3 : 4 | MS | 464215 |
| FA4G4L1S3 | 6063.41 | 32.9 | 3 : 4 | MS/MS and MS | 1581611 |
| A4G4L1S4 | 6208.45 | 33.3 | 3 | MS | 203223 |
| FA4G4L1S4 | 6354.51 | 33.9 | 3 : 4 | MS | 3532675 |
| **N-glycan site β- Asn 24** | | | | | |
| **Glycosylation** | **Calc. MS** | **RT** | **z** | **Coverage** | **Integrated Max. Int.** |
| FA2G2 | 2762.10 | 21.8 | 2 : 3 | MS | 14122 |
| A2G2S1 | 2907.14 | 22.6 | 2 | MS | 1267 |
| FA2G2S1 | 3053.19 | 23.7 | 2 : 3 | MS | 288260 |
| FA2G2S1_OAC1 | 3095.20 | 20.7 | 2 | MS | 594 |
| FA4G2 | 3168.26 | 27.0 | 2 : 3 | MS | 5584 |
| A2G2S2 | 3198.23 | 24.4 | 2 | MS | 2539 |
| FA2G2S2 | 3344.29 | 25.6 | 2 : 3 | MS/MS and MS | 1391254 |
| FA2G2S2_OAC1 | 3386.30 | 22.5 | 2 | MS | 17612 |
| FA2G2S2_OAC2 | 3428.31 | 19.8 | 2 : 3 | MS | 6369 |
| FA2G2S2_OAC2 | 3428.31 | 21.4 | 2 : 3 | MS | 15124 |
| FA4G2S1 | 3459.35 | 28.8 | 2 : 3 | MS | 17633 |
| F2A2G2S2 | 3490.35 | 27.5 | 2 : 3 | MS | 32712 |
| FA3G3S2 | 3709.42 | 28.6 | 2 : 3 | MS/MS and MS | 48555 |
| FA3G3S3 | 4000.52 | 30.1 | 2 : 3 | MS/MS and MS | 153764 |
| FA3G3S3_OAC1 | 4042.53 | 27.3 | 2 : 3 | MS | 4586 |
| FA4G4S3 | 4365.65 | 33.0 | 3 | MS | 4738 |
| FA4G4S3 | 4365.65 | 33.5 | 3 | MS | 5288 |
| FA4G4S4 | 4656.74 | 34.1 | 2 : 3 | MS | 9801 |
| FA4G4L1S4 | 5021.88 | 36.9 | 3 | MS | 375 |

As for example, the A2G2S2 glycosylation attribution on LVQK$N_{52}$VTSESTC peptide is reported in Figure 3.3.

| Range | Peptide | Modifications | Mod. Locations | Glycans | Calc. Mass | RT | Charge States | Delta [ppm] | Score |
|-------|---------|---------------|----------------|---------|-----------|-----|---------------|-------------|-------|
| Alpha 48-59 | LVQKNVTSESTC | A2G2S2 | [N52] | | 3569.43 | 27.9 | 2 : 3 | 5.9 | 205.1 |

*Figure 3.3 Extracted ion chromatogram (A)of m/z=1190.82 peak (B) wit z=3; calculated precursor mass=3569.43 attributed to the biantennary sialylated glycopeptide LVQKNVTSESTC by interpreting the relative fragment spectrum (C).*

In a similar way, all the glycopeptides in DS_FSH_T0 were attributed. After that, a glycopeptides library containing precursor mass, modification and retention time of each identified glycopeptide was built. Glycopeptide mapping profiles and custom glycopeptides libraries-MS were generated for each variant using the same workflow above mentioned and processed semi-automatically by Expressionist software.

### 3.3.2 Temperature stressed variant – Glycopeptides mapping

LC-MS/MS data was analysed to obtain the distribution of glycans based on the N-glycosylation site in DS_FSH thermal stressed variant. The relative abundance of detected species is reported in the subsequent tables for each N-glycan species compared to untreated DS_FSH T0 sample. Moreover, Fucosylation, Sialylation and Galactosylation distributions were calculated for each glycosylation site of the molecule (Asn52 and Asn78 for the α subunit, Asn7 and Asn24 for the β subunit) by grouping glycans according to the function related to structural features. The grouping gives an estimation of the most abundant glycan types per each glycosites. All the species, within each glycosite, are grouped as per the number of sialic acids and hyper O-acetylated sialic acid, fucose and exposed galactose present as schematized in table hereafter.

*Table 3.2 Grouping schematization*
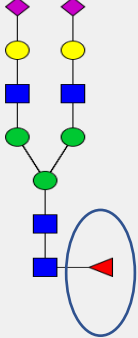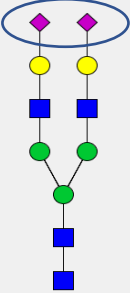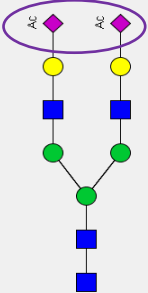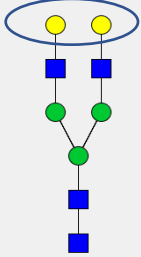
| Fucosylation | Sialylation | O-Acetylation | Galactosylation |
|---|---|---|---|
| All species containing fucose are grouped for each glycosite:<br><br>**FnAxGySz**<br>x=1,2,3,4<br>y=0,1,2,3,4<br>z=0,1,2,3,4 | Species are grouped per number of sialic acid present for each glycosite:<br><br>**FnAxGySz**<br>x=1,2,3,4<br>y=0,1,2,3,4<br>n=0,1,2 | Species are grouped per number of hyper O-acetylated sialic acid present for each glycosite:<br><br>**FnAxGySz_Ac**<br>x=1,2,3,4<br>y=0,1,2,3,4<br>n=0,1,2 | Species are grouped per number of exposed terminal galactose for each glycosite:<br><br>**FnAxGySz**<br>x=1,2,3,4<br>n=0,1,2<br>z=0 |
| **Fucosylated:** n=1,2<br>**A-fucosylated:** n=0 | **Non –sialylated:** z=0<br>**Mono-sialylated:** z=1<br>**Di-sialylated:** z=2<br>**Tri-sialylated:** z=3<br>**Tetra-sialylated:** z=4 | **FnAxGySz_Ac1**<br>Ac=1<br><br>**FnAxGySz_Ac2**<br>Ac=2 | **G0:** y=0<br>**G1:** y=1<br>**G2:** y=2<br>**G3:** y=3<br>**G4:** y=4 |

*Table 3.3 N-Glycan distribution at Asn 52 of FSH α - subunit.*

| Asn52 - α N-Glycan site | | | | | | |
|---|---|---|---|---|---|---|
| **Glycans** | **T0** | | | **Thermal stress** | | |
| | **Batch1** | **Batch2** | **Batch3** | **Batch1** | **Batch2** | **Batch3** |
| **A2G2** | 1.2 | 1.2 | 1.1 | 1.1 | 1.1 | 1.1 |
| **A2G2S1** | 22.3 | 22.5 | 21.5 | 22.0 | 22.6 | 22.1 |
| **A2G2S2** | 57.6 | 58.8 | 59.3 | 58.7 | 60.6 | 57.8 |
| **A2G2S2_1Ac** | 2.1 | 2.1 | 2.0 | 1.9 | 1.3 | 2.0 |
| **A3G3S1** | 0.5 | 0.4 | 0.4 | 0.4 | 0.1 | 0.5 |
| **A3G3S2** | 4.7 | 4.4 | 4.6 | 4.7 | 4.2 | 4.8 |
| **A3G3S3** | 8.6 | 8.1 | 8.5 | 8.5 | 8.4 | 8.7 |
| **A4G4S3** | 0.7 | 0.5 | 0.5 | 0.6 | 0.4 | 0.7 |
| **A4G4S4** | 0.6 | 0.4 | 0.5 | 0.5 | 0.3 | 0.6 |
| **FA2G2S2** | 1.6 | 1.5 | 1.5 | 1.5 | 1.1 | 1.6 |

*Table 3.4 N-Glycan distribution at Asn 78 of FSH α - subunit.*

| Asn78 - α N-Glycan site | | | | | | |
|---|---|---|---|---|---|---|
| **Glycans** | **T0** | | | **Thermal stress** | | |
| | **Batch1** | **Batch2** | **Batch3** | **Batch1** | **Batch2** | **Batch3** |
| **A2G2** | 0.8 | 0.8 | 0.7 | 0.8 | 0.7 | 0.8 |
| **A2G2S1** | 18.6 | 18.5 | 18.3 | 19.2 | 19.9 | 18.6 |
| **A2G2S2** | 62.0 | 63.0 | 62.9 | 62.5 | 63.8 | 62.7 |
| **A2G2S2_1Ac** | 2.2 | 2.3 | 2.2 | 2.2 | 2.1 | 2.1 |
| **A2G2S2_2Ac** | 4.3 | 4.2 | 4.2 | 3.6 | 3.5 | 3.7 |
| **A3G3S1** | 0.8 | 0.7 | 0.7 | 0.8 | 0.6 | 0.8 |
| **A3G3S2** | 4.0 | 3.8 | 4.0 | 4.0 | 3.5 | 4.1 |
| **A3G3S3** | 5.5 | 5.0 | 5.3 | 5.3 | 4.5 | 5.4 |
| **A4G4S3** | 0.3 | 0.2 | 0.3 | 0.3 | 0.1 | 0.3 |
| **FA2G2S2** | 1.5 | 1.4 | 1.4 | 1.4 | 1.2 | 1.2 |

*Table 3.5 N-Glycan distribution at Asn 7 of FSH β - subunit.*

| Asn7 - β N-Glycan site | | | | | | |
|---|---|---|---|---|---|---|
| **Glycans** | **T0** | | | **Thermal stress** | | |
| | **Batch1** | **Batch2** | **Batch3** | **Batch1** | **Batch2** | **Batch3** |
| **A3G3S1** | 0.2 | 0.2 | 0.2 | 0.2 | 0.2 | 0.2 |
| **A3G3S2** | 1.9 | 1.9 | 1.8 | 2.1 | 2.6 | 2.1 |
| **A3G3S3** | 9.4 | 9.6 | 9.5 | 9.5 | 9.4 | 9.2 |
| **A4G4S2** | 0.4 | 0.3 | 0.3 | 0.4 | 0.4 | 0.4 |
| **A4G4S3** | 2.9 | 2.9 | 2.9 | 3.0 | 3.0 | 3.1 |
| **A4G4S4** | 3.5 | 3.5 | 3.5 | 3.2 | 2.7 | 3.4 |
| **A5G5S4** | 0.6 | 0.5 | 0.5 | 0.4 | 0.2 | 0.5 |
| **FA2G2S1** | 0.8 | 0.8 | 0.8 | 0.9 | 1.0 | 0.9 |
| **FA2G2S2** | 5.4 | 5.3 | 5.4 | 5.4 | 5.5 | 5.3 |
| **FA3G3S1** | 0.6 | 0.5 | 0.5 | 0.6 | 0.8 | 0.6 |
| **FA3G3S2** | 6.5 | 6.6 | 6.5 | 7.6 | 9.6 | 7.1 |
| **FA3G3S3** | 30.1 | 30.4 | 30.3 | 29.5 | 30.1 | 29.3 |
| **FA3G3S3_1Ac** | 1.6 | 1.6 | 1.5 | 1.4 | 1.2 | 1.4 |
| **FA3G3S3_2Ac** | 1.5 | 1.5 | 1.5 | 1.1 | 0.8 | 1.1 |
| **FA4G4S3** | 11.9 | 12.0 | 12.0 | 13.0 | 14.2 | 13.1 |
| **FA4G4S4** | 15.8 | 16.0 | 16.2 | 15.7 | 13.8 | 15.8 |
| **FA4G4S4_1Ac** | 1.3 | 1.2 | 1.2 | 1.1 | 0.6 | 1.1 |
| **FA4G4S4_2Ac** | 0.2 | 0.2 | 0.2 | 0.1 | 0.1 | 0.2 |
| **FA5G5S3** | 1.7 | 1.5 | 1.6 | 1.6 | 1.2 | 1.8 |
| **FA5G5S4** | 3.5 | 3.4 | 3.5 | 3.2 | 2.6 | 3.4 |

*Table 3.6 N-Glycan distribution at Asn 24 of FSH β - subunit.*

| Asn24 - β N-Glycan site | | | | | | |
|---|---|---|---|---|---|---|
| **Glycans** | **T0** | | | **Thermal stress** | | |
| | **Batch1** | **Batch2** | **Batch3** | **Batch1** | **Batch2** | **Batch3** |
| **A2G2S2** | 0.1 | 0.1 | 0.1 | 0.1 | 0.1 | 0.1 |
| **FA2G2** | 0.4 | 0.4 | 0.4 | 0.5 | 0.6 | 0.5 |
| **FA2G2S1** | 11.7 | 10.9 | 11.3 | 12.7 | 14.4 | 12.8 |
| **FA2G2S1_1Ac** | 0.1 | 0.1 | 0.1 | 0.1 | 0.1 | 0.1 |
| **FA2G2S2** | 62.7 | 64.1 | 63.9 | 62.7 | 61.5 | 61.0 |
| **FA2G2S2_1Ac** | 3.5 | 3.5 | 3.5 | 3.3 | 3.4 | 3.5 |
| **FA2G2S2_2Ac** | 4.8 | 4.8 | 4.9 | 4.3 | 4.5 | 4.6 |
| **F2A2G2S2** | 2.1 | 2.0 | 2.0 | 2.0 | 1.9 | 2.0 |
| **FA3G3S2** | 1.8 | 1.7 | 1.8 | 2.0 | 2.4 | 2.0 |

| | | | | | | |
|---|---|---|---|---|---|---|
| **FA3G3S3** | 7.9 | 7.7 | 7.8 | 7.4 | 6.9 | 7.9 |
| **FA3G3S3_1Ac** | 0.7 | 0.6 | 0.6 | 0.6 | 0.5 | 0.7 |
| **FA4G2** | 0.4 | 0.4 | 0.3 | 0.5 | 0.4 | 0.5 |
| **FA4G2S1** | 1.4 | 1.6 | 1.2 | 1.8 | 1.5 | 1.9 |
| **FA4G4S3** | 0.9 | 0.7 | 0.8 | 0.8 | 0.7 | 0.9 |
| **FA4G4S4** | 1.2 | 1.1 | 1.1 | 1.1 | 0.8 | 1.2 |
| **FA5G5S4** | 0.2 | 0.1 | 0.1 | 0.1 | 0.1 | 0.2 |

## Grouping: Fucosylation

The table below shows the relative abundance (%) of non fucosylated (a-fucosylated) and fucosylated species for each glycosylation site.

*Table 3.7 Grouping of fucosylation of thermal stressed variants for each N-glycosylation site*

| Asn52 - α N-Glycan site | | | | | | |
|---|---|---|---|---|---|---|
| **Fucosylation** | **T0** | | | **Thermal stress** | | |
| | **Batch1** | **Batch2** | **Batch3** | **Batch1** | **Batch2** | **Batch3** |
| a-fucosylated | 98.4 | 98.5 | 98.5 | 98.5 | 98.9 | 98.4 |
| Fucosylated | 1.6 | 1.5 | 1.5 | 1.5 | 1.1 | 1.6 |
| **Asn78 - α N-Glycan site** | | | | | | |
| **Fucosylation** | **T0** | | | **Thermal stress** | | |
| | **Batch1** | **Batch2** | **Batch3** | **Batch1** | **Batch2** | **Batch3** |
| a-fucosylated | 98.5 | 98.6 | 98.6 | 98.6 | 98.8 | 98.5 |
| Fucosylated | 1.5 | 1.4 | 1.4 | 1.4 | 1.2 | 1.5 |
| **Asn7 - b N-Glycan site** | | | | | | |
| **Fucosylation** | **T0** | | | **Thermal stress** | | |
| | **Batch1** | **Batch2** | **Batch3** | **Batch1** | **Batch2** | **Batch3** |
| a-fucosylated | 18.9 | 18.8 | 18.8 | 18.8 | 18.4 | 18.8 |
| Fucosylated | 81.1 | 81.2 | 81.2 | 81.2 | 81.6 | 81.2 |
| **Asn24 - b N-Glycan site** | | | | | | |
| **Fucosylation** | **T0** | | | **Thermal stress** | | |
| | **Batch1** | **Batch2** | **Batch3** | **Batch1** | **Batch2** | **Batch3** |
| a-fucosylated | 0.2 | 0.2 | 0.2 | 0.2 | 0.1 | 0.2 |
| Fucosylated | 97.8 | 97.8 | 97.8 | 97.8 | 97.8 | 97.8 |
| di-fucosylated | 2.1 | 2.0 | 2.0 | 2.0 | 1.9 | 2.0 |

As shown in Figure 3.4, thermal stress did not affect fucosylation of the molecule.
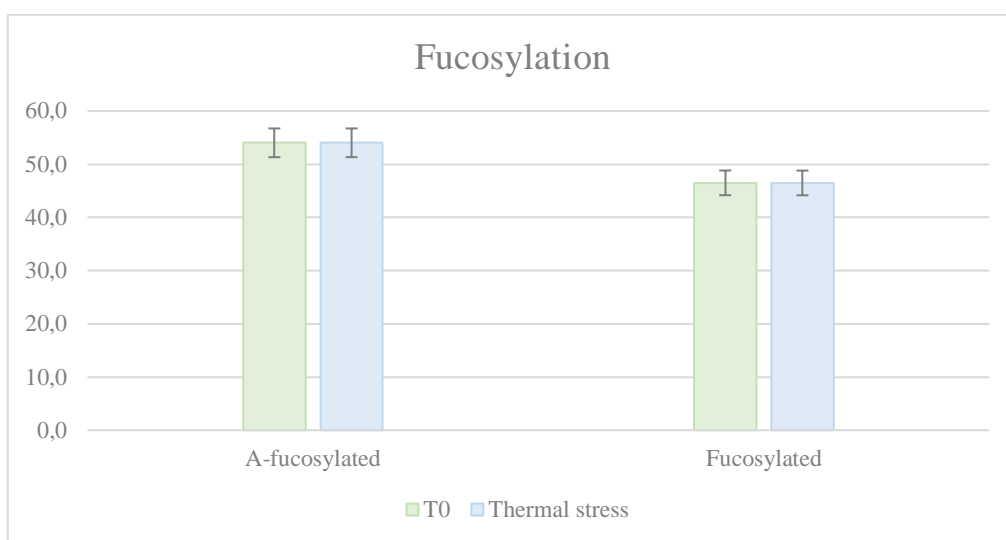
*Figure 3.4 Grouping of fucosylation of thermal stressed variants in overall molecule. The A-fucosylated and Fucosylated values of each variant is shown as an average value over the three batches*

## Grouping: Sialylation

The table below shows the relative abundance (%) of the sialylation degree (non, mono-, di-, tri- and tetra-sialylated species) for each glycosylation site.

*Table 3.8 Grouping of sialylations of thermal stressed variants for each N-glycosylation site*

| Asn52 - α N-Glycan site | | | | | | |
|---|---|---|---|---|---|---|
| **Sialylation** | **T0** | | | **Thermal stress** | | |
| | **Batch1** | **Batch2** | **Batch3** | **Batch1** | **Batch2** | **Batch3** |
| Non sialylated | 1.2 | 1.2 | 1.1 | 1.1 | 1.1 | 1.1 |
| Mono-sialylated | 22.8 | 22.9 | 22.0 | 22.4 | 22.7 | 22.6 |
| Di-sialylated | 66.0 | 66.8 | 67.4 | 66.9 | 67.2 | 66.3 |
| Tri-sialylated | 9.4 | 8.7 | 9.1 | 9.2 | 8.7 | 9.4 |
| Tetra-sialylated | 0.6 | 0.4 | 0.5 | 0.5 | 0.3 | 0.6 |
| Acetylated | 2.1 | 2.1 | 2.0 | 1.9 | 1.3 | 2.0 |
| Asn78 - α N-Glycan site | | | | | | |
| **Sialylation** | **T0** | | | **Thermal stress** | | |
| | **Batch1** | **Batch2** | **Batch3** | **Batch1** | **Batch2** | **Batch3** |
| Non sialylated | 0.8 | 0.8 | 0.7 | 0.8 | 0.7 | 0.8 |
| Mono-sialylated | 19.4 | 19.3 | 19.1 | 20.0 | 20.5 | 19.4 |
| Di-sialylated | 73.9 | 74.7 | 74.6 | 73.7 | 74.1 | 74.1 |
| Tri-sialylated | 5.8 | 5.3 | 5.6 | 5.6 | 4.6 | 5.7 |
| Acetylated | 6.5 | 6.5 | 6.4 | 5.8 | 5.6 | 5.9 |

| Asn7 - b N-Glycan site | | | | | | |
|---|---|---|---|---|---|---|
| **Sialylation** | **T0** | | | **Thermal stress** | | |
| | **Batch1** | **Batch2** | **Batch3** | **Batch1** | **Batch2** | **Batch3** |
| Non sialylated | N.D. | N.D. | N.D. | N.D. | N.D. | 0.0 |
| Mono-sialylated | 1.6 | 1.5 | 1.5 | 1.7 | 2.0 | 1.8 |
| Di-sialylated | 14.2 | 14.1 | 14.1 | 15.5 | 18.2 | 14.9 |
| Tri-sialylated | 59.2 | 59.5 | 59.2 | 59.1 | 60.0 | 58.9 |
| Tetra-sialylated | 25.0 | 24.8 | 25.2 | 23.7 | 19.8 | 24.4 |
| Acetylated | 4.7 | 4.5 | 4.5 | 3.7 | 2.7 | 3.8 |
| Asn24 - b N-Glycan site | | | | | | |
| **Sialylation** | **T0** | | | **Thermal stress** | | |
| | **Batch1** | **Batch2** | **Batch3** | **Batch1** | **Batch2** | **Batch3** |
| Non sialylated | 0.8 | 0.8 | 0.7 | 1.0 | 1.0 | 1.0 |
| Mono-sialylated | 13.3 | 12.6 | 12.6 | 14.7 | 16.1 | 15.0 |
| Di-sialylated | 75.1 | 76.3 | 76.2 | 74.4 | 73.9 | 73.2 |
| Tri-sialylated | 9.5 | 9.1 | 9.2 | 8.8 | 8.1 | 9.5 |
| Tetra-sialylated | 1.4 | 1.2 | 1.3 | 1.2 | 0.9 | 1.3 |
| Acetylated | 9.2 | 9.0 | 9.1 | 8.4 | 8.5 | 8.8 |

To visualize the sialylation trend on the entire molecule in the untreated (T0) and thermal stressed variants a sialylation index (S-index) was calculated for each N-glycosylation site as follow:

$$\text{S-index} = \sum \frac{\% \, Glycans \; relative \; abundance * N° \; of \; sialic \; acid \; content}{100}$$

*Table 3.9 S-index calculated on each N-glycan site*

| S-index | T0 | | | Thermal stress | | |
|---|---|---|---|---|---|---|
| | **Batch1** | **Batch2** | **Batch3** | **Batch1** | **Batch2** | **Batch3** |
| Asn52 - α | 1.9 | 1.8 | 1.9 | 1.9 | 1.8 | 1.9 |
| Asn78 - α | 1.8 | 1.8 | 1.9 | 1.8 | 1.8 | 1.8 |
| Asn7 - β | 3.1 | 3.1 | 3.1 | 3.0 | 3.0 | 3.1 |
| Asn24 - β | 1.9 | 1.9 | 1.9 | 2.0 | 2.0 | 2.0 |
| All sites | 2.2 | 2.2 | 2.2 | 2.2 | 2.2 | 2.2 |

As shown in Figure 3.5, the sialylation of the molecule is not affected by thermal stress.
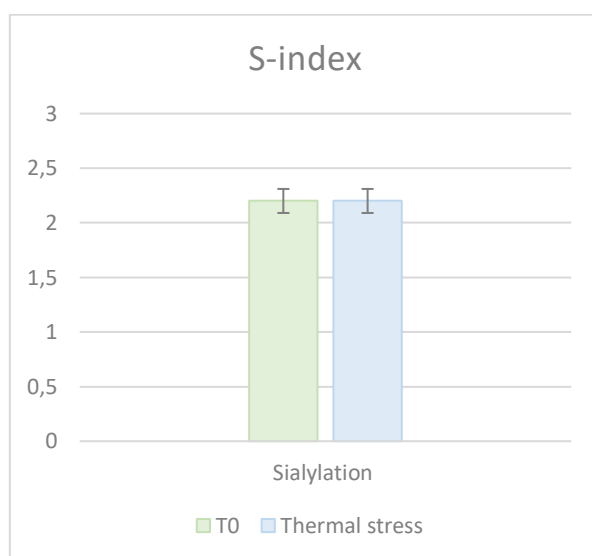
*Figure 3.5 S-Index of thermal stressed variants in overall molecule (mediated for 3 batches).*

The total sialic content of untreated DS_FSH (T0) remained unchanged after thermal stress, moreover, as shown by the Ac-index, the acetylation content of the sialic acid is also not affected by thermal stress. Acetylated index (Ac-index) is calculated as follow:

$$Ac\text{-}index = \sum \% \: relative \: abundance \: of \: Acetylated \: glycans * N° \: of \: acetylation$$

*Table 3.10 Ac-index calculated on each N-glycan site*

| Acetylated index | T0 | | | Thermal stress | | |
|---|---|---|---|---|---|---|
| | Batch1 | Batch2 | Batch3 | Batch1 | Batch2 | Batch3 |
| **Asn52 - α** | 2.1 | 2.1 | 2.0 | 1.9 | 1.3 | 2.0 |
| **Asn78 - α** | 10.8 | 10.6 | 10.6 | 9.4 | 9.1 | 9.6 |
| **Asn7 - β** | 6.0 | 6.2 | 5.7 | 4.6 | 3.6 | 5.1 |
| **Asn24 - β** | 14.0 | 13.8 | 14.0 | 12.7 | 13.0 | 13.4 |
| **All sites** | 8.2 | 8.2 | 8.1 | 7.1 | 6.8 | 7.5 |

The Figure 3.6 shows the Ac-indexes average of all N-glycan sites of the molecule for both T0 and thermal stressed variant compared to the reference.

108
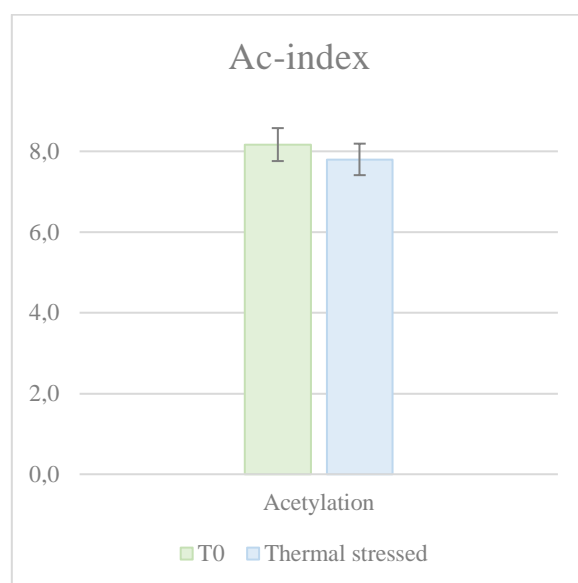
*Figure 3.6 Ac-Index of thermal stressed variants in overall molecule.*

## Grouping: Galactosylation

The table below shows the relative abundance (%) of species carrying expose galactose residues for each glycosylation site.

*Table 3.11 Grouping of galactosylation of thermal stressed variants for each N-glycoan site*

| Asn52 - α N-Glycan site | | | | | | |
|---|---|---|---|---|---|---|
| **Galactosylation** | **T0** | | | **Thermal stress** | | |
| | **Batch1** | **Batch2** | **Batch3** | **Batch1** | **Batch2** | **Batch3** |
| G2 | 1.2 | 1.2 | 1.1 | 1.1 | 1.1 | 1.1 |
| Asn78 - α N-Glycan site | | | | | | |
| **Galactosylation** | **T0** | | | **Thermal stress** | | |
| | **Batch1** | **Batch2** | **Batch3** | **Batch1** | **Batch2** | **Batch3** |
| G2 | 0.8 | 0.8 | 0.7 | 0.8 | 0.7 | 0.8 |
| Asn7 - b N-Glycan site | | | | | | |
| **Galactosylation** | **T0** | | | **Thermal stress** | | |
| | **Batch1** | **Batch2** | **Batch3** | **Batch1** | **Batch2** | **Batch3** |
| G1 | N.D. | N.D. | N.D. | N.D. | N.D. | N.D. |
| G2 | N.D. | N.D. | N.D. | N.D. | N.D. | N.D. |
| Asn24 - b N-Glycan site | | | | | | |
| **Galactosylation** | **T0** | | | **Thermal stress** | | |

|  | Batch1 | Batch2 | Batch3 | Batch1 | Batch2 | Batch3 |
| --- | --- | --- | --- | --- | --- | --- |
| G2 | 0.8 | 0.8 | 0.7 | 1.0 | 1.0 | 1.0 |

Only the G2 species were considered as they are the only galactosylated species found. As shown in Figure 3.7 the galactosylation is not affected by thermal stress.
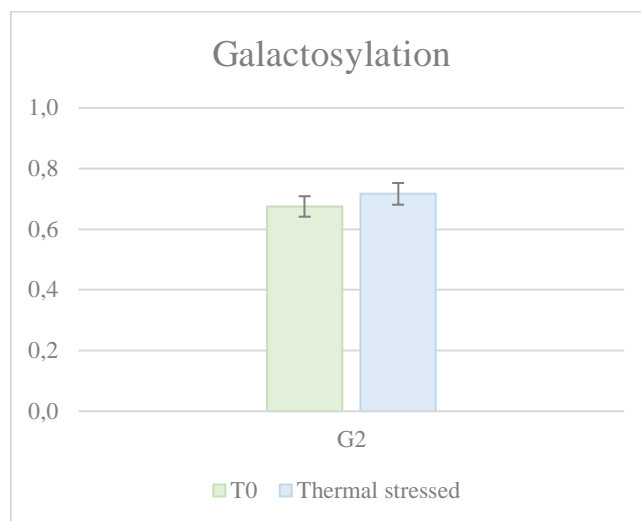


*Figure 3.7 Grouping of galactosylation of thermal stressed variant in overall molecule. The G1, G2, G3, G4 values of each variant is shown as an average value over the three batches.*

### 3.3.3 Acidic/basic variants – Glycopeptides mapping

LC-MS/MS data were analysed to obtain the distribution of glycans based on the N-glycosylation site of DS_FSH enriched acidic and basic variants. Glycopeptide mapping profiles were processed semi-automatically by Expressionist software. The relative distribution of detected species is reported in the subsequent tables; for each N-glycan species, relative abundance was calculated; moreover, three distinctive groups (Fucosylation, Sialylation and Galactosylation) were obtained for each glycosylation site of the molecule (Asn52 and Asn78 for the α subunit, Asn7 and Asn24 for the β subunit) by grouping glycans according to the function related to structural features as shown in Table 3.2.

*Table 3.12 N-Glycan distribution at Asn 52 of FSH a - subunit. (N.D.: not detected)*

| Asn52 - α N-Glycan site | | | | | | |
|---|---|---|---|---|---|---|
| **Glycans** | **Acidic fraction** | | | **Basic fraction** | | |
| | **Batch1** | **Batch2** | **Batch3** | **Batch1** | **Batch2** | **Batch3** |
| **A2G1S1** | N.D. | N.D. | N.D. | 0.1 | 0.1 | 0.2 |
| **A2G2** | N.D. | N.D. | N.D. | 5.0 | 5.1 | 5.4 |
| **A2G2S1** | 3.0 | 3.2 | 3.0 | 52.2 | 52.6 | 53.7 |
| **A2G2S2** | 33.4 | 33.6 | 32.3 | 37.7 | 37.6 | 35.9 |
| **A3G3S1** | N.D. | N.D. | N.D. | 1.6 | 1.5 | 1.7 |
| **A3G3S2** | 5.1 | 5.3 | 5.2 | 1.4 | 1.4 | 1.4 |
| **A3G3S3** | 36.9 | 35.2 | 36.0 | 0.1 | 0.1 | 0.1 |
| **A4G4S3** | 5.5 | 5.7 | 5.9 | N.D. | N.D. | N.D. |
| **A4G4S4** | 15.4 | 15.9 | 16.5 | N.D. | N.D. | N.D. |
| **FA2G2S2** | N.D. | N.D. | N.D. | 1.3 | 1.2 | 1.2 |
| **FA2G2S1** | N.D. | 0.1 | 0.1 | 0.5 | 0.4 | 0.5 |
| **A4G4L1Sa4** | 0.6 | 0.8 | 1.0 | N.D. | N.D. | N.D. |

*Table 3.13 N-Glycan distribution at Asn 78 of FSH a - subunit. (N.D.: not detected)*

| Asn78 - α N-Glycan site | | | | | | |
|---|---|---|---|---|---|---|
| **Glycans** | **Acidic fraction** | | | **Basic fraction** | | |
| | **Batch1** | **Batch2** | **Batch3** | **Batch1** | **Batch2** | **Batch3** |
| **A2G1S1** | N.D. | N.D. | N.D. | 0.1 | 0.1 | 0.2 |
| **A2G2** | N.D. | N.D. | N.D. | 3.7 | 3.8 | 4.1 |
| **A2G2S1** | 5.1 | 5.6 | 5.2 | 42.6 | 43.2 | 44.9 |
| **A2G2S2** | 54.7 | 53.4 | 53.3 | 42.1 | 41.2 | 39.8 |
| **A2G2S2_1Ac** | 1.4 | 1.4 | 1.3 | 1.5 | 1.6 | 1.4 |
| **A2G2S2_2Ac** | 1.8 | 2.0 | 1.9 | 3.0 | 3.0 | 2.6 |
| **A3G3S1** | 0.1 | 0.1 | 0.1 | 2.3 | 2.3 | 2.4 |
| **A3G3S2** | 5.1 | 5.3 | 5.3 | 2.5 | 2.6 | 2.4 |
| **A3G3S3** | 26.1 | 26.2 | 26.6 | 0.3 | 0.3 | 0.3 |
| **A4G4S2** | 0.1 | 0.1 | 0.1 | 0.1 | 0.1 | 0.1 |
| **A4G4S3** | 1.8 | 1.9 | 1.9 | N.D. | N.D. | N.D. |
| **FA2G2S2** | 0.5 | 0.5 | 0.5 | 1.3 | 1.3 | 1.2 |
| **FA2G2S1** | N.D. | N.D. | N.D. | 0.5 | 0.5 | 0.5 |
| **A3G3** | N.D. | N.D. | N.D. | 0.1 | 0.1 | 0.1 |
| **A4G4S4** | 3.2 | 3.3 | 3.4 | N.D. | N.D. | N.D. |
| **A4G4L1Sa4** | 0.3 | 0.3 | 0.3 | N.D. | N.D. | N.D. |

*Table 3.14 N-Glycan distribution at Asn 7 of FSH β - subunit. (N.D.: not detected)*

| Asn7 - β N-Glycan site | | | | | | |
|---|---|---|---|---|---|---|
| **Glycans** | **Acidic fraction** | | | **Basic fraction** | | |
| | **Batch1** | **Batch2** | **Batch3** | **Batch1** | **Batch2** | **Batch3** |
| **A3G3** | N.D. | N.D. | N.D. | 0.2 | 0.2 | 0.1 |
| **A3G3S1** | N.D. | N.D. | N.D. | 1.7 | 1.8 | 1.9 |
| **A3G3S2** | 0.1 | 0.1 | 0.1 | 5.4 | 5.5 | 5.8 |
| **A3G3S3** | 2.4 | 2.5 | 2.6 | 7.4 | 7.0 | 6.9 |
| **A4G4** | N.D. | N.D. | N.D. | N.D. | N.D. | N.D. |
| **A4G4S2** | N.D. | N.D. | N.D. | 1.5 | 1.7 | 1.8 |
| **A4G4S3** | 2.1 | 2.2 | 2.2 | 1.6 | 1.6 | 1.6 |
| **A4G4S4** | 12.3 | 12.0 | 11.9 | 0.1 | 0.1 | N.D. |
| **A5G5S4** | 0.8 | 0.9 | 0.9 | N.D. | N.D. | N.D. |
| **FA2G2S1** | N.D. | N.D. | N.D. | 4.1 | 4.3 | 4.4 |
| **FA2G2S2** | 1.1 | 1.1 | 1.2 | 8.7 | 9.0 | 9.2 |
| **FA3G3S1** | N.D. | N.D. | N.D. | 4.3 | 4.5 | 4.9 |
| **FA3G3S2** | 0.8 | 0.9 | 0.9 | 16.3 | 16.5 | 17.9 |

| | | | | | | |
|---|---|---|---|---|---|---|
| **FA3G3S3** | 10.0 | 10.2 | 10.4 | 24.4 | 23.2 | 20.8 |
| **FA3G3S3_1Ac** | 0.2 | 0.2 | 0.2 | 1.2 | 1.2 | 1.1 |
| **FA3G3S3_2Ac** | 0.1 | 0.1 | 0.1 | 0.8 | 0.8 | 0.7 |
| **FA4G4S3** | 10.9 | 10.9 | 10.9 | 10.6 | 10.5 | 10.3 |
| **FA4G4S4** | 50.3 | 49.4 | 49.2 | 0.9 | 1.1 | 0.8 |
| **FA4G4S4_1Ac** | 3.1 | 3.4 | 3.2 | N.D. | N.D. | N.D. |
| **FA5G5S3** | 0.5 | 0.6 | 0.7 | 2.2 | 2.2 | 2.3 |
| **FA5G5S4** | 5.0 | 4.9 | 5.0 | 0.6 | 0.6 | 0.4 |
| **FA4G4** | N.D. | N.D. | N.D. | 0.1 | N.D. | 0.1 |
| **FA4G4S2** | 0.3 | 0.4 | 0.5 | 6.3 | 6.7 | 7.2 |
| **A2G2S1** | N.D. | N.D. | N.D. | 1.0 | 1.0 | 1.0 |
| **FA2G2** | N.D. | N.D. | N.D. | 0.3 | 0.3 | 0.3 |
| **FA3G3** | N.D. | N.D. | N.D. | 0.4 | 0.4 | 0.4 |

*Table 3.15 N-Glycan distribution at Asn 24 of FSH β - subunit. (N.D.: not detected)*

| Asn24 - β N-Glycan site | | | | | | |
|---|---|---|---|---|---|---|
| **Glycans** | **Acidic fraction** | | | **Basic fraction** | | |
| | **Batch1** | **Batch2** | **Batch3** | **Batch1** | **Batch2** | **Batch3** |
| **A2G2S1** | N.D. | N.D. | N.D. | 0.3 | 0.3 | 0.3 |
| **A2G2S2** | N.D. | N.D. | N.D. | 0.4 | 0.3 | 0.4 |
| **FA1G1S1** | N.D. | N.D. | N.D. | 0.1 | N.D. | 0.1 |
| **FA2G2** | N.D. | N.D. | N.D. | 3.1 | 3.2 | 3.5 |
| **FA2G2S1** | 1.8 | 1.9 | 2.0 | 36.3 | 37.6 | 37.7 |
| **FA2G2S1_1Ac** | N.D. | N.D. | N.D. | 0.5 | 0.5 | 0.5 |
| **FA2G2S2** | 32.8 | 31.8 | 31.5 | 47.2 | 46.6 | 47.0 |
| **FA2G2S2_1Ac** | 1.0 | 1.1 | 1.0 | 2.2 | 2.2 | 2.1 |
| **FA2G2S2_2Ac** | 1.2 | 1.3 | 1.2 | 2.8 | 2.7 | 2.5 |
| **F2A2G2S2** | 0.5 | 0.5 | 0.5 | 1.8 | 1.7 | 1.8 |
| **FA3G3S2** | 3.1 | 3.3 | 3.3 | 1.3 | 1.2 | 1.3 |
| **FA3G3S3** | 30.1 | 29.9 | 29.8 | 1.1 | 1.0 | 1.0 |
| **FA3G3S3_1Ac** | 0.7 | 0.8 | 0.7 | N.D. | N.D. | N.D. |
| **FA4G2** | N.D. | N.D. | N.D. | 1.0 | 0.9 | 0.4 |
| **FA4G2S1** | 0.8 | 0.9 | 0.9 | 0.9 | 0.8 | 0.3 |
| **FA4G4S3** | 4.7 | 4.8 | 4.9 | 0.3 | 0.3 | 0.3 |
| **FA4G4S4** | 21.5 | 21.6 | 22.3 | N.D. | N.D. | N.D. |
| **FA5G5S4** | 1.0 | 1.0 | 1.1 | N.D. | N.D. | N.D. |
| **FA2G1** | N.D. | N.D. | N.D. | 0.2 | 0.2 | 0.2 |
| **FA3G3S1** | N.D. | N.D. | N.D. | 0.5 | 0.5 | 0.5 |
| **FA4G4S4_OAc2** | 0.8 | 0.8 | 0.8 | N.D. | N.D. | N.D. |

**Grouping: Fucosylation**

The table below shows the relative abundance (%) of non fucosylated (a-fucosylated) and fucosylated species for each glycosylation site. As a reference, the T0 fucosylation is reported as a range of the 3 batches values.

*Table 3.16 Grouping of fucosylation of Acid/Basic fractions variants for each N-glycosylation site.* (N.D.: not detected)

| Asn52 - α N-Glycan site | | | | | | | |
|---|---|---|---|---|---|---|---|
| **Fucosylation** | **T0** | **Acidic fraction** | | | **Basic fraction** | | |
| | **Range** | **Batch1** | **Batch2** | **Batch3** | **Batch1** | **Batch2** | **Batch3** |
| a-fucosylated | 98.4-98.5 | 100.0 | 99.9 | 99.8 | 98.2 | 98.4 | 98.3 |
| Fucosylated | 1.5-1-6 | N.D. | 0.1 | 0.2 | 1.8 | 1.6 | 1.7 |
| Asn78 - α N-Glycan site | | | | | | | |
| **Fucosylation** | **T0** | **Acidic fraction** | | | **Basic fraction** | | |
| | **Range** | **Batch1** | **Batch2** | **Batch3** | **Batch1** | **Batch2** | **Batch3** |
| a-fucosylated | 98.5-98.6 | 99.5 | 99.5 | 99.5 | 98.2 | 98.2 | 98.3 |
| Fucosylated | 1.4-1.5 | 0.5 | 0.5 | 0.5 | 1.8 | 1.8 | 1.7 |
| Asn7 - β N-Glycan site | | | | | | | |
| **Fucosylation** | **T0** | **Acidic fraction** | | | **Basic fraction** | | |
| | **Range** | **Batch1** | **Batch2** | **Batch3** | **Batch1** | **Batch2** | **Batch3** |
| a-fucosylated | 18.8-18.9 | 17.6 | 17.7 | 17.7 | 18.9 | 18.8 | 19.3 |
| Fucosylated | 81.1-81.2 | 82.4 | 82.3 | 82.3 | 81.1 | 81.2 | 80.7 |
| Asn24 - β N-Glycan site | | | | | | | |
| **Fucosylation** | **T0** | **Acidic fraction** | | | **Basic fraction** | | |
| | **Range** | **Batch1** | **Batch2** | **Batch3** | **Batch1** | **Batch2** | **Batch3** |
| A-fucosylated | 0.2-0.2 | N.D. | N.D. | N.D. | 0.7 | 0.7 | 0.8 |
| Fucosylated | 99.8-98.8 | 99.5 | 99.5 | 99.5 | 97.5 | 97.6 | 97.4 |
| di-fucosylated | 2.0-2.1 | 0.5 | 0.5 | 0.5 | 1.8 | 1.7 | 1.8 |

Overall, the degree of fucosylation, averaged out on 4 glycosylation sites of the molecule, appears to be the same in both acid/base enriched fractions, as we can see from the graph below.
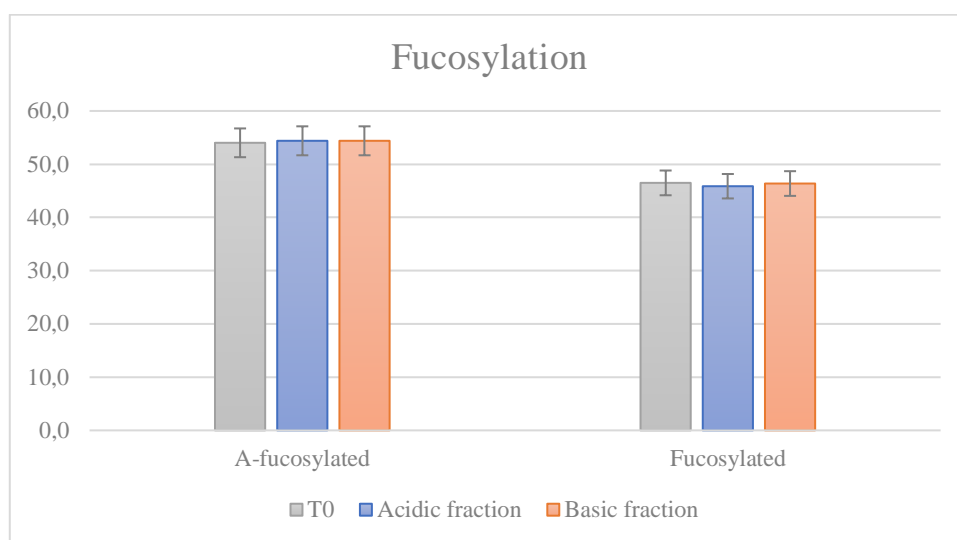
Figure 3.8 Grouping of fucosylation of Acid/Basic fractions variants in overall molecule. The A-fucosylated and Fucosylated values of each variant is shown as an average value over the three batches.

## Grouping: Sialylation

The table below shows the relative abundance (%) of sialylation (non, mono-, di-, tri- and tetra-sialylated species) for each glycosylation site. As a reference, the sialylation of T0 is reported as a range of the 3 batches values.

Table 3.17 Grouping of sialylations of Acid/Basic fractions variants for each N-glycosylation site. (N.D.: not detected)

| Asn52 - α N-Glycan site | | | | | | | |
|---|---|---|---|---|---|---|---|
| **Sialylation** | **T0** | **Acidic fraction** | | | **Basic fraction** | | |
| | **Range** | **Batch1** | **Batch2** | **Batch3** | **Batch1** | **Batch2** | **Batch3** |
| Non sialylated | 1.1-1.2 | N.D. | N.D. | N.D. | 5.0 | 5.1 | 5.4 |
| Mono-sialylated | 22.0-22.9 | 3.1 | 3.3 | 3.1 | 54.5 | 54.6 | 56.0 |
| Di-sialylated | 66.0-67.4 | 38.6 | 39.0 | 37.6 | 40.4 | 40.2 | 38.4 |
| Tri-sialylated | 8.7-9.4 | 42.4 | 41.0 | 41.9 | 0.1 | 0.1 | 0.1 |
| Tetra-sialylated | 0.4-0.6 | 16.0 | 16.8 | 17.5 | N.D. | N.D. | N.D. |
| Acetylated | 2.0-2.1 | N.D. | N.D. | N.D. | N.D. | N.D. | N.D. |
| Asn78 - α N-Glycan site | | | | | | | |
| **Sialylation** | **T0** | **Acidic fraction** | | | **Basic fraction** | | |
| | **Range** | **Batch1** | **Batch2** | **Batch3** | **Batch1** | **Batch2** | **Batch3** |
| Non sialylated | 0.7-0.8 | N.D. | N.D. | N.D. | 3.8 | 3.8 | 4.2 |
| Mono-sialylated | 19.1-19.4 | 5.2 | 5.7 | 5.4 | 45.5 | 46.1 | 48.0 |

| | 73.9-74.7 | 63.5 | 62.7 | 62.5 | 50.4 | 49.7 | 47.5 |
|---|---|---|---|---|---|---|---|
| Di-sialylated | 73.9-74.7 | 63.5 | 62.7 | 62.5 | 50.4 | 49.7 | 47.5 |
| Tri-sialylated | 5.3-5.8 | 27.8 | 28.1 | 28.5 | 0.3 | 0.3 | 0.3 |
| Acetylated | 6.4-6.5 | 3.1 | 3.4 | 3.2 | 4.5 | 4.5 | 4.0 |

| Asn7 - b N-Glycan site | | | | | | | |
|---|---|---|---|---|---|---|---|
| **Sialylation** | **T0** | **Acidic fraction** | | | **Basic fraction** | | |
| | **Range** | **Batch1** | **Batch2** | **Batch3** | **Batch1** | **Batch2** | **Batch3** |
| Non sialylated | N.D. | N.D. | N.D. | N.D. | 0.9 | 0.9 | 1.0 |
| Mono-sialylated | 1.5-1.6 | N.D. | N.D. | N.D. | 11.1 | 11.6 | 12.2 |
| Di-sialylated | 14.1-14.2 | 2.2 | 2.6 | 2.7 | 38.3 | 39.4 | 41.9 |
| Tri-sialylated | 59.2-59.5 | 26.2 | 26.8 | 27.0 | 48.2 | 46.4 | 43.7 |
| Tetra-sialylated | 24.8-25.2 | 71.6 | 70.6 | 70.3 | 1.6 | 1.7 | 1.2 |
| Acetylated | 4.5-4.7 | 3.4 | 3.8 | 3.6 | 2.0 | 1.9 | 1.8 |

| Asn24 - b N-Glycan site | | | | | | | |
|---|---|---|---|---|---|---|---|
| **Sialylation** | **T0** | **Acidic fraction** | | | **Basic fraction** | | |
| | **Range** | **Batch1** | **Batch2** | **Batch3** | **Batch1** | **Batch2** | **Batch3** |
| Non sialylated | 0.7-0.8 | N.D. | N.D. | N.D. | 4.3 | 4.2 | 4.1 |
| Mono-sialylated | 12.6-13.3 | 2.6 | 2.9 | 2.9 | 38.6 | 39.8 | 39.5 |
| Di-sialylated | 75.1-76.3 | 38.5 | 38.1 | 37.5 | 55.7 | 54.7 | 55.1 |
| Tri-sialylated | 9.1-9.5 | 35.5 | 35.5 | 35.4 | 1.4 | 1.3 | 1.3 |
| Tetra-sialylated | 1.2-1.4 | 23.3 | 23.5 | 24.2 | N.D. | N.D. | N.D. |
| Acetylated | 9.0-9.2 | 3.7 | 4.1 | 3.8 | 5.5 | 5.3 | 5.1 |

To better visualize the sialylation trend on the entire molecule in the two enriched Acid/base variants, the sialylation index (S-index) was calculated for each N-glycosylation site as follow:

$$\text{S-index} = \sum \frac{\% \, Glycans \, relative \, abundance * N° \, of \, sialic \, acid \, content}{100}$$

*Table 3.18 S-index calculated on each N-glycan site*

| **S-index** | **T0** | **Acidic fraction** | | | **Basic fraction** | | |
|---|---|---|---|---|---|---|---|
| | **Range** | **Batch1** | **Batch2** | **Batch3** | **Batch1** | **Batch2** | **Batch3** |
| Asn52 - α | 1.8-1.9 | 2.7 | 2.7 | 2.7 | 1.4 | 1.4 | 1.3 |
| Asn78 - α | 1.8-1.9 | 2.3 | 2.3 | 2.3 | 1.5 | 1.5 | 1.4 |
| Asn7 - β | 3.1-3.1 | 3.7 | 3.7 | 3.7 | 2.4 | 2.4 | 2.3 |
| Asn24 - β | 1.9-1.9 | 2.8 | 2.8 | 2.8 | 1.5 | 1.5 | 1.5 |
| All sites | 2.2-2.2 | 2.9 | 2.9 | 2.9 | 1.7 | 1.7 | 1.7 |

The Figure 3.9 shows the average S-indexes of all N-glycosylation sites of the molecule for both acidic/basic enriched fractions, compared to each other's and to the reference one.
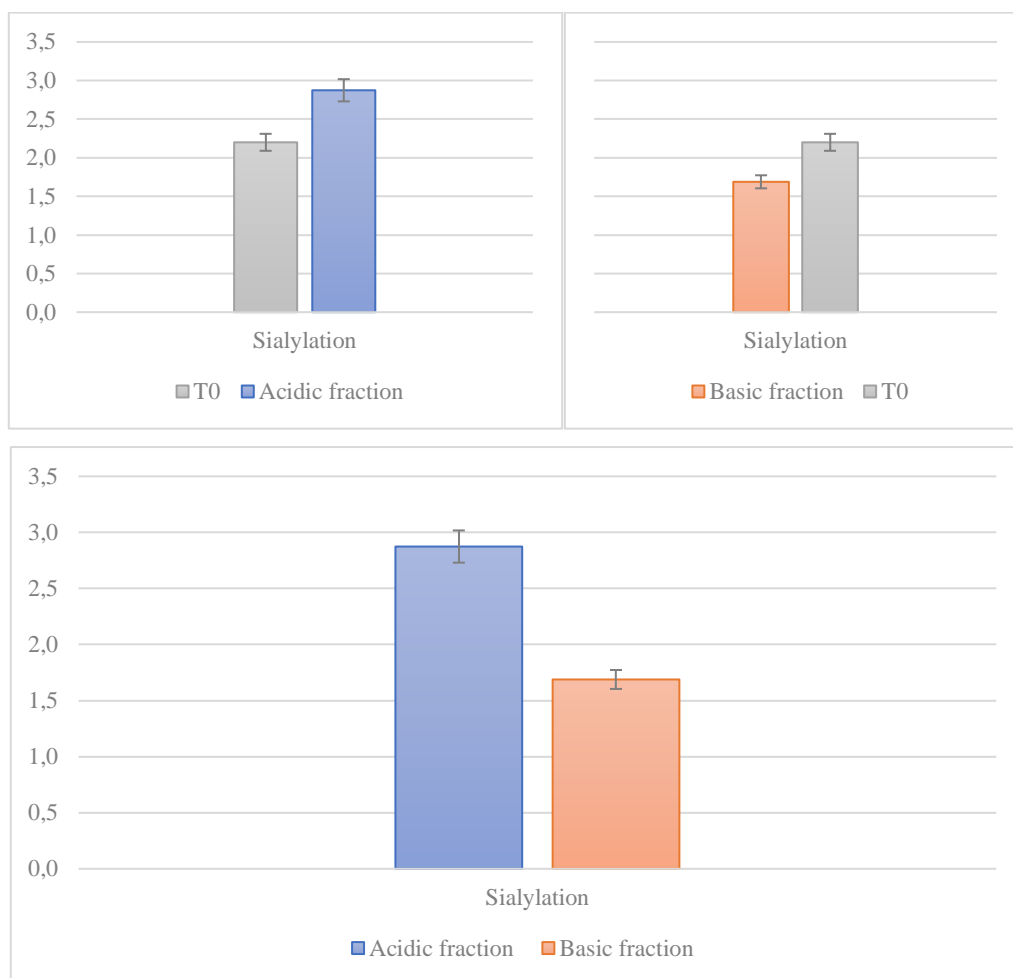


*Figure 3.9 S-Index of Acidic/Basic fractions variants in overall molecule.*

As expected, the sialic acid content is different in the two fractions. The acidic fraction is richer in species with a high sialic acid content than the basic one, therefore it shows a higher S-index.
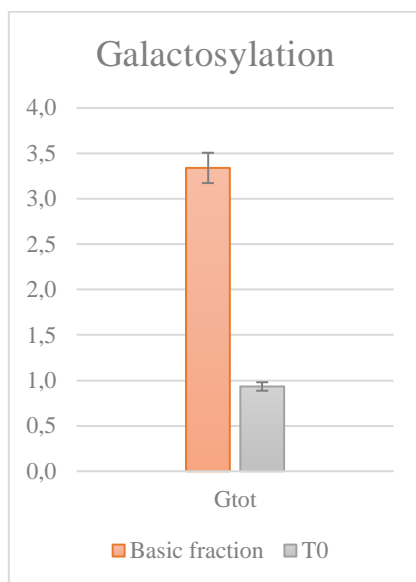
**Grouping: Galactosylation**

In the table below only the species containing exposed terminal galactose were grouped per number of exposed galactose content (G0, G1, G2, G3 etc.) and their relative abundance (%) is reported hereafter.

*Table 3.19 Grouping of galactosylation of Acidic/Basic fractions variants for each N-glycosylation site.* (N.D.: not detected)

| Asn52 - α N-Glycan site | | | | | | | |
|---|---|---|---|---|---|---|---|
| **Galactosylation** | **T0** | **Acidic fraction** | | | **Basic fraction** | | |
| | **Range** | **Batch1** | **Batch2** | **Batch3** | **Batch1** | **Batch2** | **Batch3** |
| G1 | N.D. | N.D. | N.D. | N.D. | N.D. | N.D. | N.D. |
| G2 | 1.1-1.2 | N.D. | N.D. | N.D. | 5.0 | 5.1 | 5.4 |
| **Asn78 - α N-Glycan site** | | | | | | | |
| **Galactosylation** | **T0** | **Acidic fraction** | | | **Basic fraction** | | |
| | **Range** | **Batch1** | **Batch2** | **Batch3** | **Batch1** | **Batch2** | **Batch3** |
| G1 | N.D. | N.D. | N.D. | N.D. | N.D. | N.D. | N.D. |
| G2 | 0.7-0.8 | N.D. | N.D. | N.D. | 3.7 | 3.8 | 4.1 |
| G3 | N.D. | N.D. | N.D. | N.D. | 0.1 | 0.1 | 0.1 |
| **Asn7 - b N-Glycan site** | | | | | | | |
| **Galactosylation** | **T0** | **Acidic fraction** | | | **Basic fraction** | | |
| | **Range** | **Batch1** | **Batch2** | **Batch3** | **Batch1** | **Batch2** | **Batch3** |
| G1 | N.D. | N.D. | N.D. | N.D. | N.D. | N.D. | N.D. |
| G2 | N.D. | N.D. | N.D. | N.D. | 0.3 | 0.3 | 0.3 |
| G3 | N.D. | N.D. | N.D. | N.D. | 0.5 | 0.5 | 0.6 |
| G4 | N.D. | N.D. | N.D. | N.D. | 0.1 | N.D. | 0.1 |
| **Asn24 - b N-Glycan site** | | | | | | | |
| **Galactosylation** | **T0** | **Acidic fraction** | | | **Basic fraction** | | |
| | **Range** | **Batch1** | **Batch2** | **Batch3** | **Batch1** | **Batch2** | **Batch3** |
| G1 | N.D. | N.D. | N.D. | N.D. | 0.2 | 0.2 | 0.2 |
| G2 | 0.7-0.8 | N.D. | N.D. | N.D. | 4.1 | 4.0 | 3.9 |

In the acidic fraction there is no species exposing galactose as it is rich in highly sialylated end-capping species. Galactosylated species are observed only in the in the basic fraction. The overall galactosylation in the basic fraction is greater than that observed in the untreated sample (Figure 3.10).

*Figure 3.10 Grouping of total galactosylation of Basic fractions variants in overall molecule compared to the untreated sample T0.*

### 3.3.4 Acidic/Basic pH – Glycopeptides mapping

LC-MS/MS data were analysed to obtain the distribution of glycans based on the N-glycosylation site of acid and basic pH stressed variants. Glycopeptide mapping profiles were processed semi-automatically by Expressionist software. The relative abundance of detected species is reported in the subsequent tables for each N-glycan species. Moreover, as for the previous variant, Fucosylation, Sialylation and Galactosylation were obtained for each glycosylation site of the molecule (Asn52 and Asn78 for the α subunit, Asn7 and Asn24 for the β subunit) by grouping glycans according to the function related to structural features as shown in Table 3.2.

*Table 3.20 N-Glycan distribution at Asn 52 of FSH a - subunit. (N.D.: not detected)*

| Asn52 - α N-Glycan site | | | | | | |
|---|---|---|---|---|---|---|
| **Glycans** | **Acidic pH** | | | **Basic pH** | | |
| | **Batch1** | **Batch2** | **Batch3** | **Batch1** | **Batch2** | **Batch3** |
| **A2G2** | 2.6 | 2.8 | 2.8 | 1.1 | 1.2 | 1.2 |
| **A2G2S1** | 27.9 | 27.8 | 27.8 | 21.6 | 21.3 | 21.7 |
| **A2G2S2** | 51.6 | 51.0 | 50.3 | 58.7 | 58.4 | 58.5 |
| **A2G2S2_1Ac** | 2.0 | 2.1 | 2.0 | 1.5 | 1.6 | 1.7 |
| **A3G3S1** | 1.1 | 1.3 | 1.3 | 0.6 | 0.7 | 0.7 |
| **A3G3S2** | 5.7 | 5.7 | 5.7 | 4.6 | 4.7 | 4.5 |
| **A3G3S3** | 6.6 | 6.7 | 7.1 | 8.9 | 9.0 | 8.8 |
| **A4G4S3** | 0.7 | 0.8 | 1.0 | 0.7 | 0.9 | 0.8 |
| **A4G4S4** | 0.3 | 0.3 | 0.4 | 0.5 | 0.6 | 0.5 |
| **FA2G2S2** | 1.5 | 1.6 | 1.6 | 1.6 | 1.7 | 1.6 |

*Table 3.21 N-Glycan distribution at Asn 78 of FSH a - subunit. (N.D.: not detected)*

| Asn78 - α N-Glycan site | | | | | | |
|---|---|---|---|---|---|---|
| **Glycans** | **Acidic pH** | | | **Basic pH** | | |
| | **Batch1** | **Batch2** | **Batch3** | **Batch1** | **Batch2** | **Batch3** |
| **A2G2** | 2.0 | 2.2 | 2.2 | 0.7 | 0.8 | 0.8 |
| **A2G2S1** | 24.6 | 24.9 | 25.2 | 18.4 | 17.9 | 18.2 |
| **A2G2S2** | 55.9 | 55.0 | 54.2 | 65.6 | 65.4 | 65.2 |

| | | | | | | |
|---|---|---|---|---|---|---|
| **A2G2S2_1Ac** | 2.1 | 2.1 | 2.0 | 2.0 | 2.0 | 2.2 |
| **A2G2S2_2Ac** | 3.7 | 4.0 | 3.9 | 0.6 | 0.7 | 0.8 |
| **A3G3S1** | 1.3 | 1.4 | 1.5 | 0.8 | 0.9 | 0.8 |
| **A3G3S2** | 4.5 | 4.5 | 4.6 | 4.2 | 4.4 | 4.2 |
| **A3G3S3** | 4.1 | 4.1 | 4.3 | 5.7 | 5.7 | 5.6 |
| **A4G4S2** | 0.2 | 0.2 | 0.2 | 0.1 | 0.1 | 0.1 |
| **A4G4S3** | 0.2 | 0.3 | 0.3 | 0.3 | 0.4 | 0.3 |
| **FA2G2S2** | 1.3 | 1.4 | 1.5 | 1.6 | 1.7 | 1.7 |

*Table 3.22 N-Glycan distribution at Asn 7 of FSH β - subunit. (N.D.: not detected)*

| Asn7 - β N-Glycan site | | | | | | |
|---|---|---|---|---|---|---|
| **Glycans** | **Acidic pH** | | | **Basic pH** | | |
| | **Batch1** | **Batch2** | **Batch3** | **Batch1** | **Batch2** | **Batch3** |
| **A3G3S1** | 0.7 | 0.8 | 0.8 | 0.2 | 0.2 | 0.2 |
| **A3G3S2** | 4.0 | 4.0 | 4.0 | 2.0 | 2.1 | 2.1 |
| **A3G3S3** | 8.7 | 8.6 | 8.5 | 10.0 | 9.8 | 10.0 |
| **A4G4S2** | 1.1 | 1.1 | 1.1 | 0.4 | 0.4 | 0.4 |
| **A4G4S3** | 4.0 | 4.1 | 4.0 | 3.4 | 3.5 | 3.4 |
| **A4G4S4** | 2.8 | 2.8 | 2.9 | 3.8 | 3.8 | 3.8 |
| **A5G5S4** | 0.4 | 0.4 | 0.5 | 0.5 | 0.6 | 0.5 |
| **FA2G2S1** | 1.7 | 1.9 | 1.9 | 0.7 | 0.8 | 0.8 |
| **FA2G2S2** | 5.1 | 5.3 | 5.4 | 5.2 | 5.2 | 5.4 |
| **FA3G3S1** | 2.0 | 2.2 | 2.2 | 0.5 | 0.6 | 0.6 |
| **FA3G3S2** | 12.5 | 12.6 | 12.5 | 6.9 | 6.7 | 7.0 |
| **FA3G3S3** | 27.1 | 26.5 | 26.0 | 32.5 | 31.3 | 31.4 |
| **FA3G3S3_1Ac** | 1.5 | 1.5 | 1.5 | 1.6 | 1.5 | 1.6 |
| **FA3G3S3_2Ac** | 1.4 | 1.5 | 1.4 | 0.2 | 0.2 | 0.2 |
| **FA4G4S3** | 8.0 | 8.2 | 8.0 | 8.8 | 9.1 | 9.2 |
| **FA4G4S4** | 12.6 | 12.2 | 12.4 | 16.9 | 17.2 | 16.7 |
| **FA4G4S4_1Ac** | 0.9 | 0.9 | 1.0 | 1.0 | 1.1 | 1.1 |
| **FA4G4S4_2Ac** | 0.2 | 0.2 | 0.2 | N.D. | N.D. | N.D. |
| **FA5G5S3** | 2.6 | 2.7 | 2.8 | 1.7 | 2.0 | 1.8 |
| **FA5G5S4** | 2.7 | 2.7 | 2.9 | 3.5 | 3.7 | 3.6 |

*Table 3.23 N-Glycan distribution at Asn 24 of FSH β - subunit. (N.D.: not detected)*

| Asn24 - β N-Glycan site | | | | | | |
|---|---|---|---|---|---|---|
| **Glycans** | **Acidic pH** | | | **Basic pH** | | |
| | **Batch1** | **Batch2** | **Batch3** | **Batch1** | **Batch2** | **Batch3** |
| **A2G2S1** | 0.1 | 0.1 | 0.1 | N.D. | N.D. | N.D. |
| **A2G2S2** | 0.1 | 0.1 | 0.1 | 0.1 | 0.2 | 0.2 |
| **FA2G2** | 1.2 | 1.3 | 1.3 | 0.4 | 0.4 | 0.4 |
| **FA2G2S1** | 18.5 | 19.3 | 19.4 | 10.6 | 10.6 | 10.8 |
| **FA2G2S1_1Ac** | 0.3 | 0.3 | 0.3 | 0.1 | 0.1 | 0.1 |
| **FA2G2S2** | 57.9 | 55.4 | 56.3 | 67.1 | 67.6 | 67.1 |
| **FA2G2S2_1Ac** | 2.9 | 3.1 | 2.9 | 3.0 | 3.0 | 3.2 |
| **FA2G2S2_2Ac** | 3.8 | 4.1 | 3.9 | 0.7 | 0.8 | 0.8 |
| **F2A2G2S2** | 2.0 | 2.0 | 2.0 | 2.4 | 2.3 | 2.3 |
| **FA3G3S2** | 3.3 | 3.3 | 3.2 | 2.1 | 2.1 | 2.1 |
| **FA3G3S3** | 6.0 | 6.1 | 6.1 | 8.7 | 8.5 | 8.5 |
| **FA3G3S3_1Ac** | 0.6 | 0.6 | 0.6 | 0.6 | 0.7 | 0.7 |
| **FA4G2** | 0.4 | 0.7 | 0.6 | 0.3 | 0.3 | 0.3 |
| **FA4G2S1** | 1.1 | 1.6 | 1.3 | 1.5 | 1.1 | 1.1 |
| **FA4G4S3** | 1.0 | 1.1 | 1.1 | 0.9 | 1.0 | 1.0 |
| **FA4G4S4** | 0.7 | 0.7 | 0.8 | 1.2 | 1.2 | 1.2 |
| **FA5G5S4** | 0.1 | 0.1 | 0.1 | 0.1 | 0.2 | 0.1 |
| **A2G2S1** | 0.1 | 0.1 | 0.1 | N.D. | N.D. | N.D. |
| **A2G2S2** | 0.1 | 0.1 | 0.1 | 0.1 | 0.2 | 0.2 |

## Grouping: Fucosylation

The table below shows the relative abundance (%) of non fucosylated (a-fucosylated) and fucosylated species for each glycosylation site.

*Table 3.24 Grouping of fucosylation of pH Acid/Basic stressed variants for each N-glycosylation site. (N.D.: not detected)*

| Asn52 - α N-Glycan site | | | | | | |
|---|---|---|---|---|---|---|
| **Fucosylation** | **T0** | **Acidic pH** | | | **Basic pH** | | |
| | **Range** | **Batch1** | **Batch2** | **Batch3** | **Batch1** | **Batch2** | **Batch3** |
| a-fucosylated | 98.4-98.5 | 98.5 | 98.4 | 98.4 | 98.4 | 98.3 | 98.4 |
| Fucosylated | 1.5-1-6 | 1.5 | 1.6 | 1.6 | 1.6 | 1.7 | 1.6 |
| Asn78 - α N-Glycan site | | | | | | |
| **Fucosylation** | **T0** | **Acidic pH** | | | **Basic pH** | | |

|  | Range | Batch1 | Batch2 | Batch3 | Batch1 | Batch2 | Batch3 |
|---|---|---|---|---|---|---|---|
| a-fucosylated | 98.5-98.6 | 98.7 | 98.6 | 98.5 | 98.4 | 98.3 | 98.3 |
| Fucosylated | 1.4-1.5 | 1.3 | 1.4 | 1.5 | 1.6 | 1.7 | 1.7 |

| Asn7 - b N-Glycan site | | | | | | | |
|---|---|---|---|---|---|---|---|
| **Fucosylation** | **T0** | **Acidic pH** | | | **Basic pH** | | |
|  | **Range** | **Batch1** | **Batch2** | **Batch3** | **Batch1** | **Batch2** | **Batch3** |
| a-fucosylated | 18.8-18.9 | 21.6 | 21.7 | 21.9 | 20.4 | 20.4 | 20.4 |
| Fucosylated | 81.1-81.2 | 78.4 | 78.3 | 78.1 | 79.6 | 79.6 | 79.6 |

| Asn24 - b N-Glycan site | | | | | | | |
|---|---|---|---|---|---|---|---|
| **Fucosylation** | **T0** | **Acidic pH** | | | **Basic pH** | | |
|  | **Range** | **Batch1** | **Batch2** | **Batch3** | **Batch1** | **Batch2** | **Batch3** |
| a-fucosylated | 0.2-0.2 | 0.2 | 0.2 | 0.2 | 0.2 | 0.2 | 0.2 |
| Fucosylated | 99.8-98.8 | 97.8 | 97.8 | 97.8 | 97.4 | 97.5 | 97.5 |
| di-fucosylated | 2.0-2.1 | 2.0 | 2.0 | 2.0 | 2.4 | 2.3 | 2.3 |

pH stress appears to have no significant effect on the molecule's fucosylation as shown in Figure 3.11.
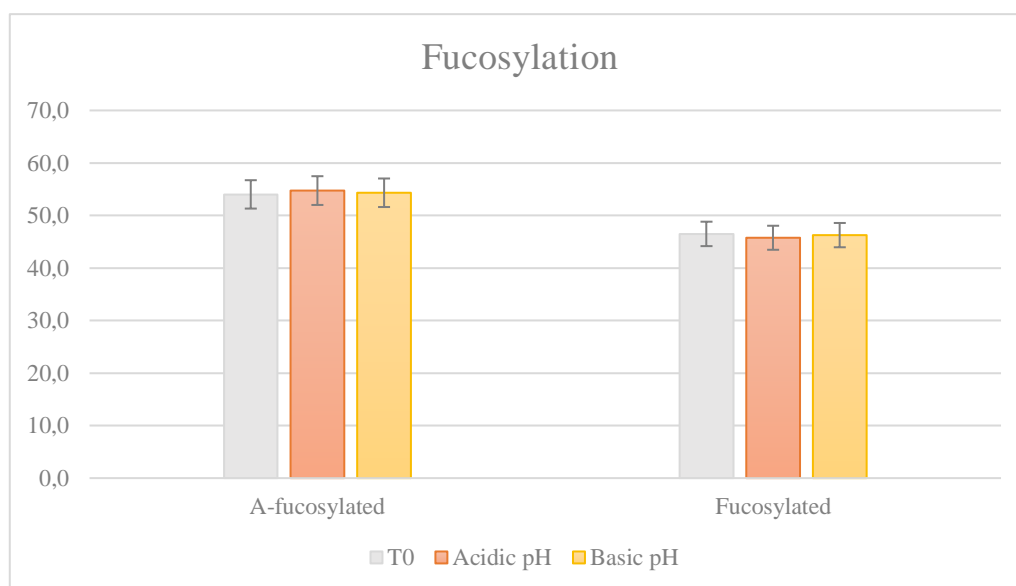


*Figure 3.11 Grouping of fucosylation of pH Acid/Basic stressed variants in overall molecule. The A-fucosylated and Fucosylated values of each variant is shown as an average value over the three batches*

**Grouping: Sialylation**

The table below shows the relative abundance (%) of the sialylation degree (non, mono-, di-, tri- and tetra-sialylated species) for each glycosylation site.

*Table 3.25 Grouping of sialylations of pH Acid/Basic stressed variants for each N-glycosylation site.*

| Asn52 - α N-Glycan site | | | | | | | |
|---|---|---|---|---|---|---|---|
| **Sialylation** | **T0** | **Acidic pH** | | | **Basic pH** | | |
| | **Range** | **Batch1** | **Batch2** | **Batch3** | **Batch1** | **Batch2** | **Batch3** |
| Non sialylated | 1.1-1.2 | 2.6 | 2.8 | 2.8 | 1.1 | 1.2 | 1.2 |
| Mono-sialylated | 22.0-22.9 | 29.0 | 29.1 | 29.1 | 22.2 | 22.0 | 22.4 |
| Di-sialylated | 66.0-67.4 | 60.8 | 60.3 | 59.7 | 66.5 | 66.4 | 66.3 |
| Tri-sialylated | 8.7-9.4 | 7.4 | 7.5 | 8.0 | 9.6 | 9.8 | 9.6 |
| Tetra-sialylated | 0.4-0.6 | 0.3 | 0.3 | 0.4 | 0.5 | 0.6 | 0.5 |
| Acetylated | 2.0-2.1 | 2.0 | 2.1 | 2.0 | 1.5 | 1.6 | 1.7 |
| Asn78 - α N-Glycan site | | | | | | | |
| **Sialylation** | **T0** | **Acidic pH** | | | **Basic pH** | | |
| | **Range** | **Batch1** | **Batch2** | **Batch3** | **Batch1** | **Batch2** | **Batch3** |
| Non sialylated | 0.7-0.8 | 2.0 | 2.2 | 2.2 | 0.7 | 0.8 | 0.8 |
| Mono-sialylated | 19.1-19.4 | 25.9 | 26.3 | 26.7 | 19.2 | 18.8 | 19.1 |
| Di-sialylated | 73.9-74.7 | 67.7 | 67.2 | 66.4 | 74.1 | 74.4 | 74.2 |
| Tri-sialylated | 5.3-5.8 | 4.3 | 4.3 | 4.6 | 6.0 | 6.0 | 5.9 |
| Acetylated | 6.4-6.5 | 5.8 | 6.0 | 5.9 | 2.6 | 2.8 | 3.0 |
| Asn7 - β N-Glycan site | | | | | | | |
| **Sialylation** | **T0** | **Acidic pH** | | | **Basic pH** | | |
| | **Range** | **Batch1** | **Batch2** | **Batch3** | **Batch1** | **Batch2** | **Batch3** |
| Non sialylated | N.D. | N.D. | N.D. | N.D. | N.D. | N.D. | N.D. |
| Mono-sialylated | 1.5-1.6 | 4.4 | 4.9 | 4.8 | 1.4 | 1.6 | 1.7 |
| Di-sialylated | 14.1-14.2 | 22.7 | 23.0 | 23.0 | 14.5 | 14.4 | 14.9 |
| Tri-sialylated | 59.2-59.5 | 53.3 | 53.0 | 52.3 | 58.1 | 57.4 | 57.6 |
| Tetra-sialylated | 24.8-25.2 | 19.6 | 19.1 | 19.9 | 25.9 | 26.6 | 25.7 |
| Acetylated | 4.5-4.7 | 3.9 | 4.1 | 4.1 | 2.8 | 2.9 | 2.9 |
| Asn24 - β N-Glycan site | | | | | | | |
| **Sialylation** | **T0** | **Acidic pH** | | | **Basic pH** | | |
| | **Range** | **Batch1** | **Batch2** | **Batch3** | **Batch1** | **Batch2** | **Batch3** |
| Non sialylated | 0.7-0.8 | 1.7 | 2.0 | 1.9 | 0.7 | 0.7 | 0.7 |
| Mono-sialylated | 12.6-13.3 | 19.9 | 21.3 | 21.1 | 12.2 | 11.9 | 12.0 |
| Di-sialylated | 75.1-76.3 | 70.0 | 68.1 | 68.4 | 75.5 | 75.9 | 75.8 |
| Tri-sialylated | 9.1-9.5 | 7.6 | 7.8 | 7.8 | 10.3 | 10.2 | 10.2 |
| Tetra-sialylated | 1.2-1.4 | 0.8 | 0.8 | 0.9 | 1.3 | 1.3 | 1.3 |
| Acetylated | 9.0-9.2 | 7.5 | 8.2 | 7.7 | 4.5 | 4.5 | 4.8 |

As done for other variants, the sialylation index (S-index) was calculated for each N-glycosylation site to better visualize the sialylation trend on the entire molecule in the two pH stressed variants in comparison with untreated sample:

$$\text{S-index} = \sum \frac{\% \ Glycans \ relative \ abundance * N° \ of \ sialic \ acid \ content}{100}$$

*Table 3.26 S-index calculated on each N-glycan site*

| S-index | T0 | Acidic pH | | | Basic pH | | |
|---|---|---|---|---|---|---|---|
| | Range | Batch1 | Batch2 | Batch3 | Batch1 | Batch2 | Batch3 |
| Asn52 - α | 1.8-1.9 | 1.7 | 1.7 | 1.7 | 1.9 | 1.9 | 1.9 |
| Asn78 - α | 1.8-1.9 | 1.7 | 1.7 | 1.7 | 1.9 | 1.9 | 1.8 |
| Asn7 - β | 3.1-3.1 | 2.9 | 2.9 | 2.9 | 3.1 | 3.1 | 3.1 |
| Asn24 - β | 1.9-1.9 | 1.9 | 1.8 | 1.8 | 2.0 | 2.0 | 2.0 |
| All sites | 2.2-2.2 | 2.1 | 2.0 | 2.0 | 2.2 | 2.2 | 2.2 |

The Figure 3.12 shows the average S-indexes of all N-glycosylation sites of the molecule for both pH Acid/basic stressed variants compared to the reference standard one.
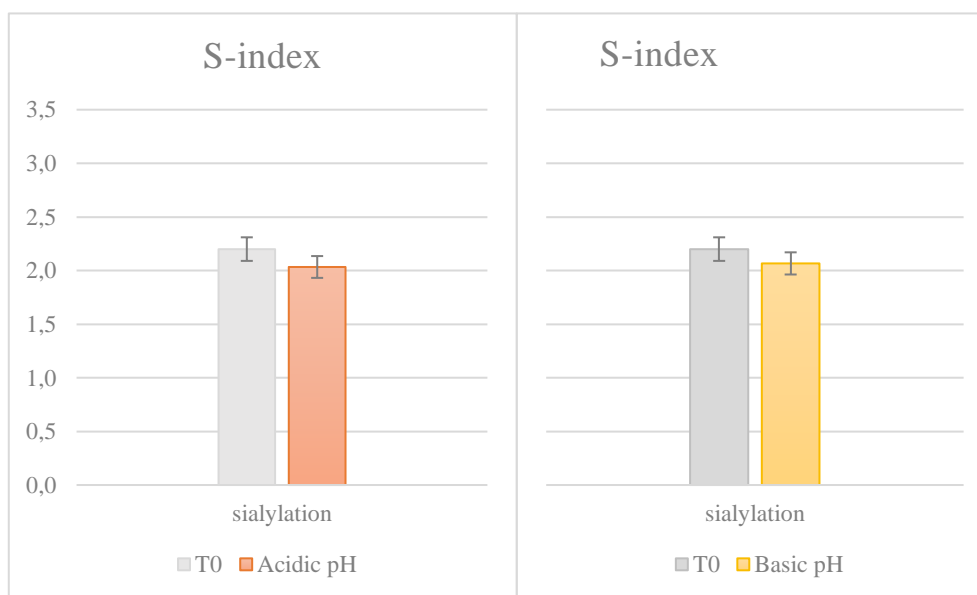


*Figure 3.12 S-Index of Acid/Basic pH stressed variants in overall molecule.*

The sialic acid content is slightly affected by pH. In both pH stressed variants, the S-index is lower than untreated one. Coherently with the decrease of sialic

acid, the level of acetylation is also decreased in the stressed variants as we can see from the Ac-index values. Acetylation on sialic acid is highly stable at pH 3.0–10.0, but nevertheless small variations in acetylation are observed after acidic-basic pH treatment. In particular, the decrease in acetylation of Basic pH (pH=9.0) stressed variant is greater than that observed in the Acidic pH (pH=3.0) variant as, in such mild pH conditions, O-acetylated residues was more susceptible to alkaline treatment [25, 26, 27] (Figure 3.13).

Acetylated index (Ac-index) is calculated.

$$\text{Ac-index} = \sum \frac{\% \: Glycans \: relative \: abundance * N° \: of \: sialic \: acid \: acetylation \: content}{100}$$

*Table 3.27 Ac-index calculated on each N-glycan site*

| Ac. index | T0 | Acidic pH | | | Basic pH | | |
|---|---|---|---|---|---|---|---|
| | Range | Batch1 | Batch2 | Batch3 | Batch1 | Batch2 | Batch3 |
| Asn52 - α | 2.0-2.1 | 2.0 | 2.1 | 2.0 | 1.5 | 1.6 | 1.7 |
| Asn78 - α | 10.6-10.8 | 9.5 | 10.0 | 9.8 | 3.2 | 3.5 | 3.9 |
| Asn7 - β | 5.7-6.2 | 5.5 | 5.4 | 5.3 | 3.0 | 3.2 | 3.2 |
| Asn24 - β | 13.8-14.0 | 11.3 | 12.3 | 11.6 | 5.2 | 5.2 | 5.7 |
| All sites | 8.1-8.2 | 7.1 | 7.4 | 7.2 | 3.2 | 3.4 | 3.6 |

The Figure 3.13 shows the average Ac-indexes of all N-glycosylation sites of the molecule for both pH Acid/basic stressed variants compared to each other's and to the untreated one.

*Figure 3.13 Ac-Index of Acid/Basic pH stressed variants in overall molecule.*

**Grouping: Galactosylation**

In the table below only the species containing exposed terminal galactose were grouped per number of exposed galactose content (G1, G2, G3 and G4) and their relative abundance (%) is reported hereafter.

*Table 3.28 Grouping of galactosylation of pH Acid/Basic stressed variants for each N-glycosite.* (N.D.: not detected)

| Asn52 - α N-Glycan site | | | | | | | |
|---|---|---|---|---|---|---|---|
| Galactosylation | T0 | Acidic pH | | | Basic pH | | |
| | Range | Batch1 | Batch2 | Batch3 | Batch1 | Batch2 | Batch3 |
| G1 | N.D. | N.D. | N.D. | N.D. | N.D. | N.D. | N.D. |
| G2 | 1.1-1.2 | 2.6 | 2.8 | 2.8 | 1.1 | 1.2 | 1.2 |
| **Asn78 - α N-Glycan site** | | | | | | | |
| Galactosylation | T0 | Acidic pH | | | Basic pH | | |
| | Range | Batch1 | Batch2 | Batch3 | Batch1 | Batch2 | Batch3 |
| G1 | N.D. | N.D. | N.D. | N.D. | N.D. | N.D. | N.D. |
| G2 | 0.7-0.8 | 2.0 | 2.2 | 2.2 | 0.7 | 0.8 | 0.8 |
| G3 | N.D. | N.D. | N.D. | N.D. | N.D. | N.D. | N.D. |
| **Asn7 - b N-Glycan site** | | | | | | | |
| Galactosylation | T0 | Acidic pH | | | Basic pH | | |
| | Range | Batch1 | Batch2 | Batch3 | Batch1 | Batch2 | Batch3 |
| G1 | N.D. | N.D. | N.D. | N.D. | N.D. | N.D. | N.D. |
| G2 | N.D. | N.D. | N.D. | N.D. | N.D. | N.D. | N.D. |
| G3 | N.D. | N.D. | N.D. | N.D. | N.D. | N.D. | N.D. |
| G4 | N.D. | N.D. | N.D. | N.D. | N.D. | N.D. | N.D. |
| **Asn24 - b N-Glycan site** | | | | | | | |
| Galactosylation | T0 | Acidic pH | | | Basic pH | | |
| | Average | Batch1 | Batch2 | Batch3 | Batch1 | Batch2 | Batch3 |
| G1 | N.D. | N.D. | N.D. | N.D. | N.D. | N.D. | N.D. |
| G2 | 0.8 | 1.7 | 2.0 | 1.9 | 0.7 | 0.7 | 0.7 |

The galactosylated species, especially the one containing two exposed galactoses (G2), increase in the Acid pH stressed consistently with the lower level of sialylation detected for this variant. In fact, the removal of sialic acid by acid hydrolysis exposes the saccharide residue bound to it, which in the case of N-glycans is galactose.
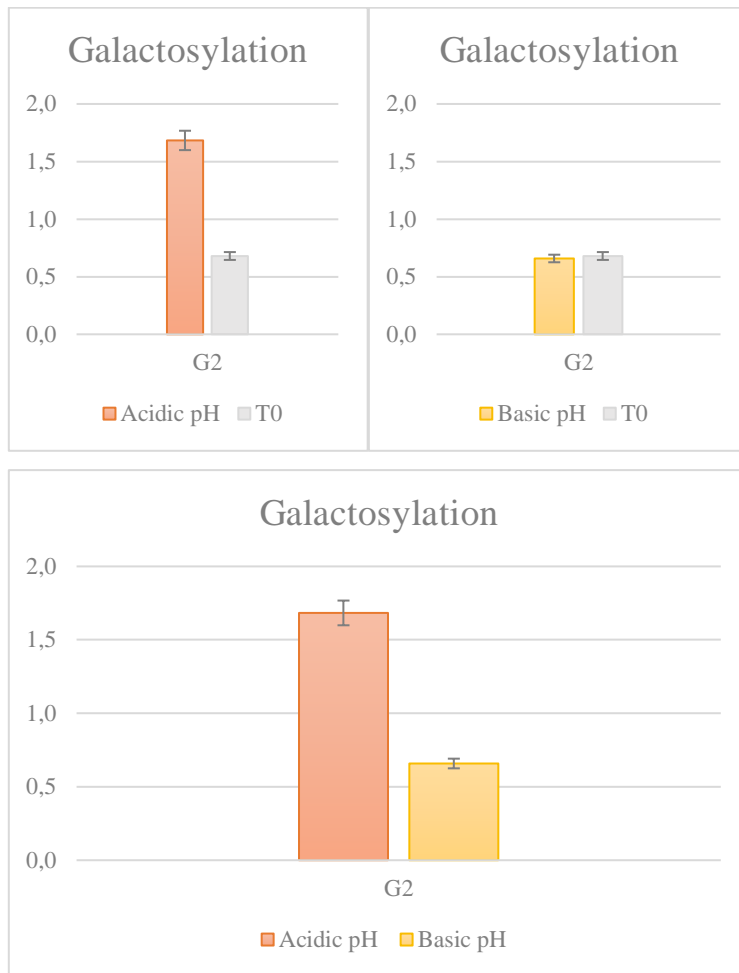
*Figure 3.14 Grouping of galactosylation of pH Acidic/Basic stressed variants in overall molecule. The G2 values of each variant is shown as an average value over the three batches.*

### 3.3.5 De-sialylated – Glycopeptides mapping

LC-MS/MS data were analysed to obtain the distribution of glycans based on the N-glycosylation site of De-sialylated variant compared with that obtained from untreated one. Glycopeptide mapping profiles were processed semi-automatically by Expressionist software. The relative abundance of detected species was calculated and reported in the subsequent tables; Fucosylation, Sialylation and Galactosylation grouping were obtained for each glycosylation site of the molecule (Asn52 and Asn78 for the α subunit, Asn7 and Asn24 for the β subunit) as described in Table 3.2.

*Table 3.29 N-Glycan distribution at Asn 52 of FSH α - subunit*

| Asn52 - α N-Glycan site | | | |
|---|---|---|---|
| **Glycans** | **De-sialylated** | | |
| | **Batch1** | **Batch2** | **Batch3** |
| A2G2 | 77.8 | 75.9 | 76.0 |
| A2G2S1 | 0.2 | 0.1 | 0.1 |
| A2G2S2 | 0.9 | 0.9 | 0.8 |
| A4G4 | 3.1 | 3.4 | 3.3 |
| FA2G2 | 1.6 | 1.8 | 1.9 |
| A3G3 | 16.4 | 17.8 | 17.8 |

*Table 3.30 N-Glycan distribution at Asn 78 of FSH a - subunit*

| Asn78 - α N-Glycan site | | | |
|---|---|---|---|
| **Glycans** | **De-sialylated** | | |
| | **Batch1** | **Batch2** | **Batch3** |
| A2G1 | 0.1 | 0.1 | 0.1 |
| A2G2 | 84.2 | 83.3 | 82.1 |
| A3G3 | 12.6 | 13.3 | 14.1 |
| A4G4 | 1.1 | 1.2 | 1.4 |
| FA2G2 | 1.9 | 2.0 | 2.2 |

*Table 3.31 N-Glycan distribution at Asn 7 of FSH β - subunit.* (N.D.: not detected)

| Asn7 - β N-Glycan site | | | |
|---|---|---|---|
| **Glycans** | **De-sialylated** | | |
| | **Batch1** | **Batch2** | **Batch3** |
| A3G3 | 11.6 | 11.4 | 11.6 |
| FA2G2S1 | 0.1 | 0.1 | 0.1 |
| FA2G2S2 | 0.1 | N.D. | N.D. |
| FA3G3S3 | 0.1 | 0.1 | 0.1 |
| FA4G4 | 35.9 | 36.0 | 36.0 |
| FA2G2 | 5.0 | 5.3 | 5.3 |
| FA3G3 | 38.0 | 38.3 | 37.4 |
| A2G2 | 0.8 | 0.8 | 0.8 |
| A4G4 | 8.4 | 8.0 | 8.7 |

*Table 3.32 N-Glycan distribution at Asn 24 of FSH β - subunit*

| Asn24 - β N-Glycan site | | | |
|---|---|---|---|
| **Glycans** | **De-sialylated** | | |
| | **Batch1** | **Batch2** | **Batch3** |
| FA2G2 | 69.4 | 74.0 | 77.0 |
| FA3G3 | 23.6 | 19.8 | 17.5 |
| FA4G4 | 6.5 | 5.8 | 5.1 |
| A2G2 | 0.4 | 0.4 | 0.3 |

## Grouping: Fucosylation

The table below shows the relative abundance (%) of non fucosylated (a-fucosylated) and fucosylated species for each glycosylation site.

*Table 3.33 Grouping of fucosylation of De-sialylated variants for each N-glycosylation site.* (N.D.: not detected)

| Asn52 - α N-Glycan site | | | | |
|---|---|---|---|---|
| **Fucosylation** | **T0** | **De-sialylated** | | |
| | **Range** | **Batch1** | **Batch2** | **Batch3** |
| a-fucosylated | 98.4-98.5 | 98.4 | 98.2 | 98.1 |
| Fucosylated | 1.5-1-6 | 1.6 | 1.8 | 1.9 |
| Asn78 - α N-Glycan site | | | | |
| **Fucosylation** | **T0** | **De-sialylated** | | |

| | Range | Batch1 | Batch2 | Batch3 |
|---|---|---|---|---|
| a-fucosylated | 98.5-98.6 | 98.1 | 98.0 | 97.8 |
| Fucosylated | 1.4-1.5 | 1.9 | 2.0 | 2.2 |
| **Asn7 - β N-Glycan site** | | | | |
| **Fucosylation** | **T0** | **De-sialylated** | | |
| | Range | Batch1 | Batch2 | Batch3 |
| a-fucosylated | 18.8-18.9 | 20.8 | 20.2 | 21.1 |
| Fucosylated | 81.1-81.2 | 79.2 | 79.8 | 78.9 |
| **Asn24 - β N-Glycan site** | | | | |
| **Fucosylation** | **T0** | **De-sialylated** | | |
| | Range | Batch1 | Batch2 | Batch3 |
| a-fucosylated | 0.2-0.2 | 0.4 | 0.4 | 0.3 |
| Fucosylated | 99.8-98.8 | 99.6 | 99.6 | 99.7 |
| di-fucosylated | 2.0-2.1 | N.D. | N.D. | N.D. |

Overall, the De-sialylated variant shows a degree of fucosylation, averaged out on 4 glycosylation sites of the molecule, very close to that of the reference sample, as expected, as Sialidase is a specific enzyme for the hydrolysis of sialic acid, hence it has no effect on fucosylation (Figure 3.15).
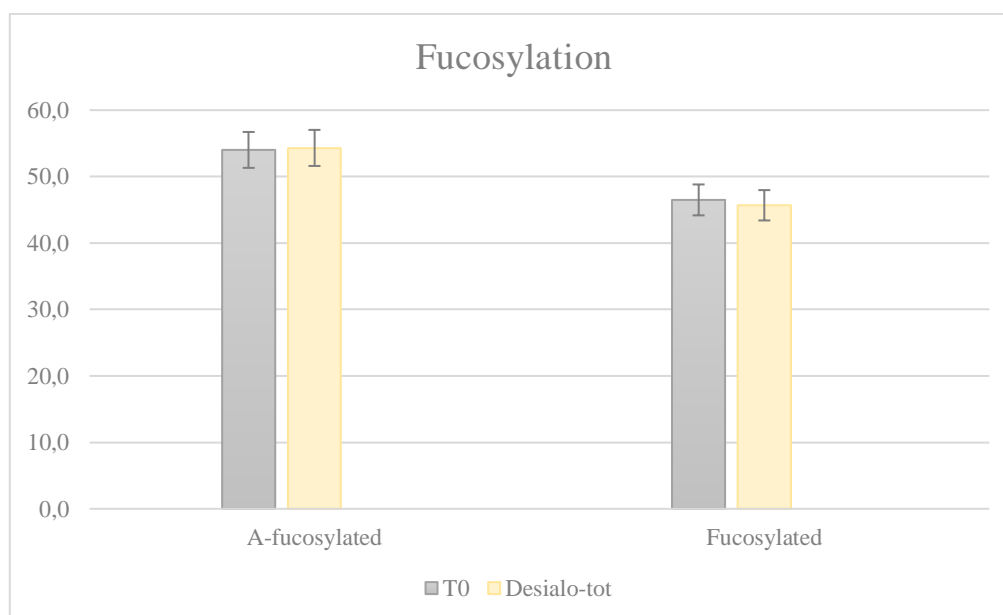


*Figure 3.15 Grouping of fucosylation of De-sialylated variants in overall molecule. The A-fucosylated and Fucosylated values of each variant is shown as an average value over the three batches*

## Grouping: Sialylation

The table below shows the relative abundance (%) of the sialylation degree (mono-, di-, tri- and tetra-sialylated species) for each glycosylation site.

*Table 22 Grouping of sialylations of De-sialylated variants for each N-glycosylation site. (N.D.: not detected)*

| Asn52 - α N-Glycan site | | | | |
|---|---|---|---|---|
| **Sialylation** | **T0** | **De-sialylated** | | |
| | **Range** | **Batch1** | **Batch2** | **Batch3** |
| Non sialylated | 1.1-1.2 | 98.9 | 98.9 | 99.0 |
| Mono-sialylated | 22.0-22.9 | 0.2 | 0.1 | 0.1 |
| Di-sialylated | 66.0-67.4 | 0.9 | 0.9 | 0.8 |
| Tri-sialylated | 8.7-9.4 | N.D. | N.D. | N.D. |
| Tetra-sialylated | 0.4-0.6 | N.D. | N.D. | N.D. |
| Acetylated | 2.0-2.1 | N.D. | N.D. | N.D. |
| Asn78 - α N-Glycan site | | | | |
| **Sialylation** | **T0** | **De-sialylated** | | |
| | **Range** | **Batch1** | **Batch2** | **Batch3** |
| Non sialylated | 0.7-0.8 | 99.9 | 99.9 | 99.9 |
| Mono-sialylated | 19.1-19.4 | N.D. | N.D. | N.D. |
| Di-sialylated | 73.9-74.7 | N.D. | N.D. | N.D. |
| Tri-sialylated | 5.3-5.8 | N.D. | N.D. | N.D. |
| Acetylated | 6.4-6.5 | N.D. | N.D. | N.D. |
| Asn7 - b N-Glycan site | | | | |
| **Sialylation** | **T0** | **De-sialylated** | | |
| | **Range** | **Batch1** | **Batch2** | **Batch3** |
| Non sialylated | N.D. | 99.6 | 99.7 | 99.7 |
| Mono-sialylated | 1.5-1.6 | 0.1 | 0.1 | 0.1 |
| Di-sialylated | 14.1-14.2 | 0.1 | N.D. | N.D. |
| Tri-sialylated | 59.2-59.5 | 0.1 | 0.1 | 0.1 |
| Tetra-sialylated | 24.8-25.2 | N.D. | N.D. | N.D. |
| Acetylated | 4.5-4.7 | N.D. | N.D. | N.D. |
| Asn24 - b N-Glycan site | | | | |
| **Sialylation** | **T0** | **De-sialylated** | | |
| | **Range** | **Batch1** | **Batch2** | **Batch3** |
| Non sialylated | 0.7-0.8 | 100.0 | 100.0 | 100.0 |

| | | | |
|---|---|---|---|---|
| Mono-sialylated | 12.6-13.3 | N.D. | N.D. | N.D. |
| Di-sialylated | 75.1-76.3 | N.D. | N.D. | N.D. |
| Tri-sialylated | 9.1-9.5 | N.D. | N.D. | N.D. |
| Tetra-sialylated | 1.2-1.4 | N.D. | N.D. | N.D. |
| Acetylated | 9.0-9.2 | N.D. | N.D. | N.D. |

As we can see from sialylation grouping table, no species containing sialic acid was found in De-sialylated variants; This data confirms that the generated variant is completely de-sialylated.

## Grouping: Galactosylation

In the table below only the species containing exposed terminal galactose were grouped per number of exposed galactose content (G0, G1, G2, G3 etc.) and their relative abundance (%) is reported hereafter.

*Table 3.34 Grouping of galactosylation of De-sialylated variants for each N-glycosylation site.* (N.D.: not detected)

| Asn52 - α N-Glycan site | | | | |
|---|---|---|---|---|
| **Galactosylation** | **T0** | **De-sialylated** | | |
| | **Range** | **Batch1** | **Batch2** | **Batch3** |
| G1 | N.D. | N.D. | N.D. | N.D. |
| G2 | 1.1-1.2 | 79.4 | 77.7 | 77.9 |
| G3 | N.D. | 16.4 | 17.8 | 17.8 |
| G4 | N.D. | 3.1 | 3.4 | 3.3 |
| **Asn78 - α N-Glycan site** | | | | |
| **Galactosylation** | **T0** | **De-sialylated** | | |
| | **Range** | **Batch1** | **Batch2** | **Batch3** |
| G1 | N.D. | 0.1 | 0.1 | 0.1 |
| G2 | 0.7-0.8 | 86.1 | 85.3 | 84.3 |
| G3 | N.D. | 12.6 | 13.3 | 14.1 |
| G4 | N.D. | 1.1 | 1.2 | 1.4 |
| **Asn7 - b N-Glycan site** | | | | |
| **Galactosylation** | **T0** | **De-sialylated** | | |
| | **Range** | **Batch1** | **Batch2** | **Batch3** |
| G1 | N.D. | N.D. | N.D. | N.D. |
| G2 | N.D. | 5.8 | 6.1 | 6.1 |

| Galactosylation | T0 | De-sialylated | | |
|:---:|:---:|:---:|:---:|:---:|
| | | | | |
| G3 | N.D. | 49.6 | 49.6 | 49.0 |
| G4 | N.D. | 44.3 | 44.0 | 44.6 |
| **Asn24 - b N-Glycan site** | | | | |
| **Galactosylation** | **T0** | **De-sialylated** | | |
| | **Average** | **Batch1** | **Batch2** | **Batch3** |
| G1 | N.D. | N.D. | N.D. | N.D. |
| G2 | 0.7-0.8 | 69.8 | 74.3 | 77.4 |
| G3 | N.D. | 23.6 | 19.8 | 17.5 |
| G4 | N.D. | 6.5 | 5.8 | 5.1 |

The relative abundance of galactosylated species (Figure 3.16) has the same trend as glycan antennarity of FSH molecule (Figure 3.17). G2 is the most abundant species since biantennary glycans are predominant in the molecule. Likewise, G3 relative abundance reflects the abundance of tri-antennary glycans and G4 of tetra-antennary ones.
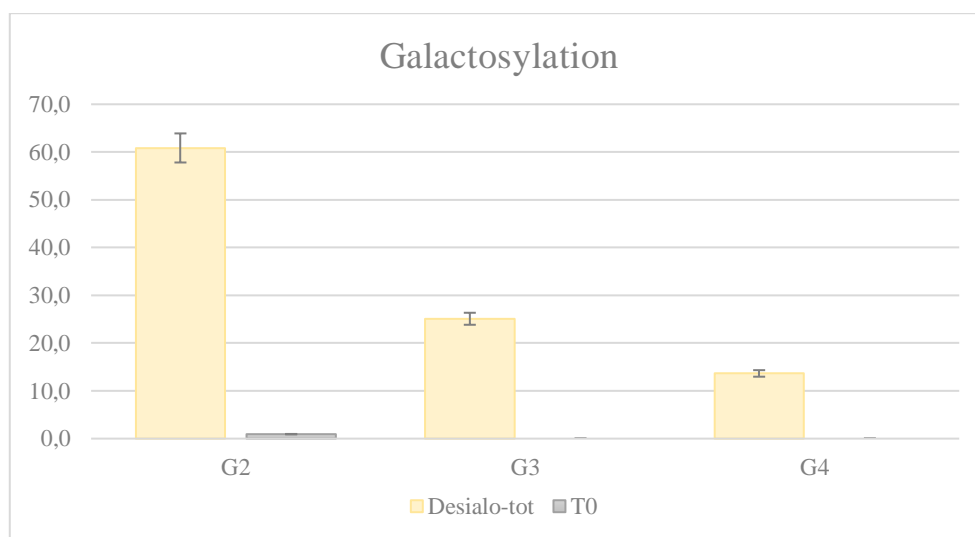


*Figure 3.16 Grouping of galactosylation of De-sialylated variants in overall molecule. The G2, G3, G4 values of each variant is shown as an average value over the three batches.*

135
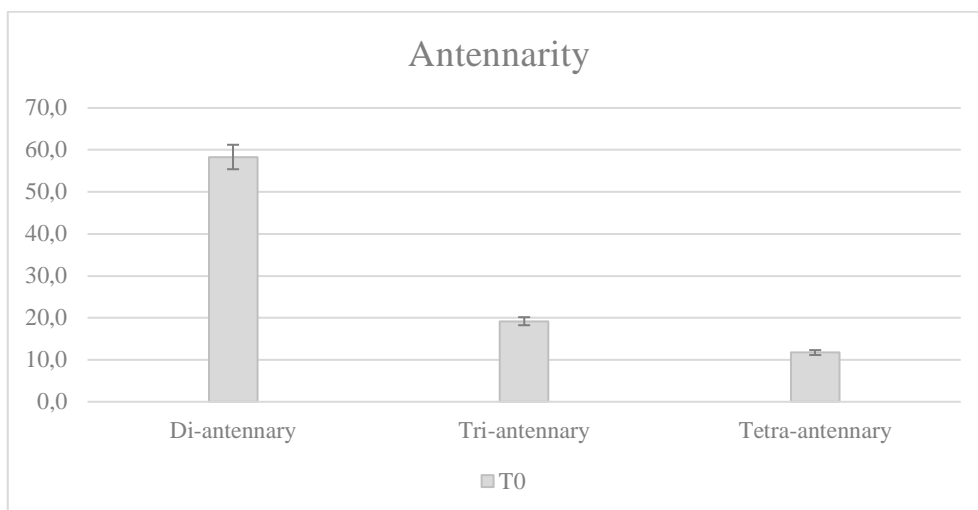
*Figure 3.17 shows the percentage of relative abundance of di-, tri- and tetra-antennary species in the overall molecule of the untreated DS_FSH T0*

### 3.3.6 De-sialylated/de-galactosylated – Glycopeptides mapping

LC-MS/MS data were analysed to obtain the distribution of glycans based on the N-glycosylation site of De-sialylated and de-galactosylated (De-sialo/De-gal) variant. Glycopeptide mapping profiles were processed semi-automatically by Expressionist software. The relative distribution of detected species is reported in the subsequent tables; for each N-glycan species, relative abundance was calculated; moreover, three distinctive groups (Fucosylation, Sialylation and Galactosylation) were obtained for each glycosylation site of the molecule (Asn52 and Asn78 for the α subunit, Asn7 and Asn24 for the β subunit) by grouping glycans according to the function related to structural features as shown in Table 3.2.

*Table 3.35 N-Glycan distribution at Asn 52 of FSH α – subunit.* (N.D.: not detected)

| Asn52 - α N-Glycan site | | | |
|---|---|---|---|
| **Glycans** | **De-sialo/De-gal** | | |
| | **Batch1** | **Batch2** | **Batch3** |
| **A2G2S1** | N.D. | 0.2 | 0.2 |
| **FA2G0** | 1.0 | 1.1 | 1.0 |
| **A2G0** | 84.6 | 84.0 | 78.3 |
| **A2G1** | 9.3 | 9.5 | 16.6 |
| **A3G0** | 2.5 | 2.5 | 2.4 |
| **A4G0** | 1.8 | 1.9 | 1.4 |
| **A4G1** | 0.8 | 0.9 | 0.2 |

*Table 3.36 N-Glycan distribution at Asn 78 of FSH α – subunit. (N.D.: not detected).*

| Asn78 - α N-Glycan site | | | |
|---|---|---|---|
| **Glycans** | **De-sialo/De-gal** | | |
| | **Batch1** | **Batch2** | **Batch3** |
| **A3G0** | 10.4 | 10.8 | 9.6 |
| **A2G0** | 89.2 | 88.8 | 90.0 |
| **A4G0** | 0.1 | 0.1 | N.D. |
| **A4G1** | 0.2 | 0.2 | 0.2 |
| **A4G2** | 0.1 | 0.1 | N.D. |

*Table 3.37 N-Glycan distribution at Asn 7 of FSH β - subunit. (N.D.: not detected)*

| Asn7 - β N-Glycan site | | | |
|---|---|---|---|
| **Glycans** | **De-sialo/De-gal** | | |
| | **Batch1** | **Batch2** | **Batch3** |
| **FA4G4** | 0.1 | N.D. | N.D. |
| **A3G1** | 2.2 | 2.4 | 3.7 |
| **A3G2** | 0.2 | 0.4 | 1.4 |
| **A4G0** | 4.1 | 4.1 | 2.6 |
| **A4G1** | 6.4 | 6.5 | 6.0 |
| **A4G2** | 5.3 | 5.1 | 5.6 |
| **FA2G0** | 0.3 | 0.3 | 0.3 |
| **FA3G0** | 1.2 | 1.2 | 0.7 |
| **FA3G1** | 18.5 | 18.9 | 19.6 |
| **FA4G0** | 19.0 | 19.4 | 16.6 |
| **FA4G1** | 24.9 | 24.9 | 24.3 |
| **FA4G2** | 17.7 | 16.8 | 19.0 |

*Table 3.38 N-Glycan distribution at Asn 24 of FSH β - subunit.*

| Asn24 - β N-Glycan site | | | |
|---|---|---|---|
| **Glycans** | **De-sialo/De-gal** | | |
| | **Batch1** | **Batch2** | **Batch3** |
| **FA4G2** | 0.5 | 0.7 | 0.6 |
| **FA2G0** | 74.6 | 62.1 | 61.1 |
| **FA3G0** | 14.1 | 21.1 | 20.3 |
| **FA3G1** | 5.4 | 7.8 | 9.6 |
| **FA4G0** | 2.5 | 3.9 | 3.6 |
| **FA4G1** | 1.8 | 2.7 | 3.1 |
| **A2G0** | 1.0 | 1.4 | 1.4 |
| **A3G0** | 0.1 | 0.1 | 0.1 |

## Grouping: Fucosylation

The table below shows the relative abundance (%) of non fucosylated (a-fucosylated) and fucosylated species for each glycosylation site.

*Table 3.39 Grouping of fucosylation of De-sialo/De-gal variant for each N-glycosylation site.* (N.D.: not detected)

| Asn52 - α N-Glycan site | | | | |
|---|---|---|---|---|
| **Fucosylation** | **T0** | **De-sialo/De-gal** | | |
| | **Range** | **Batch1** | **Batch2** | **Batch3** |
| a-fucosylated | 98.4-98.5 | 99.0 | 99.0 | 99.0 |
| Fucosylated | 1.5-1-6 | 1.0 | 1.1 | 1.0 |
| **Asn78 - α N-Glycan site** | | | | |
| **Fucosylation** | **T0** | **De-sialo/De-gal** | | |
| | **Range** | **Batch1** | **Batch2** | **Batch3** |
| a-fucosylated | 98.5-98.6 | 100.0 | 100.0 | 100.0 |
| Fucosylated | 1.4-1.5 | N.D. | N.D. | N.D. |
| **Asn7 - β N-Glycan site** | | | | |
| **Fucosylation** | **T0** | **De-sialo/De-gal** | | |
| | **Range** | **Batch1** | **Batch2** | **Batch3** |
| a-fucosylated | 18.8-18.9 | 18.3 | 18.4 | 19.4 |
| Fucosylated | 81.1-81.2 | 81.7 | 81.6 | 80.6 |
| **Asn24 - β N-Glycan site** | | | | |
| **Fucosylation** | **T0** | **De-sialo/De-gal** | | |
| | **Range** | **Batch1** | **Batch2** | **Batch3** |
| a-fucosylated | 0.2-0.2 | 1.0 | 1.6 | 1.5 |
| Fucosylated | 99.8-98.8 | 99.0 | 98.4 | 98.5 |
| di-fucosylated | 2.0-2.1 | N.D. | N.D. | N.D. |

Sialidase and galactosidase are specific enzymes for the hydrolysis of sialic acid and galactose, therefore they have no effect on fucosylation (Figure 3.18).
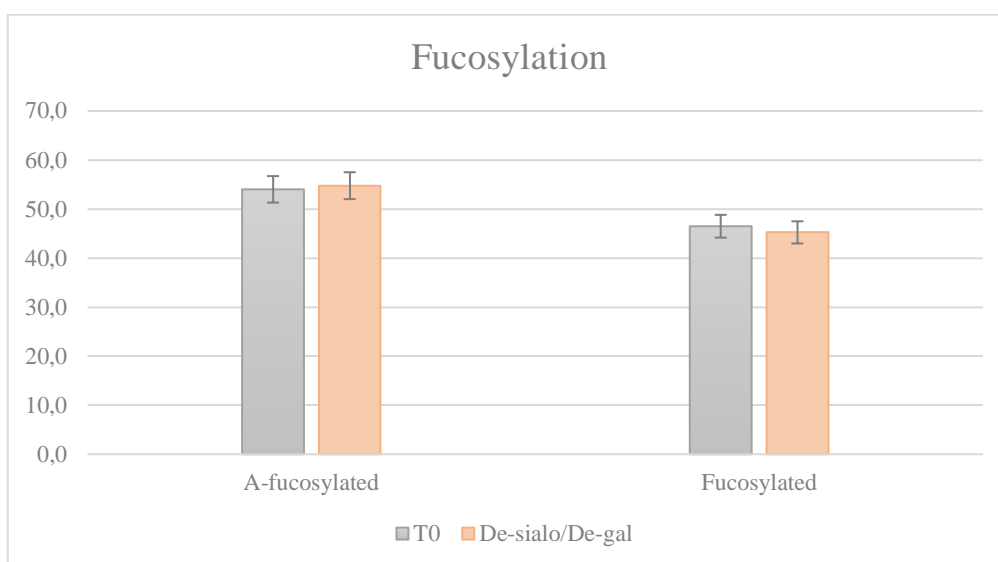
*Figure 3.18 Grouping of fucosylation of De-sialo/De-gal variant in overall molecule. The A-fucosylated and Fucosylated values of each variant is shown as an average value over the three batches*

## Grouping: Sialylation

The table below shows the relative abundance (%) of the sialylation degree (mono-, di-, tri- and tetra-sialylated species) for each glycosylation site.

*Table 3.40 Grouping of sialylations of De-sialo/De-gal variant for each N-glycosylation site.* (N.D.: not detected)

| Asn52 - α N-Glycan site | | | | |
|---|---|---|---|---|
| **Sialylation** | **T0** | **De-sialo/De-gal** | | |
| | **Range** | **Batch1** | **Batch2** | **Batch3** |
| Non sialylated | 1.1-1.2 | 100.0 | 99.8 | 99.8 |
| Mono-sialylated | 22.0-22.9 | N.D. | 0.2 | 0.2 |
| Di-sialylated | 66.0-67.4 | N.D. | N.D. | N.D. |
| Tri-sialylated | 8.7-9.4 | N.D. | N.D. | N.D. |
| Tetra-sialylated | 0.4-0.6 | N.D. | N.D. | N.D. |
| Acetylated | 2.0-2.1 | N.D. | N.D. | N.D. |
| Asn78 - α N-Glycan site | | | | |
| **Sialylation** | **T0** | **De-sialo/De-gal** | | |
| | **Range** | **Batch1** | **Batch2** | **Batch3** |
| Non sialylated | 0.7-0.8 | 100.0 | 100.0 | 99.8 |
| Mono-sialylated | 19.1-19.4 | N.D. | N.D. | 0.2 |
| Di-sialylated | 73.9-74.7 | N.D. | N.D. | N.D. |

| | | | | |
|---|---|---|---|---|
| Tri-sialylated | 5.3-5.8 | N.D. | N.D. | N.D. |
| Acetylated | 6.4-6.5 | N.D. | N.D. | N.D. |

| Asn7 - β N-Glycan site | | | | |
|---|---|---|---|---|
| **Sialylation** | **T0** | **De-sialo/De-gal** | | |
| | **Range** | **Batch1** | **Batch2** | **Batch3** |
| Non sialylated | N.D. | 100.0 | 100.0 | 99.9 |
| Mono-sialylated | 1.5-1.6 | N.D. | N.D. | N.D. |
| Di-sialylated | 14.1-14.2 | N.D. | N.D. | N.D. |
| Tri-sialylated | 59.2-59.5 | N.D. | N.D. | N.D. |
| Tetra-sialylated | 24.8-25.2 | N.D. | N.D. | N.D. |
| Acetylated | 4.5-4.7 | N.D. | N.D. | N.D. |

| Asn24 - β N-Glycan site | | | | |
|---|---|---|---|---|
| **Sialylation** | **T0** | **De-sialo/De-gal** | | |
| | **Range** | **Batch1** | **Batch2** | **Batch3** |
| Non sialylated | 0.7-0.8 | 100.0 | 100.0 | 100.0 |
| Mono-sialylated | 12.6-13.3 | N.D. | N.D. | N.D. |
| Di-sialylated | 75.1-76.3 | N.D. | N.D. | N.D. |
| Tri-sialylated | 9.1-9.5 | N.D. | N.D. | N.D. |
| Tetra-sialylated | 1.2-1.4 | N.D. | N.D. | N.D. |
| Acetylated | 9.0-9.2 | N.D. | N.D. | N.D. |

As we can see from grouping table, no species containing sialic acid was found in De-sialo/De-gal; This data confirms that the generated variant is completely free of sialic acid.

## Grouping: Galactosylation

The table below shows the relative abundance (%) of species carrying expose galactose residues (G0, G1, G2, G3 and G4) for each glycosylation site.

*Table 3.41 Grouping of galactosylation of De-sialo/De-gal variant for each N-glycosylation site.* (N.D.: not detected)

| Asn52 - α N-Glycan site | | | | |
|---|---|---|---|---|
| **Galactosylation** | **T0** | **De-sialo/De-gal** | | |
| | **Range** | **Batch1** | **Batch2** | **Batch3** |
| G0 | N.D. | 89.9 | 89.4 | 83.1 |
| G1 | N.D. | 10.1 | 10.4 | 16.7 |
| G2 | 1.1-1.2 | N.D. | N.D. | N.D. |

| Asn78 - α N-Glycan site | | | | |
|---|---|---|---|---|
| Galactosylation | T0 | De-sialo/De-gal | | |
| | Range | Batch1 | Batch2 | Batch3 |
| G0 | N.D. | 99.7 | 99.7 | 99.6 |
| G1 | N.D. | 0.2 | 0.2 | 0.2 |
| G2 | 0.7-0.8 | 0.1 | 0.1 | N.D. |
| G3 | N.D. | N.D. | N.D. | N.D. |
| Asn7 - β N-Glycan site | | | | |
| Galactosylation | T0 | De-sialo/De-gal | | |
| | Range | Batch1 | Batch2 | Batch3 |
| G0 | N.D. | 24.6 | 25.0 | 20.3 |
| G1 | N.D. | 33.6 | 33.8 | 34.1 |
| G2 | N.D. | 23.2 | 22.3 | 26.0 |
| G3 | N.D. | N.D. | N.D. | N.D. |
| G4 | N.D. | 0.1 | N.D. | N.D. |
| Asn24 - β N-Glycan site | | | | |
| Galactosylation | T0 | De-sialo/De-gal | | |
| | Range | Batch1 | Batch2 | Batch3 |
| G0 | N.D. | 92.2 | 88.8 | 86.7 |
| G1 | N.D. | 7.2 | 10.5 | 12.7 |
| G2 | 0.7-0.8 | 0.5 | 0.8 | 0.6 |

The de-galactosylation reaction shows a good yield (90-100%) on all N-Glycan sites except for Asn 7 of β-subunit. Asn 7 shows a wide microheterogeneity of highly sialylated glycoforms with a high degree of antennarity (in DS_FSH untreated T0 the tri- and tetra- sialylated N-glycans are the most abundant species on Asn 7 as reported in Table 3.40). The reaction yield of Galactosidase on this site is very low, in fact we find species with one (G1) or two (G2) galactoses exposed in percentages comparable to those of the completely de-galactosylated species. Overall, the de-galactosylation reaction is complete on roughly 75% of the glycoforms.
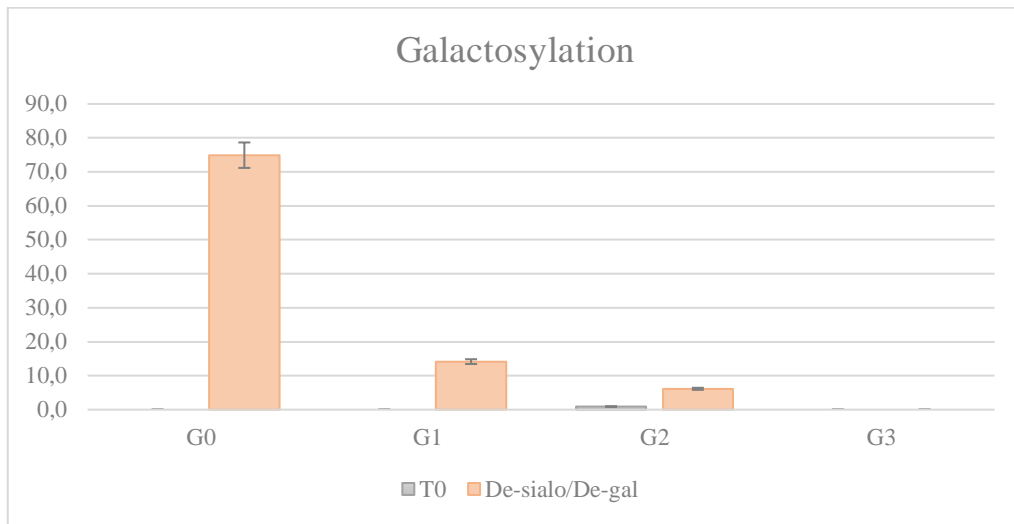
*Figure 3.19 Grouping of galactosylation of De-sialo/De-gal variant in overall molecule. The G1, G2, G3, G4 values of each variant is shown as an average value over the three batches.*

## 3.4 Conclusions

FSH glycosylation is extremely important since greatly affects the biological activity of the hormone, and there are several FSH-based drugs for *in vitro* fertilization approaches [28-31] available on the market exhibiting different glycosylation profiles [32-34].

In the studies reported in this chapter, 3 distinct DS_FSH (Merck) batches were differently treated to generate 7 variants. Alterations in the glycosidic profile of the molecule were monitored by glycopeptide mass mapping for each variant generated. Representative estimation of the most abundant glycan types for each glycosite was obtained by the evaluation of glycan grouping (fucosylation, sialylation, galactosylation and hyper sialic O-acetylation) for each variant.

The main findings on the glycovariants are:

1. The fucosylation of the molecule appears to be unaltered by any of the stress conditions adopted in this study.
2. The thermal stressed variant has not showed significant alteration of glycosylation profile.
3. The acidic fraction variant resulted to have a higher level of heavily sialylated end-capping species than the basic one thus showing a higher S-index (Table 3.18). No species exposing galactose were detected in the acidic fraction.
4. Galactosylated species were observed only in the basic fraction with an overall galactosylation level higher than the untreated sample (Figure 3.10).
5. All variants obtained by pH stress showed a slight alteration of the sialylation content because of a mild hydrolysis due to the pH conditions. The removal of the end-capping sialic acid by acid/basic hydrolysis exposes the subsequent saccharide residue (galactose), thus increasing the galactosylation of the molecule. The hyper O-acetylated sialic

content was also decreased in both stressed variants (Table 3.27). Unexpectedly, the sialic hyper O-acetylation reduction in Basic pH stressed variant was higher when compared to the Acidic pH variant. This anomalous trend could occur in such mild pH conditions, where O-acetylated sialic residues was more susceptible to alkaline treatment [27].

6. In both the de-sialylated and de-sialylated/De-galactosylated variants no species containing sialic acid was found evidencing the completeness of the enzymatic reaction. Moreover, de-galactosylation reaction shows a good yield (90-100%) on all N-Glycan sites except for Asn 7 of β-subunit.

The characterisation of glycosylation profiles of variants obtained will be of precious utility for further *in vivo* and *in vitro* experiments (the experiments will be performed by Merck KGaA). These studies will be part of a larger project finalized to highlight the effect of such glycovariants on ovarian follicular maturation and growth. The ultimate goal will be to define the relationship between the FSH glycosylation pattern and its biological activity and try to understand how the efficacy of commercial drug product (GONAL-F ®) could be improved.

## 3.5    References

1.      Smitz, Johan, et al. "*Follicle-stimulating hormone: a review of form and function in the treatment of infertility.*" Reproductive Sciences, 2016. 23 (6): 706-716..

2.      Elsaid, R., et al., *Hypo-glycosylated Human Follicle–Stimulating Hormone (hFSH21) Exhibits Higher Endocytic Rate for Recombinant hFSH Receptor than Fully-glycosylated hFSH (hFSH24).* The FASEB Journal, 2015. **29**(1_supplement): p. 890.8.

3.      Loreti, N., et al., *The glycan structure in recombinant human FSH affects endocrine activity and global gene expression in human granulosa cells.* Molecular and cellular endocrinology, 2013. **366**(1): p. 68-80.

4.      Loreti, R.N., et al., *Effect of sialylation and complexity of FSH oligosaccharides on inhibin production by granulosa cells.* 2013.

5.      Zambrano, E., et al., *Receptor binding activity and in vitro biological activity of the human FSH charge isoforms as disclosed by heterologous and homologous assay systems.* Endocrine, 1999. **10**(2): p. 113-121.

6.      Leão, R.B.F. and S.C. Esteves, *Gonadotropin therapy in assisted reproduction: an evolutionary perspective from biologics to biotech.* Clinics, 2014. **69**(4): p. 279-293.

7.      Wide, L. and O. Bakos, *More basic forms of both human follicle-stimulating hormone and luteinizing hormone in serum at midcycle compared with the follicular or luteal phase.* The Journal of Clinical Endocrinology & Metabolism, 1993. **76**(4): p. 885-889.

8.      Anobile, C., et al., *Glycoform composition of serum gonadotrophins through the normal menstrual cycle and in the post-menopausal state.* Molecular human reproduction, 1998. **4**(7): p. 631-639.

9.      Palermo, R., *Differential actions of FSH and LH during folliculogenesis.* Reproductive biomedicine online, 2007. **15**(3): p. 326-337.

10.     Davis, J.S., et al., *Naturally occurring follicle-stimulating hormone glycosylation variants.* Journal of glycomics & lipidomics, 2014. **4**(1): p. e117.

11.     Bousfield, G.R., et al., *Macro-and micro-heterogeneity in pituitary and urinary follicle-stimulating hormone glycosylation.* Journal of glycomics & lipidomics, 2014. **4**.

12.     Grass, J., et al., *Analysis of recombinant human follicle-stimulating hormone (FSH) by mass spectrometric approaches.* Analytical and bioanalytical chemistry, 2011. **400**(8): p. 2427.

13.     Mastrangeli, R., et al., *In-vivo biological activity and glycosylation analysis of a biosimilar recombinant human follicle-stimulating hormone product (Bemfola) compared with its reference medicinal product (GONAL-f).* PloS one, 2017. **12**(9).

14.     Wang, H., et al., *A human FSHB transgene encoding the double N-glycosylation mutant (Asn7Δ Asn24Δ) FSHβ subunit fails to rescue Fshb null mice.* Molecular and cellular endocrinology, 2016. **426**: p. 113-124.

15.     Zariñán, T., et al., *In vitro impact of FSH glycosylation variants on FSH receptor-stimulated signal transduction and functional selectivity.* Journal of the Endocrine Society, 2020. **4**(5): p. bvaa019.

16. Wide, L. and B. Hobson, *Influence of the assay method used on the selection of the most active forms of FSH from the human pituitary.* European Journal of Endocrinology, 1986. **113**(1): p. 17-22.

17. Ulloa-Aguirre, A., J.A. Dias, and G.R. Bousfield, *Gonadotropins and the Importance of Glycosylation.* Post-translational Modification of Protein Biopharmaceuticals, 2009: p. 109-147.

18. Bishop, L.A., et al., *Specific roles for the asparagine-linked carbohydrate residues of recombinant human follicle stimulating hormone in receptor binding and signal transduction.* Molecular Endocrinology, 1994. **8**(6): p. 722-731.

19. Barrios-De-Tomasi, J., et al., *Assessment of the in vitro and in vivo biological activities of the human follicle-stimulating isohormones.* Molecular and cellular endocrinology, 2002. **186**(2): p. 189-198.

20. Ulloa-Aguirre, A., et al., *Impact of carbohydrate heterogeneity in function of follicle-stimulating hormone: studies derived from in vitro and in vivo models.* Biology of reproduction, 2003. **69**(2): p. 379-389.

21. Bishop, L., T.V. Nguyen, and P.R. Schofield, *Both of the beta-subunit carbohydrate residues of follicle-stimulating hormone determine the metabolic clearance rate and in vivo potency.* Endocrinology, 1995. **136**(6): p. 2635-2640.

22. Sairam, M. and G. Bhargavi, *A role for glycosylation of the alpha subunit in transduction of biological signal in glycoprotein hormones.* Science, 1985. **229**(4708): p. 65-67.

23. Jiang, X., J.A. Dias, and X. He, *Structural biology of glycoprotein hormones and their receptors: insights to signaling.* Molecular and cellular endocrinology, 2014. **382**(1): p. 424-451.

24. Jiang, X., et al., *Evidence for follicle-stimulating hormone receptor as a functional trimer.* Journal of Biological Chemistry, 2014. **289**(20): p. 14273-14282.

25. Toba, S., M. Tenno, and A. Kurosaka, *An O-acetylated sialyl-Tn is involved in ovarian cancer-associated antigenicity.* Biochemical and Biophysical Research Communications, 2000. **271**(2): p. 281-286.

26. Argüeso, P. and M. Sumiyoshi, *Characterisation of a carbohydrate epitope defined by the monoclonal antibody H185: sialic acid O-acetylation on epithelial cell-surface mucins.* Glycobiology, 2006. **16**(12): p. 1219-1228.

27. Zhu, W., et al., *Degradation Kinetics and Shelf Life of N-acetylneuraminic Acid at Different pH Values*. Molecules, 2020. **25**(21): p. 5141.

28. Lunenfeld B, Bilger W, Longobardi S, Alam V, D'Hooghe T, Sunkara SK. *The Development of Gonadotropins for Clinical Use in the Treatment of Infertility.* Front Endocrinol (Lausanne). 2019 Jul **3**;10:429. doi: 10.3389/fendo.2019.00429. PMID: 31333582; PMCID: PMC6616070.

29. Hofmann, Glen E., et al. "*High-dose follicle-stimulating hormone (FSH) ovarian stimulation in low-responder patients for in vitro fertilization*." Journal of in vitro Fertilization and Embryo Transfer 1989. **6** (5): 285-289.

30. Lyu, Sang Woo, et al. "*Impact of high basal FSH/LH ratio in women with normal FSH levels on in vitro fertilization outcomes.*" Gynecological Endocrinology, 2013. **29** (5): 424-429.

31. Conforti, A., et al., *Pharmacogenetics of FSH Action in the Female*. Frontiers in Endocrinology, 2019. **10**: p. 398

32. Bergandi L, Canosa S, Carosso AR, Paschero C, Gennarelli G, Silvagno F, Benedetto C, Revelli A. *Human Recombinant FSH and Its Biosimilars: Clinical Efficacy, Safety, and Cost-Effectiveness in Controlled Ovarian Stimulation for In Vitro Fertilization.* Pharmaceuticals (Basel). 2020 Jun 27;**13**(7):136. doi: 10.3390/ph13070136. PMID: 32605133; PMCID: PMC7407829.

33. Gizzo S., Garcia-Velasco J.A., Heiman F., Ripellino C., Bühler K. *A cost-effectiveness evaluation comparing originator follitropin alfa to the biosimilar for the treatment of infertility.* Int. J. Womens Health. 2016;**8**:683–689. doi: 10.2147/IJWH.S118687.

34. Christianson M.S., Shoham G., Tobler K.J., Zhao Y., Monseur B., Leong M., Shoham Z. *Use of various gonadotropin and biosimilar formulations for in vitro fertilization cycles: Results of a worldwide Web-based survey.* J. Assist. Reprod. Genet. 2017;**34**:1059–1066. doi: 10.1007/s10815-017-0952-0

## Chapter 4     Development of MRM mass spectrometry method for absolute quantitative analysis of hFSH in woman sera

### 4.1 Introduction

Gonadotropins are the key regulators of follicular development in female reproductive system; the amount of all gonadotropins, and secretion rate widely varies at different ages and at different times during the menstrual cycle. The secretion of FSH is very low before puberty (0 to 4.0 mIU/ml); following puberty, more FSH is secreted (0.3 to 10.0 mIU/ml); during the menstrual cycle there is a dramatic increase of FSH serum concentrations in ovulation phase, and the secretion of the hormone increases 10- to 15-fold in postmenopausal women ( 25.8 to 134.8 mIU/ml) [1]. Due to their role in follicular development, exogenous FSH is used in therapeutic practice in assisted reproductive technology clinics. Actually, hormonal stimulation is an essential part of modern assisted reproductive technology (ART). By stimulating folliculogenesis, these hormones increase the number of oocytes or embryos to be used for ART. The determination of the serum values of gonadotropins is fundamental in assessing the functional status of the hypothalamus-ovarian axis. Altered levels of these hormones in the course of life for women and men indicate the onset of many diseases or problems on the reproductive system [2]. Numerous studies have been conducted on female infertility defined by decreasing levels of FSH [2, 3]. The monitoring of FSH levels, and other hormones, in sera samples is a common clinic practice. The evaluation of the various causes of anovulation can be complex; typically, measurement of body weight and height, and measurement of serum FSH, prolactin, thyroid-stimulating hormone (TSH), and androgens, can help identify the cause of the anovulation. A progestin withdrawal test may be helpful to evaluate the degree of hypogonadism present and may help guide treatment choices.

The absolute quantification of circulating FSH is a great challenge for the scientists due to its considerable heterogeneity, extensively discussed in previous chapters, and its very low concentrations in serum [1]. Since exogenous FSH is used clinically to induce ovulation in women as part of ART treatments (i.e. ovulation induction and controlled ovarian stimulation), the accurate measurement and characterisation of serum hFSH is essential to monitor the hormonal response of patient to the performed stimulation. Nowadays, Enzyme-linked immunosorbent assay (ELISA) is the reference method to quantify circulating FSH. Although ELISA tests are successfully used in diagnostics for the detection of many proteins, this method is limited by the availability of high-quality antibodies for each biomarker candidates and to the specificity of the antigen-antibody-affinity [4]. An alternative to ELISA is mass spectrometry (MS)-based targeted protein assays, providing the use of LC-MS/MS tandem in Multiple Reaction Monitoring (MS-MRM) ion mode. Previous studies have demonstrated the robustness, accuracy and precision of MRM-MS analytical methodology for protein quantification [5-8]. Furthermore, the MRM-MS approach has the advantage of identifying and quantifying multiple proteins in a single run, saving time and costs, compared to ELISA tests.

MRM-MS improves the sensitivity performances of tandem mass spectrometry by monitoring predetermined m/z precursor ions and m/z of specific fragment ions, allowing the detection of low abundant proteins in a very complex matrix; in MRM mode the detection limit for peptides decreases up to 100-fold (compared to full scan MS/MS analysis) due to rapid and continuous monitoring of specific ions of interest. The goal of this study was to set-up an analytical method based on tandem mass spectrometry in MRM ion mode, by using the external standard method. The MRM optimized method was then applied to serum samples to quantify FSH in women sera of different ages.

## 4.2 Materials and method

### 4.2.1 Chemicals and reagents

FSH standards were purchased from Sigma-Aldrich. Dithiothreitol (DTT), ethylenediaminetetraacetate (EDTA), trypsin, iodoacetamide (IAM), ammonium bicarbonate (AMBIC) were purchased from Sigma-Aldrich. Formic acid (HCOOH), methanol, acetonitrile (ACN) are from J.T. Baker. Water ($H_2O$), Stage Tip C18 47mm were purchased from Sigma-Aldrich (Saint Louis, USA).

### 4.2.2 Development MRM-MS method to quantify circulating FSH

#### 4.2.2.1 Sample treatment: standard protein and real samples

Standard proteins (10μg) were subjected to *in solution* hydrolysis protocol by using trypsin as hydrolytic enzyme. Reduction and alkylation were carried out before tryptic hydrolysis by using 10μl of 100mM dithiothreitol (DTT) and incubated at 60°C for 1 h; then, 10μl of IAM (100mM) was added to perform carboamidomethylation. The mixture was then incubated in the dark at room temperature for 1 h. Protein digestion was carried out using trypsin at a 50:1 (protein:enzyme) mass ratio in 10mM AMBIC. The sample was incubated at 37°C for 16 h. The sample clean-up was performed by using homemade C18 reverse phase chromatography. Samples were dried under vacuum and resuspended in 50μL of 98% $H_2O$/2% ACN/ 0.1% HCOOH to perform the MRM-MS analysis. Standard FSH tryptic peptides solutions were prepared by 8 serial dilution from stock solutions and used for calibration curves (0.002-0.002-0.02-1-2-10-20-40 ng/μl). With the same protocol, 10μL of 10 sera from women of fertile age were treated.

### 4.2.2.2 LC-MRM/MS analysis

Skyline software (3.7, 64-bit version MacCoss Lab Software, University of Washington, USA) was used for the *in silico* selection of peptides with unique sequence for each target protein. Proteotypic fully tryptic peptides, with no missed cleavages, with a length between 5 and 30 amino acids and devoid of methionine and cysteine residues, if possible, were chosen for MRM assays development. In addition, sequences that may cause incomplete digestion, such as continuous sequences of arginine (R) or lysine (K) and a proline (P) at the C-terminal side of R or K were also excluded since partial tryptic hydrolysis at the peptide bond is often observed in MS/MS. The top 3 or 5 transitions were selected for method development based on the presence of abundant y ions at m/z greater than that of the precursor. In the absence of high m/z y ions, the most abundant fragment b ions were selected. For each peptide, m/z precursor ion, m/z product ions and relative collision energy were provided by Skyline. Peptide mixture was analysed by LC-MS/MS analysis using a Xevo TQ-S (Waters) equipped with an Ion-Key UPLC Microflow Source coupled to an UPLC Acquity System (Waters).

### 4.2.2.3 LC-MRM/MS method

1μL peptide mixture was injected and separated on a iKey Peptide BEH C18 Separation Device, 130Å, 1.7 μm, 150 μm X 50 mm (Waters, Milford, MA, USA) at 45°C with flow rate of 3μl/min using 0.1% HCOOH in water (LC-MS grade) as eluent A and 0.1% HCOOH in ACN as eluent B. Peptides were eluted (starting 1 min after injection) with a linear gradient of eluent B in A from 7% to 95% in 55 min. The column was re-equilibrated at initial conditions for 4 min. The MRM mass spectrometric analyses were performed in positive ion mode using MRM detection window of 0.5-1.6 min per peptide; the duty cycle was set to automatic and dwell times were minimal 5 ms. Cone voltage was set to 35 V.

#### 4.2.2.4 Method validation limit of detection and quantitation

The limits of detection (LODs) were defined as the lower limit of concentration below which the sample could not be revealed and were determined by making 5 replicate measurements of blank samples. Calculations were made according to the following formula:

LOD = 3 x SD

where SD was the standard deviation of the Blank.

The limit of quantitation (LOQ) is the lowest concentration at which the analyte can not only be reliably detected but also quantitated. The LOQ may be equivalent to the LOD or higher. LOQ was determined as the analyte concentration giving a signal to noise ratio (S/N) of 3 times higher than the LOD. Possible matrix effects were evaluated by comparing standard and matrix matched calibration curves for each analyte. Standard solutions were prepared as described before. The calibration curves were repeated three times. Matrix effects were evaluated by comparing five points standard and matrix-matched calibration curves.

### 4.3 Results and discussions

#### 4.3.1 Serum FSH absolute quantification by MRM/MS based method

The aim of this work was to the setup of an analytical method based on tandem mass spectrometry in MRM ion mode, to quantify hFSH in women sera of different ages. Since FSH has an α-polypeptide chain identical to other gonadotropins, for the development of the MRM method, only peptides belonging to the β chain were selected. Thanks to skyline software, the best precursor ion → product ion transitions and the relative collision energies were selected for each prototypic peptide in order to develop the MRM method. The following table shows the list of transitions monitored for each selected peptide

of β chain FSH and the optimized collision energy used for the fragmentation of each peptide.

*Table 4.1 List of target peptide for FSH protein with precursor m/z, product ions m/z and Collision Energy (CE, V)*

| Protein | Peptide | Precursor Ion (m/z) | Product Ion (m/z) | CE |
|---------|---------|---------------------|-------------------|-----|
| β-FSH | R.DLVYK.D [53-57] | 319.1814$^{++}$ | 552.3286$^{+}$ | 11 |
| | | | 409.2445$^{+}$ | |
| | | | 310.1761$^{+}$ | |
| | | | 491.2500$^{+}$ | |
| | K.DPARPK.I [58-63] | 342.1954$^{++}$ | 568.3566$^{+}$ | 12 |
| | | | 471.3038$^{+}$ | |
| | | | 400.2667$^{+}$ | |
| | K.ELVYETVR.V [72-79] | 504.7717$^{++}$ | 879.4934$^{+}$ | 18 |
| | | | 766.4094$^{+}$ | |
| | | | 667.3410$^{+}$ | |

### 4.3.2 Quantitative analysis: external standard method

The quantitative analysis has been performed by the external standard method by using FSH standard protein. The standard protein was treated as reported in the materials and methods paragraph. Solution at known concentration of peptide mixture were analysed in triplicate to build up the calibration curves. As an example, the β-FSH [72-79] peptide MRM analysis is reported below.
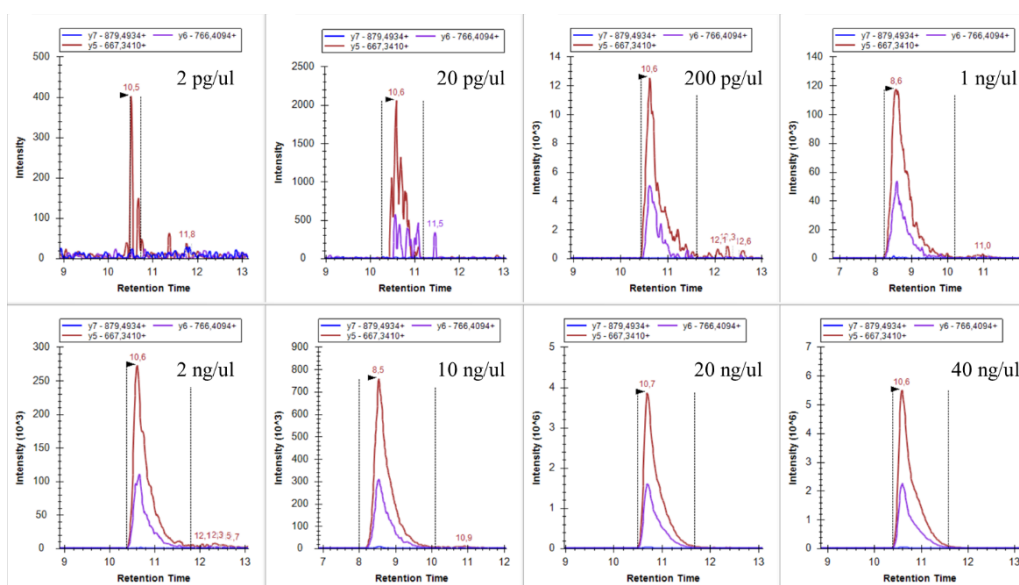
*Figure 4.1 XIC of target peptide K.ELVYETVR.V [72-79] for FSH standard protein.*

Figure 4.1 shows the extracted-ion chromatograms (XIC) of target peptide K.ELVYETVR.V [72-79] for FSH standard protein for each dilution: 0.002 - 0.02 - 0.2 - 1 - 2 - 10 - 20 - 40 ng/μL. XIC show the perfect co-elution of precursor ion → product ion transitions: y7- 504,7717$^{++}$ → 879,4934$^{+}$ in blue; y6- 504,7717$^{++}$ → 766,404094$^{+}$ in purple; y5- 504,7717$^{++}$ → 667,3410$^{+}$ in red.

The peptide shown in Figure 4.1 displays the highest ion current for all the monitored transitions so that it was defined as a Quantifier while the other/s monitored peptides were used as Qualifiers to confirm the specificity in protein recognition in a real sample. Therefore this  K.ELVYETVR.V[72-79] peptide is used as a stoichiometric representation of the protein in quantitative analysis.

To build up the calibration curves, the areas underlying the Quantifier peak of the MRM transitions were reported against the relative protein concentrations, since these latter are proportional to the XIC peak areas (Figure 4.2).
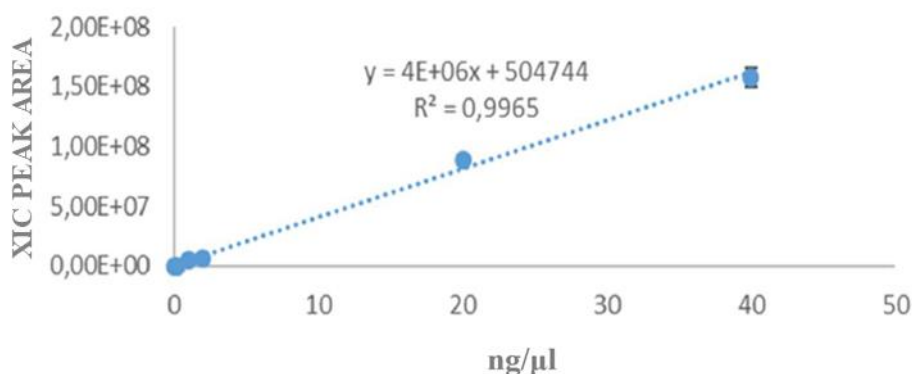
*Figure 4.2 Calibration Curves of FSH target peptide.*

Figure 4.2 shows the calibration curve of peptide K.ELVYETVR.V [72-79] of the FSH protein in concentration range 0.02-40 ng/μL, CV 5.5%. The analysis was performed in triplicate for each point of the calibration curves to evaluate the reproducibility of the developed method. The coefficients of determination ($R^2$) were greater than 0.99, and the calculated CV values were under the 10%. In Table 4.2 the analytical parameters determined for the target protein are reported.

*Table 4.2 Analytical parameters*

| Linearity range | LOD | LOQ | CV |
|---|---|---|---|
| 20 pg/μl-40 ng/ul | 20 pg/μl | 60 pg/μl | <10% |

This method shows detection and quantification limits comparable to ELISAs assays [9]. In order to define the recovery in matrix, a standard solution is prepared by spiking known amount of standard proteins in mice pooled sera. The amount of spiked standard FSH is at least 3-4 degree of magnitude higher than the human endogenous one, therefore, the contribution of the endogenous FSH of the mice serum can certainly be assumed to be zero. Recovery is calculated by comparing the area of extracted peptide K.ELVYETVR.V [72-79] yield before and after the addition of the different FSH standard solution into the mice pooled sera. In Figure 4.3 is reported the XIC of 72-79 peptide monitored in serum not spiked, 0.5 μg/μl and 1 μg/μl spiked samples.
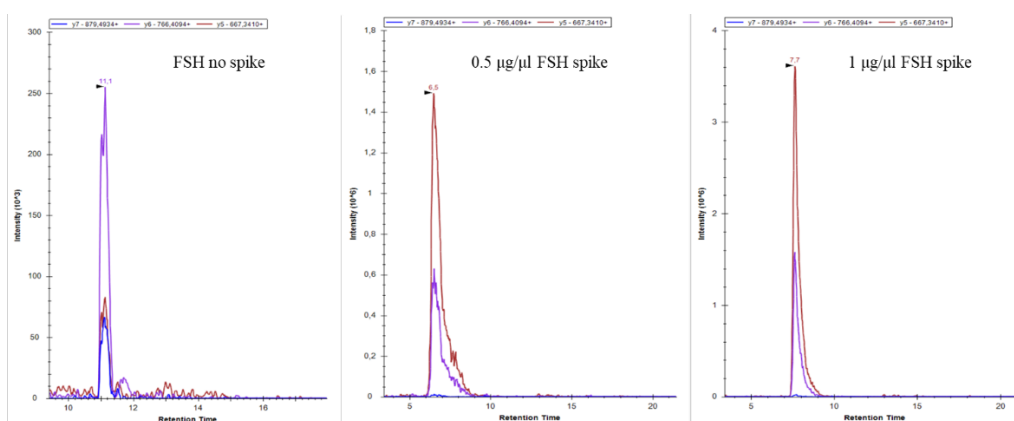
*Figure 4.3 XIC of K.ELVYETVR.V [72-79] in not spiked serum, 0.5 and 1 μg/μl FSH spiked serum samples.*

Quantitation was achieved by using the calibration curves. Recovery value ranged between 75-90%. This strategy has demonstrated high reproducibility and linearity.

### 4.3.3 Application to real samples

In order to reduce the biological variability and evaluate the instrumental response five pools of sera from fertile women were analysed. Pooling of samples in proteomics experiments might help overcome resource constraints when many individuals are analysed (as the measured biological variation should be reduced) [10]. The pools were prepared by taking 2μl from 10 sera of different women. The pools were treated with the same protocol described in the paragraph 4.2.2. For each pool, 1μl of the obtained peptide mixture was analysed with the developed MRM method. As an example, the XIC for the quantifier peptide of the pool number one is reported and the concentration of target protein was evaluated.
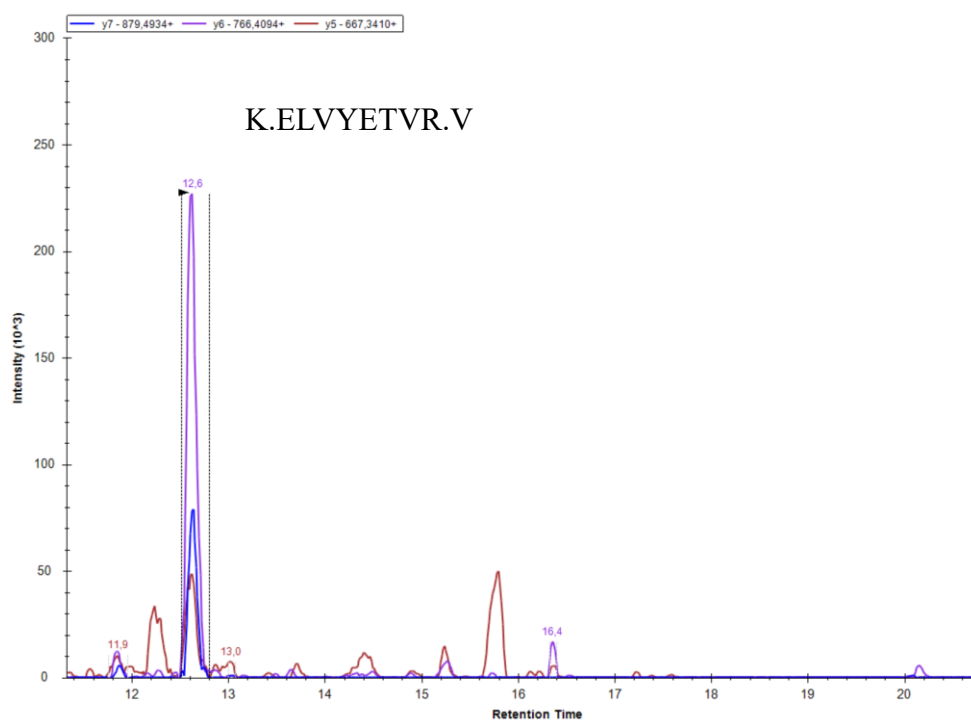
*Figure 4.4. XIC of K.ELVYETVR.V [72-79] peptide monitored in serum pool #1*

Figure 4.4 shows the XIC profile for Quantifier peptide, K.ELVYETVR.V [72-79]. The transitions $504,7717^{++} \rightarrow 667,3410^{+}$; $504,7717^{++} \rightarrow 766,4094^{+}$; $504,7717^{++} \rightarrow 879,4934^{+}$ are presented in red, violet end blue, respectively, and they are co-eluted at minute 12.6 min. By integrating the quantifier's areas underlying the peaks on the calibration lines, the concentration of the protein target in the sera pool was measured. FSH concentrations in sera pool resulted to be 0.227 ng/ml. Once the instrumental response and biological variability were assessed, ten sera from women of different fertility ages were treated with the same protocol and subjected to LC-MRM/MS analysis. The table below shows the values of FSH determined for the analysed samples.

*Table 4.3 FSH quantification in the analysed serum samples.*

| Serum sample | FSH (ng/ml) |
|:---:|:---:|
| #1 | 0.963 |
| #2 | 0.673 |
| #3 | 0.865 |
| #4 | 4.804 |
| #5 | 0.049 |
| #6 | 1.530 |
| #7 | 2.131 |
| #8 | 0.663 |
| #9 | <LOD |
| #10 | <LOD |

Although the clinical history of patients was not known, the really low values of FSH recorded for some sera (#5, #9, #10) suggest a potential pathological condition of hypothyroidism, as well as FSH high values recorded for #4 serum suggest polycystic ovary syndrome, but such hypotheses should be more thoroughly investigated toward the integrative analysis of other hormones to get a clearer diagnosis and a more complete understanding of infertility conditions [11]. Finally, more samples are needed to define a normal range of physiological dosages of women of childbearing age to better detect pathological conditions.

## 4.4 Conclusions

Tandem mass spectrometry in MRM mode is proposed today as an alternative and complementary method to immune-enzymatic assays for the quantification of hormonal proteins contained in very low amount in biological fluids, thanks to the comparable sensitivities (< 1ng/mL) and the high selectivity and specificity. In this study an MRM-MS method to quantify FSH in serum sample was developed. The method involves the identification and quantification of protein by monitoring 3 FSH specific proteotypic peptides (precursor-ion) and 3 or 4 best fragments (product-ion), and it was successfully applied to the analysis of 10 women sera samples of different ages showing a great variability of the amount of the hormone. Results show comparable sensitivity to ELISAs assays,

with the advantage of being faster and more selective. Furthermore, this method can be easily implemented for a simultaneous analysis of several serum proteins. These aspects make the MRM-MS technique advantageous in terms of time and cost.

## 4.5 References

1. Anobile, C.J., et al., *Glycoform composition of serum gonadotrophins through the normal menstrual cycle and in the post-menopausal state.* Molecular Human Reproduction, 1998. **4**(7): p. 631-639.
2. Smith, S., S.M. Pfeifer, and J.A. Collins, *Diagnosis and management of female infertility.* Jama, 2003. **290**(13): p. 1767-1770.
3. Barbieri, R.L., *Female infertility*, in *Yen and Jaffe's Reproductive Endocrinology*. 2019, Elsevier. p. 556-581. e7.
4. Munro, C. J., et al. *ELISA for the measurement of serum and urinary chorionic gonadotropin concentrations in the laboratory macaque.* American journal of primatology, 1997. **41**(4): p. 307-322.
5. Prakash, A., et al., *Platform for establishing interlaboratory reproducibility of selected reaction monitoring-based mass spectrometry peptide assays.* Journal of proteome research, 2010. **9**(12): p. 6678-6688.
6. Illiano, A., et al., *Multiple reaction monitoring tandem mass spectrometry approach for the identification of biological fluids at crime scene investigations.* Analytical chemistry, 2018. **90**(9): p. 5627-5636.
7. Fontanarosa, C., et al., *Quantitative determination of free D-Asp, L-Asp and N-methyl-D-aspartate in mouse brain tissues by chiral separation and Multiple Reaction Monitoring tandem mass spectrometry.* PloS one, 2017. **12**(6).
8. Rauniyar, N., *Parallel reaction monitoring: a targeted experiment performed using high resolution and high mass accuracy mass spectrometry.* International journal of molecular sciences, 2015. **16**(12): p. 28566-28581.
9. Güzel, Coşkun, et al. *Comparison of targeted mass spectrometry techniques with an immunoassay: a case study for HSP90α.* PROTEOMICS–Clinical Applications, 2018. **12**(1): p. 1700107.
10. Diz, Angel P., Manuela Truebano, and David OF Skibinski. *The consequences of sample pooling in proteomics: an empirical study*. Electrophoresis, 2009. **30**(17): p. 2967-2975.

11. Catteau, Aurore, et al. *Abnormally Elevated Follicle-Stimulating Hormone (FSH) Level in an Infertile Woman.* Case reports in endocrinology, 2019. (2019).

# Chapter 5    Identification and relative quantification of hFSH glycoforms in women sera by PRM-MS based approach

## 5.1    Introduction

The characterisation of human Follicle Stimulating Hormone (hFSH) glycosylation is essential to elucidate reproductive physiology regulating fertility, and to improve the diagnose and treat disorders of reproduction [1]. Although the hFSH glycovariants-mediated signaling network is not yet fully deciphered, it is well known that glycosylation at Asn-78 on the α-subunit significantly increases the receptor binding affinity, likewise glycosylation at Asn-52 was found to play an essential role in signal transduction because its removal resulted in a significantly decreased potency [2, 3]. Overall, the glycosylation level (di- try- or tetra-glycosylated hFSH) of the hormone has a significant impact on binding and functional assays [4, 5]. Furthermore, the wide microheterogeneity of carbohydrates on each glycosite increase the hormone complexity, generating an heterogeneous population of different glycoforms, which changes under different physiological and pathological conditions [6, 7]. Moreover, recent studies showed that the relative concentrations of these glycoforms vary over the course of the menstrual cycle and the lifetime [8], indicating that these differences in glycosylation may have physiologic relevance [9-11]. As discussed before, exogenous FSH is used clinically in ART (i.e. ovulation induction and controlled ovarian stimulation in women) and to treat infertility connected to gonadotropin deficiency in men [12]. The accurate measurement and characterisation of serum hFSH is essential for patient condition monitoring before and following the treatment. Furthermore, the careful characterisation of exogenous FSH used for patients' treatment is essential for safe and successful treatments [13]. ELISA is currently the reference method for quantifying serum hFSH but, with this methodology, some clinically important targets remain reliably undetectable due to the lack of specific antibodies capable of recognizing the different proteoforms generated by post-

translational modifications. In fact, ELISA-based methods are unable to determine the different hFSH glycoforms, a crucial point in preparation and characterisation of standards and therapeutic products since the variation of hFSH glycosylation have physiologic relevance [9-11] (i.e. the degree of terminal sialyation determines the half-life of the hormone and therefore affects its *in vivo* bioactivity [13, 14]). An alternative to ELISA is mass spectrometry (MS)-based targeted protein assays, such as MS in Multiple Reaction Monitoring (MRM-MS), described in previous chapter, and MS in Parallel Reaction Monitoring (PRM-MS). PRM-based targeted methods has been successfully applied in the validation of relative abundance of proteins and their post-translational modifications (PTMs) such as glycosylation [15, 16]. The aim of the last part of this Ph.D. project, carried out at the INSERM institute in Paris at the proteomics platform SFR Necker, was focused in the development of PRM-MS method for the identification and relative quantification of hFSH glycoforms in women sera sample. The development of reliable analytical MS methods to quantify and characterize the circulating glycoforms of this hormone is a great challenge not only because the huge heterogeneity and the very low concentration in serum, but also because of the low ionization efficiency; furthermore, the lack of isotope–labeled synthetic glycopeptides used as a standards make their absolute quantification complicated [17]. For these reasons, many strategies are currently developed to make glycoprotein MS characterisation more effective [18-20]. However, very often enrichment procedure (such as affinity chromatography with lectins or immunoprecipitation procedures) are required to facilitate the analysis [21-25]. Lectin affinity chromatography only allows the separation of non-glycoproteins from glycoproteins, therefore (since a considerable number of serum proteins are glycosylated); this approach is not particularly useful to purify a specific glycoprotein [18]. Immunoaffinity chromatography offers potentially greater specificity as it has already been successfully applied to purify several non-glycoprotein hormones [26-29].

## 5.2 Materials and method

### 5.2.1 Chemicals and reagents

Recombinant FSH standards were provided from Merck KGaA. Dithiothreitol (DTT), ethylendiaminetetraacetate (EDTA), trypsin, iodoacetamide (IAM), ammonium bicarbonate (AMBIC) were purchased from Sigma-Aldrich. Formic acid (HCOOH), methanol, acetonitrile (ACN) are from J.T. Baker. Water ($H_2O$), Stage Tip C18 47mm were purchased from Sigma-Aldrich (Saint Louis, USA). Concanavalina A (ConA), wheat germ agglutinin (WGA), Ricinus communis agglutinin (RCA) were purchased from Sigma.

### 5.2.2 Development of PRM-MS method for hFSH glycoforms analysis

#### 5.2.2.1 Immunopurification hFSH from serum

As hFSH glycosylation differ significantly in the different stages of women's fertility [8], we choose to analyse a serum pool to test the applicability of the method on a range of glycoforms as wide as possible. To this purpose, 54 fertile women serum samples (age range: $29 \pm 3$), supplied by Istituto Nazionale Tumori (INT) "Fondazione G. Pascale" of Naples, were pooled and hFSH immunopurification protocol was carried out. 500 µl of anti hFSH sepharose CL-4B (Merck KGaA) immunoresin were washed in binding buffer (BB: Tris-HCl 0.1M, NaCl 0.5M, NaN3 0.02%). 500 µl of serum pool were diluted 5 times in BB and incubated in batch with immunoresin at 4°C for 18h. The immunoresin was washed 5 times with AMBIC 0.1 M and eluted in $NH_4OH$ 1M. The eluted fraction was desalted with PD-10 Desalting Columns contain Sephadex G-25, and lyophilized.

### 5.2.2.2 Reduction, alkylation and enzymatic digestion

Dried pellet from immunopurification was treated with 100 µl of 8 M GuCl, 1mM EDTA, 130 mM Tris-HCl, pH 7.6. The disulfuric bond reduction was performed by adding 6 µl of 500 mM Dithiothreitol under stirring for 60 minutes at 37°C. The samples were subsequently alkylated by adding 12 µl of 500 mM Iodoacetamide in the dark for 30 minutes at room temperature. The reduced and alkylated sample was then purified from excess of reagent by $CHCl_3/CH_3OH/H_2O$ precipitation. Enzymatic hydrolysis was carried out in 50 µl digestion buffer (2M urea, 50 mM Tris-HCl at pH 8.0) using Chymotrypsin (Roche) with an enzyme:substrate ratio 1:50 at 37°C for 16 h. As positive control, 30 µg of recombinant standard FSH (RHS_FSH), were digested following the same protocol described above. The peptide and glycopeptides mixture from hFSH immunopurified from serum (IP_hFSH) and standard RHS_FSH were resuspended in 10% ACN, 0.1% (v/v) TFA for LC-MS/MS analysis.

### 5.2.2.3 Mass spectrometry analysis

1/20 of chymotryptic digested samples (IP_hFSH and RHS_FSH) were analysed using a nano-RSLC-Q Exactive PLUS (Dionex RSLC Ultimate 3000, Thermo Scientific, Waltham, MA, USA). Peptides were separated on a 50 cm reversed-phase liquid chromatographic column (Pepmap C18, Dionex) with 60 min gradient. Peptides eluting from the column were analysed by data dependent MS/MS, using top ten acquisition method. Instrument settings: resolution was set to 70 000 for MS scans and 17 500 for the MS/MS scans. The MS automatic gain control (AGC) target was set to $3.10^6$ counts with 60 ms for the injection time, while MS/MS AGC target was set to $1.10^5$ with 120 ms for the injection time. The MS scan range was from 400 to 2000 m/z. Dynamic exclusion was set to 30 s. The mass spectrometry data were analysed by Proteome Discoverer v2.4 (Thermo Scientific) against human subset from UniProtKB/Swiss-Prot complete proteome database using the Byonic node search engine (v3.6.0 Protein Metrics, PMI-Suite) with following settings:

Fixed and variable modification: Carbamidomethyl (+57.021464 @ C |fixed), Oxidation (+15.994915 @ M |variable), Gln->pyro-Glu (-17.026549 @ NTerm Q |variable); Glycan modification: N-glycan 132 human.txt @ NGlycan; Cleavage residues: FLWY; Digest cutter: C-terminal cutter; Maximum number of missed cleavages: 2; Fragmentation type: QTOF/HCD; Precursor tolerance: 5.0 ppm; Fragment tolerance: 0.05 Da.

Byonic peptides attribution quality criteria:

Byonic automatically filters both peptide-spectrum matches (PSMs) and proteins by default; PSMs were filtered by checking the "Automatic score cut" box. The threshold will be in the range 200 – 400. The protein FDR was set at 1 % (or 20 reverse count).

### 5.2.2.4   PRM method

For the PRM method, two scan events were used for each cycle: one full scan followed by a time-scheduled PRM scan. The full scan was acquired at resolution of 35 000 at m/z 200. The AGC target was set to $3.10^6$ counts with 60 ms for the injection time. The MS range was from 150 to 2000 m/z. The time-scheduled PRM scans were acquired at resolution of 17500 at m/z 200. Other parameters included an AGC target of $2.10^5$ counts with 100 ms for the injection time. Fragmentation was performed using a normalized collision energy of 35 eV. The inclusion list containing the precursor ions was used to trigger the acquisition with a 5 min scheduled scan windows. To determine the ion's start and end scan times, we used the data acquired in full scan mode. The cycle time was 2 s and the number of points per chromatographic peak were around ten.

### 5.2.2.5   Glycopeptides Enrichment

The remaining part of the IP_hFSH sample was enriched by array of lectins using modified *Yang, Z. & Hancock W. S. (2004)* multi-lectin affinity column protocol. ConA, WGA and RCA were chosen to enrich all common human FSH N-glycans

[30, 31]. The sample was resuspended in 40 µl Binding Buffer (BB: 20mM Tris-HCl pH 7.6, 1 mM MnCl2 1 mM CaCl2, 150 mM NaCl) and transferred on a filer unit (Micron YM-30). 36 µl of C.W.R. (15µl of [ConA]= 6 mg/ml + 15 µl of [WGA]= 6mg/ml + 6 µl of [RCA]= 18.89 mg/ml) mix were added to the filter unit, mixed at 600 rpm in a thermo-mixer for 1 min and incubate without mixing for 60 min at room temperature. The peptides were washed 4 times by adding 200 µl BB and discarded centrifuging at 14000 x g for 10 min. Glycopeptides were eluted in 200 µl elution buffer (EB: 20 mM Tris, 0.5 M NaCl, 0.17 M methyl-α-D-mannopyranoside, 0.17 M N-Acetyl-glucosamine and 0.27 M galactose, PH= 7.4). The buffer of eluted fraction containing glycopeptides was then exchanged against water and concentrated using Amicon Ultracel 3k centrifugal filters. 1/30 of eluted enriched glycopeptides was injected in the mass spectrometer nano RSLC-Q Exactive PLUS (Dionex RSLC Ultimate 3000, Thermo Scientific, Waltham, MA, USA) described before and analysed using PRM modality.

## 5.3    Results and discussions

### 5.3.1    Workflow of PRM-MS

This study presents a MS-PRM based strategy to identify and quantify serum hFSH glycoforms. A first step of hormone immunopurification from human serum was carried out before LC-MS/MS analysis in PRM mode. The overall workflow of PRM experiment is summarized in Figure 5.1.
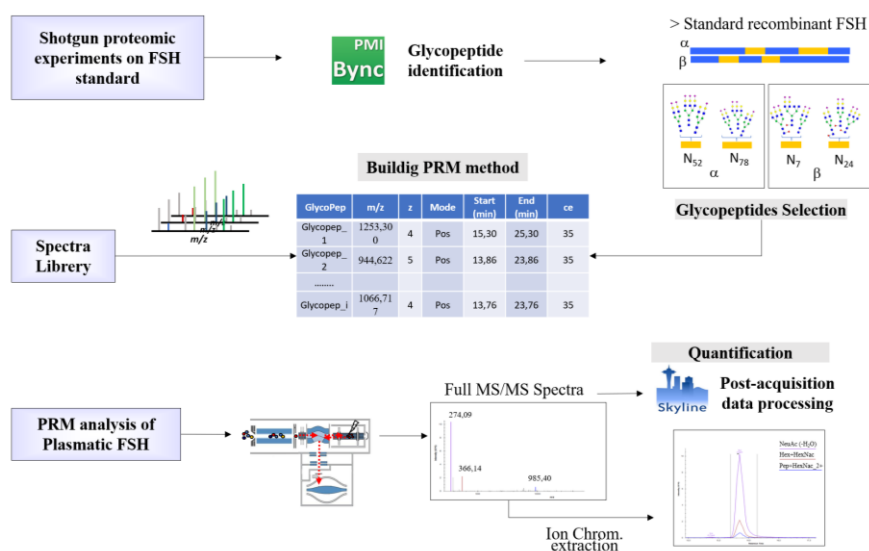


*Figure 5.1 Workflow of PRM experiment carried out to identify and quantify h-FSH glycoforms in serum sample.*

### 5.3.2    PRM method development

In order to develop the targeted data acquisition methods (MS-PRM), we have used preliminary discovery data acquisition from shotgun experiments of recombinant standard human (RSH) FSH. To this purpose, key information about the target glycopeptides (m/z of the precursor ions, charge state and start-end elution time) were obtained from classical LC-MS/MS full scan analysis of RHS_FSH chymotryptic digestion. Proteins identification was performed by Byonic search engine and each recognized glycopeptide was then validated by manually interpreting MS/MS spectra. As an example, the identification of a glycopeptide (m/z =1253.30) derived from α-FSH is showed in Figure 5.2. The

full scan LC-MS/MS TIC chromatogram is reported in Figure 5.2 A). The signal at m/z $1253.30^{+4}$, found in the full MS spectrum at 20.28 min (Figure 5.2B), is assigned to $RSKKTMLVQKN_{52}VTSESTCCVAKSY$ peptide with complex biantennary fully sialylated glycosylation at Asn52 of the α-subunit. We obtained an excellent HCD MS2 fragmentation spectrum containing signals attributable to both the peptide and the glycan moiety, thus confirming the assigned modified peptide (Figure 5.2C).
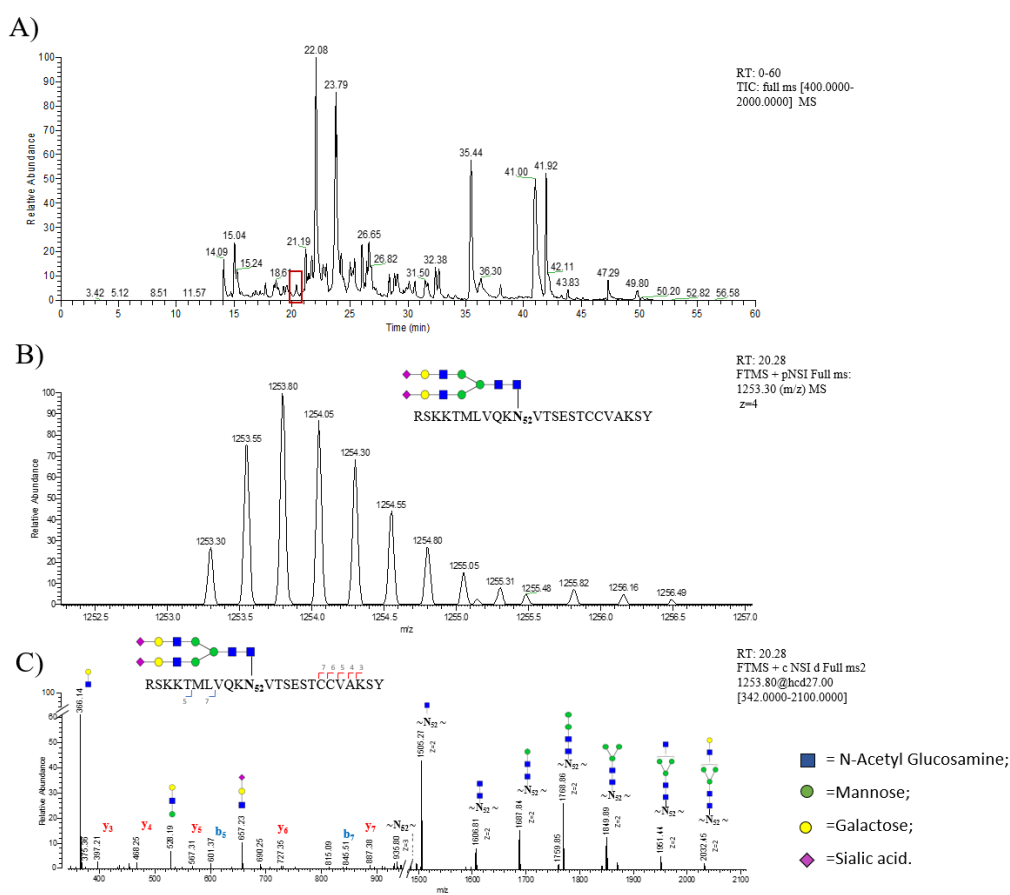


*Figure 5.2 Characterisation of RHS_FSH α-chain glycopeptide: RSKKTMLVQKN₅₂VTSESTCCVAKSY. A) LC/MS full ms TIC chromatogram; B) full mass scan at RT:20.28 and 1252.5- 1257.0 m/z zoom; C) MS2 spectrum of the glycopeptide ion is consistent with the proposed glycan structure.*

In the same way glycans moiety of 3 N-glycan sites was characterized: α-subunit Asn52-78 and Asn24 of β-subunit (supporting data Table 5.5). Unfortunately, no

glycopeptides on the other β-subunit N-glycan site (Asn7) were identified by Byonic software, probably due to the greater number of branched structures on the β-subunit at Asn7 compared with at Asn24. In fact, highly branched tetra-antennary glycan structures have only been reported at Asn7 [8, 32]. This result suggests a higher signal suppression of such highly sialylated structures on Asn7 making their identification extremely complex. However, manual spectra interpretation led to the identification of 4 possible signals attributable to the peptide L.TN$_7$ITIAIEKEECRF.C with 4 different glycovariants (Table 5.5 in supporting data). Due to the very low ionization efficiency of these ions, no fragmentation spectra could be detected, therefore the identification of the 4 glycoforms on Asn 7 could be performed on the basis of the exclusive molecular weight of the entire glycopeptide and of the retention time relative to the unmodified peptide. These results are a key step in our workflow as they will constitute a library of retention time, MS and MS$^2$ spectra of target glycopeptides included into the target list for the development of PRM-MS method to identify serum hFSH glycopeptides.

### 5.3.3  Full scan LC-MS/MS analysis of hFSH purified from serum

Human serum is a very complex matrix due to its high proteins content (~ 60-80 mg/mL). The glycoproteome is a major sub proteome present in human serum, in fact, about 50% of all serum proteins are glycosylated [33]. The glycoprotein hormone hFSH is present in very low levels in serum, as in fertile women its average level in cycle phases is 5,8 mIU/mL (~ 0,39 ng/mL) [34]. To detect and characterize this glycoprotein in such a complex biological matrix, purification and enrichment steps before MS analysis are essential [20]. Anti - hFSH sepharose CL-4B Immunoaffinity purification was carried out on three aliquots of sera pooled samples as described in material and method section, and untargeted LC-MS/MS analysis was performed to verify the efficiency of purification. Although the immunopurified sample led to the identification of other human serum proteins (supporting data Table 5.6), the applied strategy has

proved to be successful in purifying and identifying both α and β-subunit of serum hFSH with 18 % and 34 % of sequence coverage, respectively (Figure 5.3).
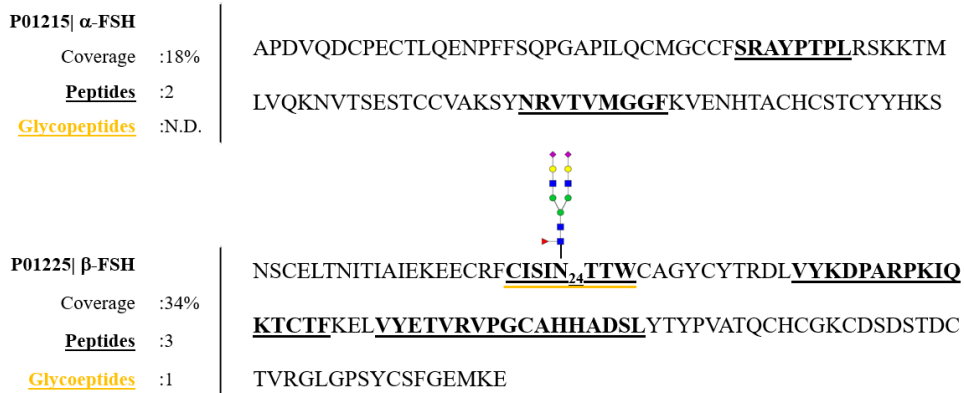


**P01215| α-FSH**
Coverage :18%
**Peptides** :2
Glycopeptides :N.D.

APDVQDCPECTLQENPFFSQPGAPILQCMGCCF**SRAYPTPL**RSKKTM

LVQKNVTSESTCCVAKSY**NRVTVMGGF**KVENHTACHCSTCYYHKS

**P01225| β-FSH**
Coverage :34%
**Peptides** :3
Glycoeptides :1

NSCELTNITIAIEKEECRF**CISIN₂₄TTW**CAGYCYTRDL**VYKDPARPKIQ**

**KTCTF**KEL**VYETVRVPGCAHHADSL**YTYPVATQCHCGKCDSDSTDC

TVRGLGPSYCSFGEMKE

*Figure 5.3 sequence coverage on both hFSH α and β-subunits in serum sample by LC-MS/MS. One glycopeptide (glycan composition: Fuc(1)HexNac(4)Hex(5)NeuAc(2) on β-FSH has been recognized.*

The low percentage of the hFSH sequence coverage is due to its wide heterogeneity of glycovariants and to the presence of other more abundant serum proteins whose signals might hinder the detection of hFSH glycopeptides. Nonetheless, we could identify one glycopeptide, the complex biantennary glycan structure on site Asn24 of β-subunit (Table 5.1).

*Table 5.1 LC-MS/MS peptides details of hFSH identification in serum sample.*

| Master Protein Accessions | PSMs | Annotated Sequence | Start-End | Theo. MH+ [Da] | Glycan Composition | Score |
|---|---|---|---|---|---|---|
| P01215 (α-FSH) | 13 | [F].SRAYPTPL.[R] | [58-65] | 904.4886 | - | 455 |
| | 7 | [Y].NRVTVMGGF.[K] | [90-98] | 996.4931 | - | 331 |
| P01225 (β-FSH) | 5 | [Y].KDPARPKIQKTCTF.[K] | [58-71] | 1689.9104 | _ | 506 |
| | 8 | [Y].ETVRVPGCAHHADSLY.[T] | [77-92] | 1811.8493 | - | 376 |
| | 1 | [F].CISIN*TTW.[C] | [38-45] | 3345.2965 | Fuc(1)HexNac(4)Hex(5)NeuAc(2) | 269 |
| | 2 | [F].CISINTTW.[C] | [38-45] | 994.4662 | - | 342 |

The presence of the unmodified peptide CISINTTW confirms the partial glycosylation of β-FSH [4]. The extracted-ion chromatogram (XIC) relative to the glycopeptide CISIN*TTW m/z=1115.7695 (RT: 49.55 min) compared with

unmodified one m/z=994.4565 (RT: 48.48 min) shows that unmodified peptide
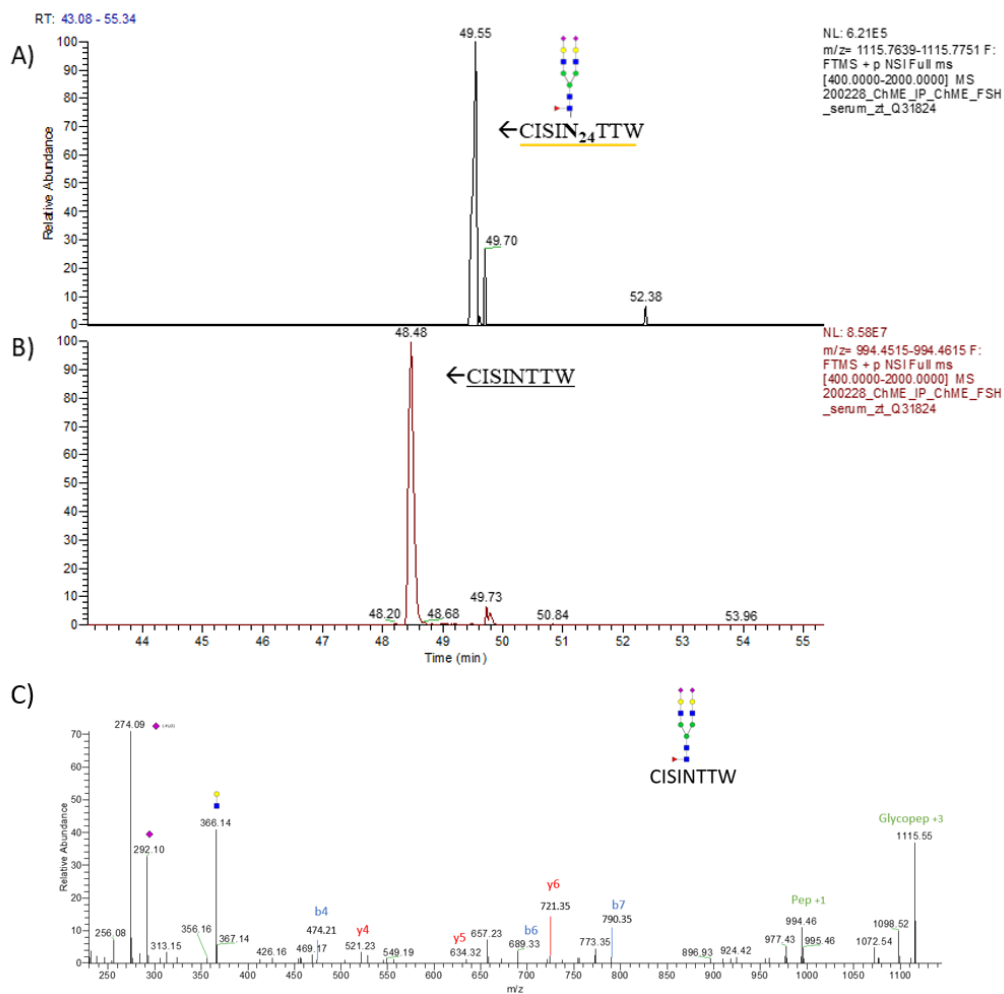intensity is approximately 100-fold higher than glycosylated one (Figure 5.4).



*Figure 5.4 LC-MS/MS full scan analysis of serum IP_hFSH. Extracted-ion chromatogram (XIC)
comparison between: A) glycopeptide CISIN\*TTW m/z=1115.7695 (RT: 49.55 min; NL: 6.21E5), B)
unmodified one m/z=994.4565 (RT: 48.48 min; NL: 8.58E7) and C) glycopeptide CISIN\*TTW
m/z=1115.77 MS2 spectrum.*

### 5.3.4 MS-PRM analysis of hFSH glycoforms purified from serum

To improve the glycopeptides identification and to obtain their relative quantification, the same three aliquots of serum IP_hFSH was re-analysed with MS-PRM target method. We applied the MS-PRM method developed on standard RHS_FSH to the analysis of hFSH purified from serum. Our method included all FSH N-glycan sites identified (4) in 24 different glycoforms (Figure 5.5) and one non-glycosylated peptide (ETVRVPGCAHHADSLY) representative of the overall concentration of FSH. MS-PRM analysis has improved the first untargeted LC-MS/MS glycopeptides identification from serum sample because when analysing complex peptides mixtures, data-dependent acquisition method suffers from a bias toward sampling and identification of most abundant peptides, resulting in poor and inconsistent detection of low level glycopeptides. Since glycosylated peptides have lower ionization efficiencies than relative unmodified one, most abundant peptides MS signals could hide the IP_hFSH glycopeptides confirming the above-mentioned hypothesis.
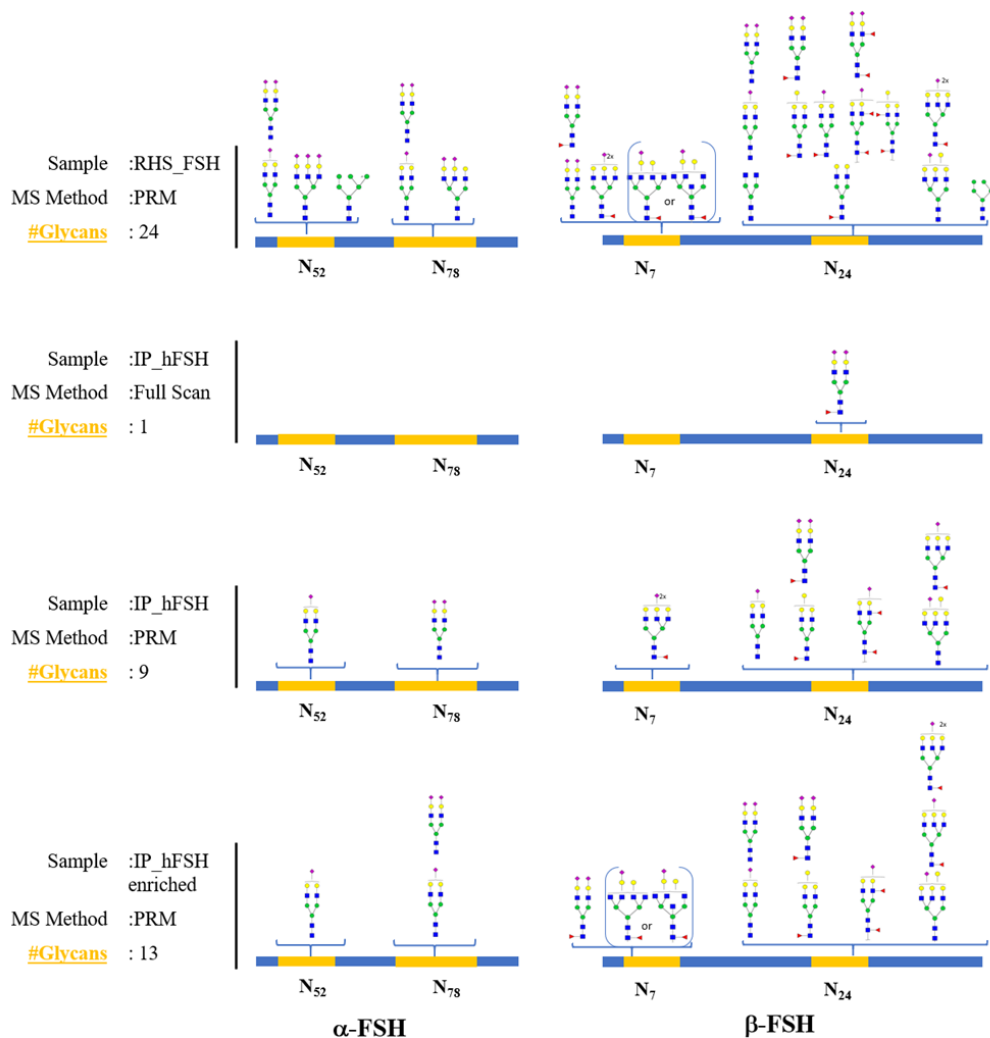
*Figure 5.5 Comparative description of the glycans present on each of the glycosylation sites in standard RHS_FSH and plasmatic IP_hFSH in all performed experiments: Full scan LC-MS/MS and MS-PRM on IP_hFSH and MS-PRM on IP_hFSH_Glycoenrich.*

PRM allows an in-depth analysis of all potential glycopeptides present in the sample because only one predetermined charge state of a given glycopeptide will be selected for fragmentation, allowing us to identify, in IP_hFSH serum samples, the great microheterogeneity of glycosylation, which characterizes the human follicle stimulating hormone. This method allowed us to identify 9 glycopeptides instead of only one obtained with the untargeted analysis, proving the great efficiency of the MS-PRM method (Table 5.7).

Moreover, the recovery of the immunopurification was calculated. In order to detect FSH even in not depleted serum (input), three new aliquots of pooled sera were spiked with 1 mg/ml of FSH before immunoprecipitation. The input, the flow through (FT) and the immunoprocipitated (IP) spiked sera samples were analysed by both Full scan LC-MS/MS and PRM-MS developed method. No FSH peptides were recorded in FT sample. Recovery value (89%) was calculated (Table 5.2) on the Quantifier (best transition: b10-1107.5364) by comparing the average area of extracted peptide ETVRVPGCAHHADSLY (906.4309 m/z, z=2) yield before and after the immuprecipitation (Figure 5.6).

*Table 5.2 Recovery calculation. The area in IP_Spike is calculated taking into consideration the dilutions and the injected quantity.*

| PRM experiments: | Peptide: ETVRVPGCAHHADSLY | m/z=906.4309 | z=2 |
|---|---|---|---|
| **SAMPLES** | **Quantifier AREA (Average)** | | |
| Input_spiked samples | 77273 | **Recovery: 89%** | |
| IP_spiked samples | 68978 | | |



β-FSH peptide      :ETVRVPGCAHHADSLY
Precursor m/z    : 906.4309 (2+)
Elution time      20.00
CE      35
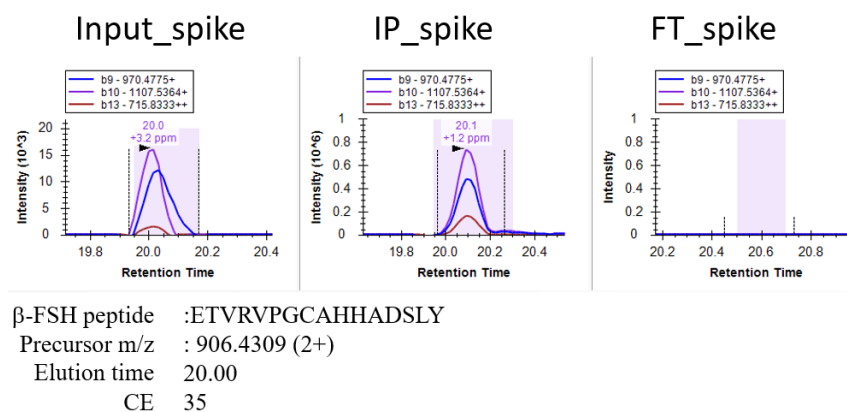
*Figure 5.6 Skyline chromatographic traces post-acquisition extraction the of best 3 selected fragment ion from peptide ETVRVPGCAHHADSLY (906.4309 m/z, z=2) yield before (Input_spike), after the immuprecipitation (IP_spike) and in the unbound (FT_spike). The best transition $906.43092^+ \rightarrow 1107.5364^+$ is selected as quantifier.*

Furthermore, the reproducibility of the immunoprecipitation was assessed by untargeted LC-MS/MS analysis on triplicate spiked samples. In figure below proteins identified after IP was shown in a Venn diagram. The number of proteins

identified in the three samples are very similar; 108 proteins in common for all samples IP_spike 1, 2 and 3 were identified. The reproducibility of LC-MS/MS was calculated on the basis of intensities of peptides of interest of FSH ( Table 5.3). The peak area of the same peptide was extracted from the entire chromatograms (XIC) and normalize against the peak height.

*Table 5.3 Full scan experiment on 3 replicates IP_spike samples. Peak area, intensity and normalization value was report for a representative FSH peptide (ETVRVPGCAHHADSLY)*

| Full Scan experiment: | XIC Peptide: ETVRVPGCAHHADSLY | m/z =906.4309 | z=2 |
|---|---|---|---|
| **SAMPLE** | **PEAK AREA** | **PEAK HEIGHT** | **NORM.** |
| IP_FSH_spike_1 | 14355779 | 2626075 | 5.5 |
| IP_FSH _spike_2 | 76709242 | 11756726 | 6.5 |
| IP_FSH _spike_3 | 8494184 | 1491645 | 5.7 |
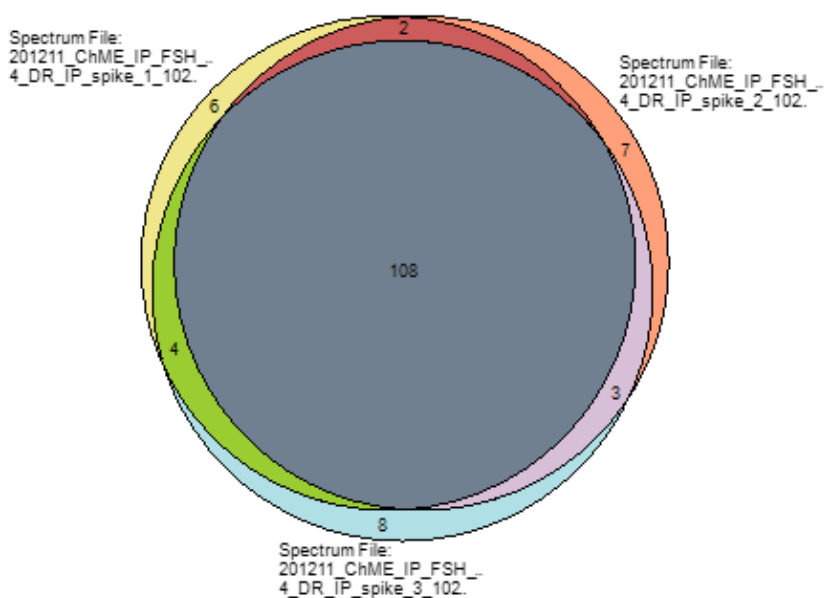


*Figure 5.7 All identified proteins in each sample were visualized by Venn diagram showing the high similarity of  the 3 replicates.*

Finally, the reproducibility of the enzymatic digestion was assessed by measuring the distribution of missed cleavage on all the proteins identified and the by the identification of the same FSH peptide (ETVRVPGCAHHADSLY) in all replicates analyses. The distribution of missed cleavages, representative of the completeness of the enzymatic digestion in all samples was very reproducible, as shown in the histograms in Figure 5.8.
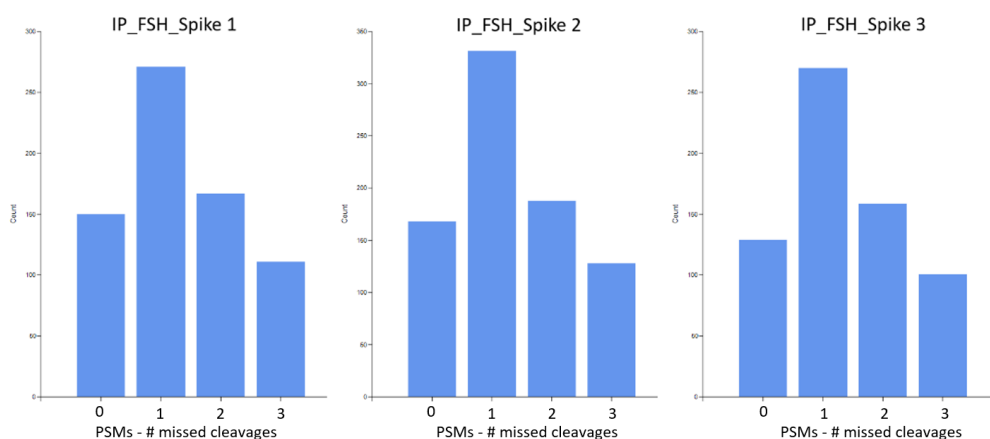


*Figure 5.8 Distribution missed cleavages in IP_FSH_spike 1, 2 and 3 samples.*

### 5.3.5 Glycopeptides enrichment prior PRM analysis

To investigate the presence of unidentified glycopeptides in serum, IP_hFSH sample was subjected to glycopeptide enrichment and a further MS-PRM experiment. Enriched glycopeptides PRM data analysis allowed the identification of 4 additional minority glycoforms that were not identified with previous MS-PRM experiment. Figure 5.5 reports the comparison of identified glycopeptides between all performed experiments on IP_hFSH serum sample (Full scan LC-MS/MS and MS-PRM, MS-PRM on enriched sample) and the standard RHS_FSH target glycopeptides (see  in supporting data paragraph for details). However, this analysis has the disadvantage of a further purification step, which could negatively affects the quantitative analysis. Therefore, we

processed the data obtained from the MS-PRM analysis performed on the non-enriched IP_hFSH sample for their subsequent quantitative analysis.

### 5.3.6   Comparison of glycoprofile of RHS_FSH and hFSH form serum

The preliminary building of a reference library containing target glycopeptides MS$^2$ spectra plays a central role in the workflow as both data acquisition and data processing methods. MS-PRM data acquisition method has the advantages to be relatively easy to build because *a priori* selection of target transitions is not required as in MS-MRM. The absolute quantification of serum hFSH glycopeptides is extremely difficult as there is no availability of isotope–labeled standard glycopeptides on the market. However, a relative quantification of hFSH glycosylation can be obtained thanks to the high specificity of PRM-MS. The high resolution allows the chromatographic traces to be extracted using tight mass tolerances, which improves the differentiation of analyte signals from co-eluting interferences, resulting in increased confidence and better analytical performance. The quantification is only applied to the glycopeptides with confirmed identities by MS$^2$ spectrum (Asn52-78 of α-subunit and Asn24 of β-subunit), and is performed by extracting, post-acquisition, the chromatographic traces of specific selected fragment using Skyline v20.1.0.76 software. Table 5.8 shows the method imported in skyline software for the relative quantification of FSH glycoprofile. Since the skyline software does not support glycosylation in the peptide modification setting, for each glycopeptide the aminoacidic sequence was manually modified on the asparagine by adding the glycan mass. As well as the fragment ions (including oxonium ions), already known from the spectra library previously built, were manually inserted in the transitions settings and then monitored.

The extracted ion currents (XIC) associated to the different transitions (precursor-ion → fragment-ion) for the same glycopeptide were recorded at the same retention time. For each glycopeptide, the most intense transition was

selected as quantifier as shown in Figure 5.9, which reports the chromatographic trace of the post-acquisition extraction of the best 3 fragment ion selected from the glycopeptide MS/MS spectrum.
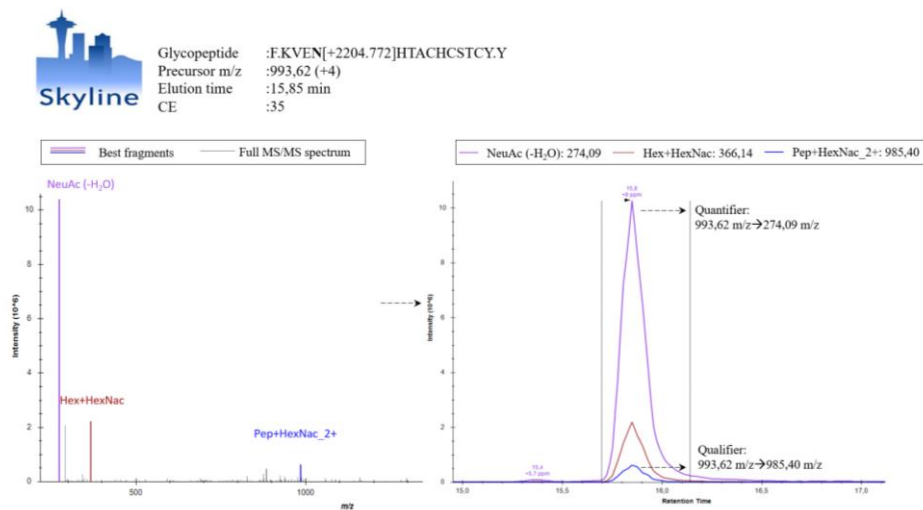


*Figure 5.9 Skyline chromatographic traces post-acquisition extraction the of best 3 selected fragment ion from glycopeptide F.KVEN[+2204.772]HTACHCSTCY.Y; the transitions 993,62 m/z → 274,09 m/z and 993,62 m/z → 985,40 m/z are respectively used for quantify and qualify the glycopeptide*

This data ensured the unequivocal identification and allowed us to disclose the relative distribution of glycans on each N-glycan site unless Asn7, as no MS$^2$ spectra were recorded at this site. As a proof of concept of the utility of the reported method, we compared the RHS_FSH with the hFSH from a pool of woman's sera (Table 5.4).

*Table 5.4 Glycopeptides relative quantification in RHS_FSH and IP_hFSH serum sample. The quantifier (Table 5.8) peak area from chromatographic traces post-acquisition extraction are recorded by skyline software. For each N-glycan site is reported the relative percentage abundance.*

| RT (min) | Peptides Sequence | N-Glycan Site | Glycan notation | RHS_FSH (Area) | IP_hFSH (Area) | Relative abundance (%) | Relative abundance (%) |
|---|---|---|---|---|---|---|---|
| **GLHA_HUMAN Glycoprotein hormones alpha chain** | | | | | | | |
| 23,24 | RSKKTMLVQKN[+1914]VTSESTCCVAKSY | | A2G2S1 | 2,00E+06 | N.D. | 20,64 | N.D. |
| 24,16 | RSKKTMLVQKN[+2205]VTSESTCCVAKSY | | A2G2S2 | 1,97E+06 | N.D. | 20,33 | N.D. |
| 20,96 | VQKN[+1458.4]VTSESTCCVAKSY | N52 | MAN6_Ph | 2,11E+05 | N.D. | 2,18 | N.D. |
| 24,26 | VQKN[+2205]VTSESTCCVAKSY | | A2G2S2 | 1,83E+06 | N.D. | 18,88 | N.D. |
| 20,84 | VQKN[+1914]VTSESTCCVAKSY | | A2G2S1 | 3,68E+06 | 7,48E+05 | 37,97 | 100 |
| 27,12 | NRVTVMGGFKVEN[+2205]HTACHCSTCY | | A2G2S2 | 1,77E+06 | N.D. | 1,47 | N.D. |
| 14,43 | KVEN[+1914]HTACHCSTCY | N78 | A2G2S1 | 9,92E+06 | N.D. | 8,25 | N.D. |
| 15,85 | KVEN[+2205]HTACHCSTCY | | A2G2S2 | 1,06E+08 | 2,70E+05 | 88,14 | 100 |
| 15,78 | KVEN[+2569.9]HTACHCSTCY | | A3G3S2 | 2,57E+06 | N.D. | 2,14 | N.D. |
| **FSHB_HUMAN Follitropin subunit beta** | | | | | | | |
| 47,54 | C[+57]ISIN[+2205]TTW | | F2A2G2S1 | 4,75E+05 | 8,32E+05 | 1,01 | 45,34 |
| 47,5 | C[+57]ISIN[+2716]TTW | | FA3G3S2 | 9,56E+05 | N.D. | 1,50 | N.D. |
| 47,51 | C[+57]ISIN[+2642.9]TTW | | A3G4S1 | 7,18E+05 | 1,04E+05 | 1,13 | 7,25 |
| 49,61 | C[+57]ISIN[+2351]TTW | | FA2G2S2 | 5,60E+07 | 4,29E+05 | 87,99 | 29,90 |
| 45,09 | C[+57]ISIN[+1930.7]TTW | | FA2G2_Gal1 | 4,77E+04 | 7,93E+04 | 0,07 | 5,53 |
| 47,51 | C[+57]ISIN[+2496.9]TTW | | F2A2G2S2 | 7,97E+05 | N.D. | 1,25 | N.D. |
| 44,08 | C[+57]ISIN[+1298.5]TTW | N24 | A2G0 | 2,75E+05 | N.D. | 0,43 | N.D. |
| 47,54 | C[+57]ISIN[+2205]TTW | | A2G2S2 | 4,75E+05 | N.D. | 0,75 | N.D. |
| 44,72 | C[+57]ISIN[+2076.8]TTW | | F2A2G3 | 4,28E+05 | N.D. | 0,67 | N.D. |
| 47,52 | C[+57]ISIN[+1914]TTW | | A2G2S1 | 2,20E+05 | 1,72E+05 | 0,35 | 11,99 |
| 36,93 | C[+57]ISIN[+1768.6]TTW | | FA2G2 | 1,26E+06 | N.D. | 1,98 | N.D. |
| 47,49 | C[+57]ISIN[+1458.4]TTW | | MAN6_Ph | 1,47E+05 | N.D. | 0,23 | N.D. |
| 44,7 | C[+57]ISIN[+2060]TTW | | FA2G2S1 | 1,68E+06 | N.D. | 2,64 | N.D. |

Figure 5.10 shows the PRM-MS relative quantification results expressed as relative percentage distribution of glycan moiety on each glycosylation site of both recombinant and serum purified hormone.
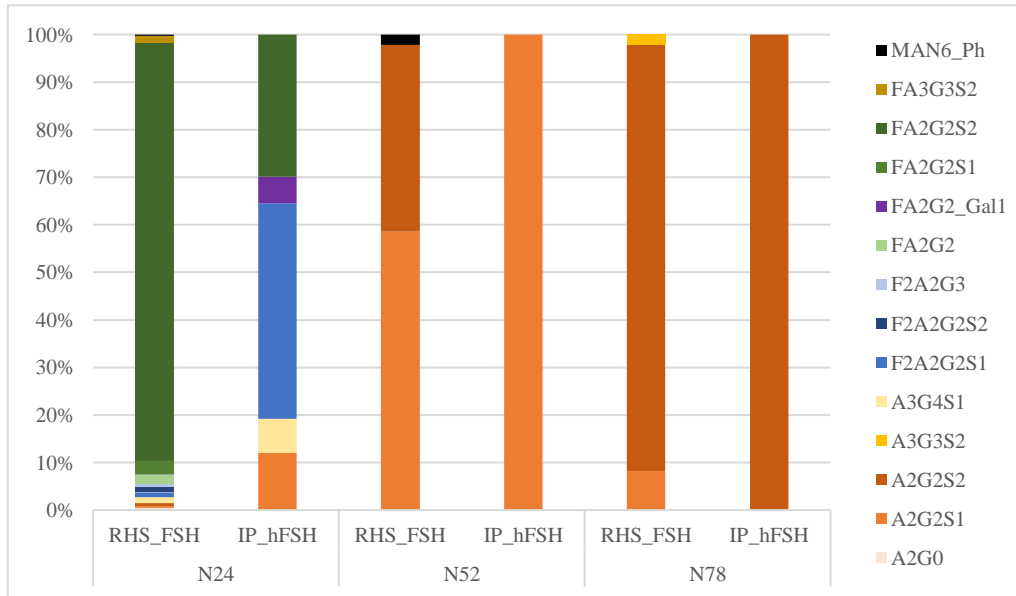


*Figure 5.10 PRM-MS relative quantification results expressed as relative percentage distribution of glycan moiety on each glycosylation site of both recombinant (RHS_FSH) and plasmatic (IP_hFSH) hormones.*

Some differences could be appreciated. In serum IP_hFSH sample we can observe the absence of high mannose (Man6) glycans structures on α-FSH Asn 52 and β-FSH Asn 24, while there is a similar high abundance of complex glycans which are largely coated with sialic acid; in both samples we found on Asn52 the predominance of the mono-sialylated complex biantennary structure (A2G2S1) while on Asn 78 the relative most abundant structure is the complete sialylated ones (A2G2S2). Asn 24 of the β-subunit shows more heterogeneity in glycan structures. Comparing the glycan structure abundance with serum IP_hFSH, recombinant protein exhibited higher number/level of biantennary glycans, most of them with core-fucosylation with respect to the serum FSH. This observation is in agreement with a previous study, which also suggests an

additional feature of recombinant FSH oligosaccharide structures, the presence of antenna-fucosylation which is rare in serum hFSH [4, 35].

## 5.4    Conclusions

Glycopeptide analysis is a very challenging task that requires expertise in sample preparation, high sensitive and specific analytical methods as long as dedicated software to analyse data. In the last part of this thesis project, PRM-MS based methodology was developed for the identification and relative site-specific quantification of serum hFSH glycan moiety.

Untargeted full scan LC-MS/MS analysis of serum hFSH, preceded by immunopurification steps, allowed the identification of a single FA2G2S2 glycopeptide (CISIN$_{24}$TTW) of the beta chain. The developed MS-PRM method improved glycopeptide untargeted analysis results leading to the identification of 9 most abundant serum hFSH glycopeptides on each of the four hFSH N-glycosites (Figure 5.5).

Moreover, this methodology allowed the relative quantification, expressed as relative percentage distribution of glycan moiety, of each identified glycopeptides by extracting post-data acquisition from the chromatographic traces the most intense glycopeptide fragments (Figure 5.10). In both recombinant and serum hFSH the same most abundant glycoforms were founded: i) the biantennary fully sialylated (A2G2S2) on glycosite α-Asn52; ii) the mono-sialylated (A2G2S1) on α-Asn78 and iii) fucosylated biantennary (FA2G2S2 and FA2G2S1) structures on β-Asn 24 glycosite. It was not possible to quantify glycoforms on β-Asn 7 because of the low quality of the fragmentation spectra whit not sufficient quality criteria for quantitative analysis.

This method, if applied to a large number of serum samples from women in different physiological conditions and age, could provide useful data to clarify the functional role of hFSH glycans in human fertility. These speculations will

lead to a step forward in delineating the clinical relevance of some glycoforms rather than others with the ultimate goal of developing or improving drugs for the treatment of infertility.

An important advantage of this methodology is its ability to be implemented with other target proteins by detecting multiple analytes simultaneously without affecting the specificity and sensitivity of the method [36], unlike clinical multiplexed ELISA. Given the numerous advantages of mass spectrometry in PRM mode, in the near future, the proposed method could be complementary to ELISA tests commonly used as a reference for the detection of FSH in the clinical setting, obviating its major limitation, the discrimination of different glycoforms of this hormone.

## 5.5 Supporting data

*Table 5.5 Full LC-MS/MS identification of RHS_FSH standard. Each identified glycopeptides were further validate by interpreting MS/MS spectrum. In green is highlighted the glycopeptides used for the validation example showed in Figure 5.2*

| N-Glycan site | Peptide | Glycan notation | Exp m/z | z | Exp (M+H) | Th mass (M+H) | Scan Time | MS/MS |
|---|---|---|---|---|---|---|---|---|
| **GLHA_HUMAN Glycoprotein hormones alpha chain** | | | | | | | | |
| N78 | F.KVEN[+1913,677]HTACHCSTCY.Y | A2G2S1 | 1227,47 | 3 | 3680,38 | 3680,38 | 14,52 | X |
| N78 | F.KVEN[+2204,772]HTACHCSTCY.Y | A2G2S2 | 993,63 | 4 | 3971,48 | 3971,48 | 15,24 | X |
| N78 | F.KVEN[+2569,905]HTACHCSTCY.Y | A3G3S2 | 1084,91 | 4 | 4336,61 | 4336,61 | 15,24 | X |
| N52 | L.VQKN[+1913,677]VTSESTCCVAKSY.N | A2G2S1 | 969,40 | 4 | 3874,59 | 3874,59 | 20,17 | X |
| N52 | L.VQKN[+1458,442]VTSESTCCVAKSY.N | A3G3S3 | 1140,46 | 3 | 3419,35 | 3419,35 | 20,27 | |
| N52 | L.RSKKTMLVQKN[+2204,772]VTSESTCCVAKSY.N | A2G2S2 | 1253,30 | 4 | 5010,18 | 5010,18 | 20,28 | X |
| N52 | L.RSKKTMLVQKN[+1458,442]VTSESTCCVAKSY.N | Man6_P | 853,58 | 5 | 4263,85 | 4263,85 | 18,67 | X |
| N52 | L.VQKN[+2204,772]VTSESTCCVAKSY.N | A2G2S2 | 1389,23 | 3 | 4165,68 | 4165,68 | 23,38 | X |
| N78 | Y.NRVTVMGGFKVEN[+2204,772]HTACHCSTCY.Y | A2G2S2 | 1234,00 | 4 | 4932,96 | 4932,96 | 25,31 | X |
| **FSHB_HUMAN Follitropin subunit beta** | | | | | | | | |
| N-Glycan site | Peptide | Short name | Observed m/z | z | Observed (M+H) | Calc. mass (M+H) | Scan Time | MS/MS |
| N24 | F.CISIN[+1768,640]TTW.C | FA2G2 | 1382,05 | 2 | 2763,10 | 2763,11 | 35,58 | X |
| N24 | F.CISIN[+2205,793]TTW.C | F2A2G2S1 | 1067,43 | 3 | 3200,26 | 3200,26 | 42,85 | X |
| N25 | F.C[+57,021]ISIN[+1458,442]TTW.C | Man6_P | 1226,96 | 2 | 2452,91 | 2452,91 | 45,07 | X |
| N24 | F.CISIN[+2059,735]TTW.C | FA2G2S1 | 1527,60 | 2 | 3054,20 | 3054,20 | 43,11 | X |
| N24 | F.CISIN[+2076,750]TTW.C | F2A2G3 | 1024,41 | 3 | 3071,23 | 3071,22 | 43,6 | X |
| N24 | F.CISIN[+1913,677]TTW.C | A2G2S1 | 1454,58 | 2 | 2908,15 | 2908,14 | 43,64 | X |
| N24 | F.CISIN[+1930,692]TTW.C | FA2G3 | 975,73 | 3 | 2925,17 | 2925,16 | 43,67 | X |
| N24 | F.CISIN[+2715,963]TTW.C | FA3G3S2 | 1237,48 | 3 | 3710,43 | 3710,43 | 48,88 | X |
| N24 | F.CISIN[+2643.941]TTW.C | A3G4S1 | 1819,20 | 2 | 3637,40 | 3637,41 | 49,09 | X |
| N24 | F.CISIN[+2496,888]TTW.C | F2A2G2S2 | 1164,46 | 3 | 3491,35 | 3491,36 | 49,78 | X |
| N24 | F.CISIN[+2350,830]TTW.C | FA2G2S2 | 1115,77 | 3 | 3345,30 | 3345,30 | 49,98 | X |
| N24 | F.CISIN[+2204,772]TTW.C | A2G2S2 | 1067,09 | 3 | 3199,24 | 3199,24 | 49,99 | X |
| N24 | F.CISIN[+1298.476]TTW.C | A2G0 | 1146,97 | 2 | 2292,94 | 2292,94 | 43,41 | X |
| N7 | L.TN[+2715.962]ITIAIEKEECRF.C | FA3G3S2 | 1480,89 | 3 | 4440,67 | 4439,83 | 29,44 | |
| N7 | L.TN[+2465.894]ITIAIEKEECRF.C | FA4G2S1 | 1397,27 | 3 | 4189,81 | 4189,69 | 26,37 | |
| N7 | L.TN[+2204.772]ITIAIEKEECRF.C | A2G2S2 | 1310,17 | 3 | 3928,51 | 3928,63 | 17,41 | |
| N7 | L.TN[+2350,830]ITIAIEKEECRF.C | FA2G2S2 | 1358,90 | 3 | 4074,70 | 4074,70 | 22,91 | |

*Table 5.6 LC-MS/MS identification of IP_hFSH serum sample. Mass spectrometry data were analysed by Proteome Discoverer v2.4 (Thermo Scientific) against human subset from UniProtKB/Swiss-Prot complete proteome database using the Byonic node search engine (v3.6.0 Protein Metrics, PMI-Suite).*

| Accession number | Protein Description | # Peptides | # PSMs | # Unique Peptides | # AAs | MW [kDa] |
|---|---|---|---|---|---|---|
| P08603 | Complement factor | 114 | 312 | 116 | 1231 | 139.0 |
| P02768 | Serum albumin | 78 | 302 | 96 | 609 | 69.3 |
| P0DOX5 | Immunoglobulin gamma-1 heavy chain | 44 | 295 | 46 | 449 | 49.3 |
| P01871 | Immunoglobulin heavy constant mu | 52 | 240 | 53 | 453 | 49.4 |
| P0DOX7 | Immunoglobulin kappa light chain | 25 | 125 | 25 | 214 | 23.4 |
| P01023 | Alpha-2-macroglobulin | 55 | 106 | 62 | 1474 | 163.2 |
| P01859 | Immunoglobulin heavy constant gamma 2 | 15 | 98 | 16 | 326 | 35.9 |
| P01876 | Immunoglobulin heavy constant alpha 1 | 37 | 98 | 37 | 353 | 37.6 |
| P02787 | Serotransferrin | 41 | 97 | 47 | 698 | 77.0 |
| P07358 | Complement component C8 beta chain | 45 | 90 | 45 | 591 | 67.0 |
| P02746 | Complement C1q subcomponent subunit B | 20 | 88 | 21 | 253 | 26.7 |
| P0DOY2 | Immunoglobulin lambda constant 2 | 15 | 76 | 15 | 106 | 11.3 |
| P07357 | Complement component C8 alpha chain | 39 | 76 | 39 | 584 | 65.1 |
| P01024 | Complement C3 | 43 | 70 | 48 | 1663 | 187 |
| P17538 | Chymotrypsinogen B | 2 | 68 | 7 | 263 | 27.7 |
| Q6GPI1 | Chymotrypsinogen B2 | 2 | 68 | 7 | 263 | 27.9 |
| P02751 | Fibronectin | 47 | 67 | 48 | 2386 | 262.5 |
| P00450 | Ceruloplasmin | 38 | 56 | 42 | 1065 | 122.1 |
| P36955 | Pigment epithelium-derived factor | 29 | 54 | 29 | 418 | 46.3 |
| O43866 | CD5 antigen-like | 30 | 52 | 30 | 347 | 38.1 |
| P02747 | Complement C1q subcomponent subunit C | 16 | 50 | 16 | 245 | 25.8 |
| P01860 | Immunoglobulin heavy constant gamma 3 | 11 | 45 | 11 | 377 | 41.3 |
| P02745 | Complement C1q subcomponent subunit A | 14 | 39 | 15 | 245 | 26.0 |
| **P01215** | **Glycoprotein hormones alpha chain** | **2** | **34** | **9** | **116** | **13.1** |
| P01009 | Alpha-1-antitrypsin | 17 | 34 | 18 | 418 | 46.7 |
| P0DOX2 | Immunoglobulin alpha-2 heavy chain | 13 | 33 | 13 | 455 | 48.9 |
| P00738 | Haptoglobin | 15 | 30 | 15 | 406 | 45.2 |
| A0A0A0MS14 | Immunoglobulin heavy variable 1-45 | 5 | 30 | 5 | 117 | 13.5 |
| P01619 | Immunoglobulin kappa variable 3-20 | 11 | 30 | 8 | 116 | 12.5 |
| P01701 | Immunoglobulin lambda variable 1-51 | 9 | 30 | 9 | 117 | 12.2 |
| P0DOX8 | Immunoglobulin lambda-1 light chain | 10 | 27 | 10 | 216 | 22.8 |
| Q03591 | Complement factor H-related protein 1 | 19 | 27 | 19 | 330 | 37.6 |
| **P01225** | **Follitropin subunit beta** | **3** | **26** | **10** | **129** | **14.7** |
| P02790 | Hemopexin | 11 | 23 | 14 | 462 | 51.6 |
| P02647 | Apolipoprotein A-I | 15 | 21 | 16 | 267 | 30.8 |
| P07360 | Complement component C8 gamma chain | 14 | 21 | 15 | 202 | 22.3 |

| | | | | | |
|---|---|---|---|---|---|
| A0A075B6P5 | Immunoglobulin kappa variable 2-28 | 12 | 20 | 12 | 120 | 12.9 |
| P02743 | Serum amyloid P-component | 12 | 18 | 13 | 223 | 25.4 |
| P06312 | Immunoglobulin kappa variable 4-1 | 10 | 18 | 10 | 121 | 13.4 |
| P01704 | Immunoglobulin lambda variable 2-14 | 6 | 16 | 6 | 120 | 12.6 |
| A0A0A0MRZ8 | Immunoglobulin kappa variable 3D-11 | 5 | 14 | 6 | 115 | 12.6 |
| P04433 | Immunoglobulin kappa variable 3-11 | 5 | 14 | 6 | 115 | 12.6 |
| P01624 | Immunoglobulin kappa variable 3-15 | 7 | 13 | 7 | 115 | 12.5 |
| P02749 | Beta-2-glycoprotein 1 | 9 | 12 | 9 | 345 | 38.3 |
| A0A0B4J1X5 | Immunoglobulin heavy variable 3-74 | 3 | 12 | 3 | 117 | 12.8 |
| A0A0B4J1V1 | Immunoglobulin heavy variable 3-21 | 5 | 12 | 4 | 117 | 12.8 |
| P23083 | Immunoglobulin heavy variable 1-2 | 5 | 12 | 5 | 117 | 13.1 |
| P01700 | Immunoglobulin lambda variable 1-47 | 6 | 10 | 6 | 117 | 12.3 |
| Q969X2 | Alpha-N-acetylgalactosaminide alpha-2.6-sialyltransferase 6 | 2 | 10 | 2 | 333 | 38.0 |
| P02766 | Transthyretin | 6 | 10 | 6 | 147 | 15.9 |
| P00751 | Complement factor B | 7 | 9 | 8 | 764 | 85.5 |
| P01011 | Alpha-1-antichymotrypsin | 4 | 7 | 6 | 423 | 47.6 |
| A0A0B4J1V6 | Immunoglobulin heavy variable 3-73 | 2 | 5 | 3 | 119 | 12.8 |
| P04217 | Alpha-1B-glycoprotein | 3 | 5 | 3 | 495 | 54.2 |

*Table 5.7 Comparison of identified glycopeptides between all performed experiments: Full scan LC-MS/MS and MS-PRM on IP_hFSH and MS-PRM on IP_hFSH_Glycoenrich*

| GLHA_HUMAN Glycoprotein hormones alpha chain | | | | | | | |
|---|---|---|---|---|---|---|---|
| Peptide sequence | Glycan site | N-glycan | RHS_FSH | IP_hFHS | | IP_hFHS enrich | Note |
| | | short name | PRM | Full scan | PRM | PRM | |
| L.RSKKTMLVQKN*VTSESTCCVAKSY.N | N52 | A2G2S2 | x | | | | |
| | | A2G2S1 | x | | | | |
| L.VQKN*VTSESTCCVAKSY.N | | A2G2S1 | x | | x | x | |
| | | A3G3S3 | x | | | | NO MS/MS |
| | | A2G2S2 | x | | | | |
| | | Man6_Ph | x | | | | |
| Y.NRVTVMGGFKVEN*HTACHCSTCY.Y | N78 | A2G2S2 | x | | | | |
| F.KVEN*HTACHCSTCY.Y | | A3G3S2 | x | | | | |
| | | A2G2S2 | x | | x | x | |
| | | A2G2S1 | x | | | x | |

| FSHB_HUMAN Follitropin subunit beta | | | | | | | |
|---|---|---|---|---|---|---|---|
| Peptide sequence | Glycan site | N-glycan | RHS_FSH | IP_hFHS | | IP_hFHS_enrich | Note |
| | | short name | Full scan | Full scan | PRM | PRM | |
| L.TN*ITIAIEKEECRF.C | N7 | FA3G3S2 | x | | x | | NO MS/MS |
| | | FA4G2S1 | x | | | x | NO MS/MS |
| | | A2G2S2 | x | | | | NO MS/MS |
| | | FA2G2S2 | x | | | x | NO MS/MS |
| F.CISIN*TTW.C | N24 | A2G2S2 | x | | | x | |
| | | F2A2G2S1 | x | | x | x | |
| | | FA3G3S2 | x | | | x | |
| | | A3G4S1 | x | | x | x | |
| | | FA2G2S2 | x | x | x | x | |
| | | FA2G2Gal_1 | x | | x | x | |
| | | F2A2G2S2 | x | | | | |
| | | FA2G2S1 | x | | | | |
| | | A2G0 | x | | | | |
| | | F2A2G2Gal_1 | x | | | | |
| | | A2G2S1 | x | | x | x | |
| | | FA2G2 | x | | | | |
| | | FA3G3S1 | | | x | x | NO MS/MS |
| | | Man3_Ph | x | | | | |

186

*Table 5.8 Glycopeptides transitions (Precursor Mz → Product Mz) were selected post-data acquisition by PRM-MS analysis. Glycopeptides were identified by coelution of at least 3 transitions and the highest one (**bolded***) was selected as quantifier.*

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| **GLHA_HUMAN Glycoprotein hormones alpha chain** | | | | | | | |
| N-glyco site | Peptide Modified Sequence | Glycan | Precursor Mz | Precursor Charge | Product Mz | Product Charge | Fragment Ion description |
| N52 | RSKKTMLVQKN[+1914]VTSESTCCVAKSY | A2G2S1 | 944,6869 | 5 | 292,08 | 1 | NeuAc |
| | | | | | **366,14*** | 1 | Hex+HexNac |
| | | | | | 528,19 | 1 | HexNac+2Hex |
| | | | | | 1505,27 | 2 | Pep+HexNac |
| | RSKKTMLVQKN[+2205]VTSESTCCVAKSY | A2G2S2 | 1253,357 | 4 | 292,08 | 1 | NeuAc |
| | | | | | **366,14*** | 1 | Hex+HexNac |
| | | | | | 528,19 | 1 | HexNac+2Hex |
| | | | | | 1505,27 | 2 | Pep+HexNac |
| | VQKN[+1458.4]VTSESTCCVAKSY | MAN6_P | 1140,456 | 3 | 204,09 | 1 | HexNac |
| | | | | | 243,03 | 1 | Hex_P |
| | | | | | **366,14*** | 1 | Hex+HexNac |
| | | | | | 405,08 | 1 | 2Hex_P |
| | VQKN[+2205]VTSESTCCVAKSY | A2G2S2 | 1389,308 | 3 | 204,09 | 1 | HexNac |
| | | | | | 292,08 | 1 | NeuAc |
| | | | | | **366,14*** | 1 | Hex+HexNac |
| | | | | | 1391,58 | 2 | y12 |
| | VQKN[+1914]VTSESTCCVAKSY | A2G2S1 | 1292,308 | 3 | 204,09 | 1 | HexNac |
| | | | | | **274,09*** | 1 | NeuAc-18 |
| | | | | | 292,08 | 1 | NeuAc |
| | | | | | 366,14 | 1 | Hex+HexNac |
| N78 | NRVTVMGGFKVEN[+2205]HTACHCSTCY | A2G2S2 | 1234,052 | 4 | 292,08 | 1 | NeuAc |
| | | | | | **366,14*** | 1 | Hex+HexNac |
| | | | | | 512,2 | 1 | Fuc+Hex+HexNac |
| | | | | | 1465,66 | 2 | Pep+HexNac |
| | KVEN[+1914]HTACHCSTCY | A2G2S1 | 920,9315 | 4 | 204,09 | 1 | HexNac |
| | | | | | **274,09*** | 1 | NeuAc-18 |
| | | | | | 366,14 | 1 | Hex+HexNac |
| | | | | | 985,4 | 2 | Pep+HexNac |
| | KVEN[+2205]HTACHCSTCY | A2G2S2 | 993,6815 | 4 | **274,09*** | 1 | NeuAc-18 |
| | | | | | 292,08 | 1 | NeuAc |
| | | | | | 366,14 | 1 | Hex+HexNac |
| | | | | | 985,4 | 2 | Pep+HexNac |
| | KVEN[+2569.9]HTACHCSTCY | A3G3S2 | 1084,908 | 4 | **366,14*** | 1 | Hex+HexNac |
| | | | | | 528,19 | 1 | HexNac+2Hex |
| | | | | | 985,4 | 2 | Pep+HexNac |

| FSHB_HUMAN Follitropin subunit beta | | | | | | |
|---|---|---|---|---|---|---|
| **N-glyco site** | **Peptide Modified Sequence** | **Glycan** | **Precursor Mz** | **Precursor Charge** | **Product Mz** | **Product Charge** | **Fragment Ion description** |
| N24 | CISIN[+2205.8]TTW | F2A2G2S1 | 1067,424 | 3 | **274,09*** | 1 | NeuAc-18 |
| | | | | | 366,14 | 1 | Hex+HexNac |
| | | | | | 671,31 | 2 | Pep+HexNac+Fuc |
| | | | | | 1067,42 | 3 | precursor |
| | CISIN[+2716]TTW | FA3G3S2 | 1855,717 | 2 | 274,09 | 1 | NeuAc-18 |
| | | | | | **366,14*** | 1 | Hex+HexNac |
| | | | | | 671,31 | 2 | Pep+HexNac+Fuc |
| | CISIN[+2642.9]TTW | A3G4S1 | 1819,207 | 2 | **274,09*** | 1 | NeuAc-18 |
| | | | | | 292,08 | 1 | NeuAc |
| | | | | | 366,14 | 1 | Hex+HexNac |
| | CISIN[+2351]TTW | FA2G2S2 | 1115,827 | 3 | **274,09*** | 1 | NeuAc-18 |
| | | | | | 366,14 | 1 | Hex+HexNac |
| | | | | | 671,31 | 2 | Pep+HexNac+Fuc |
| | CISIN[+1930.7]TTW | FA2G2_Gal1 | 975,7243 | 3 | 204,09 | 1 | HexNac |
| | | | | | **274,12*** | 1 | b2 |
| | | | | | 366,14 | 1 | Hex+HexNac |
| | | | | | 512,2 | 1 | Fuc+Hex+HexNac |
| | CISIN[+2496.9]TTW | F2A2G2S2 | 1746,181 | 2 | **274,09*** | 1 | NeuAc-18 |
| | | | | | 292,08 | 1 | NeuAc |
| | | | | | 366,14 | 1 | Hex+HexNac |
| | CISIN[+1298.5]TTW | A2G0 | 1146,975 | 2 | 204,09 | 1 | HexNac |
| | | | | | **274,12*** | 1 | b2 |
| | | | | | 366,14 | 1 | Hex+HexNac |
| | CISIN[+2076.8]TTW | F2A2G3 | 1536,112 | 2 | 274,12 | 1 | b2 |
| | | | | | 361,15 | 1 | b3 |
| | | | | | **366,14*** | 1 | Hex+HexNac |
| | CISIN[+1914]TTW | A2G2S1 | 1454,737 | 2 | 204,09 | 1 | HexNac |
| | | | | | 292,08 | 1 | NeuAc |
| | | | | | **366,14*** | 1 | Hex+HexNac |
| | CISIN[+1768.6]TTW | FA2G2 | 1382,057 | 2 | **366,14*** | 1 | Hex+HexNac |
| | | | | | 512,2 | 1 | Fuc+Hex+HexNac |
| | | | | | 671,31 | 2 | Pep+HexNac+Fuc |
| | | | | | 1382,05 | 2 | precursor |
| | CISIN[+1458.4]TTW | MAN6_P | 1226,958 | 2 | 243,03 | 1 | Hex_P |
| | | | | | **366,14*** | 1 | Hex+HexNac |
| | | | | | 405,08 | 1 | 2Hex_P |
| | | | | | 1226,96 | 2 | precursor |

| | | | | 274,09* | 1 | NeuAc-18 |
|---|---|---|---|---|---|---|
| CISIN[+2060]TTW | FA2G2S1 | 1527,737 | 2 | 292,08 | 1 | NeuAc |
| | | | | 366,14 | 1 | Hex+HexNac |
| | | | | 671,31 | 2 | Pep+HexNac+Fuc |
| ETVRVPGCAHHADSLY | - | 906.4309 | 2 | 970, 48 | 1 | b9 |
| | | | | 1107,54 | 1 | b10 |
| | | | | 715,83 | 2 | b13 |

## 5.6    References

1.    Rose, M.P., R.E. Gaines Das, and A.H. Balen, *Definition and measurement of follicle stimulating hormone.* Endocrine Reviews, 2000. **21**(1): p. 5-22.
2.    Bishop, L.A., et al., *Specific roles for the asparagine-linked carbohydrate residues of recombinant human follicle stimulating hormone in receptor binding and signal transduction.* Molecular Endocrinology, 1994. **8**(6): p. 722-731.
3.    Dias, J.A., et al., *Human follicle-stimulating hormone structure-activity relationships.* Biology of reproduction, 1998. **58**(6): p. 1331-1336.
4.    Bousfield, G.R., et al., *Hypo-glycosylated human follicle-stimulating hormone (hFSH21/18) is much more active in vitro than fully-glycosylated hFSH (hFSH24).* Molecular and cellular endocrinology, 2014. **382**(2): p. 989-997.
5.    Jiang, C., et al., *Hypoglycosylated hFSH has greater bioactivity than fully glycosylated recombinant hFSH in human granulosa cells.* The Journal of Clinical Endocrinology & Metabolism, 2015. **100**(6): p. E852-E860.
6.    D'Antonio, M., et al., *Biological characterisation of recombinant human follicle stimulating hormone isoforms.* Human Reproduction, 1999. **14**(5): p. 1160-1167.
7.    Emperaire, J.C., *Review of physiology, Ovulation stimulation with gonadotropins.* 2015, Geneva: Springer International Publishing.
8.    Davis, J.S., et al., *Naturally occurring follicle-stimulating hormone glycosylation variants.* Journal of glycomics & lipidomics, 2014. **4**(1): p. e117.
9.    Li, H. and M. d'Anjou, *Pharmacological significance of glycosylation in therapeutic proteins.* Current opinion in biotechnology, 2009. **20**(6): p. 678-684.
10.    Sinclair, A.M. and S. Elliott, *Glycoengineering: the effect of glycosylation on the properties of therapeutic proteins.* Journal of pharmaceutical sciences, 2005. **94**(8): p. 1626-1635.
11.    Ulloa-Aguirre, A., et al., *Role of glycosylation in function of follicle-stimulating hormone.* Endocrine, 1999. **11**(3): p. 205-215.
12.    Daya, S., *Follicle-Stimulating Hormone in Clinical Practice.* Treatments in endocrinology, 2004. **3**(3): p. 161-171.
13.    Mastrangeli, R., et al., *In-vivo biological activity and glycosylation analysis of a biosimilar recombinant human follicle-stimulating hormone product (Bemfola) compared with its reference medicinal product (GONAL-f).* PloS one, 2017. **12**(9).

14. Scheele, F. and J. Schoernaker, *The role of follicle-stimulating hormone in the selection of follicles in human ovaries: a survey of the literature and a proposed model.* Gynecological Endocrinology, 1996. **10**(1): p. 55-66.

15. Harvey, D.J., et al., *Differentiation between isomeric triantennary N-linked glycans by negative ion tandem mass spectrometry and confirmation of glycans containing galactose attached to the bisecting (β1-4-GlcNAc) residue in N-glycans from IgG.* Rapid Communications in Mass Spectrometry, 2008. **22**(7): p. 1047-1052.

16. Srikanth, J., R. Agalyadevi, and P. Babu, *Targeted, Site-specific quantitation of N-and O-glycopeptides using 18 O–labeling and product ion based mass spectrometry.* Glycoconjugate journal, 2017. **34**(1): p. 95-105.

17. Kettenbach, A.N., J. Rush, and S.A. Gerber, *Absolute quantification of protein and post-translational modification abundance with stable isotope–labeled synthetic peptides.* Nature Protocols, 2011. **6**(2): p. 175-186.

18. Butnev, V.Y., et al., *Production, purification, and characterisation of recombinant hFSH glycoforms for functional studies.* Molecular and cellular endocrinology, 2015. **405**: p. 42-51.

19. PEKONEN, F., D.M. WILLIAMS, and B.D. WEINTRAUB, *Purification of thyrotropin and other glycoprotein hormones by immunoaffinity chromatography.* Endocrinology, 1980. **106**(5): p. 1327-1332.

20. Wang, Y., S.-l. Wu, and W.S. Hancock, *Approaches to the study of N-linked glycoproteins in human plasma using lectin affinity chromatography and nano-HPLC coupled to electrospray linear ion trap—Fourier transform mass spectrometry.* Glycobiology, 2006. **16**(6): p. 514-523.

21. Dufau, M., T. Tsuruhara, and K. Catt, *Interaction of glycoprotein hormones with agarose-concanavalin A.* Biochimica et Biophysica Acta (BBA)-Protein Structure, 1972. **278**(2): p. 281-292.

22. Idler, D., L. Bazar, and S. Hwang, *Fish gonadotropin (s). II. Isolation of gonadotropin (s) from chum salmon pituitary glands using affinity chromatography.* Endocrine research communications, 1975. **2**(3): p. 215-235.

23. Ui, N., et al., *Bioaffinity chromatography of thyrotropin using immobilized concanavalin A.* Biochimica et Biophysica Acta (BBA)-General Subjects, 1977. **497**(3): p. 812-815.

24. Bloomfield, G.A., M.R. Faith, and J.G. Pierce, *Sepharose-linked concanavalin A in the purification and characterisation of glycoprotein hormones of the bovine pituitary.* Biochimica et biophysica acta, 1978. **533**(2): p. 371-382.

25. NISULA, B.C. and J.-P. LOUVET, *Radioimmunoassay of thyrotropin concentrated from serum.* The Journal of Clinical Endocrinology & Metabolism, 1978. **46**(5): p. 729-733.

26. Weintraub, B.D., *Concentration and purification of human chorionic somato-mammotropin (HCS) by affinity chromatography: application to radioimmunoassay.* Biochemical and biophysical research communications, 1970. **39**(1): p. 83-89.

27. Akanuma, Y., et al., *Immunological reactivity of insulin to Sepharose coupled with insulin-antibody—its use for the extraction of insulin from serum.* Biochemical and biophysical research communications, 1970. **38**(5): p. 947-953.

28. Murphy, R., D. Elmore, and K. Buchanan, *Isolation of glucagon-like immunoreactivity of gut by affinity chromatography.* Biochemical Journal, 1971. **125**(3): p. 61P.

29. Murphy, R.F., K.D. Buchanan, and D.T. Elmore, *Isolation of glucagon-like immunoreactivity of gut by affinity chromatography on anti-glucagon antibodies coupled to Sepharose 4B.* Biochimica et Biophysica Acta (BBA)-Protein Structure, 1973. **303**(1): p. 118-127.

30. Yang, Z. and W.S. Hancock, *Approach to the comprehensive analysis of glycoproteins isolated from human serum using a multi-lectin affinity column.* Journal of chromatography A, 2004. **1053**(1-2): p. 79-88.

31. Dalpathado, D.S., et al., *Comparative glycomics of the glycoprotein follicle stimulating hormone: glycopeptide analysis of isolates from two mammalian species.* Biochemistry, 2006. **45**(28): p. 8665-8673.

32. Singh, V., et al., *Characterisation of human FSH isoforms reveals a nonglycosylated b-subunit in addition to the conventional glycosylated b-subunit.* 2010.

33. Yang, Z., et al., *A study of glycoproteins in human serum and plasma reference standards (HUPO) using multilectin affinity chromatography coupled with RPLC-MS/MS.* PROTEOMICS, 2005. **5**(13): p. 3353-3366.

34. Anobile, C.J., et al., *Glycoform composition of serum gonadotrophins through the normal menstrual cycle and in the post-menopausal state.* Molecular Human Reproduction, 1998. **4**(7): p. 631-639.

35. Bousfield, G.R., et al., *Comparison of Follicle-Stimulating Hormone Glycosylation Microheterogenity by Quantitative Negative Mode Nano-Electrospray Mass Spectrometry of Peptide-N Glycanase-Released Oligosaccharides.* J Glycomics Lipidomics, 2015. **5**(1).

36. Gallien, S. and B. Domon, *Detection and quantification of proteins in clinical samples using high resolution mass spectrometry.* Methods, 2015. **81**: p. 15-23.

# Chapter 6    Final Conclusions

Glycosylation has a high impact on the biological activity of the hormones [1-4]. Follicle Stimulating Hormone (FSH) is one of the most important glycohormone involved in human fertility and its glycosylation plays a decisive role in the biological activity of this hormone. FSH glycosylation is known to affect the binding affinity to its receptor and to have a crucial role in metabolic clearance [5-9], but the correlation between the glycosylation structures and the hormone activity is still not fully understood [10-13]. Therefore, glycomic studies on FSH could bring key information to understanding the molecular mechanisms of action of FSH and its role in human fertility.

Glycomic studies on circulating hFSH suffer several limitations: they are very tricky because of the huge heterogeneity of glycoforms combined with their very low concentration in serum. Furthermore, quantification and characterization of the circulating glycoforms of this hormone present analytical issues, like the glycopeptide low ionization efficiency and the lack of isotope–labeled synthetic glycopeptides to be used as standards for the absolute quantification. Lastly, the interpretation of MS/MS spectra from glycopeptides is not fully automated, and the reconstruction of the full glycan structure not always possible. Altogether, these factors render the development of MS-based detection methods a great challenge. On the other hand, MS analysis can bring a specificity in distinguishing each glycoform that is superior to immunodetection based methods.

The main achievement of this PhD project is the development of an innovative MS-based targeted methodology allowing the detection and the quantification of different glycoforms of circulating hFSH. This analytical development was performed in two steps. First, the characterization of different recombinant FSH molecules used in the commercial drugs was performed by untargeted proteomics and glycan analysis. Second, MRM/PRM-MS targeted methods, characterised by

an higher specificity and sensitivity, were developed to detect and quantify site specific glycoforms of endogenous hFSH in serum.

As concern the first part, it should be notice that there are several FSH-based drugs for *in vitro* fertilization approaches [14, 15] available on the market exhibiting different glycosylation profiles [16]. Among these FSH-based drugs, GONAL-F ® (Follitropin alpha, Merck) represents the reference medicine for all the biosimilar products [16, 17]. The analysis of site-specific glycosylation profiles of different commercial drugs (GONAL-F ®, BEMFOLA, OVALEAP, FOSTIMON) conducted in this thesis allowed to identify some possibly crucial biochemical differences in glycosylation of these biosimilars. BEMFOLA and OVALEAP showed the presence of bulkier glycan structures and greater sialylation extent than GONAL-F ®. The urinary commercial FSH (FOSTIMON) showed a glycosylation pattern with an even higher level of antennary and higher extent of sialylation compared to GONAL-F ® and the other two biosimilars BEMFOLA and OVALEAP. These results are in agreement with existing literature data [18, 19] on different r-hFSH preparations compared to GONAL-F ® displaying nominally identical polypeptide chains but a somewhat different glycosylation pattern [16]. Since glycosylation has a direct impact on the biological activity of the hormone, the differences detected in glycan moieties of biosimilars could be related to the different efficacy played by the drugs [20]. *Mastrangeli et al.* speculates on differences in the glycosylation content between BEMFOLA and GONAL-F ®. The decrease of glomerular filtration caused by the presence of bulkier glycans could be counterbalanced by a higher proportion of non-fully sialylated antennae in BEMFOLA to achieve similar pharmacokinetic results [16].

As different glycosylation composition clearly influences the function of the hormone, and its efficiency on the treatment, seven glycovariants were artificially induced and characterised biochemically. The fucosylation of the molecule results to be unaltered by any of the stress conditions adopted in this study. In

contrast, other glycan groups -sialylation, hyper O-acetylation and galactosylation- were altered in almost each variant. Glycosylation of the molecule subject to pH stress condition results in a slight decrease of the sialic acid content. The removal of the end-capping sialic acid by acid/basic hydrolysis exposes the galactose residue, thus increasing the galactosylation of the molecule. The hyper O-acetylated sialic content was also decreased in both stressed variants, especially in basic pH stressed one. AEX-pH gradient chromatography performed on the molecule allows the separation of two fractions, an acidic fraction characterised by highly sialylated end-capping species, and a basic fraction with a lower content of sialylated species and characterized by the presence of neutral species (galactosylated). In both the de-sialylated and de-sialylated/de-galactosylated variants, no species containing sialic acid was found evidencing the completeness of the enzymatic reaction. While de-galactosylation reaction shows a good yield (90-100%) on all N-Glycan sites except for Asn 7 of β-subunit.

Altogether, these results demonstrate that different modifications on the glycosylation moieties were induced by biochemical stress and that they could be characterised with precision by MS. These results are part of a larger project aiming to highlight the effect of such glycovariants on ovarian follicular maturation and growth and ultimately correlate the FSH glycosylation pattern and its biological activity through further *in vivo* and *in vitro* experiments. This will allow to understand how to improve the efficacy of commercial drug product (GONAL-F ®).

The untargeted characterization of FSH glycosylation sites and the glycan moieties linked to each glycosite was instrumental for the development of a novel method based on targeted mass spectrometry to identify and quantify the different glycoforms of human serum FSH (hFSH). A Multiple Reaction Monitoring (MRM) for the absolute quantification of circulating hFSH molecule was optimised, validated and successfully applied to the analysis of 10 women sera

samples of different ages showing a great variability of the amount of the hormone. The method demonstrated comparable sensitivities and higher selectivity and specificity than conventional ELISA assays, and therefore it could be proposed as an alternative method for the quantification of hFSH.

However, this method will not yield more information on glycoforms than a classic ELISA, as it lacks the ability to distinguish glycoforms. To address this, a method based on mass spectrometry operating in Parallel Reaction Monitoring (PRM-MS) mode aimed to the identification and relative quantification of serum hFSH glycoforms was also developed. Because of the complexity, the heterogeneity, and the large dynamic range of the biological matrix, a first step of hormonal immunopurification from human serum was required prior to LC-MS/MS analysis. An initial untargeted LC-MS/MS analysis of immunopurified hFSH from serum allowed the identification of hFSH therefore confirming the efficiency of the immunopurification. As expected, although the aminoacidic sequence coverage was satisfactory for both the alpha and beta polypeptide chains, only one glycopeptide of beta chain could be identified. For this reason we developed a PRM-MS targeted method contained an inclusion list of 24 different glycopeptides, as well as one non-glycosylated peptide representative of the overall concentration of hFSH, which allowed high specific detection the most abundant serum hFSH glycopeptides on each of the four hFSH N-glycosites. Moreover, this analysis also allowed the relative quantification, expressed as relative percentage distribution of glycan moiety, of each identified glycopeptide by extracting post-data acquisition the chromatographic traces of best glycopeptide fragments. This is the first study aimed at the site-specific quantification of circulating h-FSH glycoforms. The PRM-MS glycopeptide method was applied to r-FSH in order to compare the site specific glycosylation profile of hFSH and r-FSH (commercialised as GONAL-F ®). Data obtained show the same trend for most abundant glycoforms: the biantennary fully sialylated (A2G2S2) and the mono-sialylated (A2G2S1) are the most abundant

glycoforms on α-Asn52 and α-Asn78; while fucosylated biantennary (FA2G2S2 and FA2G2S1) are the main glycans on β-Asn 24. Differences were observed in terms of high mannose glycan content in the minority glycoforms, that are observed only for recombinant molecule. Moreover, recombinant protein exhibited higher number/level of biantennary glycans, most of them with core-fucosylation with respect to the serum FSH.

The perspectives of this study are aligned along two axes. Fist, on a fundamental research axe, the results obtained from a large cohort of women in different physiological conditions, could provide useful insights into the functional role of hFSH glycans in human fertility. These speculations will be essential in understanding the clinical relevance of some glycoforms rather than others with the ultimate goal of developing or improving drugs for the treatment of infertility. Second, on a clinical axe, the MRM/PRM-MS method presented here could be complementary or even substitute the immunoassays commonly used for the detection of FSH and other gonadotropins in the clinical setting. This would add the possibility to determine each patient specific glycosylation profile, leading ultimately to patient stratification and personalized therapy

## 6.1 References

1. Wide, L. and K. Eriksson, *Dynamic changes in glycosylation and glycan composition of serum FSH and LH during natural ovarian stimulation.* Upsala journal of medical sciences, 2013. **118**(3): p. 153-164.

2. Zariñán, T., et al., *In vitro impact of FSH glycosylation variants on FSH receptor-stimulated signal transduction and functional selectivity.* Journal of the Endocrine Society, 2020. **4**(5): p. bvaa019.

3. Anobile, C.J., et al., *Glycoform composition of serum gonadotrophins through the normal menstrual cycle and in the post-menopausal state.* Molecular Human Reproduction, 1998. **4**(7): p. 631-639.

4. Hayes, J.M., et al., *Glycosylation and Fc Receptors*, in *Fc Receptors*, M. Daeron and F. Nimmerjahn, Editors. 2014, Springer International Publishing: Cham. p. 165-199.

5. Loreti, R.N., et al., *Effect of sialylation and complexity of FSH oligosaccharides on inhibin production by granulosa cells.* 2013.

6. PADMANABHAN, V., et al., *Modulation of Serum Follicle-Stimulating Hormone Bioactivity and Isoform Distribution by Estrogenic Steroids in Normal Women and in Gonadal Dysgenesis\*.* The Journal of Clinical Endocrinology & Metabolism, 1988. **67**(3): p. 465-473.

7. Wide, L. and B. Hobson, *Influence of the assay method used on the selection of the most active forms of FSH from the human pituitary.* European Journal of Endocrinology, 1986. **113**(1): p. 17-22.

8. Zambrano, E., et al., *Receptor binding activity and in vitro biological activity of the human FSH charge isoforms as disclosed by heterologous and homologous assay systems.* Endocrine, 1999. **10**(2): p. 113-121.

9. Meher, B.R., et al., *Glycosylation effects on FSH-FSHR interaction dynamics: a case study of different FSH glycoforms by molecular dynamics simulations.* PLoS One, 2015. **10**(9): p. e0137897.

10. Willey, K.P., *An elusive role for glycosylation in the structure and function of reproductive hormones.* Human reproduction update, 1999. **5**(4): p. 330-355.

11. Bishop, L.A., et al., *Specific roles for the asparagine-linked carbohydrate residues of recombinant human follicle stimulating hormone in receptor binding and signal transduction.* Molecular Endocrinology, 1994. **8**(6): p. 722-731.

12. Sairam, M. and G. Bhargavi, *A role for glycosylation of the alpha subunit in transduction of biological signal in glycoprotein hormones.* Science, 1985. **229**(4708): p. 65-67.

13. Ulloa-Aguirre, A., et al., *Role of glycosylation in function of follicle-stimulating hormone.* Endocrine, 1999. **11**(3): p. 205-215.

14. Lunenfeld, B., et al., *The development of gonadotropins for clinical use in the treatment of infertility.* Frontiers in endocrinology, 2019. **10**: p. 429.

15. Lyu, S.W., et al., *Impact of high basal FSH/LH ratio in women with normal FSH levels on in vitro fertilization outcomes.* Gynecological Endocrinology, 2013. **29**(5): p. 424-429.

16. Mastrangeli, R., et al., *In-vivo biological activity and glycosylation analysis of a biosimilar recombinant human follicle-stimulating hormone product (Bemfola) compared with its reference medicinal product (GONAL-f).* PloS one, 2017. **12**(9).

17.     https://www.ema.europa.eu/en/medicines/human/EPAR/gonal-f. *GONAL F*.

18.     Riccetti, L., et al., *Glycosylation pattern and in vitro bioactivity of reference follitropin alfa and biosimilars.* Frontiers in endocrinology, 2019. **10**: p. 503.

19.     Lombardi, A., et al., *Evaluation of the oligosaccharide composition of commercial follicle stimulating hormone preparations.* Electrophoresis, 2013. **34**(16): p. 2394-2406.

20.     Bergandi, L., et al., *Human Recombinant FSH and Its Biosimilars: Clinical Efficacy, Safety, and Cost-Effectiveness in Controlled Ovarian Stimulation for In Vitro Fertilization.* Pharmaceuticals, 2020. **13**(7): p. 136.

## Published papers during PhD period

1. Carpentieri, A.; et al., *Mass spectrometry based proteomics for the molecular fingerprinting of Fiano, Greco and Falanghina cultivars*. Food Research International 2019, **120**, 26-32. https://doi.org/10.1016/j.foodres.2019.02.020

2. Carpentieri, A et al., *Fiano, Greco and Falanghina grape cultivars differentiation by volatiles fingerprinting, a case study*. Heliyon 2019, **5** (8). https://doi.org/10.1016/j.heliyon.2019.e02287

3. Castiglia, D.; et al., *High-level production of single chain monellin mutants with enhanced sweetness and stability in tobacco chloroplasts.* Planta 2018, **248** (2), 465-476. https://doi.org/10.1007/s00425-018-2920-z

4. Illiano, A.; et al., *Protein Glycosylation Investigated by Mass Spectrometry: An Overview*. Cells 2020, **9** (9). https://doi.org/10.3390/cells9091986

5. Leone, S.; et al., *Structural effects of methylglyoxal glycation, a study on the model protein MNEI.* Molecular and Cellular Biochemistry 2019, **451** (1-2), 165-171. https://doi.org/10.1007/s11010-018-3403-z.

6. Melchiorre, C.; et al., *A multidisciplinary assessment to investigate a XXII dynasty wooden coffin.* International Journal of Conservation Science 2020, **11** (1), 25-38. ISSN: 2067-533X. http://ijcs.ro/public/IJCS-20-03_Melchiorre.pdf

7. Melchiorre, C.; et al., *A procedure for combining the removal and the identification of a patina on a 15Th century byzantine icon*. International Journal of Conservation Science 2019, **10** (2), 249-256. ISSN: 2067-533X. http://ijcs.ro/public/IJCS-19-22_Melchiorre.pdf

8. Melchiorre, M.; et al., *Direct and Solvent-Free Oxidative Cleavage of Double Bonds in High-Oleic Vegetable Oils.* ChemistrySelect 2020, **5** (4), 1396-1400. https://doi.org/10.1002/slct.201903516**.**

9. Pinto, G.; et al., *Quantification of Polyphenols and Metals in Chinese Tea Infusions by Mass Spectrometry.* Foods 2020, **9** (6). https://doi.org/10.3390/foods9060835.

10. Carpentieri, A., et al., *Molecular Characterisation of Olive Cultivars from Italy (Campania): A Mass Spectrometry Approach.* Journal of Food Science & Nutrition 2020. DOI:10.24966/FSN-1076/100052.

11. Melchiorre, C., et al., *Identification and relative quantification of hFSH glycoforms in women sera by PRM MS based approach*. Pharmaceutics, 2021 (submitted).