

Development of new analytical methods for the analysis at the intact level of glycoforms of hCG and other gonadotropins by nano liquid chromatography hyphenated to high resolution mass spectrometry

Amira Al Matari

▶ **To cite this version:**

Amira Al Matari. Development of new analytical methods for the analysis at the intact level of glycoforms of hCG and other gonadotropins by nano liquid chromatography hyphenated to high resolution mass spectrometry. Analytical chemistry. Université Paris sciences et lettres, 2021. English. NNT : 2021UPSLS034 . tel-03270927

HAL Id: tel-03270927
https://pastel.hal.science/tel-03270927

Submitted on 25 Jun 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

PSL UNIVERSITÉ PARIS

THÈSE DE DOCTORAT
DE L'UNIVERSITÉ PSL

Préparée à l'ESPCI Paris, au Laboratoire Sciences Analytiques, Bioanalytiques et Miniaturisation (LSABM)

Development of new analytical methods for the analysis at the intact level of glycoforms of hCG and other gonadotropins by nano liquid chromatography hyphenated to high resolution mass spectrometry

Développement de nouvelles méthodes pour l'analyse au niveau intact des glycoformes de la hCG et d'autres gonadotrophines par nano chromatographie en phase liquide couplée à la spectrométrie de masse à haute résolution

Soutenue par

Amira AL MATARI

Le 09 février 2021

Ecole doctorale n° 388

Chimie Physique et Chimie Analytique de Paris Centre

Spécialité

Chimie analytique

Confidential until 01/06/2021

ESPCI PARIS | PSL

Composition du jury:

Didier THIEBAUT Directeur de recherche, CNRS - ESPCI Paris PSL	*Président*
Myriam TAVERNA Professeur, Université Paris Saclay	*Rapporteur*
Olivier LAPREVOTE Professeur, Université de Paris, CNRS	*Rapporteur*
Eric REITER Directeur de recherche, INRAE	*Examinateur*
Thierry FOURNIER Directeur de recherche, Université de Paris, INSERM	*Examinateur*
Nathalie DELAUNAY Chargée de recherche, CNRS-ESPCI Paris PSL	*Directeur de thèse*

Table of Contents

Publications and Communications 7

List of Abbreviations 10

General Introduction 14

Chapter 1: Human Chorionic Gonadotropin Hormone and other gonadotropins 19

1.1. Generalities about hCG 19

1.1.1. *hCG production* 19

1.1.2. hCG Biological functions 21

1.1.3. hCG structure 22

1.1.4. hCG glycosylation 24

1.1.5. hCG nomenclature and the 1st international reference reagents 28

1.1.6. Metabolism of hCG 29

1.2. Other gonadotropins 30

1.2.1. FSH 31

1.2.2. LH 32

1.3. Gonadotropins for infertility treatments 33

1.4. Conclusion 35

Chapter 2: State-of-the-art about High performance analytical methods for the characterization of gonadotropins 36

2.1. Bottom-up approach 36

2.1.1. Analytical methods applied in bottom-up approach 43

2.1.2. Advantages and disadvantages of the bottom-up approach 45

2.2. Analysis of released glycans 46

2.2.1. Analytical methods used in released glycan approach 47

2.2.2. Advantages and disadvantages of the released glycan approach 49

2.3. Analysis of gonadotropins at intact level 50

2.3.1. Capillary electrophoresis 54

2.3.2. Liquid chromatography 56

2.3.3. Advantages and disadvantages of intact approaches 59

2.4. Strategies for structural assignments and structures of already identified glycans on gonadotropins 60

2.5. Conclusion 66

Chapter 3: Identification and semi-relative quantification of intact glycoforms of human chorionic gonadotropin and follicle-stimulating hormone 69

1.5. Introduction 72
1.6. Experimental Part 75
1.6.1. Materials and Reagents 75
1.6.2. Sample preparation 75
1.6.3. NanoLC-MS system 76
1.7. Results and discussion 77
1.7.1. Development of the nanoLC-MS method 77
1.7.2. Nano-LC-HRMS data treatment for the identification of glycoforms 79
1.7.3. Figures of merit of the nanoLC-HRMS method 86
1.7.4. Potential of the analytical method for batch-to-batch comparison 88
1.7.5. Potential for the analysis of another gonadotropin 89
1.7.6. rhCG vs rFSH 93
1.8. Conclusion 95
Chapter 4: Identification and semi-relative quantification of intact glycoforms of hCGα and hCGβ by nanoLC-(Orbitrap)MS 98
4.1. Introduction 101
4.2. Experimental Part 104
4.2.1. Materials and Reagents 104
4.2.2. Sample preparation 104
4.2.3. NanoLC-MS system 105
4.2.4. Nano-LC-HRMS data treatment for the identification of glycoforms 106
4.3. Results and discussion 106
4.3.1. Optimization of the injection conditions 107
4.3.2. Mobile phase gradient optimization 111
4.3.3. Effect of the nature and content of acidic additive introduced in the mobile phase 113
4.3.4. Figures of merit of the nanoLC-HRMS method 115
4.3.5. Identification of the hCGα and hCGβ glycoforms 119
4.4. Conclusion 121
Chapter 5: Analysis of different recombinant and urinary gonadotropin proteins in nanoLC-(Orbitrap)MS at the intact level 123
5.1. Introduction 127
5.2. Experimental Part 129
5.2.1. Materials and Reagents 129
5.2.2. Sample preparation 130

5.2.3. NanoLC-MS system 130
5.2.4. Nano-LC-HRMS data treatment for the identification of glycoforms 131
5.3. Results and discussion 131
5.3.1. TFA vs FA as mobile phase additive for different samples of hCG 131
5.3.2. TFA vs FA as mobile phase additive for gonadotropins other than hCG 138
5.3.3. Comparison between gonadotropin glycoforms 142
5.4. Conclusion 143
Chapter 6: synthesis and first characterization of immunsorbents dedicated to the selective extraction of hCG glycoforms 145
6.1. Characteristics of ISs 147
6.2. Immunoextraction procedure 150
6.3. Immunoextraction of hCG 151
6.4. Complementary Results 158
6.4.1. Materials and reagents 158
6.4.2. Sample preparation 159
6.4.3. Immunosorbent synthesis 159
6.4.4. Quantification of antibodies grafted on Sepharose 159
6.4.5. Imunoextraction procedure 160
6.4.6. NanoLC-MS analysis 160
6.5. Results and discussion 161
6.5.1. Preparation of ISs 161
6.5.2. First evaluations of the IS potential to extract hCG glycoforms 163
General conclusion and perspectives 171
Supplementary Data 174
List of tables and figures 219
List of figures 219
List of tabes 224
References 226

Publications and Communications

Accepted publications:

A. Al Matari, A. Combès, J. Camperi, T. Fournier, V. Pichon, N. Delaunay: Identification and semi-relative quantification of intact glycoforms by nano-LC–(Orbitrap)MS: application to the α-subunit of human chorionic gonadotropin and follicle-stimulating hormone, Anal Bioanal Chem, Jul. 2020, doi: 10.1007/s00216-020-02794-3.

A. Al Matari, A. Goumenou, A. Combès, T. Fournier, V. Pichon, N. Delaunay: Identification and semi-relative quantification of intact glycoforms of hCGα and hCGβ by nanoLC-(Orbitrap)MS, Journal of Chromatography A, 2021 Jan 28;1640:461945. doi: 10.1016/j.chroma.2021.461945

Publication in preparation:

A. Al Matari, A. Combès, T. Fournier, V. Pichon, N. Delaunay: Analysis of different recombinant and urinary gonadotropin proteins in nanoLC-(Orbitrap)MS at the intact level. Submitted to Journal of Pharmaceutical and Biomedical Analysis.

Oral communications:

International symposium

Development of Complementary Separation Modes Associated to Different Mass Spectrometers for the Characterization of the Structural Heterogeneity of an Intact Glycoprotein. J. Camperi, N. Eskenazi, A. Al Matari, A. Combès, J. Guibourdenche, T. Fournier, V. Pichon, J. Vinh, N. Delaunay, 32nd International Symposium on Chromatography (ISC 2018), 23-27 septembre 2018, Cannes-Mandelieu, France.

Separation methods hyphenated to mass spectrometry for the characterization of glycoproteins at the intact level. J. Camperi, A. Al Matari, A. Combès, J. Guibourdenche, T. Fournier, V. Pichon,

N. Delaunay, 9th International Symposium on Separation and Characterization of Natural and Synthetic Macromolecules (SCM-9), 30 janvier-1er février 2019, Amsterdam, Pays-Bas.

A. Al Matari, A. Combès, T. Fournier, J. Guibourdenche, V. Pichon, N. Delaunay : Identification and semi-quantification of glycoforms by nano-LC-Orbitrap-MS: application to the α-subunit of human chorionic gonadotropin and follicular stimulating hormone (Euroanalysis 2019), 1-5 September Istanbul (Turkey). Young scientist oral presentation price.

National Symposium

Développement de méthodes séparatives couplées à la spectrométrie de masse pour la caractérisation de glycoprotéines intactes. A. Al Matari, J. Camperi, A. Combès, J. Guibourdenche, T. Fournier, V. Pichon, N. Delaunay, 13ème congrès francophone de l'AfSep sur les sciences séparatives et les couplages (Sep 2019), 25-28 mars 2019, Paris, France.

Poster Communications:

Amira Al Matari , Audrey Combès, Jean Guibourdenche, Thierry Fournier ,Valérie Pichon, Nathalie Delaunay: Characterization and quantification by nano-LC-Orbitrap-MS of the glycoforms of several glycoprotein hormones in biological samples. Journee ED 388, 2018, Paris.

A. Al Matari, **J. Camperi**, N. Eskenazi, A. Combès, J. Guibourdenche, T. Fournier, J. Vinh, V. Pichon, Nathalie Delaunay. Assessment of structural heterogeneity of Follicular stimulating hormone in pharmaceutical preparations by nano-liquid chromatography coupled to high-resolution mass spectrometry, 11th International Symposium on Drug Analysis and 29th International Symposium on Pharmaceutical and Biomedical Analysis (DA-PBA), 09-12 september 2018, Leuven (Belgium).

Amira Al Matari, Julien Camperi, Audrey Combès, Jean Guibourdench[c], Thierry Fournier, *Valérie* Pichon, Nathalie Delaunay : Characterization and quantification by nano-LC-Orbitrap-MS of the glycoforms of several glycoprotein hormones in biological samples. 32nd

International Symposium on Chromatography (ISC), 23-27 September 2018, Cannes-Mandelieu (France)

Invited seminar

New analytical approaches to characterize the hormone specific to human pregnancy. J. Camperi, S. Perchepied, A. Al Matari, A. Combès, V. Pichon, N. Delaunay, *School of Science Research Seminar, University of Western Sydney-Australia*, 16 juillet 2020, Visioconférence.

List of Abbreviations

AA: amino acid
Ab: antibody
ABC: ammonium bicarbonate
ACN: Acetonitrile
AEX: anion exchange chromatography
Ag: antigen
Asn: Asparagine
BPC: Base Peak Chromatogram
Ca: capacity
CE: capillary electrophoresis
CGE-UV: capillary gel electrophoresis-ultraviolet
CHO : Chinese hamster ovary
CIEF: Capillary isoelectric focusing;
CNBr: Cyanogen bromide
CTP: carboxy-terminal peptide
CZE: capillary zone electrophoresis;
CZE-UV: capillary zone electrophoresis-ultraviolet detector
DAP: diaminopropane
DNA: deoxyribonucleic acid
E/S: enzyme/substrate ration
EDTA,: Ethylenediaminetetraacetic acid
EGF: epidermal growth factor
EVT: extravillous trophoblasts
FA: Formic acid
FPLC: fast protein liquid chromatography
FSH: follitropin or follicle-stimulating hormone
Fuc: fucose
Gal: galactose GalNAc: acetylgalactosamine
GalNAc: N-Acetylgalactosamine
GC-MS: gas chromatography-mass spectrometry;
GlcNAc*: N*-acetylglucosamine
HCD-pd-ETD: higher-energy collision dissociation-accurate mass-product-dependent electron transfer dissociation
hCG: human Chorionic Gonadotropin
hCG-H: hyperglycosylated hCG
hCGn: human Chorionic Gonadotropin nicked
hCGα: human Chorionic Gonadotropin alpha subunit
hCGβ: human Chorionic Gonadotropin beta subunit
hCGβcf: human Chorionic Gonadotropin beta subunit core fragment
hCGβn: human Chorionic Gonadotropin beta subunit nicked
HCl: Hydrochloric acid
Hex: hexose
HILIC-MS: Hydrophilic interaction chromatography – mass spectrometry
HILIC-UV: Hydrophilic interaction

chromatography-ultraviolet
hMG: human menopausal gonadotropin
hMG-HP: highly purified human menopausal gonadotropin
H-NMR: proton nuclear magnetic resonance
HPAE-PAD: High-Performance Anion-Exchange Chromatography-Pulsed Amperometric Detection
hPG : human pituitary gonadotropin
HRMS: high resolution mass spectrometry
IAE: immunoaffinity extraction
IFCC: International Federation of Clinical Chemistry and Laboratory Medicine
IMERs: immobilized enzyme reactors
IS: immunosorbent
IS-dSPE: immunosorbent-dispersive solid phase extraction
K: affinity constant
L C/MS: liquid chromatography-mass spectrometry;
LC: liquid chromatography
LH: lutropin or luteinizing hormone
LOD: limit of detection
LOD: limit of detection;
LOQ: limit of quantification
m/z: mass-to-charge
mAb: monoclonal Antibody
mAbs: monoclonal antibodies
MALDI-TOF: matrix-assisted laser desorption/ionization- time of flight
Man: mannose
MIP: molecularly imprinted polymer
MW: molecular weight
nanoES-ITMS: nanoelectrospray-ion trap mass spectrometry.
nanoLC-MS/MS: nano-liquid chromatography coupled to tandem mass spectrometry
NeuAc: sialic acid
pAbs: polyclonal antibodies
PBS: phosphate buffer saline
PEG: polyethylene glycol
hCGh hyperglycosylated
pI: isoelectric point
PTMs: post-translational modifications
Q-TOF: Quadrupole Time-of-Flight Mass Spectrometry
rFSH: recombinant follicle-stimulating hormone
rhCG: recombinant human chorionic gonadotropin
RIA: radioimmunoassay
RP HPLC-ESI-MS: reverse phase high performance liquid chromatography-electrospray ionization-mass spectrometry
RPLC-UV: reverse phase liquid chromatography-ultraviolet detector
Rpm: rotation per minute
RSD: relative standard deviation;
sCTB: subsyncytial cytotrophoblasts
SDS-PAGE: sodium dodecyl sulphate–polyacrylamide gel electrophoresis;

SEC-ICP:MS, size exclusion chromatography with inductively coupled plasma mass spectrometric detection
SEC-UV: size exclusion chromatography-ultraviolet detector
Ser: serine
SPE: solid phase extraction
SYN: syncytium
TFA: trifluoroacetic acid
Thr: threonine
TSH: thyrotropin or thyroid-stimulating hormone
uFSG: urinary follicle-stimulating hormone
uhCG: urinary human chorionic gonadotropin
UPLC-MS: ultra performance liquid chromatography-mass spectrometry
US FDA: US Food and Drug Administration
UV: ultra violet
w/w: weight by weight.
WB: western blot.
WG: week of gestation
XIC: extracted ion chromatogram

General Introduction

Glycosylation is the attachment of sugar moieties to a protein, so called a glycoprotein. It is one of the most common and diverse post-translational modifications (PTMs) that enhances the molecular and functional diversity of the proteome. Hence a given glycoprotein with several potential glycosylation sites that can be occupied by a large diversity of oligosaccharide structures can have many different glycosylated forms (called glycoforms), each of which may be generated only in a specific cell type or developmental stage [1]. However, the characterization of glycosylation patterns of glycoproteins is currently a major scientific interest, from the therapeutic and diagnostic point of view. Indeed, 70% of human proteins are glycosylated, and glycosylation controls numerous biological functions such as cell–cell signaling, protein stability, protein localization, and immune response [2]. Glycosylation may affect the protein structure and function and its effect is dependent on the protein itself [3]. The development and progression of various human diseases are directly related to it such as inflammatory diseases [3], Alzheimer's disease [4], auto-immune diseases [5], and cancer [6].

Human Chorionic Gonadotropin (hCG) is a highly glycosylated hormone with about one-third of its molecular weight made of carbohydrates. It is an heterodimeric protein composed of an alpha subunit (hCGα) having two N-glycosylation sites and a beta subunit (hCGβ) with two N- and four O-glycosylation sites. This hormone is the main interest of our work that was done in collaboration with the Pathophysiology & Pharmacotoxicology of the Human Placenta, pre & postnatal Microbiota lab (Université de Paris, INSERM, UMR-S1139). hCG mainly secreted by the trophoblast cells of human placenta as early as one week after fecudation [7]. hCG is also produced by various trophoblastic and non-trophoblastic tumors. It can also be detected in small amounts in the serum and urine of healthy men and non-pregnant women produced by the

pituitary gland [8]. In maternal blood, hCG is detectable two days after implantation and it is used in pregnancy diagnosis while it peaks in maternal blood around 8-10 weeks of gestation [9]. Glycosylation of hCG affects both its half-life in circulation and the signal transduction induced by this hormone [3]. A hyperglycosylated form of hCG has been inversely associated with subsequent development of early onsent preeclampsia when detected in maternal blood at low concentration during first trimester (8-13 week of gestation (WG)) [10]. Other reports stated that in some fetal aneuploidies such as Trisomy 13 and 18, hCG from maternal serum is abnormally glycosylated [9]. Therefore, it is suggested that certain hCG glycoforms could be used as maternal biomarkers for the identification of so-called "at risk" pregnancies. To confirm or infirm this hypothesis, the development of reliable analytical methods for the detection, identification, and quantification of these hCG glycoforms in biological samples is required. However, it is highly challenging because it deals with the most complex PTM, since glycosylation varies in glycan structure, number, and position leading to the gathering of a very large number of hCG glycoforms.

Various analytical strategies have been conducted for the analysis of glycoproteins [3], [11]–[19]. The mainly used strategy is based on the digestion of the protein and the analysis of the resultant glycopeptides (bottom-up approach). It provides essential information on the location and micro-heterogeneity of each glycosylation site on the protein. As an example, this approach was adapted by our team for the analysis of hCG [14]. Two immobilized enzyme reactors (IMERs) were synthesized by immobilizing either trypsin or pepsin for the digestion of hCG prior to the analysis of the peptides and glycopeptides by nano-liquid chromatography coupled to tandem mass spectrometry (nanoLC-MS/MS). It led to the identification of the structure and position of various N-glycans. A second approach is based on the analysis of the released

glycans obtained after a de-glycosylation step that enables gathering information about the glycans solitary (sequence, structure, and branching). But these two approaches result in the loss of information on the macro-heterogeneity of hCG and do not allow the determination of the number of different glycoforms and the identification of the exact combinations of the glycans on a given glycoform. To reach this objective, the analysis of hCG at the intact level is required, which is the third approach.

This third approach was poorly described for hCG in literature when a former PhD student of our team began to consider it [15]. Two high performance separation techniques were then investigated, capillary electrophoresis (CE) and liquid chromatography (LC), implementing in both cases different separation modes and using as hCG standards two hCG-based drugs, Ovitrelle® (recombinant rhCG) and Pregnyl (hCG isolated from the urine of pregnant women, uhCG). The most promising method for the separation and detection of hCG glycoforms was based on reversed-phase LC coupled to high resolution mass spectrometry (HRMS) using a Q-TOF analyzer, which led to the detection of several hCGα glycoforms of Ovitrelle® and Pregnyl®. An identification of the five main glycoforms was proposed by mass matching [16]. Unfortunately, the resolution of the Q-TOF analyzer was insufficient for more detection and identification due to the overlapping of some isotopic patterns corresponding to different glycoforms.

Therefore, the aim of our current work was to improve this method for the analysis of hCG at the intact level in an attempt to get glycoprofiles of hCG present in complex samples. Thus, this thesis manuscript first starts by a bibliographic section dedicated to the different properties and biological functions of hCG, as it is our main interest. But hCG belongs to a family of homologous proteins that are secreted by the pituitary gland of all vertebrates in addition to

placental cells of primates. These proteins include the following pituitary hormones: follitropin, also called follicle-stimulating hormone (FSH), lutropin (or luteinizing hormone, LH) and thyrotropin (thyroid-stimulating hormone, TSH), in addition to hCG that is secreted by placental cells of primates [20]. TSH is indispensable for the control of thyroid structure, function and metabolism [20], while FSH, LH, and hCG are involved in the regulation of reproductive functions of the organisms [21]. Therefore only these three similar hormones are discussed in the bibliographic part.

Then, the analytical methods used for their characterization are reviewed by summarizing all the previous work done from an analytical point of view. Next, a new nanoLC-MS method was developed. The nano format of the separation allowed an increase in sensitivity and its coupling with a high mass accuracy and high mass resolution analyzer, an Orbitrap, was successful in the detection and semi-quantification of numerous α-subunit glycoforms in hCG-based drugs, but also for another gonadotropin sharing the same α-subunit, the follicle-stimulating hormone (FSH). This nanoLC-MS method was next further optimized for better detection of both hCGα and hCGβ glycoforms. The optimization of both a preconcentration step in the nanoLC-MS set-up and of the mobile phase gradient and nature allowed the detection of hCGβ glycoforms for the first time and also of a higher number of hCGα glycoforms. Next this analytical method using either formic acid or trifluoroacetic acid as mobile phase additive was evaluated for the detection of the α- and β-subunits of different recombinant or urinary gonadotropin-based drugs. Finally, with the aim to determine hCG glycoforms in serum of pregnant women, i.e complex biological samples, rather than in concentrated and quite pure gonadotropin-based drugs, it was necessary to develop a selective extraction step of hCG glycoforms before their analysis in nanoLC-MS. This is why, immunosorbents were synthesized by immobilizing two different anti-hCG

antibodies on a solid sorbent. The primary experiments of hCG immunoextraction showed promising results.

Chapter 1: Human Chorionic Gonadotropin Hormone and other gonadotropins

hCG was describe by L. Cole as the wonder of today's science. It is wonder because of its extremeness. First it is one of the most acidic proteins having a peak isoelectric point (pI) up to pI 3.1. hCG variants are considered to be one of the the most glycosylated among glycoproteins. Finally, due to its high molecular weight (≈ 36 KDa), it has long circulating half-life of 36 hours [22]. hCG is a placental hormone, mainly produced during the full course of pregnancy rendering it the biomarker of pregnancy detection. It can be used as a biomarker for certain diseases and disorders such as Down syndrome, gestational trophoblastic diseases (disorders of pregnancy) and nontrophoblastic neoplasms [23]. The use of hCG as a biomarker required extended research to fully identify the exact variant of hCG present in each scenario. Therefore, in this manuscript we are mainly interisted in studdying hCG and its differend glycoforms. But hCG belongs to a family of closely related glycoproteins comprising LH, FSH, and hCG. They are secreted by the anterior pituitary gland as well as the placental cells. These glycoproteins are involved in the regulation of reproductive functions of the organisms [21]. Therefore in this manuscript we proceeded in studdying these three hormones that are structurally and functionally related.

1.1.Generalities about hCG

1.1.1. *hCG production*

hCG is synthesized and secreted in large amounts by the trophoblast cells of human first trimester placenta [7]. It is also produced by various trophoblastic and non-trophoblastic tumors [8]. The placenta is a transitory organ located between the mother and the fetus, which supports intrauterine life. The functions of this organ are nutritional, endocrinal, and immunological to

support fetal development [24]. The placenta is composed of a mesenchymal core containing the feto-placental vascularization surrounded by two layers: an epithelium of cytotrophoblasts and a syncytiotrophoblast, the latter being in direct contact with maternal blood within the intervillous space. These two layers constitute the placental barrier and separate the fetal circulation from the maternal blood, ensuring that no intermixing occurs between the two blood flows [25]. This is represented in **figure 1** that shows the structure and orientation of fetus and placenta in the uterus in addition to showing the maternal blood that bathes the villous trees, which are covered by syncytium (SYN) that is underlined by a population of progenitor cells called subsyncytial cytotrophoblasts (sCTB) and a basement membrane (purple). Cytotrophoblasts differentiate along the invasive pathway into extravillous trophoblasts (EVT). EVT anchor the placenta in the uterus and invade uterine arteries to allow maternal blood to flow into the intervillous space [26].

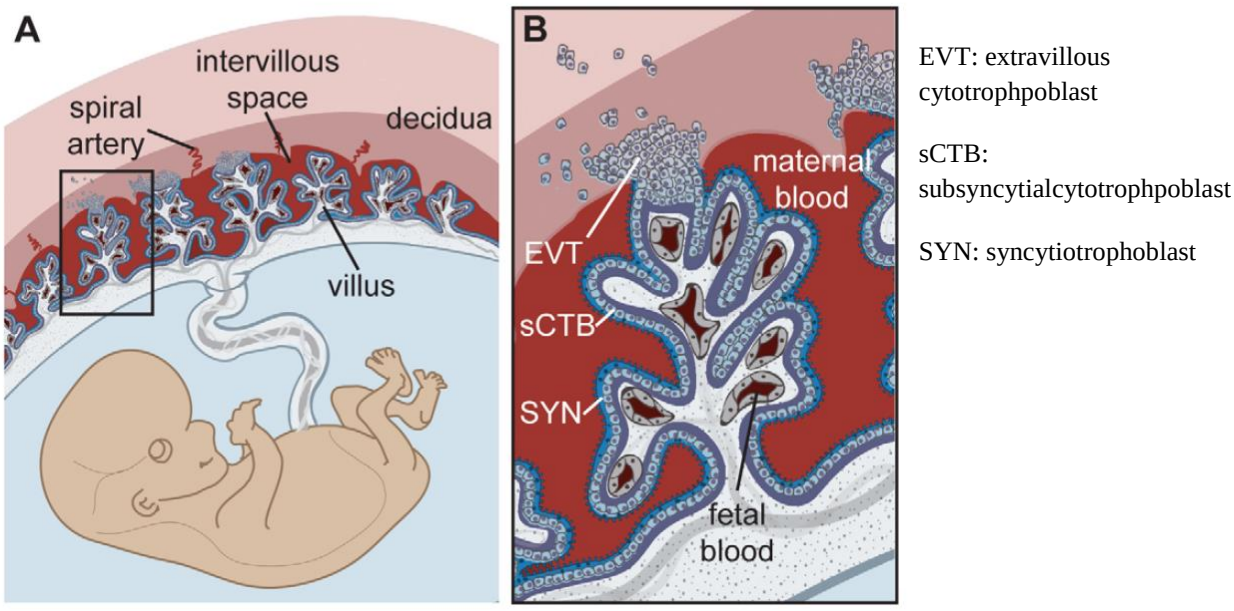


Figure 1: Human placental structure. (A) Structure and orientation of fetus and placenta in uterus. (B) Enlargement of box in panel A. [26]

Two distinct sources of hCG secretion exist within the chorionic villi which is illustrated in **figure 2**. It is clarified that hCG is mainly secreted by the syncytiotrophoblast but can also be secreted by invasive extra-villous cytotrophoblasts. hCG produced by these two different trophoblast subtypes have different states of glycosylation and thus different biological activities and functions [9]. hCG is also secreted by the pituitary gland in very small amounts in normal men and non-pregnant women. The concentration of hCG measured in the sera of normal subjects is usually less than 5 IU/L which is equivalent to 0.5 µg/L, while in the sera of pregnant women it ranges between 25- 60,000 IU/l depending on the stage of pregnancy [27]. It is important to note that 1 IU correspond to 0.1 µg of hCG [28].

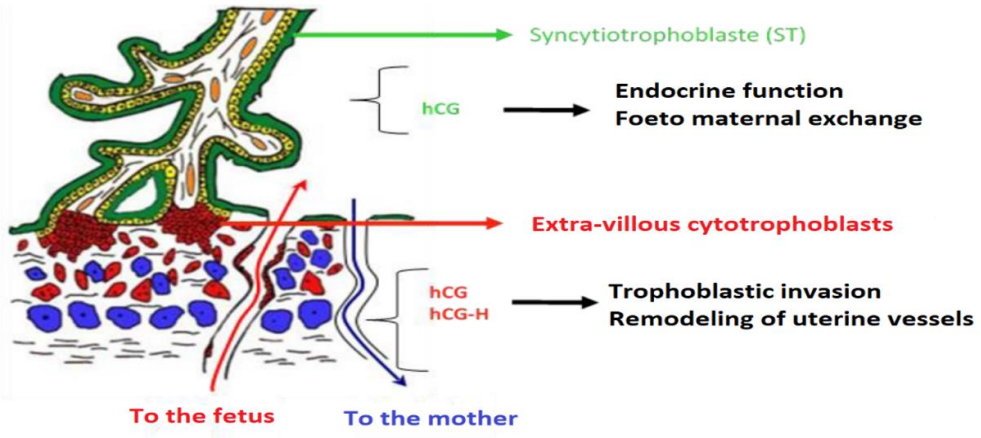


Figure 2. Different sources of hCG production: syncytiotrophoblast (ST in green) and invasive extravillous cytotrophoblast (EVT in red). Modified from [9]

1.1.2. hCG Biological functions

hCG has an incredibly wide spectrum of biological functions. They range from pregnancy implantation and placental development to hemochorial placentation, in addition to fetal and uterine growth and numerous other key functions during pregnancy. These are summarized in **table 1** regarding only pregnancy [22]. Aside from pregnancy functions, sulfated hCG is produced by the pituitary gland in women and controls steroidogenesis during the menstrual

cycle and ovulation of the oocyte (immature egg cell) [29]. Other forms of hCG denoted as hyperglycosylated hCG and hCG free ß are remarkable in human cancer biology, driving cancer growth, invasion, and malignancy. Finally hCG has a key evolutionary role which is most noted by the development of the hemochorial placentation system, that supports the development of the human brain [22]. In men hCG is produced in low concentrations by the pituitary gland in response to gonadotropin releasing hormone and it acts by stimulating the production of testosterone by the testicles [29].

Table 1: The biological functions of hCG during pregnancy. Adapted from [22]

hCG Functions
Promotion of corpus luteal progesterone production
Angiogenesis of uterine vasculature
Cytotrophoblast differentiation
Immuno-suppression and blockage of phagocytosis of invading trophoblast cells
Growth of uterus in line with fetal growth
Quiescence of uterine muscle contraction
Promotion of growth of fetal organs
Umbilical cord growth and development
Blastocysts signals uterine decidua prior to invasion regarding pending implantation
hCG in sperm and receptors found in fallopian tubes suggesting pre-pregnancy communication
hCG receptors in hippocampus and brain stem, may cause nausea and vomiting in pregnancy

1.1.3. hCG structure

hCG is composed of 237 amino acids with a molecular weight (MW) of approximately 37 kDa [9]. It is made of two subunits denoted alpha (hCGα) and beta (hCGβ) that are held together by

non-covalent and ionic interactions [30]. **Figure 3** represents the structure of the heterodimeric hCG protein. The alpha and beta subunits are represented by red and blue strands, respectively, while the carbohydrate chains are represented by light blue balls. Finally the annotation (CHO) indicated the location of the attached csrbohydrates. hCGα is a 92 amino acid polypeptide with 5 intra-chain disulfide bonds while hCGβ is a 145 amino acid polypeptide with 12 intra-molecular cysteine residues (6 disulfide bonds) [8], [31]. hCG is highly glycosylated, with glycans representing 25-41% of the total mass of hCG (25-30% in regular hCG and 35-41% in hyperglycosylated hCG, hCG-H) [22], [30]. hCGα contains two *N*-glycosylation sites on Asn 52 and Asn 78, while hCGβ contains two *N*-glycosylation sites on Asn 13 and 30 in its core region and four *O*-glycosylation sites on Ser 121, 127, 132, and 138 [8], [32]. This means that numerous hCGα and hCGβ glycoforms can exist.

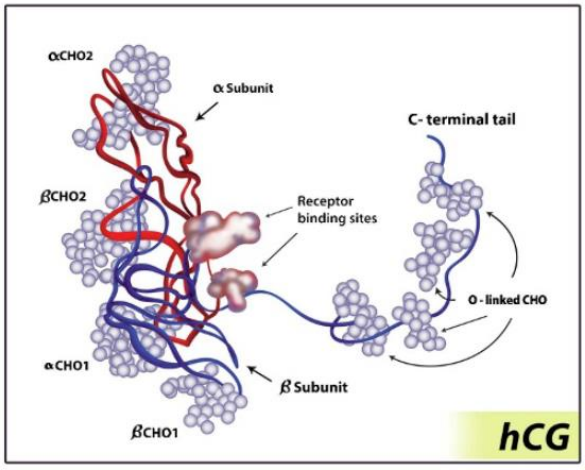


Figure 3. hCG structure with its two subunits, α in red and β in blue. Modified from **[33]**

Crystal structure shows how the β-subunit surrounds the α-subunit. It is represented in **Figure 4**. This morphology occurs because the β-subunit is folded before the subunits are combined, and the α-subunit becomes incorporated into the dimer by a mechanism termed often “threading,”

which it is passing between parts of the preassembled β-subunit. This is facilitated by transient disruption of a disulfide (termed the "tensor disulfide") between hCG residues 93 and 100 that expands this region of the molecule and permits loop α2 to pass through the seatbelt A1 [34].

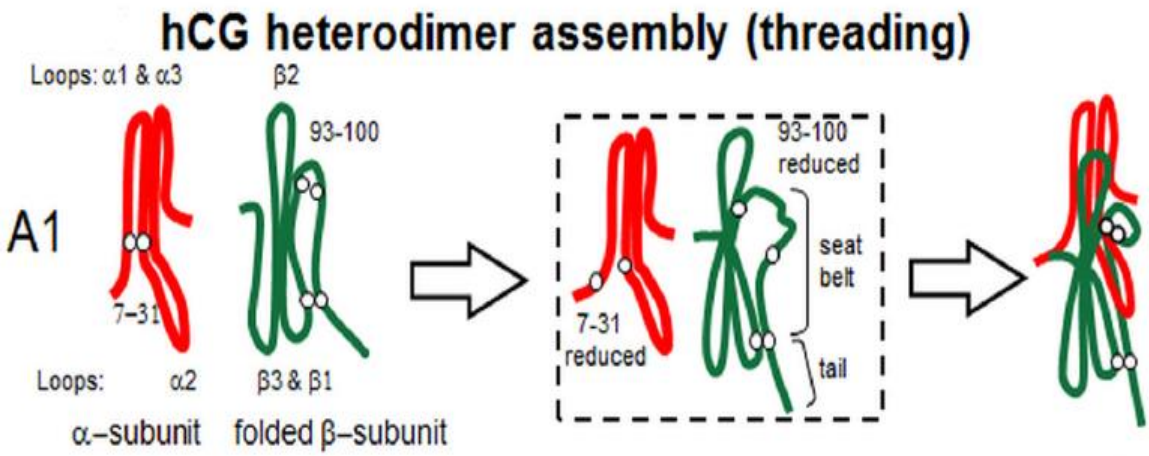


Figure 4. Scheme of hCG subunits assembly process that occurs in vivo. Modified from [34]

1.1.4. hCG glycosylation

As already mentioned, hCG has eight carbohydrate moieties in total, which accounts for about 30% carbohydrate content by weight [27]. The hCGα contains two *N*-linked carbohydrate chains while hCGβ carries two N-linked glycans [35] in addition to a 23 amino acid tail containing four additional O-linked carbohydrate side chains. **Figure 5** presents the three dimensional structure of deglycosylated hCG stressing on the location of the O- and N-linked glycans sites [22].

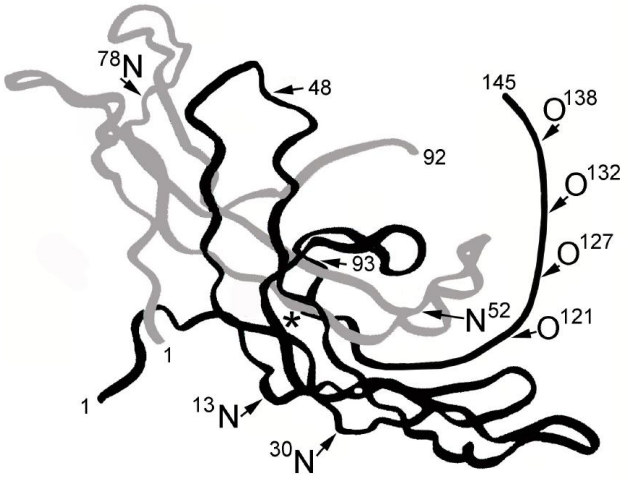
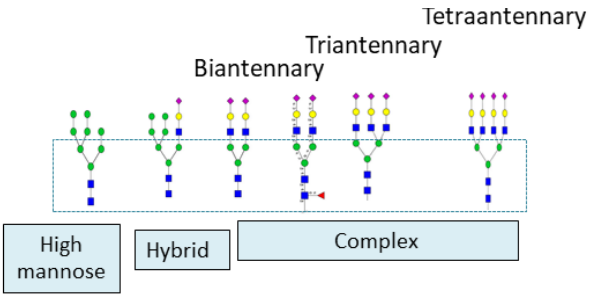


Figure 5. Three dimensional structure of deglycosylated hCG shown by X-ray defraction .*O* and *N* mark *O*- and *N*-linked oligosaccharide sites. Grey is hCGα and black is hCGß [22]

N-linked glycans are attached to a nitrogen atom of asparagine or arginine side-chains. The most common motif results from a reducing terminal *N*-acetylglucosamine (GlcNAc) linked to asparagine in a *consensus* sequence Asn-Xxx-Ser/Thr (where Xxx is any amino acid except proline). Most N-linked glycans contain a common core structure consisting of two GlcNAc and three mannose residues. The further modification of the core structure generally results in three main types of N-glycans, including high mannose, hybrid, and complex glycans which are represented in **figure 6** and are reported to exist for hCG [36]. High mannose type glycans contain only two GlcNAc and a variable number of Man. The complex type is composed of different monosaccharides, in particular GlcNAc, acetylgalactosamine (GalNAc), galactose (Gal), fucose (Fuc), and sialic acid (NeuAc) in addition to the common structure. While the hybrid type glycans combine species with high mannose content and complex species. Complex N-glycans are the main forms in humans [37].



Monosaccharide's name	Abbreviation	Symbol
Mannose	Man	
Galactose	Gal	
N-acetylglucosamine	GlcNAc	
N-acetylneuraminic acid	NeuAc	
Fucose	Fuc	

Figure 6. Representation of the N-glycan family and the common core structure.

O-linked glycans are added to an oxygen atom with amino acid side-chains without directing consensus sequence. Most O-glycans are initiated by the addition of a single monosaccharide GlcNAc (O-GlcAcylation) or GalNAc (mucin-like glycosylation) to Ser or Thr. The O-glycans have no common structure and can range from a single monosaccharide residue to extended chains similar to N-glycans. Many types of O-glycans have been identified, of which the mucin-type, where N-acetylgalactosamine (GalNAc) is linked to serine or threonine, is the most abundant. They are represented in **figure 7 A** [37]. These O-linked glycans do not share a distinct core structure similar to N-glycans, but comprise a number of different core regions with, in part, common motifs. The four most frequently observed types are illustrated in **figure 7 B** [38].

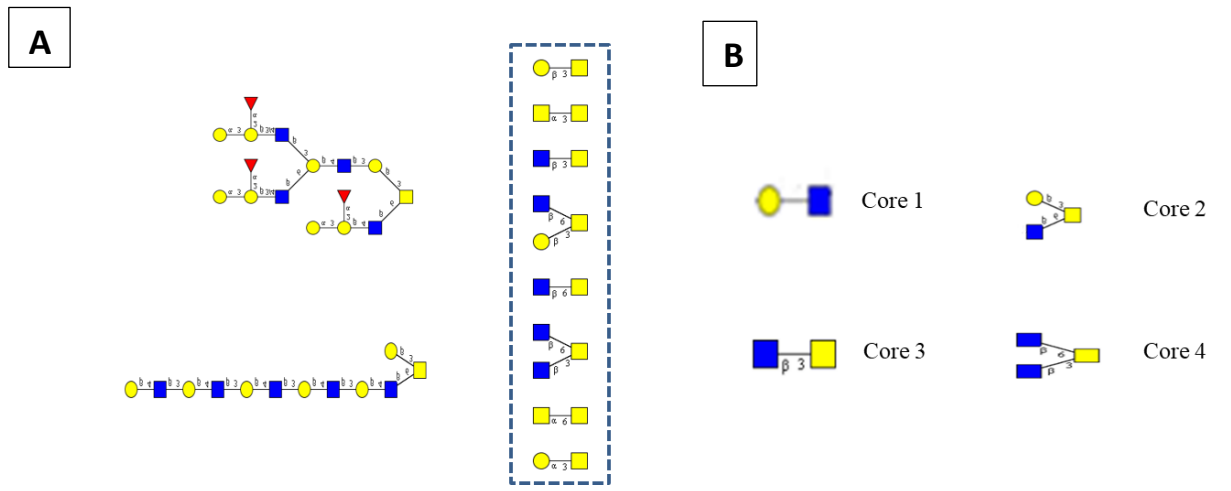


Figure 7. A: Representation of the O-glycan family. B: Frequently occurring core structures of mucin-type O-linked glycans.

It was observed that hCG glycosylation patterns change significantly during the advancement of pregnancy [22]. For example, hyper glycosylated hCG (hCG-H common to the one produced by the choriocarcinoma) is the principal variant of hCG produced in the early pregnancy. It comprises an average of 87% of the total hCG produced in serum during the third week of pregnancy, while it is around 51% during the fourth week and 43% during the fifth week of gestation [30]. During late pregnancy fucose is increasingly incorporated into the carbohydrate moieties. In addition a hyper glycosylated hCG (different from hCG-H) has been associated with fetal Down Syndrome [32] which ensures the significant change in hCG glycosylation not only during the advancement of pregnancy but also in certain malignancies. Moreover, some forms of hCG have been associated with subsequent preeclampsia when detected in low concentration (8-13 WG) in the first trimester maternal blood [10]. Other reports stated that in some fetal aneuploidies such as Trisomy 13 and 18, hCG from maternal serum is abnormally glycosylated

[9]. Therefore, it is suggested that certain hCG glycoforms could be used as maternal biomarkers for the identification of so-called "at risk" pregnancies.

1.1.5. hCG nomenclature and the 1st international reference reagents

The International Federation of Clinical Chemistry and Laboratory Medicine (IFCC) established a Working Group to define an unambiguous nomenclature [39]. The IFCC nomenclature is “user friendly” and as similar as possible to that in current use. It is summarized in **figure 8** that clearly presents the differences between intact hCG and its closely related molecules. It is clearly demonstrated that hCG refers to the intact molecule made of two subunits. Each subunit alone is designated by hCGα and hCGβ. The other related molecules are hCGn, the partially degraded hCG, hCGβn, the partially degraded β-subunit, and the hCGβcf, which is a residue of hCGβ [39].

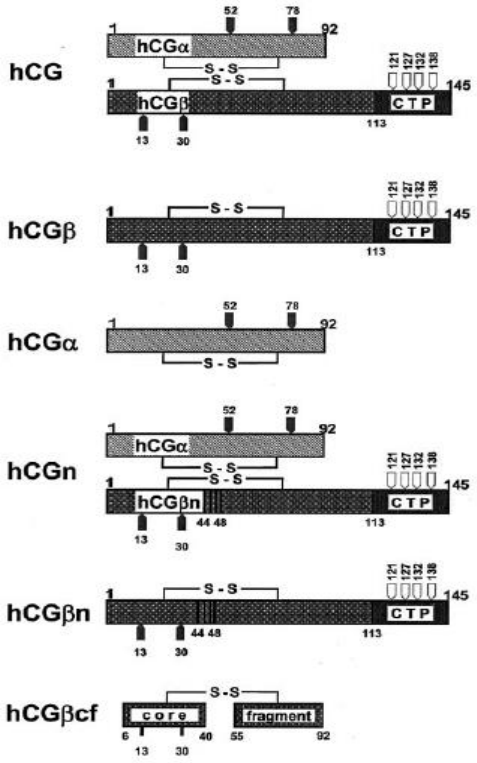


Figure 8. IFSS nomenclature of hCG and its closely related molecules. S-S represents disulfide bonds, CTP shows carboxyl-terminal peptide. N-linked carbohydrates are annotated by while O-linked carbohydrates are annotated by [39]

1.1.6. Metabolism of hCG

Metabolism of hCG means the proteolytic processing of the glycoprotein hormone from the tissue of origin, through the circulation reaching the final forms present in urine [40]. Tissues of uptake and processing of hCG are kidneys, liver, and ovaries. The kidneys aid the body to excrete intact hCG, in addition to its ability to accumulate hCG intracellularly and degrade it before its excretion as an altered form. It allows hCG dissociation and molecular processing by action of proteases and glycosidases. Fast degradation of some portion of the α-subunit into free amino acid residues takes place, while some portion of the β-subunit degrades in a slower manner [39].

However, the majority of hCG in the circulation is metabolized by the liver, whereas about only 20% is excreted by the kidneys [41]. The liver aids in the removal of sialic acids to increase the clearance of hCG [42]. hCG excretion in urine is a well-recognized metabolic fate of hCG. This fact has been utilized clinically for very long time, especially for pregnancy testing. Recently other metabolic fates for hCG have been under investigation. Indeed in a clinical trial, volunteers agreed to receive infusions and injections of hCG and the study showed that only 21.7% of the circulating hCG was excreted into urine and the remaining 78.3% are retained in the body, taken up by tissues and processed there proving that other metabolic fates do exist [43].

The metabolic fate of the placental hCG is remarkably different from that of the pituitary hCG. Placental hCG is excreted into the urine in different forms including hCG, hCGn, hCGα, hCGβ and predominantly hCGβcf in addition to hCG missing the C-terminal extension [40]. **Figure 9** summarizes the structures of hCG and related molecules in the placenta, urine, and blood [44]. On the other hand the pituitary hCG is metabolized in healthy menopausal women and non-pregnant women and then excreted in the urine as hCG partly sulfated form [40]. Therefore one

can conclude that the excretion in urine promotes various modifications of hCG molecules which makes blood the sample of choice when testing for the originally produced glycoforms of hCG and investigating about pregnancy disease biomarkers. Moreover, for more theoretical studies, the analysis of hCG present in trophoblast cell culture media gives a preliminary information about the specifications of the various glycoforms produced in the different pregnancy stages.

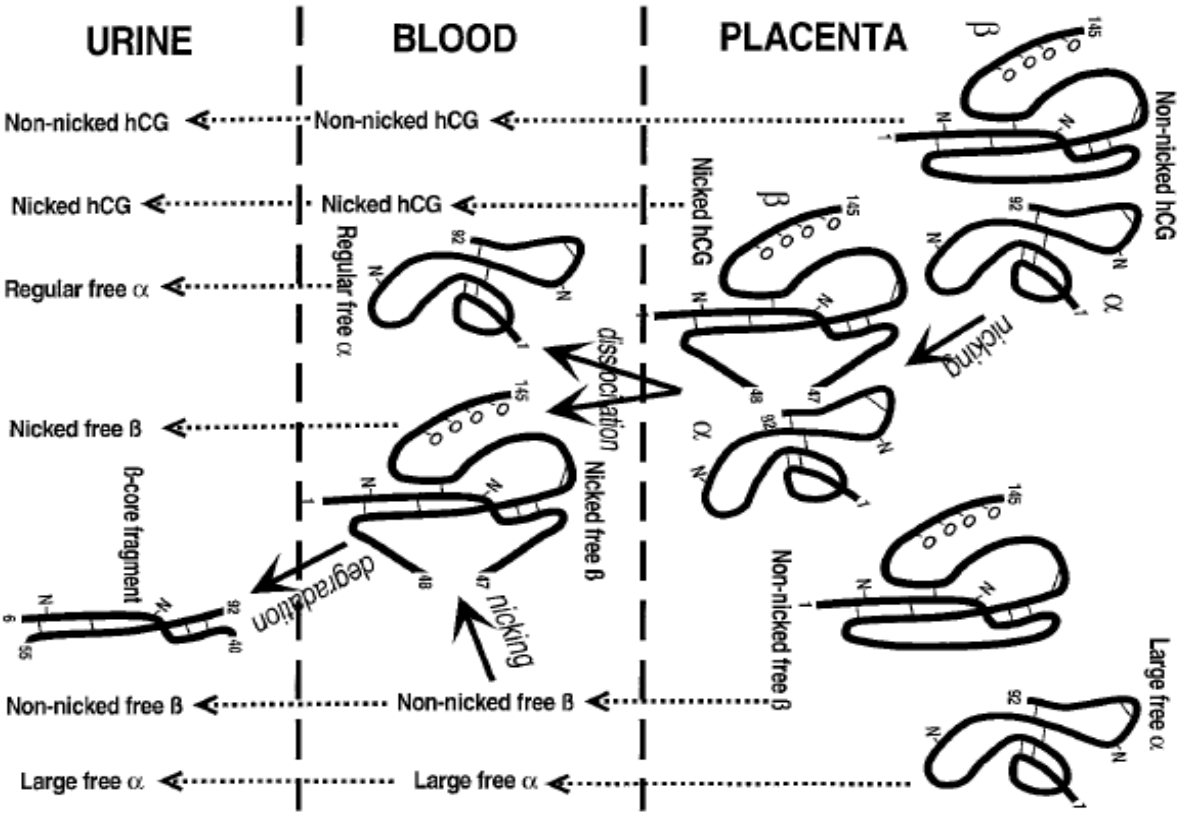


Figure 9. Representation that summarizes the structures of hCG produced by placental cells and related molecules in the placenta, urine and blood **[44]**

1.2.Other gonadotropins

hCG belongs to the gonadotropin family, composed of different dimeric glycoproteins sharing the α-subunit. The other gonadotropins are: follicle-stimulating hormone (FSH) and luteinizing

hormone (LH) [45]. There are thus structural analogs of hCG, that should have α-subunit glycoforms in common at least partially. Therefore, the analytical methods developed and used for one gonadotropin should be also appropriate for the other gonadotropins. Moreover, to develop a method for the analysis of hCG glycoforms in a given sample where other gonadotropins could be present, it is necessary to be able to discriminate them, which should not be easy as they share the same alpha subunit. Therefore, even if our interest focused on hCG, it was important to take into consideration the other gonadotropins.

1.2.1. FSH

FSH is synthesized by the anterior pituitary gland. It regulates the development, growth, pubertal maturation, and reproductive processes [46]. The activity of FSH differs over the course of the menstrual cycle and the lifetime, due to changes in its overall serum concentration and changes in the relative concentrations of its isoforms [45]. It is a heterodimeric molecule made of two non-covalently linked subunits [47]. The α-subunit is similar to hCGα (92 amino acids and two *N*-glycosylation sites at Asn52 and Asn78) while the β subunit is made up of 111 amino acids instead of 145 for hCGβ with only two *N*-glycosylation sites at Asn7 and Asn24 (no O-glycosylation sites) [48]. The overall structure of FSH is represented in **figure 10** that shows the N and C termini of each subunit, the N-glycosylated Asn residues and some important loops [48]. FSH is clinically used for controlled ovarian stimulation in women undergoing assisted reproduction, most commonly involving in vitro fertilization of retrieved oocytes. It is also used to treat anovulatory infertility in women and hypogonadotropic hypogonadism in men [49].

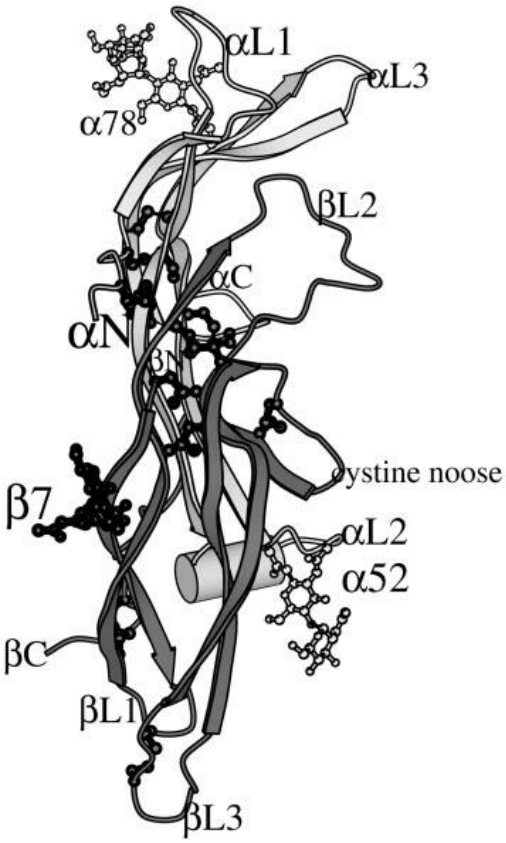


Figure 10. Cartoon Illustrating the Overall Fold of FSH [48].

1.2.2. LH

Similar to FSH, LH is synthesized by the same anterior pituitary gland and composed of two subunits [50]. LH shares the same alpha subunit with the other gonadotropins but its beta subunit is made up of 120 amino acids and it is N-glycosylated at only one position in the amino acid terminal portion (Asn-30 in LHβ) [51]. It is not only structurally but also functionally related to hCG by binding to the same receptor (LHCGR), which is a 7 loop-transmembrane glycoprotein present in the ovarian theca cells in females and in the testicular Leydig cells in males. Despite their use of the same receptor, each hormone triggers a different function. LH in particular contributes in the testicular and ovarian regulation, follicular maturation, ovulation, corpus

luteum development and maintenance, and interferes in the modification of the synthesis of steroid hormones, growth factors, and cytokines [52].

1.3.Gonadotropins for infertility treatments

Gonadotropin therapy plays an essential role in ovarian stimulation for infertility treatments. Efforts have been made over the last century to improve gonadotropin preparations. For example, commercial hCG was first available in 1931. The initial gonadotropin extraction and purification from human sources lead to the production of urinary extracted human menopausal hCG [33]. Gonadotropin derived from either animal or human tissues has complimented with clinical danger such as Creutzfeld–Jacob disease from human pituitary gonadotropin. To increase the safety and purity, recombinant DNA technology was the solution where it enables the production of specific and uniform products that reduces the need of urine collection and hormonal extraction. The initial manufacturing of recombinant hCG (rhCG) was done by transfecting non-human cell lines (Chinese hamster ovary (CHO) cells) with genetic material capable of replicating amino acid sequences identical to the human compound. This was developed as a pharmaceutical product named Ovidrel® (Merck Serono, Geneva, Switzerland). It was marketed in India initially and in France currently as Ovitrelle®. rhCG was approved in 2000 by the US Food and Drug Administration (US FDA) for the induction of final follicular maturation and early luteinization in infertile women. It is being used as part of assisted reproductive technology program in addition to the induction of ovulation and pregnancy in ovulatory infertile patients whose cause of infertility is not due to ovarian failure [31]. The difference between rhCG and hCG derived from human extract is expected to be in terms of glycosylation.

For the FSH gonadotropin hormone, in the late 1950s and 1960s, early attempts to purify this protein from pituitary extracts or urine were carried out. The development of antibody-based purification methods enhanced the purity of human urinary FSH (uFSH) preparations. Next, modern recombinant technologies allowed the development of highly purified preparations of FSH that was produced by recombinant expression systems. In 1990s recombinant FSH (rFSH)-based products entered into the gonadotropin marketplace, such as follitropin alpha (Gonal-f; EMD Serono) and follitropin beta (Follistim; Organon/Merck & Co).

An interesting timeline of major events in the gonadotropin identification and development of gonadotropin-based drugs is represented in **Figure 11** starting by the initial identification of gonadotropin hormones in 1927-1929 and moving to the extraction of human menopausal gonadotropin (hPG) and human pituitary gonadotropin (hPG) then the extraction of urinary hormones and reaching the synthesis of recombinant glycoproteins and the approval of biosimilars [53].

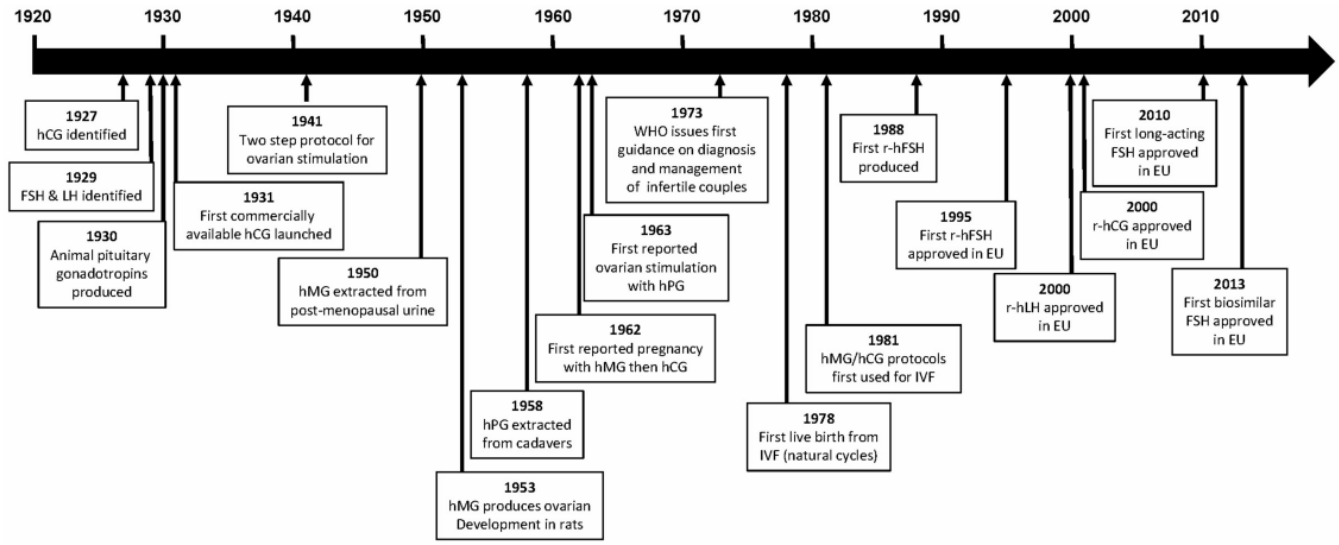


Figure 11. Representation of a timeline of major events in the development of gonadotropin-based drugs. **[53]**

1.4. Conclusion

The aim of this work was to explore and develop analytical techniques that enable the analysis of hCG in complex biological samples, by minimizing sample preparation steps and targeting the differentiation between the various glycoforms keeping them intact. Due to the fact that variations in glycoforms seems to be related to certain outcomes of pregnancies and diseases, they would be helpful in future to correlate the glycoforms identified in real biological samples to the outcomes that enables the identification of biomarkers that could be used in the future for fast and accurate diagnosis. This is why the following part will present the state-of-the-art of the analytical methods used for the characterization of hCG, but also of other gonadotropins since they share the same alpha subunit, before starting this project.

Chapter 2: State-of-the-art about High performance analytical methods for the characterization of gonadotropins

As presented in the first chapter, hCG belongs to gonadotropins, which are highly glycosylated proteins, and a correlation between its glycosylation state and some pregnancy diseases or trisomy was observed [9]. Therefore the characterization of its glycosylation pattern is essential either in biological samples of pregnant women to try to identify the glycoforms that could be biomarkers of at risk pregnancies or in the pharmaceutical hCG-based preparations as it may impact their efficiency and even toxicity. However, the characterization of hCG glycosylation faces huge analytical challenges due to the complexity of the analytes being studied, being a mixture of numerous glycoforms with close structures.

The analysis of glycoproteins and the detection and identification of their glycoforms constitute an important analytical challenge whatever the glycoprotein. It requires the development of efficient separation methods associated with a sensitive and informative detection mode, such as mass spectrometry (MS). Three main approaches exist to study the glycosylation of a protein, by carrying out analysis at the peptide (bottom-up), glycan or intact protein level. These different approaches for analyzing hCG, but also other gonadotropins that share the same alpha subunit, are detailed below.

2.1.Bottom-up approach

The bottom-up approach is based on the analysis of the glycopeptides resulting from the glycoprotein digestion. It is the most common used method to date to study glycoproteins based on the application of an endoproteolytic step to break down proteins into well-defined peptides and glycopeptides. The analysis of the resulting glycopeptides by liquid chromatography

hyphenated with mass spectrometry (LC-MS) allows the localization of the glycans on the amino acid chain by sequencing using MS/MS fragmentation. This approach gives thus information on the micro-heterogeneity of the glycoproteins. At the level of glycopeptides in particular, the majority of the studies performed a digestion step with the trypsin or pepsin protease during 3-16 h (overnight) in solution. The enzyme/substrate (E/S) ratio had to be increased as much as possible to speed up the digestion but without favoring too much the autoproteolysis of the protease. Therefore, a ratio of about 1:50 is the most commonly encountered in the literature. The following **Table 2** summarizes the different studies targeting the analysis of gonadotropins involving the bottom-up strategy.

Table 2 Compiling of the various analytical studies of gonadotropins using bottom-up approach.

Gonado-tropin	Sample	Method	Outcome	Identification of Glycans/ Glycoforms (Y: yes; N: no)	Ref
hCG	Profasi® Pharmaceutical hCG (Serono)	Tryptic digestion + MALDI-TOF	sequence coverage of 59% and 52% for the α and β subunits respectively unambiguous identification of hCG	N	[8]
hCG	urine spiked with Pregnyl®	Immunoextraction (B108 mAb) + Tryptic digestion + LC-MS/MS	hCG detected and confirmed in urine	Y (5 N-glycans)	[54]
hCG	hCG (C5297)-Sigma urine of choriocarcinoma patients and healthy subjects	Vinylpyridine, endoproteinase Glu-C or tryptic digestion + LC-MS/MS	Two nicking sites (in vivo proteolytic bond cleavage that results in minor changes in protein structure) observed in both samples. Important differences in the oligosaccharide structures, including the presence of triantennary oligosaccharides not found in hCG from healthy subjects.	Y (10 N-glycans)	[55]
hCG	hCG (C5297; 3000 IU mg~1)	LC + collection of the β subunit fraction + tryptic digestion + O-Deglycosylation + LC-MS/MS	Confirmation of both the primary sequence and the attachment of four oligosaccharide groups.	Y (7 N- and 6 O-glycans)	[56]
hCG	hCG bound to LH receptor expressed in L cells	Purification + tryptic digestion + LC-MS/MS	Six N-glycosylation sites predicted from the amino acid sequence (some are found to be deglycosylated)	Y (5 N-glycans)	[57]
hCG	urinary hCG	Immunoextraction (pAb anti-hCGβ) + tryptic	Quantitative, rapid and selective isolation of hCG from crude urine	N	[58]

		digestion + LC–MS/MS			
hCG	hCGβ from urine of patients with testicular cancer and culture medium of a choriocarcinoma cell line. urine of a normally pregnant woman and urine of a patient with invasive mole	Immunoaffinity chromatography + gel filtration + tryptic digestion + LC-MS/MS+ differential expression analysis software	N-glycans at Asn-13 and Asn-30 as well as *O*-glycans at Ser-121, Ser-127, Ser-132, and Ser-138 were characterized. In all samples, the major type of N-glycan: a biantennary complex-type structure triantennary structures linked to Asn-30 and fucosylation of the Asn-13-bound glycan increased in cancer-derived hCGβ. Significant site-specific differences in the O-glycans, with constant core-2 glycans at Ser-121, core-1 glycans at Ser-138, and putative sites unoccupied by any glycan. Core-2 glycans at either Ser-127 or Ser-132 enriched in cancer	Y (26 N- and 11 O-glycans)	[3]
hCG	WHO RRs for hCG-α and hCG-β Pregnyl®	Ion pair RP LC- MS, using a monolithic separation column	Identification of glycans	Y (31 N- and 6 O-glycans)	[32]
hCG	Pregnyl® Serum and urine of pregnant women	Immunoextraction (mAb, E27) on a 96-well plate + *in-well* tryptic digestion + LC-MS/MS	Qualifying hCG*β*cf, hCG*β*n 44/ 45, hCG*β*n 47/48 and the unmodified hCG*β* subunit in serum	N	[59]
hCG	Human serum and urine from healthy subjects receiving injections of Pregnyl® and Ovitrelle®	Immunoextraction (E27 mAb) + tryptic digestion + LC–MS/MS	Detection and differentiation of the various forms of hCG in serum and urine as a function of metabolism (hCGβ, Nicked hCG β-subunit and hCG β-core fragment). For both pharmaceuticals, only the intact hCG was detected in serum, whereas in urine the injection of Pregnyl was shown to generate a more complex hCG variant pattern compared to Ovitrelle	N	[60]
hCG	Reference	Immunoextraction (30 anti-	hCG variant composition of hCGn (1st IRR	N	[61]

	reagents for hCG obtained from the National Institute of Biological Standards and Controls	hCG mAbs)+ tryptic digestion + LC-MS/MS	99/642) and hCGβn (1st IRR 99/692) hCG variant specificity profiles and epitope recognition of hCG mAbs		
hCG	Urine after administration of Novarel® (uhCG of pregnant women) or Ovidrel® (rhCG)	Immunoextraction of hCG (β7 and β1 mAb)-conjugated on magnetic beads + tryptic digestion + LC-MS/MS	Detection of intact dimeric hCG, hCGβ and hCGβcf Quantification of total intact hCG and hCGβ	N	[62]
hCG	hCG from USBIO and plasma	Tryptic digestion + N- and partial O-deglycosylation + Jacalin-affinity chromatography + nanoLC-MS/MS	Identification of 49 O-GalNAc-linked glycopeptides from 36 glycoproteins in human plasma	N	[63]
hCG	Ovitrelle® and Pregnyl® Bovine serum albumin and cytochrome C	heat denaturation (60 ◦C) + tryptic digestion + LC–MS/MS	Application of the accelerated digestion protocol (E/S 1:1, 45 min) to realistic human serum samples containing hCG to demonstrate its applicability	N	[64]
hCG	Random male urine + urine of a healthy male receiving a single dose of Ovidrel®	Immunoextraction (β7 mAb-conjugated magnetic beads) + tryptic digestion + LC-MS/MS	Immunoisolation of hCG isoforms (hCG, hCGβ, and hCGβcf) and concentration determination	N	[29]
hCG	Ovitrelle®,	Dual immunocapture (intact	Quantification and differentiation between the	N	[65]

	Pregnyl® and Spiked serum	hCG with mAb ISOBM-414, hCGβ and other free variants with mAb ISOBM-419) + Tryptic digestion + LC-MS/MS	heterodimer hCG and its free β- subunit		
hCG	Ovitrelle® hCG spiked serum	Immunoextraction (E27 mAb-coated magnetic beads) + tryptic digestion + nanoLC–MS/MS	On-beads digestion generated less peptides, more zero missed cleavages and higher signal intensity when compared with digestion after elution from the Ab coated beads.	N	[66]
hCG	male urine	Tryptic digestion + immunoextraction (β1 and β7 mAb) + LC–MS/MS	Mean concentrations and upper reference limits of intact hCG, hCGβ and hCGβcf	N	[11]
hCG	Dried spots from whole blood, serum, plasma and urine spiked with Pregnyl®	Immunoextraction (E27 mAb) + Tryptic digestion + LC-MS/MS	LOD : pM range Linearity between 0.93 and 0.94 range 1.84-92 µg L^{-1} Repeatability RSD of the S/N ratio at the retention time of the signature peptide of the lowest and highest level of the calibration curve between 13 and 31% and accuracy between 95 and 106 for the detection of low abundant hCG	N	[67]
hCG	Bio similar: (rhCG) SB005 Ovidrel® and Ovitrelle®	TOF-MS	6 hCGα and 13 hCGβ glycoforms observed and similar in all preparations	Y (18 N-glycans)	[31]
hCG	Ovitrelle,® and Pregnyl®	Trypsin and pepsin digestion + nanoLC-MS/MS	40 N-glycans identified after trypsin digestion and 36 after pepsin digestion with 22 N-glycans in common	Y (54 N-glycans)	[14]
LH	Bovine LH	In-gel tryptic digestion + nanoLC-MS/MS	Protein identification	N	[68]
FSH	Pituitary gland extracts	Chromatofocusing + SDS-PAGE-Western blot + Glycopeptide gel filtration	Inability of Immunofocusing to separate each FSH isoform	N	[69]

		chromatography+ HPAE-PDA			
FSH	uFSH (Shangahi, China); rFSH-Puregon®	1:Tryptic digestion + LC-MS/MS 2:Tryptic digestion+ HILIC-MS 3: Tryptic digestion of hFSH + UPLC(HILIC)- MS/MS 4: Tryptic digestion+ HCD-pd-ETD	Detection of 4 N-glycosylation sites in both samples Assignment of two O-glycosylation sites in both subunits of rFSH, and 5 in both subunits of uFSH with their structural identification	Y (9 and 12 N-glycans for rFSH and uFSH)	[70]
LH, FSH, hCG	uhCG (Ferringand Organon); hMG (Ferringand Livzon); hMG-HP (Ferring); rhCG, r-hFSH , and r-hLH (Merck Serono)	in gel tryptic digestion + nanoLC-MS/MS	Identification of non-gonadotropin proteins in purified hCG preparations showing the risk of using urinary hCG due to the possible co-extraction of prion proteins	N	[71]

E/S, enzyme/substrate ration; FSH, follicle stimulating hormone; GalNAc, *N*-Acetylgalactosamine; HILIC-MS, Hydrophilic interaction chromatography – mass spectrometry; hCG, human chorionic gonadotropin; hCGβcf, human chorionic gonadotropin beta core fragment; hCGβn human chorionic gonadotropin beta nicked; hMG, human menopausal gonadotropin; hMG-HP, highly purified human menopausal gonadotropin; HCD-pd-ETD, higher-energy collision dissociation-accurate mass-product-dependent electron transfer dissociation; HPAE-PAD, High-Performance Anion-Exchange Chromatography-Pulsed Amperometric Detection; LH, luteinizing hormone; LC/MS, liquid chromatography-mass spectrometry; LOD, limit of detection; mAb, monoclonal Antibody; MALDI-TOF matrix-assisted laser desorption/ionization- time of flight; nanoLC-MS/MS, nano liquid chromatography tandem mass spectrometry; RSD, relative standard deviation; SDS-PAGE, sodium dodecyl sulphate–polyacrylamide gel electrophoresis; uhCG, urinary hCG; UPLC-MS, ultra performance liquid chromatography-mass spectrometry; w/w, weight by weight.

2.1.1. Analytical methods applied in bottom-up approach

The bottom-up approach is the most commonly applied method for the analysis of gonadotropins. The majority of these studies implemented the use of LC coupled to tandem MS, as it appears in **Table 2**. It has gained enormous popularity, mainly due to its unique possibility to obtain site-specific information and the ability to separate glycopeptides of interest, which are often hard to ionize due to either their hydrophobicity or the presence of interfering nonglycosylated peptides within the same run. This approach requires proteolytic digestion of the glycoprotein of interest that could be done directly in solution in case of high-purity glycoprotein, or in gel after electrophoretic pre-purification and even on immobilized enzyme reactors (IMERs) as shown in **Table 2**. Trypsin is the most commonly used protease for gonadotropin analysis, as it is the case for other glycoproteins in general [72]. However, a study performed in the laboratory demonstrated for the first time the complementarity of two IMERs based on trypsin and pepsin for the characterization of the hCG glycosylation [14].

The main separation mode in LC is the reversed-phase mechanism, which is applied in the majority of the bottom up approaches presented in **Table 2**. Only one study involved hydrophilic interaction liquid chromatography (HILIC). HILIC separation allows the retention of analytes based on their partitioning between water-enriched layer formed at the surface of a polar stationary phase and the hydro-organic mobile phase, and interactions with the stationary phase. The LC separation is often coupled with tandem MS (MS/MS). The molecules of a given sample are ionized in the first spectrometer and separated by their mass-to-charge ratio. Next, ions of a particular m/z-ratio coming from the first MS are selected and made to split into smaller fragment ions that are then introduced into the second mass spectrometer, which in turn separates the fragments by their m/z-ratio and detects them.

As an example, LC separation coupled to tandem MS was used in a very interesting article by Valmu et al. in 2006 for the analysis of hCGβ from urine of patients with testicular cancer and culture medium of a choriocarcinoma cell line in addition to urine of a normally pregnant woman and urine of a patient with invasive mole. LC-MS/MS enabled the characterization of the N- and O-glycosylation sites of uhCG and the structure identification of 26 N- and 11 O-glycans. **Figure 12** represents some of the identified N-glycans. It also illustrates that different glycosylation profiles were observed according to the source of hCG analyzed. Another description of two O glycosylation sites with their structural identification were reported by Wang et al. in 2016 for FSH which contradicts the know information about lacking FSH for an O-glycosylation sites.

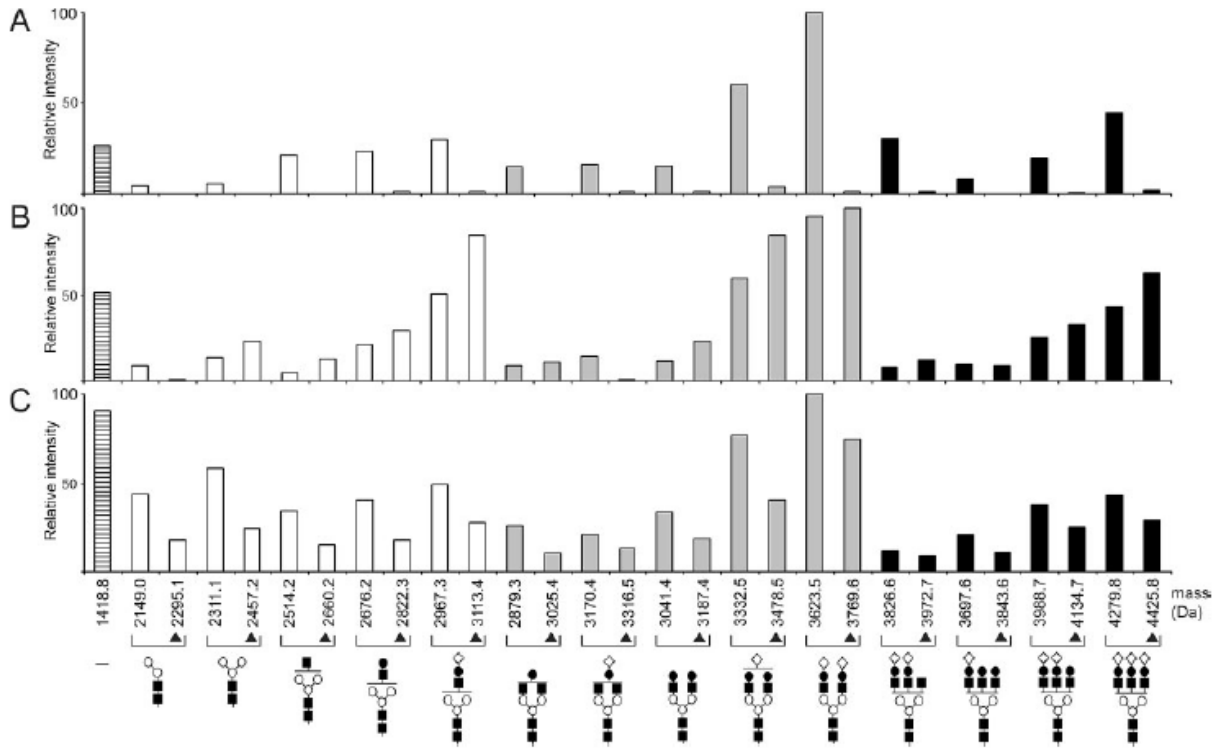


Figure 12. Structures and relative intensities of the glycans attached to Asn-13 of hCGβ. Panel A shows oligosaccharides from the urine of a normally pregnant woman at 35 weeks of pregnancy, panel B from the urine of a patient suffering from invasive mole, and panel C from the urine of a patient with testicular cancer. The proposed sugar structures have been calculated by the mass-matching approach Symbols: ■, GlcNAc; ○ Man; ●, Gal; ▲, Fuc; ◆, NeuAc. Fuc is attached to the first GlcNAc within the structures when present. [73]

Several other studies applying the same method succeeded in the identification of various N- and O-linked glycans attached to different gonadotropins [14], [32], [70], while other were only able to detect the glycoprotein of interest without the ability of the separation and identification of the glycans [68], [69], [71].

2.1.2. Advantages and disadvantages of the bottom-up approach

The advantages of the use of the bottom-up approach are numerous starting from the use of LC that provides high-resolution separation of peptides and glycopeptides with solvents that are compatible with ESI and moving to its ability to gather information on the micro-heterogenity of each glycosylation site of the analyzed glycoprotein. But there are several disadvantages such as the fact that only a fraction of the total peptide population of a given protein is identified. Therefore, information on only a portion of the protein sequence is obtained lacking the ability of global overviewing of the macro-heterogeneity of the tested proteins. Other practical limitations are the laborious additional experimental steps before the analysis (digestion plus potential reduction-alkylation) and the loss of information about low-abundance peptides in mass spectra dominated by high-abundance species. Moreover, in the case of the gonadotropin glycoforms, each glycoform is cut into small pieces which are then mixed and it is then impossible to go back to the initial structure of each glycoform, i.e. to attribute such and such glycans to a given glycoform which has several glycosylation sites. The total number of different glycoforms cannot be assessed too.

2.2. Analysis of released glycans

Another conventional approach enabling the determination of the glycosylation profile of gonadotropins is based on the release of the glycans from the protein backbone either chemically or enzymatically, followed by mass spectrometric determination of the glycan's structure. This approach generates information about the nature and diversity of glycans released from native glycoproteins. Several studies implemented this method for the analysis of gonadotropins. They are summarized in **table 3**.

Table 3 Compiling several studies of gonadotropins using released-glycans method.

Sample	Method	Outcome	Identification of Glycans/Glycoforms (Y: yes; N: no)	Ref
Purified hCGβ from pooled urine of healthy, pregnant women	PNGase-F + FPLC-Fractionation+ H-NMR and GC-MS preceded by methanolysis, N-(re)acetylation and trimethylsilylation	Finding of higher branched carbohydrate chains. Data on the possible significance of the structures of the N-linked oligosaccharides derived from hCG of patients with different types of trophoblastic diseases	Y (7 N-glycans)	[36]
BeWo cell line: a throphoblastic cell line derived from human choriocarcinoma tissue synthesizing chorionic gonadotropin	PNGase-F + gel-permeation chromatography + O-deglycosylation + AEX fractionation + HPLC subfractionation + HPAEC-PAD of some of the HPLC fractions and H-NMR	The carbohydrates of BeWo hCG-β differ from those of normal urinary hCG-β.	Y (5 N- and 3 O-glycans)	[74]
rFSH, rLH, and rhCG (CHOP)	hydrazinolysis + fluorescent labeling + WAX-MALDI-(TOF)-MS or RPLC-MS and MALDI-TOF MS (FSH, LH) Or AEX MALDI-TOF	25 FSH N-glycans identified. 12 N-complex type glycans of hCG with di-, tri-, and tetraantennarity forms and with a variable level of sialylation identified with both fucosylated and nonfucosylated species determination of the branching pattern	Y (N- and O-glycans)	[21]

	MS or RPLC-MS or MALDI-TOF MS or nanoES-ITMS (hCG)	of an unknown O-glycan All three molecules bear complex type N-glycans with di-, tri-, tetra-, and minor pentaantennary species with various levels of sialylation The analysis of more than 120 batches of drug substance over a period of 3 years showed that the sialylation level, expressed by the Z number varies by less than 3.5%.		

AEX, anion exchange chromatography; FPLC, fast protein liquid chromatography; GC-MS, gas chromatography-mass spectrometry; H-NMR, proton nuclear magnetic resonance; HPAEC-PAD, high-pH anion-exchange chromatography-pulsed amperometric detection; MALDI-(TOF), TOF matrix-assisted laser desorption/ionization- time of flight; nanoES-ITMS, nanoelectrospray-ion trap mass spectrometry.

2.2.1. Analytical methods used in released glycan approach

LC glycan analysis most commonly used mechanisms based on RP, anion exchange chromatography (AEC) and HILIC, HILIC being the gold standard. Two common detection methods are regularly associated with this LC separation, fluorescence-based detectors or mass spectrometers. Glycans do not contain fluorophores and thus must be labeled with fluorescent tags prior to analysis. As most of the glycans have a weak basicity or acidity, their ionization in MS is not straightforward and has to be improved with a derivatization step. Therefore, with fluorescence or MS detection, an additional derivatization step is required before their analysis.

LC methods were applied with fluorescent-labeled oligosaccharides for the separation of sialylated and neutral glycans from various samples [21]. However, to identify a completely unknown glycan, MS is required, such as LC-ESI-MS and MALDI-TOF MS. ESI enables the ionization of all types of glycans in both polarities but the relative intensities of neutral and sialylated glycans naturally differ when measured in positive or negative ionization mode [75].

MALDI-TOF MS offers a rapid and uncomplicated approach highly suitable for neutral or negatively charged oligosaccharides [75]. MALDI enables the formation of predominantly singly charged ions making data processing and interpretation of spectra straightforward. In summary the main strengths of MALDI-TOF-MS are its exceptional speed, low sample consumption, and high-throughput [72]. A representative summary of MALDI-TOF-MS for the analysis of released glycans is the work done by *Gervais et al.* in 2003, which succeeded in the separation and identification of the N-glycans of rFSH by HPAEC coupled to MALDI-TOF-MS. These glycans are represented in **figure 13** showing the ability of this method for the exact characterization of the glycan structures.

Another possible approach is based on gas chromatography (GC) hyphenated with an MS detection, as it appears in Table 3. However, to make the glycan volatile compounds, they have to be again derivatized before their analysis.

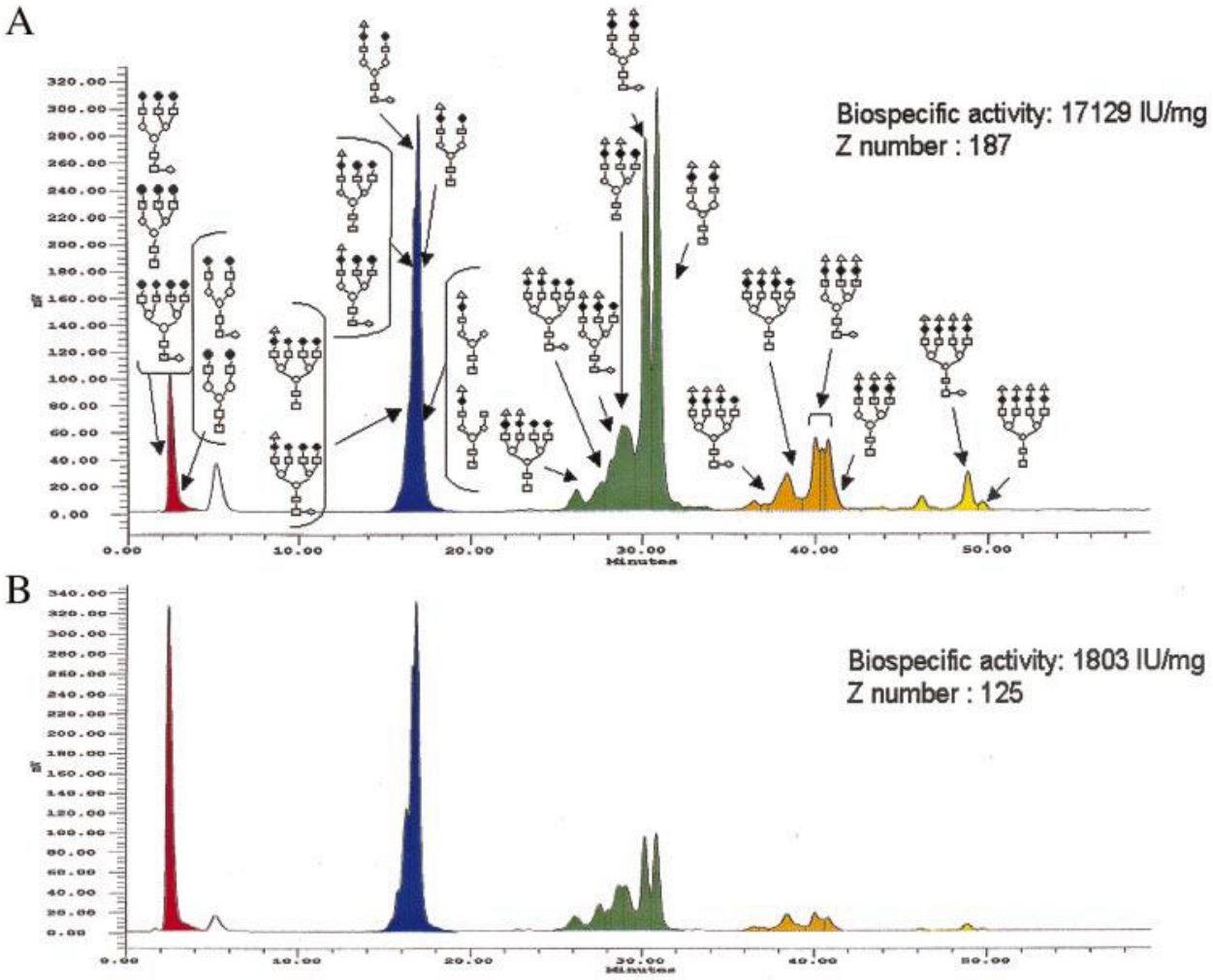


Figure 13. Separation of the N-glycans of rFSH by HPAEC-MALDI-TOF-MS. (A) rFSH batch with a high biospecific activity; (B) rFSH batch with a low biospecific activity. Triangles: sialic acid; closed circles: galactose; closed squares: N-acetylgalactosamine; open circles: mannose. The attribution of the glycan structures to the peaks is deduced from MALDI-TOF mass spectra [21].

2.2.2. Advantages and disadvantages of the released glycan approach

The main advantage of the analysis of released glycans is the precise information gathered on the attached glycans such as their sequence, structure and branching, in addition to the rapidity of this method. But this is considered a time-consuming method especially when an extra derivatization step is needed that increases the possibilities of sample modifications. The second main disadvantage is again the lack of the ability of global overviewing of the macro-heterogeneity of the tested proteins.

2.3. Analysis of gonadotropins at intact level

The previous approaches based on digestion and deglycosylation are time-consuming, because they require numerous pre-treatment steps, and they lack the ability to exactly identify combinations of multiple PTMs of a given glycoform. This is why today the analysis of intact glycoprotein is of huge interest, as it has the advantage of preserving the integrity of the analyzed glycoforms and the combination of their potential glycosylation sites [14]. The analysis of glycoproteins at their intact level has been done for years using SDS-PAGE and 2D-gel electrophoresis. These techniques allow protein separation but they lack of resolution and thus ability to identify and characterize the several glycoforms that could be present in the same sample in addition to their general low sensitivity. Therefore, alternative approaches have been developed and applied for high performance separation and sensitive detection and characterization of the glycosylation pattern of glycoproteins.

The main approaches for the analysis of intact glycoproteins are based on CE and LC in reversed-phase mode, in combination with UV or MS detection. Both methods require minimal, if any, sample preparation and allow the characterization of the macro-heterogeneity of the glycoforms. **Table 4** compiles the various studies that applied the intact approach for the analysis of gonadotropins. In several studies the analysis of the intact dissociated subunits is the main aim. Free α and β subunits, with activities that are distinct from those of the dimer hormones, can exist. These forms must be avoided in pharmaceutical preparations intended for human use since they have no specific hormone action and may lead to undesired effects. This is the main reason behind the interest in the analysis of the dissociated subunits. As an example, free circulating α-subunit has been shown to be useful when detected as a marker for gonadotropin-releasing hormone pulse generator activity, for isolated central hypothyroidism due to TSHβ gene

mutations, for hypogonadotrophic hypogonadism and, together with free FSHβ and LHβ, for pluri-hormonal pituitary tumors [51]. Moreover, various studies implemented the dissociation of the α and β subunits of gonadotropins and graded the ability of the method to separate and identify each. This allowed the characterization of the heterodimeric gonadotropin, leading to information about the alpha and beta subunit glycoforms. However, the use of UV detection or MS analyzer of low resolution has prevented until now the glycoform identification.

Table 4 Compiling of studies involving separation methods to analyze gonadotropins at the intact level.

Sample	Method	Outcome	Identification of Glycans/Glycoforms (Y: yes; N: no)	Ref
hCG CR127 Crude urinary hCG	CZE-UV	Separation of 7 to 8 glycoforms	N	[76]
Ovitrelle® and Pregnyl®	CGE-UV, CIEF-UV, CZE-UV, CZE-MS(TQ)	CGE led to a better resolution than SDS-PAGE and confirmed the large heterogeneity of hCG pI ranges of the hCG isoforms between 3.4 and 4.7, and 4.5 and 5.2 for rhCG and uhCG, respectively CZE led to at least 6 peaks, corresponding to different hCG glycoforms CZE-MS(TQ): fingerprinting	N	[17]
Purified preparations of hCG, LH, TSH	Dissociation (6 M guanidine or 10% TFA, 37°C, 1 h) + RPLC-UV	Separation of the intact form, α and β of each hormone	N	[18]
uhCG (Gonasi HP) r-hCG (Ovitrelle®)	RPLC-UV for oxidized forms, SEC-UV for content and purity	Separation and detection of α and β subunits for both rhCG and uhCG Demonstration of higher purity of rhCG in comparison to uhCG	N	[77]
Recombinant and native LH and hCG	RPLC-UV	Quantification of the different gonadotropins in the heterogeneous preparations No separation of glycoforms Hydrophobicity: hCGα < LHα < hCG < LH < hCGβ < LHβ	N	[52]
Spiked serum (or tissue cytosol) with biomarker proteins (including hCG) from Sigma	Incubation of sample with 5 Abs, each labeled with a different lanthanide + SEC-ICPMS and ELISA or RIA	SEC-ICPMS: accurate and precise determination of the complexes formed by metal-labeled antibodies with target proteins The method proved to be able to discriminate ovary and uterus tumor tissue samples from those of healthy subjects	N	[78]
WHO RRs for hCGα and hCGβ	Ion pair RPLC-MS, using a monolithic separation column	Separation of the heterogeneous glycoforms of both native hCGα and hCGβ	Y	[32]

Pregnyl®	Assignment of glycans identified by bottom up approach to the intact masses of hCGα glycoforms using bioinformatics	14 different glycoforms of hCGα, containing biantennary, partly sialylized hybrid-type glycans, including methionine-oxidized and *N*-terminally truncated forms identified Mass spectra of high quality were obtained for hCGβ, however, a database search mass accuracy of +/- 5 Da was insufficient to unambiguously assign the possible combinations of PTMs		
Bio similar: (rhCG) SB005, Ovidrel®, Ovitrelle®	SDS-PAGE and western blot , RPLC-MS, CIEF-UV, CZE-UV	No fragmentation or aggregation. Similar mobility, band pattern and molecular weights similar pI of individual isoform Similar profile and number of isoforms (not mentioned)	Y	[13]
Ovitrelle® and Pregnyl®	RPLC-MS (QTof)	More than 30 hCGα isoforms, which differ by their number and their nature in the two drugs Successful identification of five glycoforms for both rhCG and uhCG	Y	[16]
Ovitrelle® and Pregnyl®	HILIC-UV	Best HILIC separation obtained with the amide column and a mobile phase composed of water/ACN containing 0.1% of TFA 10 peaks resolved for each rhCG and uhCG.	N	[79]
CHO-, pituitary and urinary-derived FSH	RPLC-UV, SEC-UV	Detection and differentiation between recombinant, pituitary and urinary FSH based on retention times	N	[80]
uFSH and rFSH Puregon®	SEC-UV, SDS-PAGE	rFSH and uFSH exibit similar MW (43 kDa)	N	[70]
Urinary, pituitary and recombinant FSH	RPLC-UV, MALDI-TOF-MS	LC conditions preserved the heterodimer form of FSH	N	[46]
Recombinant FSH, LH and TSH. In addition to pituitary-derived preparations	Dissociation (37 ◦C overnight in acetic acid) + RPLC-UV, RPLC-MS, SDS-PAGE and MALDI-TOF	comparison of purity, hydrophobicity, molecular mass and charge distribution Relative molecular mass of heterodimers and of the subunits of native and recombinant hormones determined by MALDI-TOF	N	[51]

CZE-UV, capillary zone electrophoresis-ultraviolet detector; CGE-UV, capillary gel electrophoresis-ultraviolet; CIEF, Capillary isoelectric focusing; CZE, capillary zone electrophoresis; EGF, epidermal growth factor; EDN, eosinophil derived neurotoxin; FSH, follicle stimulating hormone; H-NMR, proton nuclear magnetic resonance; HILIC-UV, Hydrophilic interaction chromatography-ultraviolet; hCG, human chorionic gonadotropin; LH, luteinizing hormone; MALDI-TOF-MS, matrix-assisted laser desorption/ionization- time of flight; QTof, quadruple time of flight; RIA, radioimmunoassay; RPLC-UV, reverse phase liquid chromatography-ultraviolet detector; RP HPLC-ESI-MS, reverse phase high performance liquid chromatography-electrospray ionization-mass spectrometry; RPLC-MS, reverse phase liquid chromatography-mass spectrometry; SEC-ICPMS, size exclusion chromatography with inductively coupled plasma mass spectrometric detection; SEC-UV, size exclusion chromatography-ultraviolet detector; SDS-PAGE, sodium dodecyl sulphate–polyacrylamide gel electrophoresis; TFA, trifluoroacetic acid; WB, western blot.

2.3.1. Capillary electrophoresis

High performance capillary electrophoresis (CE) is simple, fast, and reproducible with high separation efficiency. Different modes of CE have been used for the analysis of proteins such as capillary zone electrophoresis (CZE) and capillary gel electrophoresis (CGE) [81]. CZE is the most commonly used method for the analysis of intact proteins, especially for the glycosylation characterization of biopharmaceuticals such as gonadotropins. It provides complex, but also very comprehensive profile revealing the overall glycosylation pattern [72]. One of the first studies using CZE succeeded in the separation of hCG glycoforms [76]. The importance of the addition of diaminopropane (DAP) to the running buffer was reported by testing different concentrations, which resulted in the UV detection of multiple, well-resolved peaks represented in **figure 14** [76]. The optimal selected concentration was 5.0 mM DAP that led to the separation of eight clearly defined peaks in less than 50 min.

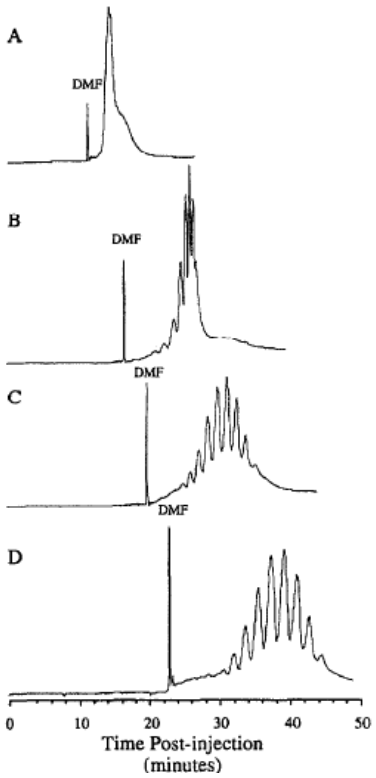


Figure 14. Effect of various concentrations of diaminopropane on the electrophoretic migration and CZE separation of hCG 2292 (4 mg/ml). The DAP concentrations used were 0 (A), 1.0 (B), 2.5 (C) and 5.0 mM (D). UV detection. **[76]**

CZE-UV was also used by *Thennati* et al. to compare the charge distribution and micro-heterogeneity between rhCG, SB005, and reference products, Ovidrel® and Ovitrelle® [13]. The profile and number of isoforms of SB005 and reference products were found to be similar, but without any identification of the glycoforms, prevented by the use of UV detection.

For the hyphenation of the CZE separation with MS, new conditions based on a volatile separation electrolyte were optimized by our team to improve the level of information that could be obtained when comparing two hCG-based drugs, Ovitrelle® (rhCG) and Pregnyl (uhCG) [17]. The electrophoregrams corresponding to the analysis of both drugs are represented in **figure 15**. For rhCG, at least 12 peaks corresponding to hCG glycoforms were detected, while

for uhCG the resolution was insufficient to distinguish individual peaks, probably due to the higher complexity of the uhCG sample.

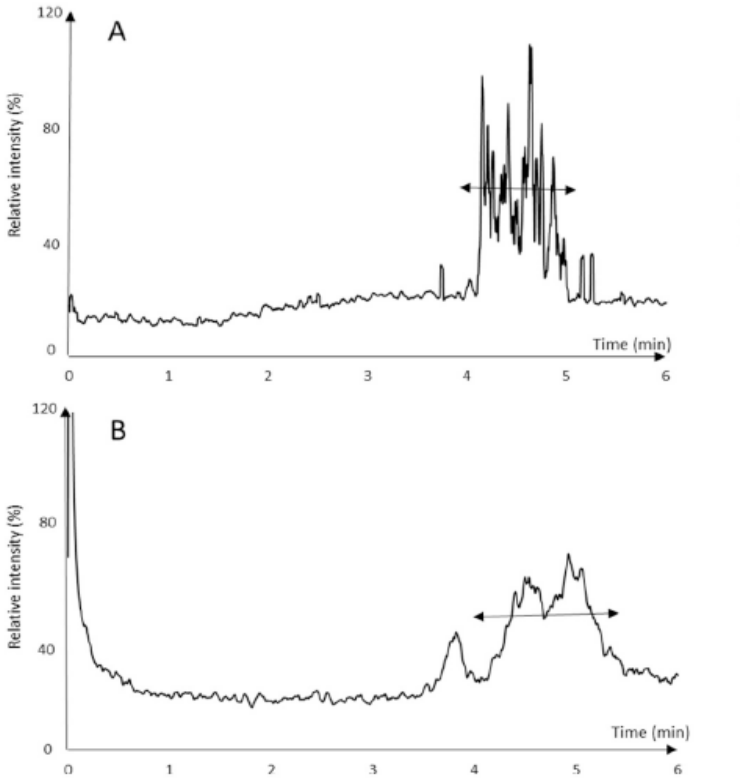


Figure 15. CZE-MS analysis of (A) rhCG and (B) uhCG [17]

Therefore it was demonstrated that CZE coupled to either UV or MS is a powerful tool for profiling of hCG-based drugs. However, the MS analyzer used, a quadrupole, prevented from the identification of the glycoforms. CE-HRMS has not been applied for the analysis of hCG but could be an interesting approach to be considered in the future.

2.3.2. Liquid chromatography

LC is a powerful method for the separation of intact proteins due to its high-resolving power, and good reproducibility. The LC separation modes are various, but the most commonly applied ones

in this field are HILIC and reversed-phase liquid chromatography (RPLC), due to their high compatibility with MS.

The main challenge facing the use of HILIC-MS for the separation of intact hCG glycoforms was the unavoidable use of TFA in mobile phase acting as an ion-pairing agent and which was essential to obtain an efficient separation, but resulted in a dramatic decrease in MS signal [79], [82].

This is why the most used LC mode for the analysis of glycoproteins at the intact level is RPLC. Indeed, several strategies have been implemented to overcome its limitations that are protein adsorption, pore exclusion, and low diffusion coefficient. They consist in using silica-based stationary phases with restricted access to residual silanols and particles with large pore sizes (>200 Å), applying elevated temperature, and adding an acidic additive such as FA in the mobile phase.

RPLC coupled to UV detection was used in various studies [46], [52], [77], [80]. These methods were successful to get profiles of hCG, FSH and LH glycoforms, but they lacked the ability of their identification. In this case, RPLC hyphenated with MS is mandatory. This is why, it is routinely used for profiling glycoforms in biopharmaceuticals and in biosimilar research, clone selection, batch-to-batch comparison, or as routine check for molecular weight and heterogeneity in biopharma industry [72]. Nevertheless, it can be highly challenging if multiple glycosylation sites are present, leading to a higher degree of heterogeneity, as it is the case for the gonadotropins. Glycoprofiling was first carried out by Toll *et al.* in 2006, which proposed identification for 14 glycoforms of hCGα and obtained mass spectra for the β-subunit as represented in **figure 16** that shows the complexity of the mass spectra [32]. The ion trap

quadrupole analyzer used in this study lacked mass accuracy and resolution, which could lead to misidentifications and prevent unambiguous assignment.

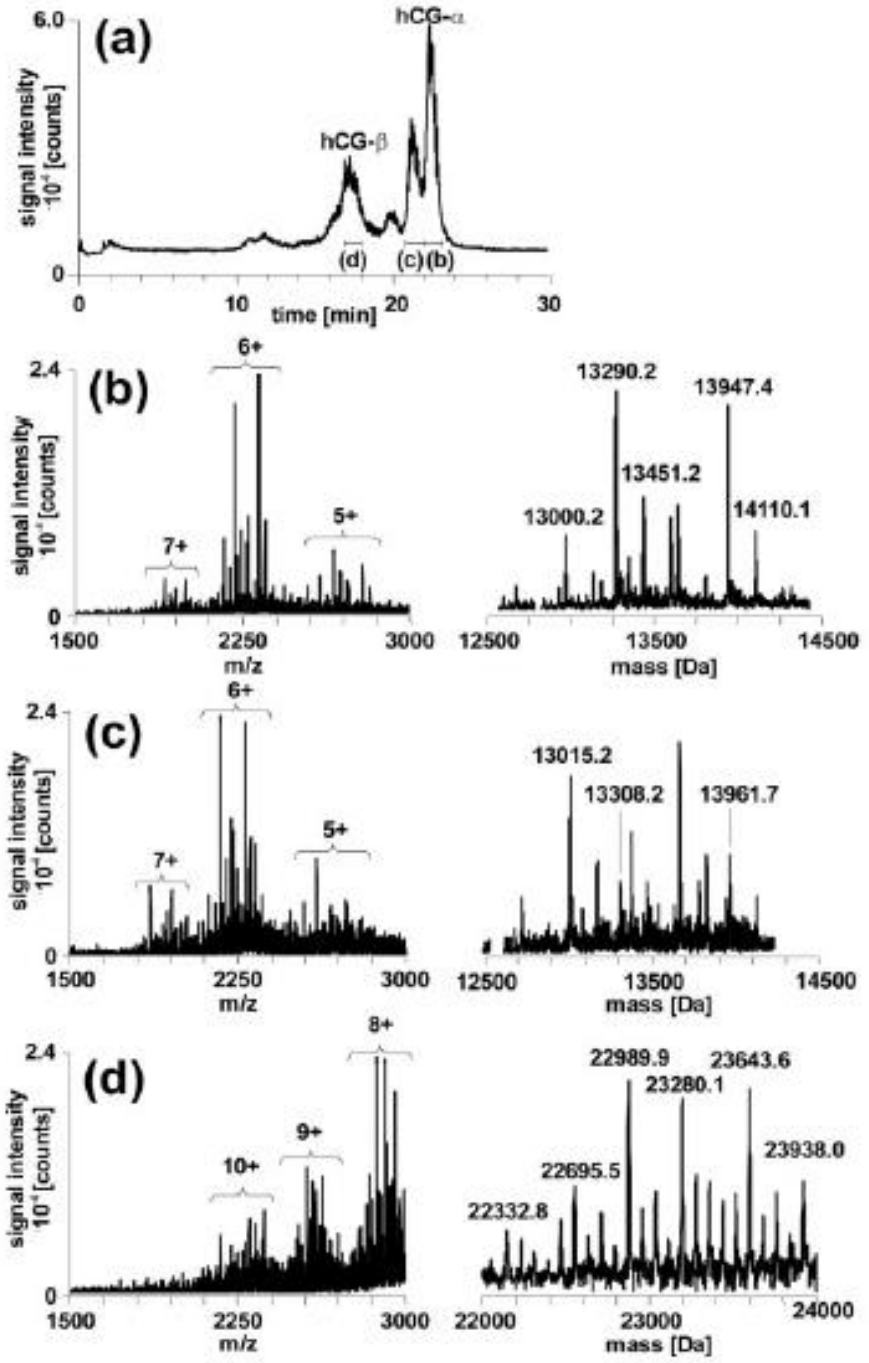


Figure 16. TIC (a) and the averaged mass spectra (b-d) extracted at the positions indicated in (a) showing the separation of hCGα and hCGβafter the RPLC-MS [32]

The previous results were the trigger for testing the coupling of RPLC with high resolution mass spectrometry. This was adapted by our team that started by developing a RPLC separation method coupled to a Q-TOF analyzer for the analysis of intact hCG [16]. This method succeeded in the detection of several hCGα glycoforms of two hCG-based drugs, Ovitrelle® and Pregnyl®, and the identification of the five main glycoforms was proposed by mass matching and represented in **figure 17**.

Unfortunately, the resolution of the Q-TOF analyzer used was still insufficient due to the overlapping of some isotopic patterns corresponding to different glycoforms.

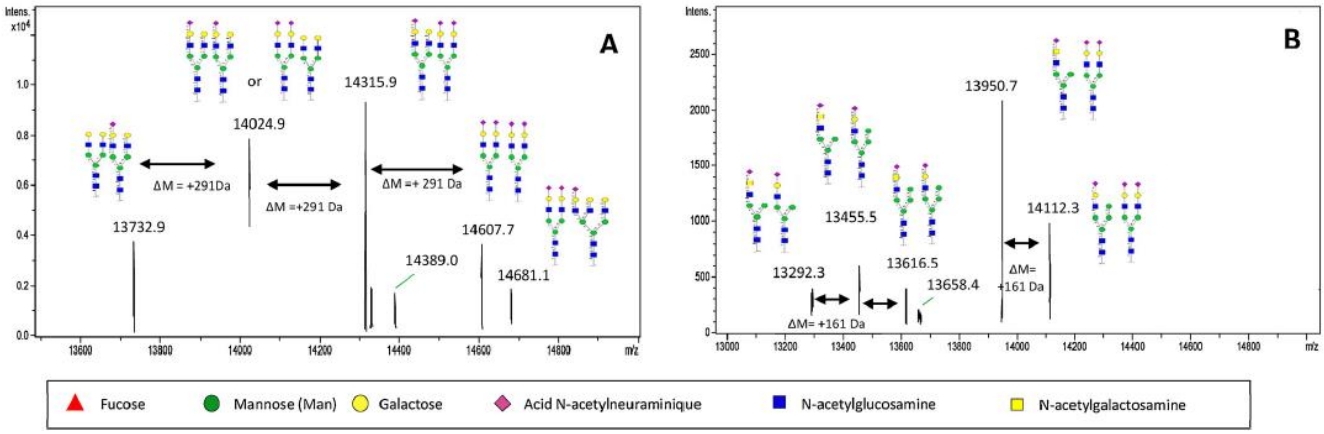


Figure 17. Deconvoluted spectrum of (A) rhCG and (B) uhCG between, with the different possible N-glycan structures obtained by RPLC-HRMS [16]

2.3.3. Advantages and disadvantages of intact approaches

The main and most essential advantage for the analysis of protein glycoforms at their intact level is the fact that it is the only way to correlate a glycoform with its different PTMs and determine the total number of glycoforms and their relative abundances in a given sample. This is very relevant when some glycoforms are expected to be biomarkers. Another pro fact is that the intact approach does not require complex sample pretreatment, therefore it is a fast method that

prevents serious modifications induced by sample handling. For all the mentioned reasons, the characterization of glycoproteins at the intact level has increased rapidly over the past decade but even though the main challenge facing the use of this approach is to achieve an efficient separation of glycoforms that have a high percentage of similarities especially for proteins with different glycosylation sites leading to a high number of glycoforms present in one sample.

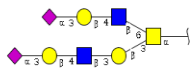
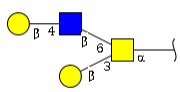
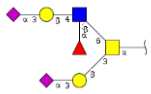
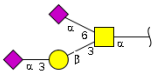
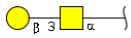
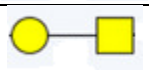
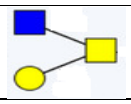
After gathering all these informations regarding the advantages and disadvantages of the methods applied for the analysis of gonadotropin glycoforms, the analysis of our protein of interest, hCG, at the intact level using LC, in the nano format to enhance sensitivity, hyphenated with a high mass resolution and a high mass accuracy Orbitrap analyzer will be investigated with the objective to develop a simple and fast method that enables the separation and detection of the numerous glycoforms expected to be present in a given sample. The next step would be precise structural assignment of each glycoform, which could help further in the identification of glycoforms.

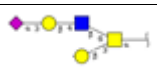
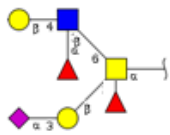
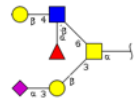
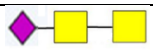
2.4. Strategies for structural assignments and structures of already identified glycans on gonadotropins

After the implementation of one of the mentioned separation and spectrometric methods, structural assignments of glycoproteins such as gonadotropins must occur by taking various parameters into consideration such as the retention time (t_R) delivered by separation techniques, which are a valuable characteristic of a particular glycoform, and of course its molecular weight (MW) that can be determined with great accuracy when using high resolution MS analyzers. MW is the most important parameter for structural assignment, by mass matching using for example the information also provided with complementary approaches such as the bottom-up

[75]. This is why, it was relevant to collect, whatever the analytical approach used, all information about O- and N- glycans identified on gonadotropins in literature. They are summarizes in **Tables 5 and 6**, respectively. Therefore, they present for all previously identified glycans, their molecular weight, composition, and the method used for their identification.

Table 5 Identified O-glycan structures reported in gonadotropin literature, their monoisotopic masses, composition, and method of detection.

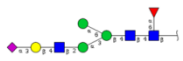
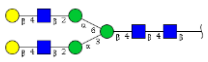
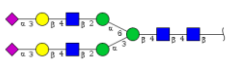
Structure	Monoisotopic mass Da *	Composition	Method	Ref
	1,695.6	Hex_3 HexNA c_3 $NeuAc_2$	Released glycans HPAEC- H NMR	[74]
	748.27	Hex_2 HexNA c_2	Bottom-up LC-MS/MS Released glycans HPAEC-PDA	[3] [84]
	1,476.52	Hex_2 HexNA c_2 $dHex_1$ Neu Ac_2	Released glycans HPAEC- H NMR	[74]
	965.33	Hex_1 HexNA c_1 $NeuAc_2$	Released glycans HPAEC- H NMR fluorescent labeling + HPAEC-Fluo Bottom-up GC-MS and LC-MS/MS	[74] [85] [3] [21]
	383.14	Hex_1 HexNA c_1	Bottom-up LC-MS/MS Released glycans HPAEC-PDA	[3] [84]
	203.07	$HexNAc_1$	Bottom-up LC- MS/MS	[3]
	203.07	$HexNAc_1$	Bottom-up LC-MS/MS	[56]
	512.16	$HexNAc_2$ Ne uAc$_1$	Bottom-up LC- MS/MS	[3]
	383.12	Hex_1 HexNA c_1	Bottom-up LC- MS/MS	[3]
	674.21	Hex_1 HexNA c_1 $NeuAc_1$	Bottom-up LC- MS/MS Released glycans fluorescent labeling + HPAEC-Fluo	[3] [21]
	877.28	Hex_1 HexNA c_2 $NeuAc_1$	Bottom-up LC- MS/MS	[3]
	586.19	Hex_1 HexNA c_2	Bottom-up LC- MS/MS	[3]

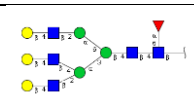
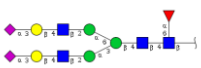
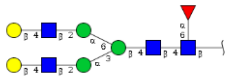
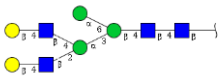
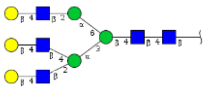
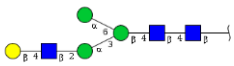
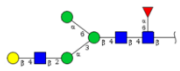
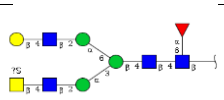
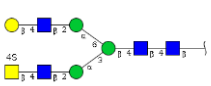
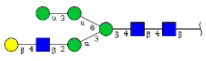
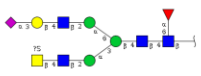
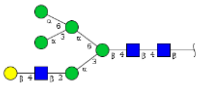
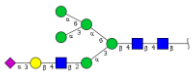
	1039.33	Hex_2 HexNAc$_2$ $NeuAc_1$	Bottom-up LC- MS/MS	[3]
	1331.43	Hex_2 HexNAc$_2$ $dHex_2$ $NeuAc_1$	Released glycans fluorescent labeling + HPAEC-Fluo	[21]
	1185.38	Hex_2 HexNAc$_2$ $dHex_1$ $NeuAc_1$	Released glycans fluorescent labeling + HPAEC-Fluo	[21]
	715.23	$HexNAc_2$ $NeuAc_1$	Released glycans fluorescent labeling + HPAEC-Fluo	[21]
	1273.4	Hex_2 HexNAc$_1$ $dHex_1$ $NeuAc_2$	Released glycans fluorescent labeling + HPAEC-Fluo	[21]

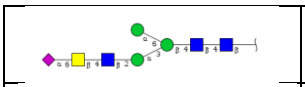
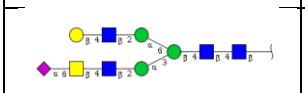
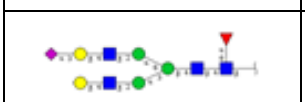
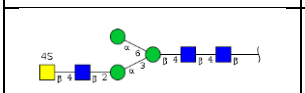
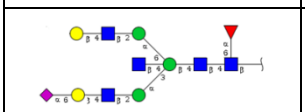
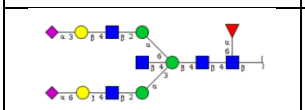
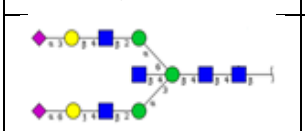
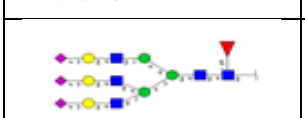
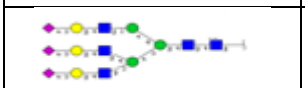
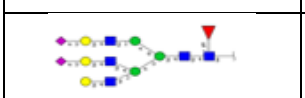
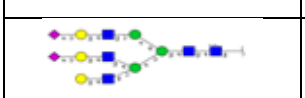
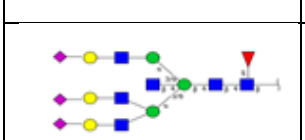
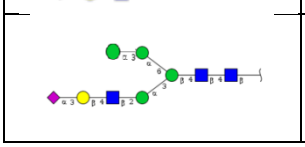
▼ Fucose ● Mannose ○ Galactose ■ N-acetylglucosamine □ N-acetylgalactosamine ◆ N-acetylneuraminic acid

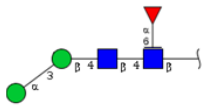
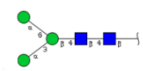
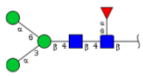
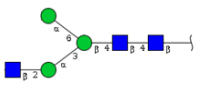
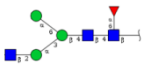
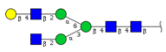
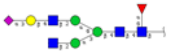
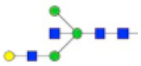
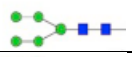
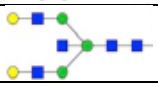
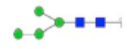
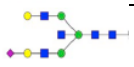
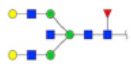
* Monoisotopic mass is given after the reduction of the mass of H_2O lost for the formation of each glycosidic bond, except for the last one

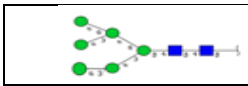
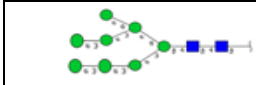
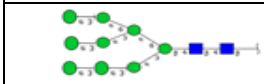
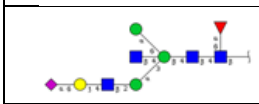
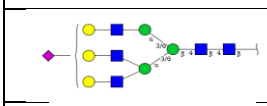
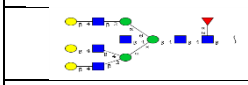
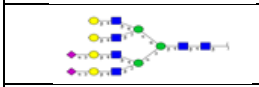
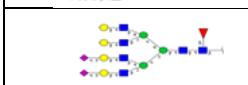
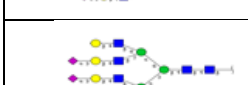
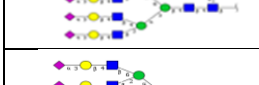
Table 6 Identified N-glycan structures reported in gonadotropin literature, their monoisotopic masses, composition, and method of detection.

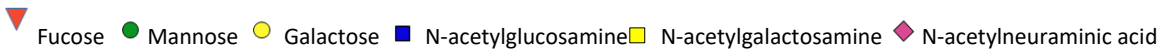
Structure	Monoistopic mass Da *	Composition	Method	Ref
	1566.5	Hex_4 $HexNAc_3$ $NeuAc_1$	Bottom-up: LC- 2D H NMR, LC-MS/MS, nanoLC-MS/MS, Immunoaffinity extraction + LC-MS Released glycans H NMR Intact RPLC-QTOF-MS	[83] [3] [14] [86][36] [16]
	1712.55	Hex_4 $HexNAc_3$ $dHex_1$ $NeuAc_1$	Bottom-up: LC-MS/MS, nanoLC-MS/MS	[3] [14]
	1478.48	Hex_5 $HexNAc_4$	Bottom-up nanoLC-MS/MS Intact RPLC-QTOF-MS Released glycans: LC-(TOF)MS, fluorescent labeling + HPAEC-Fluo, HPAEC-PDA	[14] [16] [31] [21] [84]
	2222.71	Hex_5 $HexNAc_4$ $NeuAc_2$	Intact NMR Bottom-up: LC- 2D H NMR, nanoLC-MS/MS, Immunoaffinity extraction + LC-MS Released glycans: H NMR, LC-(TOF)MS, fluorescent labeling + HPAEC-FluoIntact RPLC-QTOF-MS	[87] [83] [14] [36] [16] [31] [21] [86]

	2151.7	Hex_6 $HexNAc_5$ $dHex_1$	Released glycans HPAEC-PDA	[84]
	2368.76	Hex_5 $HexNAc_4$ $dHex_1$ $NeuAc_2$	Bottom-up: RLC- 2D H NMR, nanoLC-MS/MS, Immunoaffinity extraction + LC-MS Released glycans: HPAEC- H NMR, NMR, LC-(TOF)MS, fluorescent labeling + HPAEC-Fluo	[83] [74] [14] [36] [31] [21] [86]
	1786.58	Hex_5 $HexNAc_4$ $dHex_1$	Bottom-up nanoLC-MS/MS Released glycans: LC-(TOF)MS, fluorescent labeling + HPAEC-Fluo, HPAEC-PDA	[14] [31] [21] [84]
	1640.53	Hex_5 $HexNAc_4$	Bottom-up RLC-2D H NMR Released glycans HPAEC-PDA	[83] [84]
	2005.65	Hex_6 $HexNAc_5$	Released glycans: LC-MS/MS, HPAEC-PDA	[13] [84]
	1275.41	Hex_4 $HexNAc_3$	Bottom-up: LC-MS/MS, nanoLC-MS/MS Released glycans HPAEC-PDA	[3] [14] [84]
	1421.46	Hex_4 $HexNAc_3$ $dHex_1$	Bottom-up: LC-MS/MS, nanoLC-MS/MS	[3] [14]
	1827.6	Hex_4 $HexNAc_5$ $dHex_1$	Bottom-up nanoLC-MS/MS	[14]
	1,931.69	Hex_5 $HexNAc_4$ $NeuAc_1$	Bottom-up: nanoLC-MS/MS, Immunoaffinity extraction LC-MS Released glycans: H NMR, LC-(TOF)MS Intact RPLC-QTOF-MS, fluorescent labeling + HPAEC-Fluo	[14] [36] [16] [31] [21] [86]
	1681.55	Hex_4 $HexNAc_5$	Bottom-up nanoLC-MS/MS	[14]
	1,437.51	Hex_5 $HexNAc_3$	Bottom-up nanoLC-MS/MS	[14]
	2118.69	Hex_4 $HexNAc_5$ $dHex_1$ $NeuAc_1$	Bottom-up nanoLC-MS/MS	[14]
	1,599.57	Hex_6 $HexNAc_3$	Bottom-up nanoLC-MS/MS	[14]
	1,890.66	Hex_6 $HexNAc_3$ $NeuAc_1$	Bottom-up: LC- 2D H NMR, nanoLC-MS/MS Intact NMR	[83] [87] [14]

	1,607.58	Hex_3 $HexNAc_4$ $NeuAc_1$	Intact RPLC-QTOF-MS	[16]
	1,972.71	Hex_4 $HexNAc_5$ $NeuAc_1$	Bottom-up nanoLC-MS/MS	[14]
	2077.67	Hex_5 $HexNAc_4$ $dHex_1$ $NeuAc_1$	Bottom-up: LC-MS/MS, nanoLC-MS/MS, Immunoaffinity extraction + LC-MS Released glycans: H NMR, LC-(TOF)MS, fluorescent labeling + HPAEC-Fluo	[3] [14] [36] [31] [21] [54]
	1316.43	Hex_3 $HexNAc_4$	Bottom-up nanoLC-MS/MS	[14]
	2281.19	Hex_5 $HexNAc_5$ $dHex_1$ $NeuAc_1$	Released glycans HPAEC-H-NMR	[74]
	2571.28	Hex_5 $HexNAc_5$ $dHex_1$ $NeuAc_2$	Released glycans HPAEC- H NMR Bottom-up nanoLC-MS/MS	[74] [14]
	2425.78	Hex_5 $HexNAc_5$ $NeuAc_2$	Bottom-up: LC-MS/MS, nanoLC-MS/MS	[3] [14]
	3024.97	Hex_6 $HexNAc_5$ $dHex_1$ $NeuAc_3$	Released glycans: HPAEC- H NMR, H NMR, LC-(TOF)MS, fluorescent labeling + HPAEC-Fluo Bottom-up nanoLC-MS/MS	[74] [14] [36] [31] [21]
	2878.92	Hex_6 $HexNAc_5$ $NeuAc_3$	Bottom-up: LC-MS/MS, nanoLC-MS/MS Released glycans H NMR, LC-(TOF)MS, fluorescent labeling + HPAEC-Fluo	[3] [14] [36] [31] [21]
	2733.88	Hex_6 $HexNAc_5$ $dHex_1$ $NeuAc_2$	Bottom-up: LC-MS/MS, nanoLC-MS/MS Released glycans LC-(TOF)MS, fluorescent labeling + HPAEC-Fluo	[3] [14] [31] [21]
	2587.83	Hex_6 $HexNAc_5$ $NeuAc_1$	Bottom-up: LC-MS/MS, nanoLC-MS/MS Released glycans LC-(TOF)MS, fluorescent labeling + HPAEC-Fluo	[3] [14] [31] [21]
	3228.04	Hex_6 $HexNAc_6$ $dHex_1$ $NeuAc_3$	Released glycans HPAEC- H NMR	[74]
	1728.55	Hex_5 $HexNAc_3$ $NeuAc_1$	Bottom-up: LC- 2D H NMR, nanoLC-MS/MS	[83] [14]

	748.24	Hex_2 $HexNAc_2$	Bottom-up LC-MS/MS	[3]
	894.29	Hex_2 $HexNAc_2$ $dHex_1$	Bottom-up LC-MS/MS	[3]
	951.31	Hex_3 $HexNAc_2$	Bottom-up: LC-MS/MS, nanoLC-MS/MS	[3] [14]
	1097.36	Hex_3 $HexNAc_2$ $dHex_1$	Bottom-up LC-MS/MS	[3]
	1113.36	Hex_3 $HexNAc_3$	Bottom-up LC-MS/MS	[3]
	1259.41	Hex_3 $HexNAc_3$ $dHex_1$	Bottom-up LC-MS/MS	[3]
	1478.48	Hex_4 $HexNAc_4$	Bottom-up LC-MS/MS	[3]
	1624.53	Hex_4 $HexNAc_4$ $dHex_1$	Bottom-up LC-MS/MS	[3]
	1769.57	Hex_4 $HexNAc_4$ $NeuAc_1$	Bottom-up LC-MS/MS	[3]
	1915.62	Hex_4 $HexNAc_4$ $dHex_1$ $NeuAc_1$	Bottom-up LC-MS/MS	[3]
	2296.74	Hex_6 $HexNAc_5$ $NeuAc_1$	Bottom-up: LC-MS/MS, nanoLC-MS/MS Intact RPLC-QTOF-MS Released glycans fluorescent labeling + HPAEC-Fluo	[3] [14] [16] [21]
	2442.79	Hex_6 $HexNAc_5$ $dHex_1$ $NeuAc_1$	Bottom-up: LC-MS/MS, nanoLC-MS/MS	[3] [14]
	1478.48	Hex_4 $HexNAc_4$	Bottom-up nanoLC-MS/MS	[14]
	1234.39	Hex_5 $HexNAc_2$	Bottom-up nanoLC-MS/MS	[14]
	1843.6	Hex_5 $HexNAc_5$	Bottom-up nanoLC-MS/MS	[14]
	1113.36	Hex_3 $HexNAc_3$	Bottom-up nanoLC-MS/MS	[14]
	1072.34	Hex_4 $HexNAc_2$	Bottom-up nanoLC-MS/MS	[14]
	2134.69	Hex_5 $HexNAc_5$ $NeuAc_1$	Bottom-up nanoLC-MS/MS	[14]
	1989.65	Hex_5 $HexNAc_5$ $dHex_1$	Bottom-up nanoLC-MS/MS	[14]

	1396.44	Hex_6 $HexNAc_2$	Bottom-up nanoLC-MS/MS	[14]
	1720.54	Hex_8 $HexNAc_2$	Bottom-up nanoLC-MS/MS	[14]
	1882.59	Hex_9 $HexNAc_2$	Bottom-up nanoLC-MS/MS	[14]
	1915.62	Hex_4 $HexNAc_4$ $dHex_1$ $NeuAc_1$	Bottom-up nanoLC-MS/MS	[14]
	2296.74	Hex_6 $HexNAc_5$ $NeuAc_1$	Bottom-up nanoLC-MS/MS	[14]
	2354.77	Hex_6 $HexNAc_6$ $dHex_1$	Released glycans LC- MS/MS	[13]
	2952.95	Hex_7 $HexNAc_6$ $NeuAc_2$	Released glycans LC- (TOF)MS	[31]
	3099	Hex_7 $HexNAc_6$ $dHex_1$ $NeuAc_2$	Released glycans LC- (TOF)MS	[31]
	3244.04	Hex_7 $HexNAc_6$ $NeuAc_3$	Released glycans: LC- (TOF)MS, fluorescent labeling + HPAEC-Fluo	[31] [21]
	3390.09	Hex_7 $HexNAc_6$ $dHex_1$ $NeuAc_3$	Released glycans LC- (TOF)MS	[31]
	3535.13	Hex_7 $HexNAc_6$ $NeuAc_4$	Released glycans LC- (TOF)MS	[31]



* Monoisotopic mass is given after the reduction of the mass of H_2O lost for the formation of each glycosidic bond, except for the last one

2.5. Conclusion

LC separation hyphenated to MS detection with a high resolution and high mass accuracy analyzer appeared as the best approach for the characterization of gonadotropin glycoforms at intact level to get fast and informative analytical methods that is a major goal for clinical purposes. Therefore, in this thesis manuscript, the main goal was to focus on the development and optimization of a new analytical method based on nanoLC separation and HRMS detection

of hCG glycoforms starting by the analysis of pharmaceutical preparations and aiming for the future analysis of biological fluids as well.

Chapter 3: Identification and semi-relative quantification of intact glycoforms of human chorionic gonadotropin and follicle-stimulating hormone

The project initiation was performed previously by our team in collaboration with Thierry Fournier, University of Paris (UMR-S 1139). After having investigated CE methods, RPLC and HILIC modes were tested and optimized hyphenated with UV or MS for the detection of hCG glycoforms at the intact level. First HILIC-UV allowed observing significant different profiles analysing 2 hCG-based drugs, used as hCG standards. HILIC was next hyphenated with HRMS using a quadrupole-time-of-flight (Q-TOF) analyzer, but the indispensable use of TFA in the mobile phase to improve the LC selectivity caused a dramatic ion suppression effect in MS. A RPLC separation method coupled to HRMS using a Q-TOF analyzer was also developed, which led to the detection of several hCGα glycoforms of the two hCG-based drugs, Ovitrelle® and Pregnyl®, and an identification of the five glycoforms giving the highest MS signals was proposed by mass matching. However, the resolution of the Q-TOF analyzer was still insufficient due to the overlapping of some isotopic patterns corresponding to different glycoforms. Therefore, the objective of the PhD was first the transfer of the RPLC method from the conventional to the nano format to enhance the sensitivity and its hyphenation with a high mass resolution and high mass accuracy MS analyzer, an Orbitrap. A new non-automated way to treat the obtained data was also implemented. A recombinant hCG--based drug was used as standard. This chapter is in the form of a published article in Analytical and Bioanalytical Chemistry journal: A. Al Matari, A. Combès, J. Camperi, T. Fournier, V. Pichon, and N. Delaunay, “Identification and semi-relative quantification of intact glycoforms by nano-LC–

(Orbitrap)MS: application to the α-subunit of human chorionic gonadotropin and follicle-stimulating hormone," *Anal Bioanal Chem*, Jul. 2020, doi: 10.1007/s00216-020-02794-3.

Analytical and Bioanalytical Chemistry
https://doi.org/10.1007/s00216-020-02794-3

RESEARCH PAPER

Identification and semi-relative quantification of intact glycoforms by nano-LC–(Orbitrap)MS: application to the α-subunit of human chorionic gonadotropin and follicle-stimulating hormone

Amira Al Matari[1] · Audrey Combès[1] · Julien Camperi[1] · Thierry Fournier[2] · Valérie Pichon[1,3] · Nathalie Delaunay[1]

Received: 14 May 2020 / Revised: 16 June 2020 / Accepted: 26 June 2020
© Springer-Verlag GmbH Germany, part of Springer Nature 2020

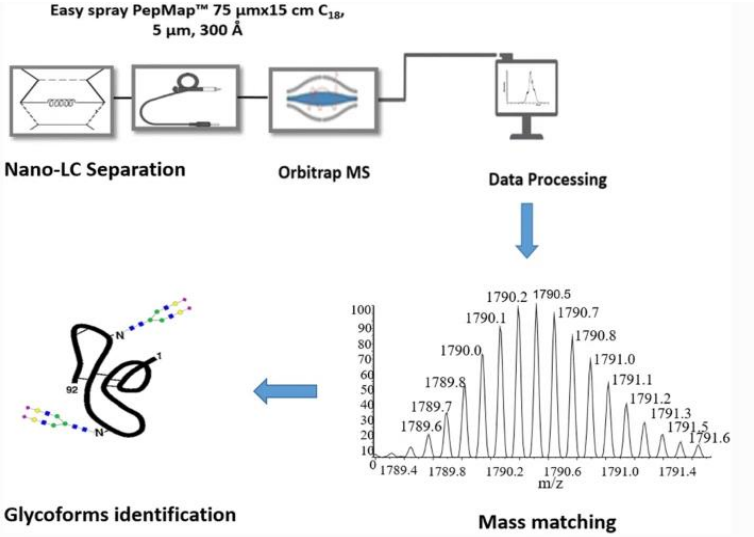
KEYWORDS

Follicle-stimulating hormone
Glycoprotein
Human chorionic gonadotropin
Intact protein
Mass spectrometry
Nano-liquid chromatography

ARTICLE INFO

Article history
Received: 14 May 2020 Revised: 16 June 2020 Accepted: 26 June 2020
Available online on 08 July 2020

GRAPHICAL ABSTRACT



ABSTRACT

Human chorionic gonadotropin (hCG) and follicle-stimulating hormone (FSH) belong to the family of glycoprotein polypeptide hormones called gonadotropins. They are heterodimers sharing the α-subunit structure that has 2 N-glycosylation sites. A method based on nano-reversed-phase liquid chromatography coupled to high-resolution mass spectrometry with an Orbitrap analyzer was developed for the first time to characterize the glycosylation state of the α-subunit at the intact level. A recombinant hCG-based drug, Ovitrelle®, was analyzed. This method combined with an appropriate data treatment allowed the detection of not only the major isoforms but also the minority ones with a high mass accuracy. More than 30 hCGα glycoforms were detected without overlapping of the isotopic patterns. The figures of merit of the method were assessed. The relative standard deviations (RSDs) of the retention time ranged between 0.1 and 6.08% (n = 3), with an average of 0.4%. The RSDs of the peak area measured on the extracted ion chromatogram of each glycoform are below 38% (n = 3), with an average of 16%, thus allowing semi-relative quantification. The ability to accurately profile glycosylated variants of hCGα was next demonstrated by comparing qualitatively and semi-quantitatively 3 batches of Ovitrelle®. The method was also used to analyze 3 batches of a recombinant FSH-based drug, Puregon®, and 30 FSHα glycoforms were detected and semi-quantified. This demonstrates the high potential of this method for fast quality control or comparison of the glycosylation of glycoprotein-based pharmaceutical preparations.

1.5.Introduction

Human chorionic gonadotropin (hCG) and follicle-stimulating hormone (FSH) belong to a family of related hormones called gonadotropins [88]. They are heterodimers sharing the α-subunit structure non-covalently bound to a specific β-subunit resulting in differences in their physiological roles and signaling. The α-subunit has 2 N-glycosylation sites, whereas the β-subunit of hCG has 2 *N*- and 4 *O*-glycosylation and of FSH contains 2 N-glycosylation sites. Similar to other hormones such as luteinizing hormone, hCG and FSH are used as therapeutic agents in reproductive medicine. However, their glycosylation state has an impact on their stability, solubility, biological activity, potential toxicity, and clearance [20]. This is why their characterization is required throughout their lifecycle (discovery, preclinical, and clinical steps), production, storage, and delivery.

However, protein glycosylation remains one of the most challenging problems in the analysis of post-translational modifications (PTMs), due to structural differences of glycans, number and types of glycosylation sites, and different degrees of sites occupancy leading to a large number of isoforms that are called glycoforms. Studies generally conducted for the analysis of glycoproteins can be done at three levels that complement each other: glycopeptides [89]–[91], released glycans [74], [91] , and intact glycoproteins [72], [81], [91]. The most commonly used method, called bottom-up, is the analysis of the glycopeptides resulting from the enzymatic digestion of the protein that allows the localization of the glycans along the protein chain in addition to the determination of the molecular weight (MW) of each glycan. The main disadvantages of this method are that it results in the loss of information regarding the different combinations of the glycopeptides together and the lack of certainty of the existence of glycoforms of the intact proteins. Other disadvantages are related to the numerous tedious

sample preparation steps such as alkylation, reduction, and enzymatic digestion. These additional steps are costly, time-consuming, and could lead to the loss or modifications of certain glycopeptides. The same disadvantages are faced upon the use of the released glycans in addition to the fact that the obtained glycan profile does not give information on the micro-heterogeneity in presence of various glycosylation sites, and one glycan cannot be attributed to a given glycoform. To avoid the mentioned disadvantages, there is an ongoing effort to develop analytical methods at the intact level to evaluate their macro-heterogeneity by combining all glycans associated to each glycoform, and to obtain a global comprehensive picture of the glycosylation pattern [81]. Nevertheless, separation and identification of intact glycoforms can be a difficult task due to their high MW, number, and structural similarity.

By focusing on hCG and FSH, high-performance capillary zone electrophoresis (CZE) was used to separate their isoforms [17], [92], [93]. For instance, our group recently developed a CZE separation method hyphenated with mass spectrometry (MS) allowing a fast glycoform fingerprinting of 2 hCG-based drugs with a turnaround of only ≈6 min [17]. Thakur *et al.* involved a CZE separation hyphenated with a linear trap quadrupole–Fourier Transform MS analyzer for assignment of intact glycoforms of 2 recombinant α-subunits of hCG [92]. This method allowed in 20 min the identification of 9 and 20 major glycoforms of the α-subunit of recombinant hCG obtained from Chinese hamster ovary (CHO) and murine cell line, respectively.

With regard to high performance liquid chromatography (LC), we recently optimized a hydrophilic interaction liquid chromatography (HILIC) separation hyphenated with UV that gave significantly different chromatographic profiles for 2 hCG-based drugs [79]. Its hyphenation with high-resolution MS (HRMS) using a quadrupole-time-of-flight (Q-TOF) analyzer was next

studied, but the use of TFA that had been shown to be necessary for the separation of glycoforms caused a dramatic ion suppression effect in MS [82]. Moreover, the biggest challenge in HILIC for the analysis of intact proteins is the injection conditions, i.e. to have sample media that do not induce a significant deformation of chromatographic peaks while allowing the solubilization of the protein without causing its precipitation. Typically, the protein samples are prepared in 100% water and a very low volume was injected on the HILIC column to prevent from the peak deformation induced by the high percentage of water. Compared to the RPLC mode where the analytes are pre-concentrated at the head of the column, the high percentage of water is a major drawback for the retention of intact proteins on the HILIC columns. Reversed-phase LC (RPLC) for glycoform separation of hCG or FSH was used far more before the last [16], [32], [46], [77], [94]. To get more information on the nature and the number of glycoforms, the coupling of RPLC with an MS detection is required as it was first carried out by Toll *et al.* who proposed an identification for 14 glycoforms of hCGα and obtained mass spectra for the β-subunit [32]. However, the ion trap quadrupole analyzer used in this study lacked mass accuracy and resolution which could lead to misidentifications and prevent unambiguous assignment. This is why we recently developed a RPLC separation method coupled to HRMS using a Q-TOF analyzer for the analysis of intact hCG [16]. It led to the detection of several hCGα glycoforms of two hCG-based drugs, Ovitrelle® and Pregnyl®. An identification of the five main glycoforms was proposed by mass matching, even if, unfortunately, the resolution of the Q-TOF analyzer was still insufficient due to the overlapping of some isotopic patterns corresponding to different glycoforms.

This is why in the current study, for the first time, a method involving a high mass resolution and a high mass accuracy Orbitrap analyzer was used for the analysis of hCG and FSH at the intact

level to detect their α-subunit glycoforms. The LC separation was performed at the nanoLC format to enhance the sensitivity. Even if nowadays nanoLC is used in routine for the bottom-up approaches in proteomics field [95], there are only a few papers reported its use for the analysis of intact glycoforms of targeted proteins [96], [97]. This nanoLC-(Orbitrap)HRMS method combined to an adapted data treatment allowed the detection of more than 30 hCGα glycoforms in the Ovitrelle® hCG-based drug (a recombinant hCG-based drug, rhCG) without overlapping of the isotopic patterns. After determining some figures of merit of the analytical method, the semi-relative quantification of each glycoform was next performed for the first time and the method was used to compare three batches of rhCG. The method was next successfully used for the qualitative and semi-quantitative glycosylation characterization of FSHα of Puregon®, a recombinant FSH-based drug (rFSH) and three batches were analyzed. The batch-to-batch studies of the glycoform heterogeneity demonstrated the utility of the nanoLC-HRMS for quality control of biopharmaceutical preparations.

1.6.Experimental Part

1.6.1. Materials and Reagents

HPLC-grade acetonitrile (ACN), formic acid (FA), and LC-MS-grade water were supplied by Carlo Erba (Val de Reuil, France). Puregon® and Ovitrelle® (Organon, Oss, The Netherlands) are drugs made of recombinant FSH and hCG from a CHO cell line; they will be written as rFSH and rhCG, respectively, in the rest of the manuscript. rFSH contains 900 UI of FSH in a 1.08 mL solution. rhCG contains 250 µg of hCG resuspended in 500 µL of water.

1.6.2. Sample preparation

The hormone-based drug stock solutions were desalted and filtered for 6 min at 4°C with ultra-centrifugal units (Microcon) (Merck, Darmstadt, Germany) with a MW cut-off of 10 kDa. This

step was performed to remove additives and excipients from the formulation before the analysis. Then the residues were washed 3 times with water. Finally, the samples were recovered from the inverted membrane by centrifugation for 4 min at 14,000 g (about 11,000 rpm) and the volumes were adjusted with water to obtain a final concentration of 160 µg mL^{-1} for each batch of FSH and 500 µg mL^{-1} for each batch of hCG. In this study, 3 batches of rFSH and 3 of rhCG were analyzed each in triplicate analysis.

1.6.3. NanoLC-MS system

The analyses were carried out using a nanoLC system (NC 3500 RS, ThermoFisher Scientific, Courtaboeuf, France) coupled to an Orbitrap MS (Exactive plus, ThermoFisher Scientific). The column used was an EASY-Spray PepMap RSLC C18 (150 x 0.075 mm internal diameter, 300 Å) packed with 5 µm particles (ThermoFisher Scientific). The mobile phase was composed of (A) 2% FA in H_2O/ACN (98/2; v/v) and (B) 2% FA in H_2O/ACN (1/9; v/v), and the column temperature was 60°C. The injection volume was set at 70 nL and the flow-rate at 300 nL min^{-1}. The elution gradient starts from 10 to 30% of ACN in 30 min, followed by a rapid increase to 62% of ACN within 5 min, then a 5 min plateau, and finally a return to the initial 10% of ACN concentration and kept for 15 min for equilibration. The EASY-spray source parameters were optimized as follows: capillary voltage, 1.7 kV; drying temperature, 280°C; and sheath, auxiliary, and sweep gas flows were set at 0. The focusing voltages were tuned weekly for the transmission of the standard mixture positive ion calibration solution (ref 88323, ThermoFisher Scientific) and the negative ion calibration solution (ref 88324, ThermoFisher Scientific).The mass spectrometer was operated in the positive mode, in full MS, using a scan range from 1000 to 2500 m/z with a scan rate of 1 spectrum/s. The mass resolution was set to 140,000 at 200 m/z, the AGC target was 3e6, and the maximum injection time was 200 ms. The parameters for

deconvolution used were mass tolerance for m/z of 30 mDa, and a minimal charge state of 5, with at least 3 different charge states to identify a compound, accuracy mass tolerance for deconvoluted MW was the 300 mDa. The deconvolution spectra were processed using Freestyle 1.3 SP2 (ThermoFisher Scientific).

1.7.Results and discussion

1.7.1. Development of the nanoLC-MS method

We previously optimized a RPLC separation of hCG glycoforms using 3.6 µm core-shell C18-bonded silica-based particles packed in a column with conventional dimensions (150 × 2.1 mm) [16]. This method was here transferred to the nano format to enhance the sensitivity and detect minor glycoforms. The nanoLC parameters were optimized starting from our previous optimized conditions in LC. Due to the limited choice of nanoLC columns, an EASY-Spray PepMap RSLC C18 (150 x 0.075 mm) column was selected because it drastically reduces the extra-column band broadening, one of the most negative effects in nanoLC, by having electrospray emitter directly connected to the column, thus minimizing post-column dead volumes. In addition, this temperature-controlled column enables a perfect stability of the column temperature during the analysis compared to other nanoLC columns, since the capillary between the column and the MS source is heated. A column temperature of 65°C was previously optimized, but the value was set here at 60°C, since it is the maximal recommended value given by the manufacturer. A similar mobile phase was used consisting of a water-ACN mixture but acidified with 2% of FA instead of 0.1% to favor the multiple ionization of hCG glycoforms. The gradient previously optimized was slightly adjusted to these new temperature and column to finally recover the chromatographic profile previously obtained with the conventional LC column (2.1 mm) [16]. The Orbitrap MS parameters were also optimized. Various spray voltage values were tested (2.5,

2.2, 2, 1.7, 1.5, and 1.2 kV) to minimize the possibility of in-source fragmentation. The final selected spray voltage was 1.7 kV, which enabled more glycoforms to be detected with the highest signal intensity and generating the fewest smaller m/z fragments while keeping a stable spray. rhCG was used as standard sample after a simple filtration step. Using these optimized conditions, the typical total ion current (TIC) obtained for the analysis of hCG is presented **Figure 18A**. hCG sample contains so many glycoforms that they cannot be baseline resolved giving rise to a large massif of peaks between 15 and 21.5 min, as it was already observed in LC) [16]. However, the separation power of the MS analyzer can partially overcome it.

The objectives of transferring the analysis to the nanoLC format were to reduce sample consumption and improve sensitivity. With respect to sample consumption, the injected volume decreased from 5 µL in LC to 70 nL in nanoLC, thus corresponding to a decrease in the injected quantity of a factor 14 as the hCG concentration was increased from 100 to 500 µg mL^{-1}. Regarding the sensitivity, a theoretical gain of 784 based on the ratio of the square column diameters was calculated. The signal to noise ratio values were calculated on the TIC for the most intense peak obtained in LC [16] and in nanoLC (**Figure 18 A**). It allowed the calculation of the experimental sensitivity enhancement and a value of about 40 was obtained, taking into account the different injected amount values, which is lower than expected but still interesting. Several hypotheses can explain this difference: (i) the relative high-injected volume in nanoLC, (ii) the presence of dead volumes in the system, both parameters causing a band broadening and then a decrease of peak heights, and (iii) the use of different MS detectors.

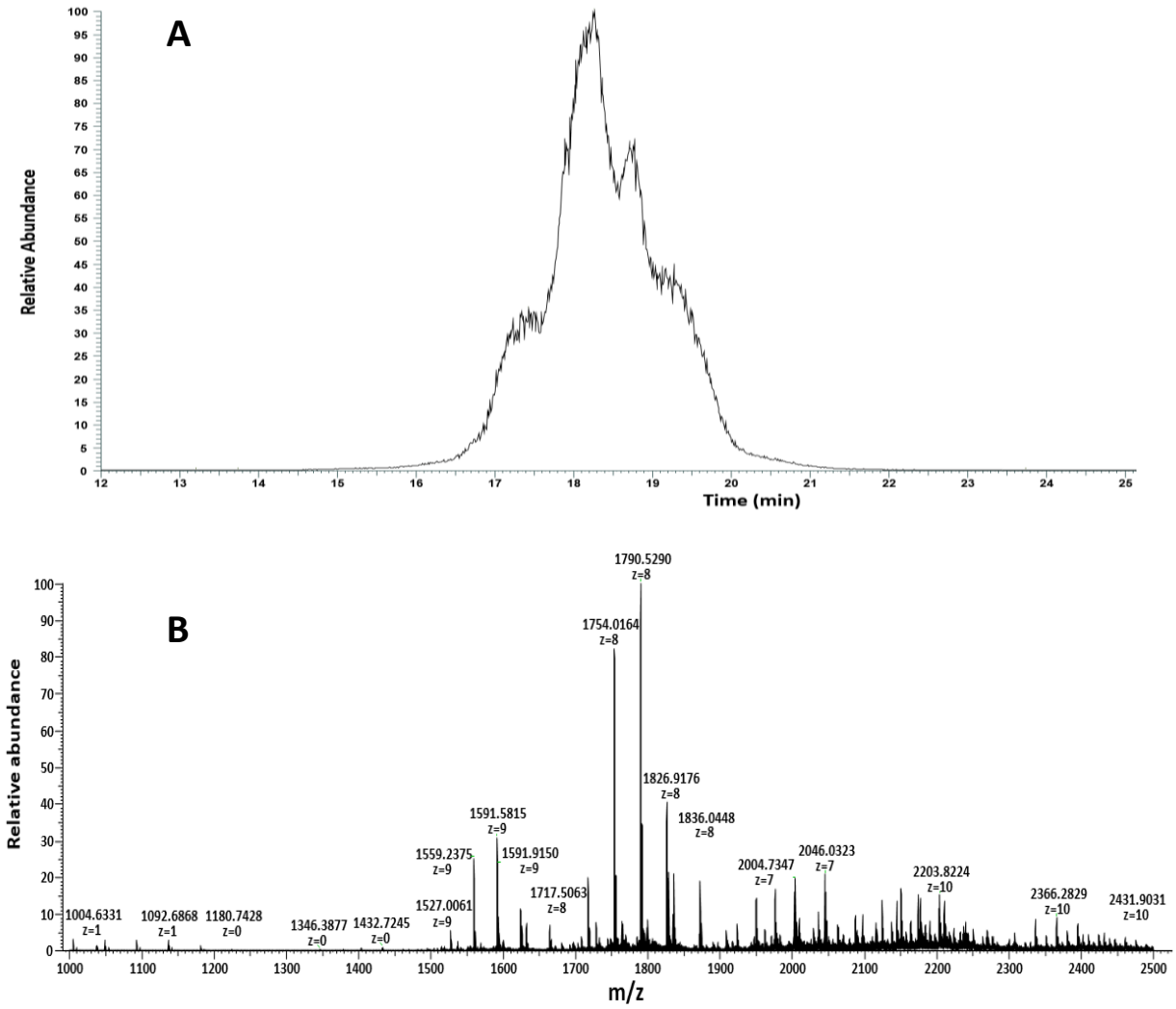


Figure 18. (A) TIC of hCG analyzed by nanoLC-HRMS. (B) Average mass spectrum corresponding to the totality of the chromatographic peak observed (retention time range: 15.5-21 min). Conditions: see section 2.3. rhCG, batch 1, run 1.

1.7.2. Nano-LC-HRMS data treatment for the identification of glycoforms

Electrospray ionization led to multiple charged values for each molecular mass and an isotopic cluster for each nominal mass-to-charge (m/z) value ratio, thus leading to highly complex mass spectra. For complex samples studied in MS-based proteomics, overlapping isotope patterns are commonly encountered and are problematic for accurate mass determination. To overcome these problems, the first step in spectral interpretation usually consists in spectral deconvolution,

which converts the complex spectrum to a list of monoisotopic deconvoluted masses [98]. In the current study, two deconvolution methods were compared with the same algorithm in Freestyle (1.3) software.

We first applied the most commonly used deconvolution method by selecting the whole chromatographic peak (retention time between 15.5 and 21.5 min) at once and the resulting average mass spectrum with the determined charge states is presented in **Figure 18B**. The deconvolution algorithm was next performed to determine the corresponding MWs. It allowed the detection of only 12 compounds with (M+H) ranging between 13,521.8 and 14,963.2 Da which correspond to hCGα glycoforms.

For the identification of glycoforms, the MWs of the detected putative glycoforms were compared to those of the primary amino acid sequence of hCGα plus the MW of all the combinations of two potential N-glycans already described in the literature for hCGα [32], [99] and to the ones that we recently identified with a bottom-up approach [14]. From a practical point of view, the hCGα backbone sequence was thus converted to its elemental composition. Ten hydrogens were eliminated from the elemental composition due to the fact that five disulfide bridges are present in this subunit. Then the elemental composition of two identical glycans bearing two sialylated antennas (HexNAc(4)Hex(5)NeuAc(2)) was determined. These glycans were chosen because they are considered to be the most commonly encountered [14], [32], [99]. The combined elemental composition of hCGα and these two glycans ($C_{84}H_{158}N_6O_{75}$) was used in a MW calculator to obtain the theoretical isotopic profile of this glycoform ([M+H] = 14,598.2 Da) [100]. This value was present in the data of the current study as one of the most intense isotopes with a mass tolerance of 0.03 Da for the m/z and mass tolerance of 0.3 Da for the MW. The experimental isotopic profile (**Figure 19**) matches with the theoretical one (**Figure**

S1) and allows the identification of the mentioned glycoform. The structure of this glycoform was the starting point of mass matching to determine the other structures. **Figure 20** represents the corresponding average deconvoluted mass spectrum with the 12 detected glycoforms with 9 of which were identified. The structures of the corresponding N-glycans present on each hCGα glycoform are also represented. The nanoLC-HRMS method allowed thus to confirm the identity of the glycoforms that could have been only assumed using our previously developed LC-(Q-TOF)MS method [16]. All these glycoforms were confirmed by a complementary bottom-up approach, which was applied to the same hCG sample [14].

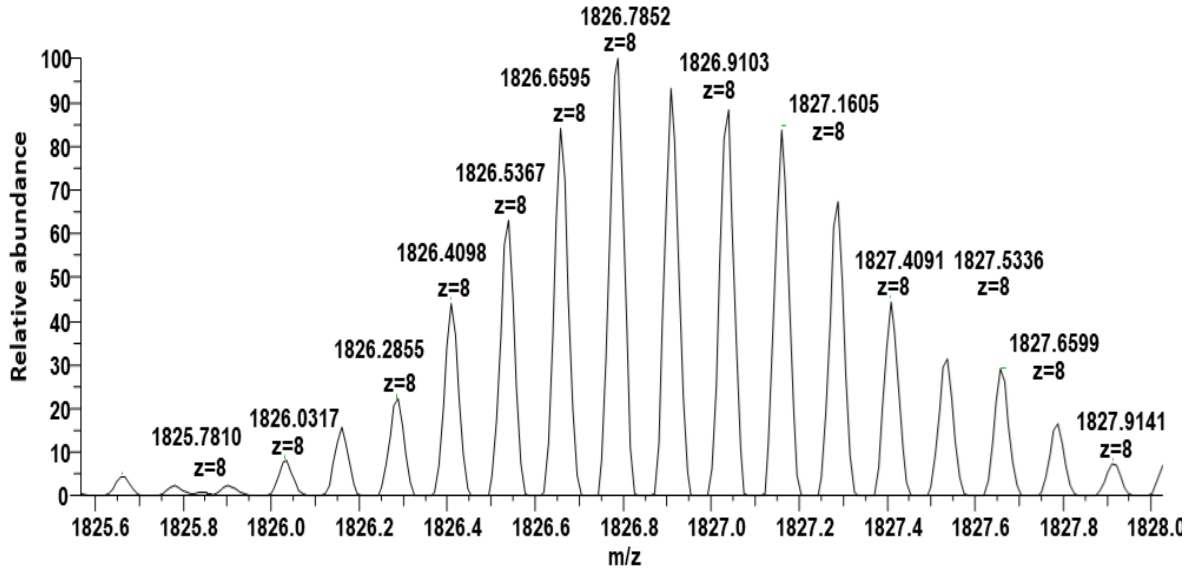


Figure 19. Zoom on average mass spectrum on the mass range [1825.6 – 1828.0 Da] showing the experimental isotopic profile (m/z = 1825.78) corresponding to the most commonly described hCGα glycoforms (hCGα with 2 HexNAc(4)Hex(5)NeuAc(2) N-glycans).

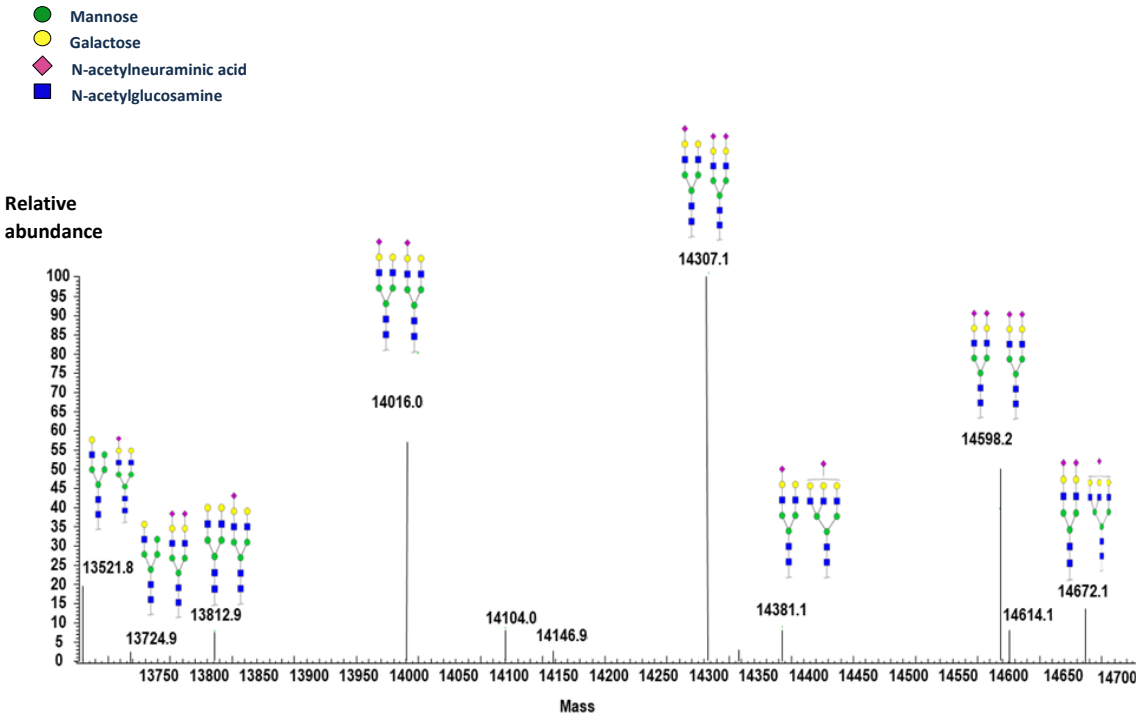


Figure 20. Deconvoluted average mass spectrum, determined over the entire chromatographic peak observed in Figure 1, representing the 12 detected hCGα glycoforms and the N-glycan structures of the 9 glycoforms that were identified among the 12.

The later average MS spectrum over 7 min of retention time enabled only the detection of the majority peaks. To improve the data processing to detect the minority forms, a second data treatment approach was considered. It consists in cutting the chromatogram in successive time slices of 30 s, starting from the retention time of 15 till 22 min, and also in slices of 50 s, starting from 15.5 till 21.5 min. The average deconvoluted mass spectrum for each slice was obtained resulting in the achievement of a list of M+H corresponding to detected glycoforms present in rhCG. Next, the results of both types of division were combined to reduce the risk of data loss at the boundaries of the performed cuts. A list of 28 M+H ranging between 13,050.6 and 15,254.4 Da corresponding to hCGα glycoforms with different retention times was obtained,

which increased by a factor of 2 the number of detected glycoforms by comparison with results from the first data treatment approach.

This second data treatment approach was applied to elucidate the structure of the two N-glycans present on each glycoform for 3 different batches of rhCG. **Table 7** presents the M+H of the detected glycoforms in these batches to compile the maximum information obtained by this method. It is worthwhile to mention that the retention time range where glycoforms were identified corresponds on the TIC to the central part of the peak massif with the highest intensities. The corresponding N-glycan structures for each hCGα glycoform were determined by mass matching and by comparison with results from the bottom-up approach [14] and with all the N-glycan structures identified for gonadotropins gathered by GlyGen database [101]. It was sometimes not possible to elucidate the glycoform structure that may result from a combination of glycosylation plus other PTMs. Of course, the previously 12 detected glycoforms were present but this second approach, even if it is non-automated and time-consuming, allowed also the detection of additional glycoforms that have sometimes a much less intense signal. Indeed, a ratio of 200 was obtained between the highest and the smallest intensity of the detected glycoforms. It thus provides a very much more detailed profile of hCGα.

Table 7 also shows all the masses detected in each run of each batch. It appears that all the detected masses are detected in each run for a given batch. There are only one exception in batch 2 and one in batch 3, for which the mass value was detected only in two runs out of the three. For batch 3, it corresponds to a glycoform of low intensity and that has a retention time similar to 4 other glycoforms (the 5 glycoforms are eluted in 1.2 s), which may explain that in one run, the ionization process of this glycoform was not sufficiently efficient to detect it. For batch 2, it is more difficult to find an explanation because it concerns a glycoform with a higher intensity that

does not co-elute with other ones. Nevertheless, regarding the total number of detected glycoforms and the complexity of the sample, it can be concluded that the direct nanoLC-HRMS analysis of the hCG-based drug in duplicate should be sufficient to get consistent hCGα glycoprofiling.

Table 7 MW of each hCGα glycoform detected after nanoLC-HRMS analysis of rhCG (3 batches, 3 runs each) and use of the second data treatment approach with glycan structure suggestions for some of them in comparison with bottom-up data (marked with *) or with GlyGen database (marked with **), while when those confirmed by both are marked with ***. x: no structure to propose. The notation 1st and 2nd glycan is not correlated to their localization on the 2 N-glycosylation sites of hCGα.

Mass (Da)	**Number of times the glycoforms was detected among the three runs**			**1st N-Glycan Proposed structure**	**2nd N-Glycan Proposed structure**	**Structure identification method**
	Batch 1	**Batch 2**	**Batch 3**			
13050.6	3/3	3/3	3/3	x	x	
13434.8	3/3	3/3	3/3	x	x	
13521.8	0/3	0/3	3/3	HexNAc(3)Hex(5)* or HexNAc(5)Hex(4)** or HexNAc(4)Hex(5)** or HexNAc(5)Hex(4)NeuAc(1)**	HexNAc(4)Hex(5)NeuAc(1)* or HexNAc(4)Hex(3)NeuAc(1)** or HexNAc(3)Hex(5)NeuAc(1)** or HexNAc(4)Hex(3)**	* ** ** **
13609.8	3/3	3/3	3/3	HexNAc(3)Hex(4)NeuAc(1) or HexNAc(3)Hex(5)NeuAc(1) or HexNAc(3)Hex(4)NeuAc(1)	HexNAc(3)Hex(6)NeuAc(1) or HexNAc(3)Hex(5)NeuAc(1) or HexNAc(3)Hex(6)NeuAc(1)	* ** **
13650.9	3/3	3/3	2/3	HexNAc(3)Hex(4) or HexNAc(4)Hex(5)NeuAc(1) or HexNAc(3)Hex(6)NeuAc(1)	HexNAc(4)Hex(5)NeuAc(2) or HexNAc(3)Hex(4)NeuAc(1) or HexNAc(4)Hex(3)NeuAc(1)	* ** **
13724.9	3/3	3/3	3/3	HexNAc(4)Hex(5) or HexNAc(3)Hex(4) or HexNAc(5)Hex(6) or HexNAc(6)Hex(4)NeuAc(1)	HexNAc(4)Hex(5)NeuAc(1) or HexNAc(5)Hex(6)NeuAc(1) or HexNAc(3)Hex(4)NeuAc(1) or HexNAc(4)Hex(3)	* ** ** **

13756.8	0/3	3/3	3/3	x	x	
13812.9	3/3	3/3	3/3	HexNAc(3)Hex(5) or HexNAc(3)Hex(5)NeuAc(1)	HexNAc(4)Hex(5)NeuAc(2) or HexNAc(4)Hex(5)NeuAc(1)	* **
13847.8	3/3	3/3	3/3	x	x	
13852.9	0/3	3/3	3/3	HexNAc(5)Hex(4)NeuAc(1)	HexNAc(3)Hex(5)NeuAc(1)	**
13942.9	3/3	3/3	3/3	HexNAc(5)Hex(4)	HexNAc(6)Hex(3)Fuc(1)	**
13974.0	3/3	3/3	3/3	HexNAc(3)Hex(6)NeuAc(1)	HexNAc(4)Hex(5)NeuAc(1)	*
14016.0	3/3	3/3	3/3	HexNAc(4)Hex(5)NeuAc(1)	HexNAc(4)Hex(5)NeuAc(1)	***
14024.1	3/3	0/3	3/3	x	x	
14032.0	3/3	3/3	3/3	HexNAc(5)Hex(4)NeuAc(1) or HexNAc(3)Hex(5)NeuAc(1) or HexNAc(4)Hex(4)Fuc(1)	HexNAc(5)Hex(4)Fuc(1) or HexNAc(5)Hex(6)Fuc(1) or HexNAc(6)Hex(4)NeuAc(1)	**
14048.9	0/3	3/3	3/3	HexNAc(3)Hex(6)	HexNAc(5)Hex(6)NeuAc(1)	**
14089.0	3/3	0/3	3/3	x	x	
14104.0	3/3	3/3	3/3	HexNAc(4)Hex(5)NeuAc(2)	HexNAc(3)Hex(5)NeuAc(1)	**
14139.0	3/3	3/3	3/3	x	x	
14148.0	3/3	3/3	3/3	x	x	
14265.1	3/3	3/3	3/3	x	x	
14307.1	3/3	3/3	3/3	HexNAc(4)Hex(5)NeuAc(1)	HexNAc(4)Hex(5)NeuAc(2)	***
14316.2	3/3	0/3	3/3	x	x	
14323.0	3/3	3/3	3/3	HexNAc(5)Hex(5)NeuAc(2) or HexNAc(5)Hex(4)	HexNAc(5)Hex(3)Fuc(1) or HexNAc(5)Hex(4)NeuAc(2)Fuc(1)	**
14339.0	0/3	3/3	3/3	x	x	
14381.1	3/3	3/3	3/3	HexNAc(4)Hex(5)NeuAc(1) or HexNAc(3)Hex(5) or HexNAc(5)Hex(4) or HexNAc(5)Hex(4)NeuAc(1)	HexNAc(5)Hex(6)NeuAc(1) or HexNAc(6)Hex(6)NeuAc(1) or HexNAc(6)Hex(4)NeuAc(2) or HexNAc(6)Hex(4)NeuAc(1)	*** ** ** **
14598.2	3/3	3/3	3/3	HexNAc(4)Hex(5)NeuAc(2)	HexNAc(4)Hex(5)NeuAc(2)	***

14607.2	3/3	3/3	3/3	HexNAc(5)Hex(5)NeuAc(2) or HexNAc(5)Hex(4)Fuc(1)	HexNAc(4)Hex(5)Fuc(1) or HexNAc(6)Hex(4)NeuAc(2)	** **
14614.2	3/3	3/3	3/3	x	x	
14640.2	3/3	3/3	3/3	x	x	
14672.2	3/3	3/3	3/3	HexNAc(4)Hex(5)NeuAc(2)	HexNAc(5)Hex(6)NeuAc(1)	***
14688.2	0/3	2/3	3/3	HexNAc(5)Hex(4)Fuc(1)	HexNAc(6)Hex(4)NeuAc(1)	**
14963.3	3/3	3/3	3/3	HexNAc(4)Hex(5)NeuAc(2)	HexNAc(5)Hex(6)NeuAc(2)	*
15254.4	3/3	3/3	3/3	HexNAc(4)Hex(5)NeuAc(2)	HexNAc(5)Hex(6)NeuAc(3)	*

1.7.3. Figures of merit of the nanoLC-HRMS method

The repeatability of the nanoLC-HRMS analysis was investigated. The 3 batches of rhCG were analyzed, each in triplicate, and the resulting detected masses and their corresponding average retention time and RDS values are compiled in **Table S1** (see supplementary data). It appears that the RSDs of retention times are all inferior to 6.08%, with an average value of 0.4%, demonstrating a good repeatability of the analysis in terms of retention times. Next, the peak area of each glycoform was measured on its corresponding extracted ion chromatogram (XIC) and normalized by the area of the glycoform leading to the highest area value measured on its own XIC which corresponds to the glycoform having a MW 14,307.1 Da. It is worthwhile to notice that the glycoform giving the highest area is not necessarily the most abundant in the initial hCG sample due to potential bias associated with the electrospray ionization process. The obtained average relative area and RSD of each detected mass are reported in **Figure 21** and all the data are compiled in **Table S1**. The RSD values are below 38%, with an average of 16%, which is acceptable for the analysis of intact proteins by nanoLC-HRMS [102].

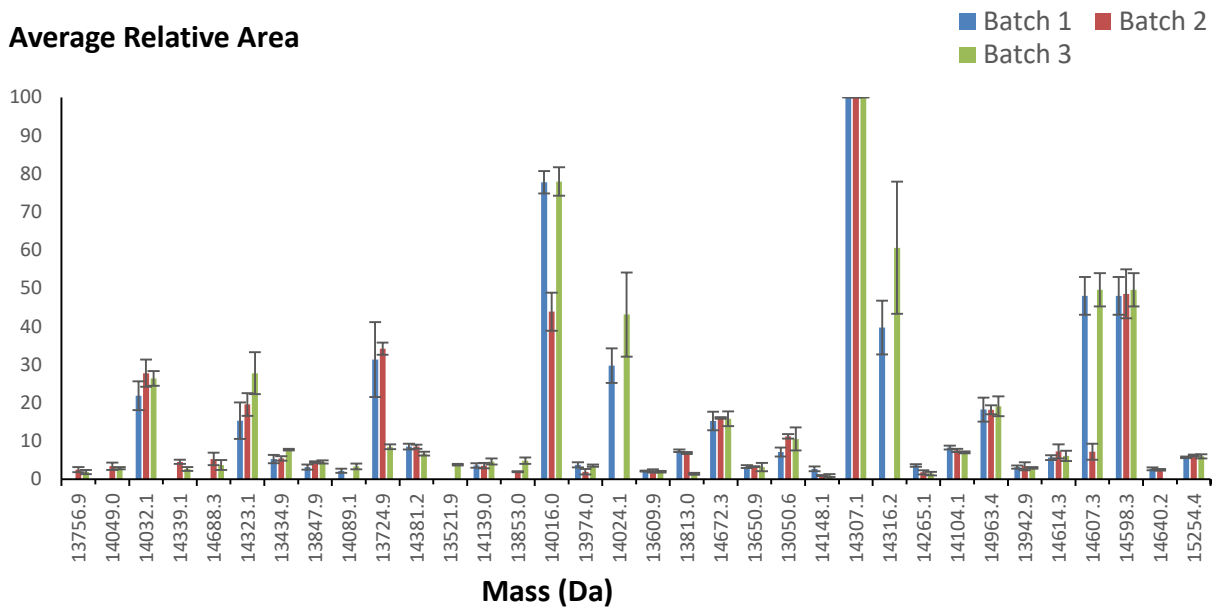


Figure 21. Comparison of the relative intensity of the glycoforms of hCGα present in 3 batches of rhCG analyzed by nanoLC-HRMS (n=3 for each). The comparison is based on the average relative area of each glycoform characterized by its MW and ordered by their retention times.

Next, the linearity of the method was also studied for the highest intensity glycoforms for 5 concentrations of hCG ranging from 7.5 to 250 µg mL^{-1} leading to correlation coefficient (R^2) values between 0.991 and 0.997. All the R^2 values and the equations of regression of each glycoform are gathered in **Table S2** (see supplementary data).

The determination of the limit of quantification (LOQ) was achieved by successive dilutions of the rhCG sample. It should be mentioned that if the global hCG concentration is known, the exact concentration of each glycoform present in the sample is unknown. By considering only the 11 glycoforms giving the highest signal in MS (fixed arbitrarily above 15% for their relative areas), they were all detected injecting 7.5 µg mL^{-1} of the global hCG content with a signal to noise ratio measured on the XIC of each glycoform above 10. However, the injection of 2.5 µg mL^{-1} of the global hCG content led to the absence of peaks. Therefore, the LOQ of the glycoforms giving the highest signal in MS is about 7.5 µg mL^{-1} of the global hCG content.

1.7.4. Potential of the analytical method for batch-to-batch comparison

The repeatability of the retention times and relative areas was assessed, which allowed the identification and semi-quantification of about 30 hCGα glycoforms. This analytical method was used to compare the 3 batches of rhCG as a proof-of-concept of its potential to characterize the glycosylation state of biopharmaceuticals in a simple and rapid manner. In **Table S1**, some qualitative differences between the three batches appear. First, the number of glycoforms detected in each batch varies: 28 in batch 1, 30 in batch 2, and 33 in batch 3. In more detail, the three batches shared 24 glycoforms with similar retention times while 10 glycoforms were only detected in one or two batches. As already mentioned, the glycoforms in a given batch are always detected in each run of the 3 repeated analysis, except for two that were detected in two runs only, which indicates that the result of 10 glycoforms that are different in the 3 batches is consistent and is not coming from a lack of repeatability of the analytical method. This difference between batches in terms of nature of some glycoforms exists.

Next, the three batches were compared semi-quantitatively based on the average relative area of each glycoform (**Table S1**). It is worthwhile to notice that only semi-quantitative comparison was possible due to the fact that the efficiency of the electrospray ionization process depends on the structure of the compounds and their co-elution. Therefore, it is not possible to directly correlate the relative area with the relative concentration of glycoforms in the sample because a glycoform could be better ionized than another one. **Figure 21** allows the comparison of the relative area of each detected glycoform ordered by increasing retention time. First the most abundant glycoform is the same in the three batches. Second**,** regarding the 11 most abundant glycoforms (≥15%), 7 have very close relative area in the three batches while 2 are not detected in batch 2. Next**,** the sum of the relative area of the 24 glycoforms that are common to the 3 batches are equal to 87, 96, and 78% of the sum of the relative areas of all the glycoforms in

batch 1, 2, and 3, respectively. Therefore, the batches are very close from a semi-quantitative point of view, but some relevant differences allow their discrimination. Finally, this simple and effective method directly applied to the intact hCG-based drug without requiring time-consuming and expensive treatment steps enables the detection of numerous hCGα glycoforms and led to the α-subunit profiling which could be promising for the comparison of other types of gonadotropin-based drugs.

1.7.5. Potential for the analysis of another gonadotropin

1.7.5.1. Identification of FSH glycoforms

In order to evaluate the applicability of the developed nanoLC-HRMS method for the characterization of the α-subunit of gonadotropin hormone, it was used for the analysis of a FSH-based drug (rFSH) by applying the same conditions previously optimized. **Figure 22A** presents the resulting TIC. The two deconvolution methods were tested starting first by the selection of the whole peak eluting from 17.5 to 20 min. **Figure 22B** presents the corresponding average deconvoluted mass spectrum. This resulted in the detection of 19 MWs corresponding to FSHα glycoforms ranging between 13,724.8 and 15,328.4 Da. With the second data treatment approach, by cutting the chromatographic peak in 30 s- and 50 s-slices and deconvoluting each slice separately, a total of 30 glycoforms was detected (see **Table S3** in supplementary data). This demonstrates again that this approach is definitively much more efficient to detect a higher number of α-subunit glycoforms present in the gonadotropin-based drug, even if it is more time consuming and not yet automated than the approach based on the one-time data treatment of the totality of the elution profile.

Due to the fact that FSH and hCG share the same alpha subunit with the same number of possible N-glycosylation sites and same origin (CHO), the obtained masses of FSH were

compared with the ones of hCG, the results that we got with the bottom-up approach done on hCG [14], and the GlyGen data base of all identified gonadotropin related N-glycans [101]. This allowed to propose N-glycans structures for 11 FSHα glycoforms (see **Table S3** in supplementary data), which demonstrates the ability of the nanoLC-HRMS method for the characterization of the glycosylation of the α-subunit of gonadotropins analyzing them at the intact level.

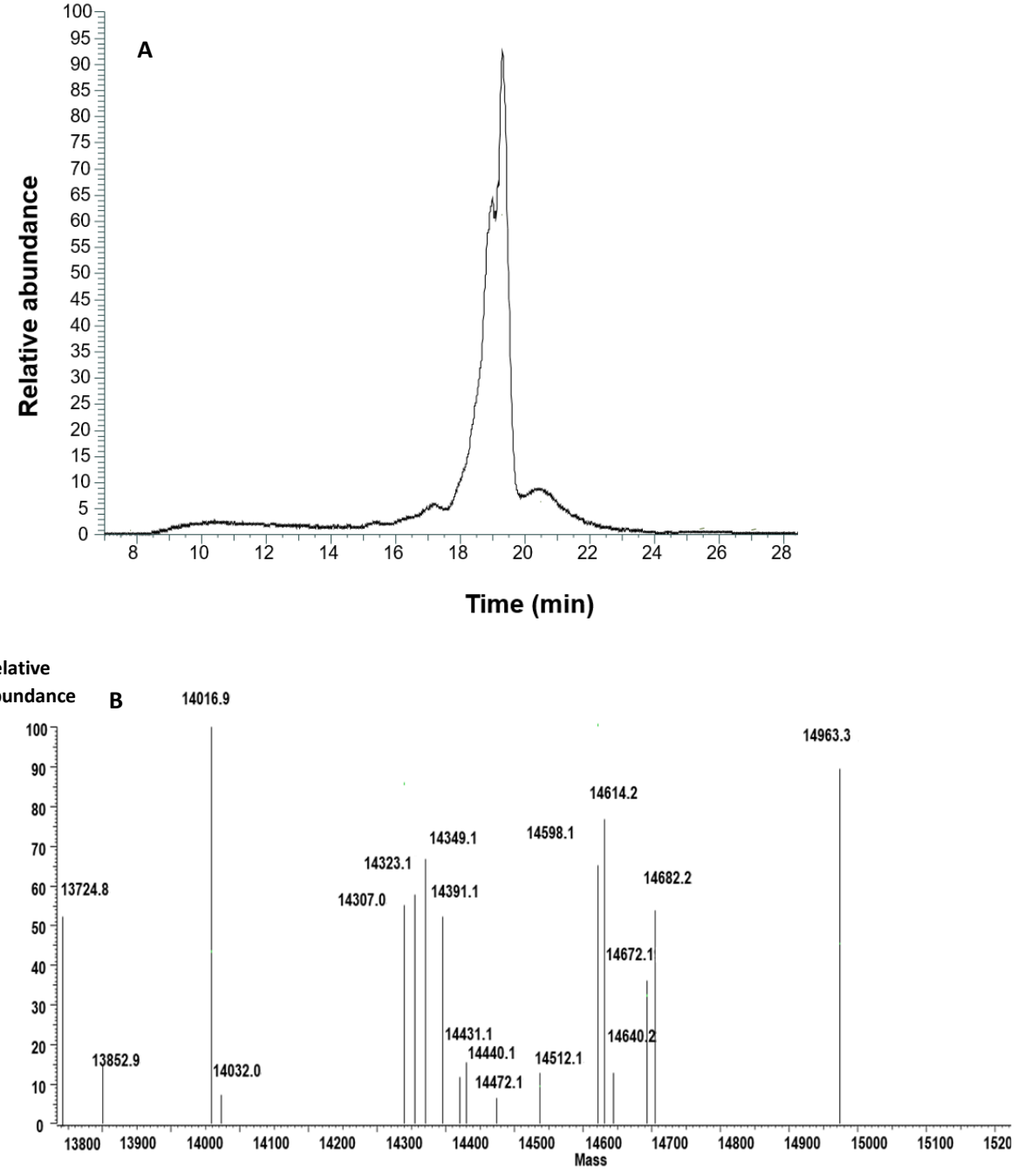


Figure 22. A: TIC of FSH obtained by nanoLC-HRMS. B: average deconvoluted mass spectrum for the 17.5-20 min range. rFSH, batch 1, run 1. Conditions: see section 2.3.

1.7.5.2.Repeatability of the nanoLC-HRMS method for the analysis of FSH

The repeatability of the method was studied by repeating the analysis of a first batch of rFSH (n=3) and involving the second data treatment approach. The number of detected masses was compared and proved to be the same in each run except for one glycoform that was detected only in two runs out of the three. It corresponds to a glycoform of low intensity (relative area of 8%) and that has a retention time close to 4 other glycoforms (the 5 glycoforms are eluted in 1.2 s), which may explain that in one run, the ionization process of this glycoform was not sufficiently efficient to detect it. For each detected mass, the retention time was assigned and the average RSD of retention times was equal to 0.3% (n=3). Next, 2 additional batches of rFSH were analyzed in triplicate. **Table S4** (see supplementary data) shows all the masses detected in each run of each batch and, for each glycoform, their corresponding average retention time and average relative area plus their RSD values are given. The RSD values of the retention times were inferior to 1.2% with an average of 0.4 %, while that of the relative area are below 39% with an average of 17%.

1.7.5.3.Comparison of rFSH batches

As the repeatability of the nanoLC-HRMS method was assessed, the comparison of the 3 batches of rFSH was performed using the same approach as the one used previously with hCG. For ease of comparison, **Figure 23** presents the average relative area of each glycoform in each batch. First, from a qualitative point of view, the 3 batches shared 28 glycoforms. One glycoform was detected only in batch 3 and another one is missing in batch 2, but these glycoforms have very low relative area values when they are detected, which may explain why they were not detected in some cases. Therefore, the 3 batches are qualitatively very similar.

It can also be noticed that the most "abundant" glycoform (*i.e.* the one having the highest average area) is the same in the 3 batches. In addition, 19 glycoforms have a relative area higher than 15% and they are all present in the three batches showing that the batches are quite homogeneous. The sums of the relative areas of the 28 glycoforms that are common to the three batches are equal to 95, 100, and 98% of the sum of the relative areas of all the glycoforms in batch 1, 2, and 3, respectively. Therefore, the batches are very close from a semi-quantitative point of view. These results demonstrate again the potential of the analysis by nanoLC-HRMS of the glycoproteins at the intact level to qualitatively compare the glycoprofiling of different samples in addition to the possibility of semi-quantitative assessment for quality control purposes for example.

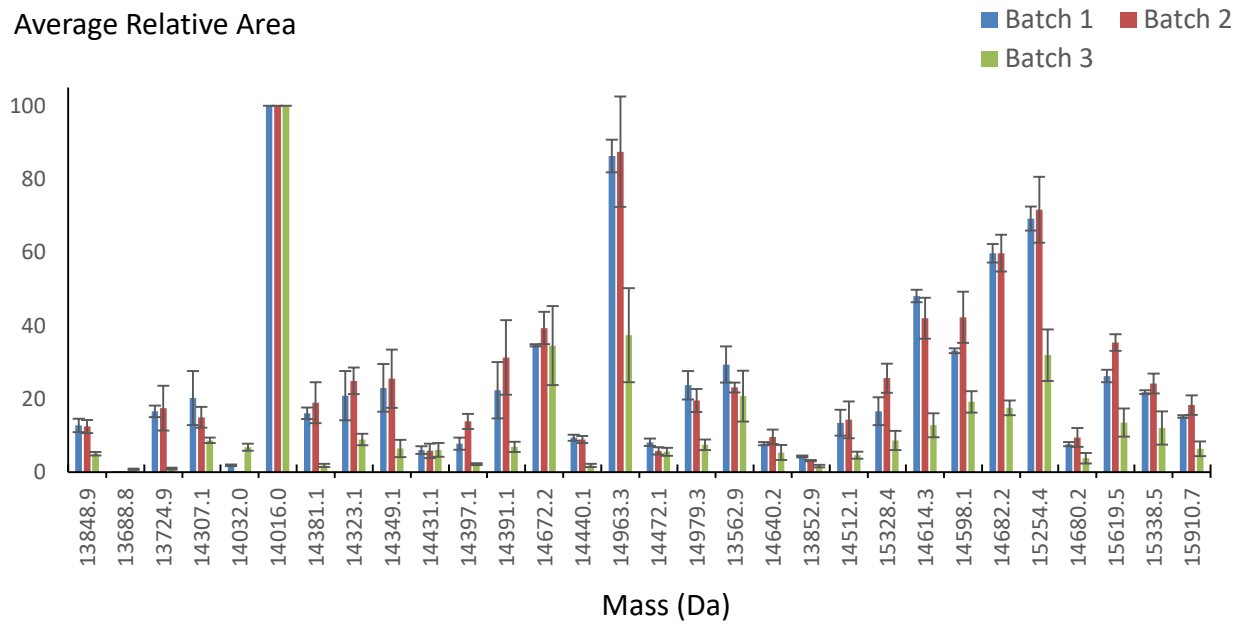


Figure 23. Comparison of the glycoforms of FSHα present in 3 batches of rFSH analyzed by nanoLC-HRMS (n=3 for each). The comparison is based on the average relative area of each glycoform characterized by its mass and ordered by their retention time.

1.7.6. rhCG vs rFSH

The nanoLC-HRMS method enabled the identification and semi-quantification of the α-subunit

glycoforms of the two gonadotropins hCG and FSH contained in rhCG and rFSH, respectively. Despite the fact that these two hormones share the same α-subunit, their analysis led to the detection of different glycoforms. Only 13 of those were in common between the two as shown in **Figure 24A**, which means less than half of them. Therefore, both pharmaceutical drugs have different glycoprofiling of the α-subunit.

A pie graph was constructed (**Figure 24B)** to compare semi-quantitatively the average relative area of each glycoform detected in the 3 batches of rhCG and rFSH analyzed in triplicate (n=9). It shows that the common glycoforms seem to be the most abundant ones, with the hypothesis that the MS signal is correlated to the initial concentration in the sample. Nevertheless, it appears that the composition of the two pharmaceutical preparations varies significantly with regard to the α-subunit glycoforms, even if the differences in MS intensities may result from differences in MS ionization. Although this hypothesis cannot be eliminated, it seems unlikely that it will lead to such different profiles. Knowing that rFSH and rhCG are recombinant gonadotropins produced using the CHO cell line and the same DNA recombinant technology, it does not eliminate the fact that there are still slight differences in their production and purification procedures, leading to non-identical preparations, with variations in glycosylation that yield different sialic acid residue compositions for example [33]. This could explain the differences between the identified isoforms.

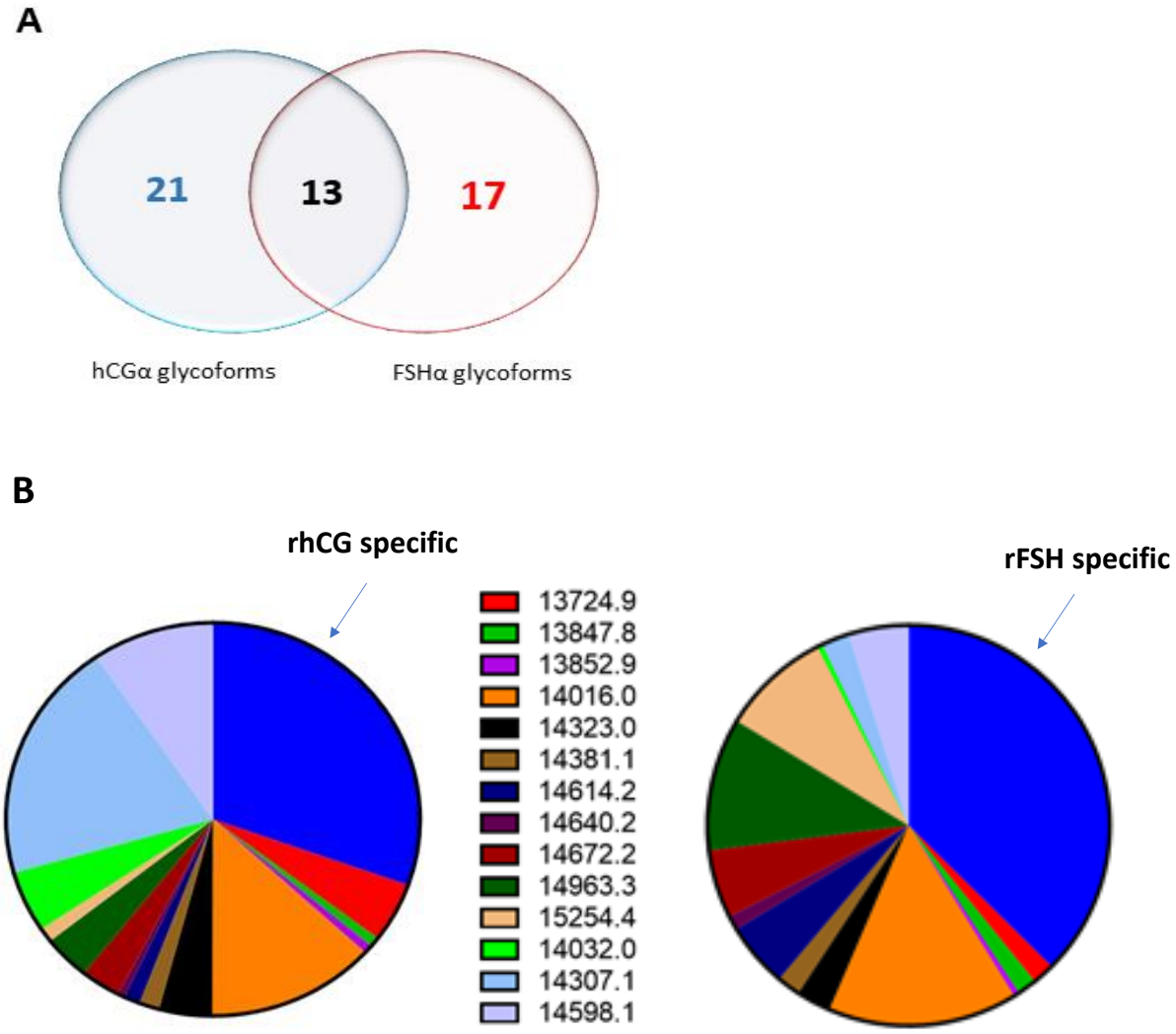


Figure 24. A: Venn diagram comparing the number of the α-subunit glycoforms identified by the analysis of rhCG and rFSH by nanoLC-HRMS. B: Pie charts representing the average relative area of each detected glycoform in (Top) rhCG and (Bottom) rFSH. For each drug, 3 batches were analyzed 3 times each (n=9).

1.8.Conclusion

Special attention is being implemented in studying glycoproteins due to the fact that glycosylation has a strong impact on the properties of a glycoprotein-based drug. This work demonstrated for the first time the great potential of nano-LC coupled to HRMS (Orbitrap) for

fast glycoprofiling of the α-subunit of two gonadotropin-based drugs analyzed at the intact level after simple filtration, without any tedious sample preparations. The good figures of merit of the method allow its use for identification and semi-quantification of the glycoforms. It could be useful at different stages of the life of these kinds of drugs, such as their development (discovery, preclinical, and clinical steps), production, storage, and delivery.

Acknowledgements

This work has received the support of "Institut Pierre-Gilles de Gennes" (Laboratoire d'excellence, "Investissements d'avenir" program ANR-10-IDEX-0001-02 PSL and ANR-10-LABX-31), of PSL University, the Conseil Régional d'Île-de-France with the DIM Analytics and of the association "Al Ihssan".

Chapter 4: Identification and semi-relative quantification of intact glycoforms of hCGα and hCGβ by nanoLC-(Orbitrap)MS

In the previous chapter, the development of a nanoLC-(Orbitrap)MS method that enables the detection and semi-quantification of numerous hCGα glycoforms contained in a recombinant hCG-based drug was detailed. However, it didn't allow the detection of hCGβ glycoforms. Therefore, it was next crucial to optimize this method aiming for the detection of both α and β subunits. For this reason several parameters were optimized, starting by the injection conditions, the mobile phase gradient, and finally the nature and the content of the acidic additive introduced in the mobile phase. All the optimization conditions are detailed in this chapter that is in the form of an article submitted to Journal of Chromatograaphy A: A. Al Matari, A. Goumenou, A. Combès, T. Fournier, V. Pichon, and N. Delaunay, "Identification and semi-relative quantification of intact glycoforms of human chorionic gonadotropin alpha and beta subunits by nanoLC-(Orbitrap)MS". 2021 Jan 28; 1640:461945. doi: 10.1016/j.chroma.2021.461945.

Article 2: Identification and semi-relative quantification of intact glycoforms of human chorionic gonadotropin alpha and beta subunits by nanoLC-(Orbitrap)MS

Amira Al Matari[1], Anastasia Goumenou[1], Audrey Combès[1], Thierry Fournier[2], Valérie Pichon[1, 3], Nathalie Delaunay[1]

[1]Laboratory of Analytical, Bioanalytical Sciences and Miniaturization, Chemistry, Biology and Innovation (CBI) UMR 8231, ESPCI Paris PSL, CNRS, PSL Research University, Paris, France

[2]Université de Paris, INSERM, UMR-S1139, «Pathophysiology & Pharmacotoxicology of the Human Placenta, pre & postnatal Microbiota», 3PHM, F-75006 Paris, France.

[3]Sorbonne Université, France

Abstract

The human chorionic gonadotropin (hCG) protein belongs to a family of glycoprotein hormones called gonadotropins. It is a heterodimer made of two non-covalently linked subunits. The α-subunit structure, hCGα, has 2 N-glycosylation sites, while the beta subunit, hCGβ, has 2 N- and 4 O-glycosylation sites. This leads to numerous glycoforms. A method based on the analysis of hCG glycoforms at the intact level by nano-reversed phase liquid chromatography coupled to high resolution mass spectrometry (nanoLC-HRMS) with an Orbitrap analyzer was previously developed using a recombinant hCG-based drug, Ovitrelle®, as standard. It allowed the detection of about 30 hCGα glycoforms, but didn't allow the detection of hCGβ glycoforms. This method was thus significantly modified by adding a pre-concentration step of the sample. Therefore, the injected sample volume was increased by a factor 14 (from 70 nl to 1 µl). Moreover, the gradient slope and the nature and content of the acidic additive in the mobile phase were optimized. The best conditions were obtained with a slope of 1.02%/min of

acetonitrile and replacing 2% of formic acid by 0.05% of trifluoroacetic acid. These conditions led to an improvement of the separation of hCGα and hCGβ glycoforms, which allowed for the first time the detection of 33 hCGβ glycoforms at intact level. In addition, a higher number of hCGα glycoforms (42 in total, *i.e.* a 40% increase) was detected. The figures of merit of this new method were next assessed. The relative standard deviations (RSDs) of the retention time ranged between 0.02 and 0.95% (n=3), with an average value of 0.36% for the alpha glycoforms and between 0.01 and 1.08% (n=3) with an average value of 0.23% for the beta glycoforms. The RSDs of the relative peak area measured on the extracted ion chromatogram of each glycoform were below 20% (n=3), with an average value of 9.8%, thus allowing semi-relative quantification. Therefore, this method has a high potential for rapid quality control aiming for the detection and comparison of glycoforms present in glycoprotein-based pharmaceutical preparations.

4.1. Introduction

Human chorionic gonadotropin (hCG) hormone is a highly glycosylated protein produced by the placenta during pregnancy and also expressed by certain cancer cells [3]. hCG consists of two non-covalently linked subunits, hCGα and hCGβ. hCGα contains 2 N-glycosylation sites while hCGβ contains 2 N- and 4 O-glycosylation sites. Glycosylation is the process by which a carbohydrate is covalently attached to a protein by a series of complex enzymatic reactions. It is one of the most important post-translational modifications (PTMs) of proteins, as approximately 70% of the human proteins are glycosylated. Glycosylation regulates numerous biological functions such as cell–cell signaling, protein localization, protein stability, and immune response [103].

The state of glycosylation of hCG depends on the term of pregnancy, its source of production and pathologies. It is well established that hCG is mainly secreted by the syncytiotrophoblast, which represents the exchange and endocrine tissue of the human placenta. The invasive extravillous trophoblast also secretes hCG, and in particular hyperglycosylated forms of hCG (hCG-H) also produced by choriocarcinoma cells. In maternal blood, hCG reaches a peak at 8-10 weeks of gestation whereas hCG-H is elevated during early first trimester corresponding to the trophoblastic cell invasion process and then decreases [104]. We have demonstrated that in contrast to hCG, hCG-H stimulates trophoblast invasion and angiogenesis by interacting with the transforming growth factor ß receptor in a luteinizing hormone/human chorionic gonadotropin receptor independent signaling pathway [105]. Therefore, glycoforms of hCG display different biological activities and functions that are essential for pregnancy outcome [9].

hCG is largely used in antenatal screening and is also integrated in gonadotropin therapy used for certain infertility treatments for ovarian stimulation. Its administration started in 1960s and it is

still applied. Nevertheless, the supply of hCG differed remarkably since then. The initial source of hCG was human pituitary extracts and blood from pregnant mares, whose use was later banned and replaced by gonadotropins extracted from human urine of postmenopausal women. Following that, advances in DNA technology enabled the development of recombinant gonadotropins readily available for the use as therapeutic solutions [33]. Therefore, precise characterization of the glycan structure of gonadotropins is significant for drug production and quality control as well as specific analysis of gonadotropins present in biological fluids.

Several methods have been developed for the analysis of glycosylated proteins in different samples. These methods are globally divided into three categories. The first, and most commonly applied approach is called bottom-up approach. It is based on the enzymatic digestion of the protein and next the analysis of the resulting glycopeptides [89]–[91]. The second category consists of the analysis of released glycans [91], [106]. Both approaches provide accurate information on micro-heterogeneity but required laborious experimental steps and lack the ability of global overviewing of the macro-heterogeneity of the studied proteins. The last category consists of the analysis of glycoproteins at the intact level [72], [81], [91]. It provides specific and complementary information compared to the two approaches described above and is the only way to know the composition of a given glycoform.

With regards to hCG, the three approaches were already reported. For instance, hCG was analyzed based on a released glycan level for glycan mapping [21], [36], but the bottom-up approach was by far the most commonly used method for the analysis of hCG [11], [14], [29], [64], [66], [67]. Different enzymes such as pepsin and trypsin were used to get the glycopeptides that were further analyzed by liquid chromatography (LC) coupled to tandem mass spectrometry (MS), allowing the localization of the glycans on the peptides. Recently increased attention has

been made towards the analysis of hCG at the intact level to preserve the integrity of the analyzed glycoforms and to assess the combinations of their potential glycosylation sites. For this, high performance separation modes were implemented, capillary zone electrophoresis [17], [92] and especially LC [13], [16], [32], [52], [77], [79]–[81]. For instance, reversed-phase LC (RPLC) was hyphenated with an MS detection to get more information on the nature and the number of glycoforms with the analysis of hCG at the intact level [32]. The identification of 14 glycoforms of hCGα was reported and some mass spectra were provided for the β-subunit, but the ion trap quadrupole analyzer used in this study lacked mass accuracy and resolution, which could lead to misidentifications and prevent unambiguous assignment. Our group also proposed the hyphenation of RPLC with high-resolution MS (HRMS) using a quadrupole-time-of-flight (Q-TOF) analyzer [16].It led to the detection of several hCGα glycoforms of two hCG-based drugs, Ovitrelle® and Pregnyl®, and the identification of the five main glycoforms by mass matching. We also optimized a hydrophilic interaction liquid chromatography (HILIC) separation hyphenated with UV for the analysis of two hCG-based drugs [79]. HILIC was then hyphenated with HRMS (Q-TOF analyzer) and was compared to the RPLC-MS (Q-TOF) method concluding that both methods were complementary for the identification of a high number of hCGα glycoforms [82]. However, the Q-TOF analyzer was still insufficient due to the overlapping of some isotopic patterns corresponding to different glycoforms. This is why, we recently proposed to transfer the RPLC separation to the nano format to improve sensitivity (x40) and to couple it to a higher mass resolution and mass accuracy MS analyzer (Orbitrap) [107]. This nano-RPLC-HRMS method associated with a non-automated and extensive data treatment was applied to the analysis at the intact level of a recombinant hCG-based drug (rhCG, Ovitrelle®) and also of a follicle stimulating hormone (FSH)-based drug, which lead to the

identification of about 30 hCGα and FSH alpha subunit glycoforms for each drug, respectively. The ability of this method to achieve the semi-quantification of various batches of glycoprotein-based drugs was provided as well. However, it did not allow the detection of any hCGβ glycoforms. This is why, in this current study, the nano-LC-HRMS method was modified by adding and optimizing a preconcentration step to significantly increase the injected sample volume. The effect of the slope gradient and the nature and content of the acidic additive in the mobile phase was also studied to favor the separation of the hCG glycoforms, and consequently their ionization. Altogether, it allowed for the first time the detection of numerous hCGβ glycoforms by nano-RPLC-(Orbitrap)MS analysis of rhCG. Once the new method optimized, some figures of merit of the developed method were determined such as repeatability, linearity, and limit of quantification (LOQ).

4.2. Experimental Part

4.2.1. Materials and Reagents

HPLC-grade acetonitrile (ACN), formic acid (FA), trifluoroacetic acid (TFA) and LC-MS-grade water were supplied by Carlo Erba (Val de Reuil, France). Ovitrelle® (Organon, Oss, The Netherlands) is a drug made of recombinant hCG from a Chinese hamster ovary cell line; it will be written as rhCG in the rest of the manuscript. rhCG contains 250 µg of hCG suspended in 500 µL of suspension liquid containing mannitol, methionine, Poloxamer 188, phosphoric acid, sodium hydroxide, and water.

4.2.2. Sample preparation

The rhCG drug stock solution was desalted and filtered for 6 min at 4°C with ultra-centrifugal units (Microcon) (Merck, Darmstadt, Germany) with a molecular weight (MW) cut-off of 10 kDa. After this step that aims to remove additives and excipients from the formulation, the

residues were washed with water and finally the samples were recovered from the inverted membrane by centrifugation for 4 min at 14,000 g (≈ 11,000 rpm) and the volumes were adjusted with water to obtain a final concentration of 500 µg mL^{-1} of hCG.

4.2.3. NanoLC-MS system

The analyses were carried out using a nano-LC system (NC 3500 RS, ThermoFisher Scientific, Courtaboeuf, France) coupled to an Orbitrap mass analyzer (Exactive plus, ThermoFisher Scientific). The precolumn was a C18 PepMap (0.3 x 5 mm internal diameter, 300 Å) packed with 5 µm particles (ThermoFisher Scientific) while the column was an EASY-Spray PepMap RSLC C18 (150 x 0.075 mm internal diameter, 300 Å) packed with 5 µm particles (ThermoFisher Scientific). The mobile phase was composed of (A) H_2O/ACN (98/2; v/v) and (B) H_2O/ACN (1/9; v/v) acidified with FA 2% or TFA at 0.05 or 0.1% v/v. The column temperature was set at 60°C. Different injection volumes (1, 1.5, and 2 µl) and injection modes (full loop, partial loop, and µl pick-up) were tested. The optimum injection volume and injection mode were 1 µL and partial loop mode, respectively. The flow-rate was set at 300 nL min^{-1}. Five gradients were tested before the selection of the optimal one which starts from 10 to 65% of ACN in 54 min (1.02%/min), followed by a rapid increase to 90% of ACN within 5 min, then a 5 min plateau, and finally a return to the initial 10% of ACN concentration and kept for 15 min for equilibration**.** The other tested gradients are gathered in **Table S5** (see supplementary data). The EASY-spray source parameters were kept the same as follows: capillary voltage, 1.7 kV; drying temperature, 280°C; and sheath, auxiliary, and sweep gas flows were set at 0. The focusing voltages were tuned on weekly bases by the transmission of positive and negative ion calibration solutions (ThermoFisher Scientific). The mass spectrometer was operated in the positive full MS mode and using a scan range from 1000 to 2500 mass-to-charge (m/z) with a

scan rate of 1 spectrum/s. The mass resolution was similarly set to 140,000 at 200 m/z, the AGC target was 3e6, and the maximum injection time was 200 ms. The deconvolution analyses were processed using Freestyle 1.3 SP2 (ThermoFisher Scientific) with the same deconvolution parameters as the previous method **[107]** that uses a mass tolerance for m/z of 30 mDa, and a minimal charge state of 5, with at least 3 different charge states to identify a compound, accuracy mass tolerance for deconvoluted MW was set at 300 mDa.

4.2.4. Nano-LC-HRMS data treatment for the identification of glycoforms

In all the experiments electrospray ionization was used, which led to multiple charged values for each glycoform and an isotopic cluster for each nominal mass-to-charge (m/z) value ratio, thus leading to highly complex mass spectra. In the current study the spectral deconvolution and data treatment method used was the same as the one previously optimized *[107]*. Briefly each chromatogram was cut in successive time slices of 30 s and next 50 s. The average deconvoluted mass spectrum for each slice was obtained resulting in the achievement of a list of MWs corresponding to the detected glycoforms present in the rhCG sample. Next, the results of both types of division were combined to reduce the risk of data loss at the boundaries of the performed cuts. The resulted list of MWs are compared to the theoretical mass of the amino acid sequence of hCGα or hCGβ added to their potential glycans.

4.3. Results and discussion

We previously optimized a nano-RPLC separation coupled with high resolution MS using an Orbitrap analyzer for the detection, without overlapping of the isotopic patterns, which led to the detection of 28 high and low intensity hCGα glycoforms when analyzing rhCG *[107]*. However, it didn't allow the detection of hCGβ glycoforms that may be less efficiently ionized due to the

presence of 2 N- and 4 O-glycosylation sites. This is why, this study aimed to improve this method in order to detect hCGβ glycoforms, starting from our previous conditions *[107]*. The same EASY-Spray PepMap RSLC C18 column was used because it has an electrospray emitter directly connected to the column, thus minimizing post-column dead volumes and its integrated column heating device enables a great stability of the temperature during the analysis. The column temperature was set at 60°C, since it was previously observed that a high temperature favors the separation of glycoforms and is the maximal recommended temperature given by the column manufacturer. The mobile phase consisted of a water-ACN mixture acidified with 2% of FA to favor the multiple ionization of the hCG glycoforms. The Orbitrap MS parameters previously optimized were kept the same here. Finally, the rhCG-based drug Ovitrelle® was kept as a "standard" solution of hCG, which allows to compare the separation and sensitivity performances obtained here with the ones previously reported.

4.3.1. Optimization of the injection conditions

The first main change in the nano-LC-HRMS method lies in the introduction of a trapping precolumn to increase the injected sample volume, and consequently the sensitivity. Nevertheless, it was necessary to determine optimal conditions allowing the trapping of all the hCG glycoforms on the precolumn and next their efficient back-flush desorption and transfer to the separation column with the mobile phase having a low content of ACN (10%) at the beginning of the elution gradient. A trapping precolumn based on C18-silica beads with the same particle (5 µm) and pore (300 Å) sizes than the separation column was selected, but with a higher internal diameter (300 *vs* 75 µm) to favor retention for large sample volume but with the risk of increasing band-broadening effects. The corresponding experimental set-up is presented in **Figure S2** (see supplementary data).

After the introduction of the trapping column, the first step was to test the quality of the coupling and the sensitivity enhancement by comparing the obtained results with those previously obtained with a simple 70 nl loop. Therefore, the same rhCG sample was used and the same amount of rhCG was injected (35 ng) but this time testing out 3 injection volumes of 1, 1.5, and 2 µl (microliter pick-up mode, see **Figure S3** in supplementary data) instead of the former 70 nl. It is worth mentioning that all the other parameters were kept the same. The preconcentration of the sample lasted for 3 min, which is the time required for the sample to be transferred from the loop to the precolumn, knowing the flow-rate and the tubing volumes. The glycoforms preconcentrated on the trapping precolumn were then back-flush transferred to the nano-LC separation column by the mobile phase and detected in MS. The average peak areas of the higher intensity hCGα glycoforms (i.e. the glycoforms with a relative area higher than 15% of that of the glycoform giving the highest peak) that were measured on their corresponding extracted ion chromatogram (XIC) were compared (n=3). **Figure 25** shows the results for the 6 glycoforms of higher intensity. First, as expected, a significant enhancement of the signal was obtained with the introduction of the trapping precolumn compared with the direct injection of the same rhCG amount without preconcentration. Second, it is clearly demonstrated that even if the same quantity of rhCG was injected, better sensitivity was achieved for these glycoforms by preconcentrating 1 µl of the sample rather 1.5 or 2 µl. These higher volumes may have induced a loss of the glycoforms by exceeding their breakthrough volume. These conclusions made with the glycoforms giving rise to the highest intensity signals were the same for all the detected hCGα glycoforms. It is worthwhile to notice that the same number of hCGα glycoforms were detected when injecting the same amount of rhCG using the 70 nl loop or by percolating 1 µl on the trapping column. An improvement in sensitivity ranging from 1.5 and 20 was obtained for all

glycoforms with an average factor of 9. This means that not all glycoforms behaved the same on the trapping precolumn, some being more retained than others, and that 1 µl seems to be the best compromise injection volume for all.

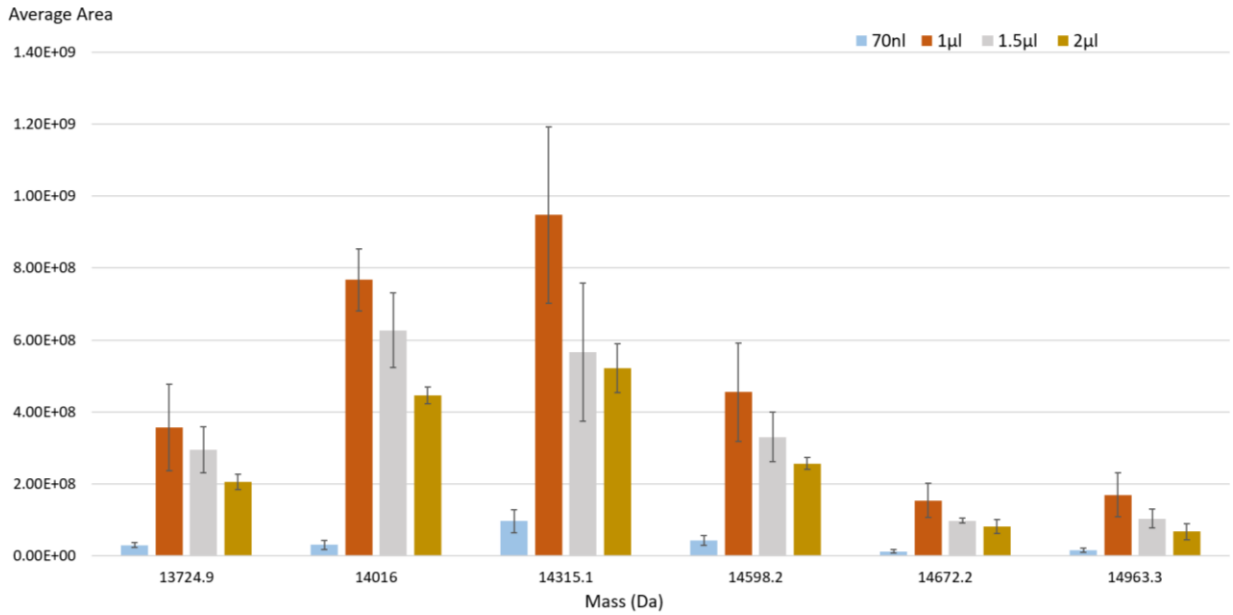


Figure 25. Mean value of peak areas (and standard deviation, n=3 injections) of 6 hCGα glycoforms obtained after injecting 35 ng of rhCG using a 70 nl loop or after the preconcentration of 1, 1.5 or 2 µl of sample on the trapping precolumn. Mobile phase: 2% formic acid in H_2O/ACN mixture (v/v). Gradient: from 10 to 30% ACN with a slope of 0.66% ACN/min. Injection mode: µl pickup. Other conditions: see section 2.3

Three modes of injection can be next implemented to inject 1 µl of sample: full-loop, partial-loop, and microliter pick-up injection modes that are schematically presented **Figure S3**. Based on the information given by ThermoFischer Scientific, the full loop injection mode provides maximum reproducibility but less accuracy, while the partial loop injection mode provides maximum accuracy and very good reproducibility, and finally the microliter pick-up injection mode offers zero sample loss and maximum accuracy but slightly diminished reproducibility. Therefore, all three injection modes were tested with the nano-LC-HRMS conditions previously developed [107] with the exception of the addition of the trapping precolumn. For each injection mode, the replacement of the sample loop was required. 25 µl, 5 µl, and 1 µl sample loops were used for the microliter pick-up, partial loop, and full loop injection modes, respectively. The

peak areas of the 9 hCGα glycoforms giving rise to the highest intensity signals were measured on their corresponding XIC and the average values with their standard deviations are represented in **Figure 26** (n=3). It clearly appears that the highest and most repeatable average areas are obtained by performing the analysis with the partial loop injection mode. This is why this injection mode was selected for the rest of the study. The sensitivity enhancement when comparing the peak area of these 9 glycoforms obtained using the optimized preconcentration method (1 µl, partial loop mode) and the initial injection of the same amount with the 70 nl loop without any preconcentration step averages 19 with a value range of 5-30, which is greater than the theoretical gain (14). This may be due to a global decrease in the band broadening effect because a volume of 70 nl directly injected on the nanoLC column of low dimension was already too high and induced some significant band broadening, but this value resulted from a compromise between sensitivity and efficiency. The 29 hCGα glycoforms detected here were also observed in the method without any preconcentration step, for which between 28 and 30 hCGα glycoforms were detected depending on the rhCG batch. This confirmed that the preconcentration on the trapping precolumn did not induce the loss of some glycoforms.

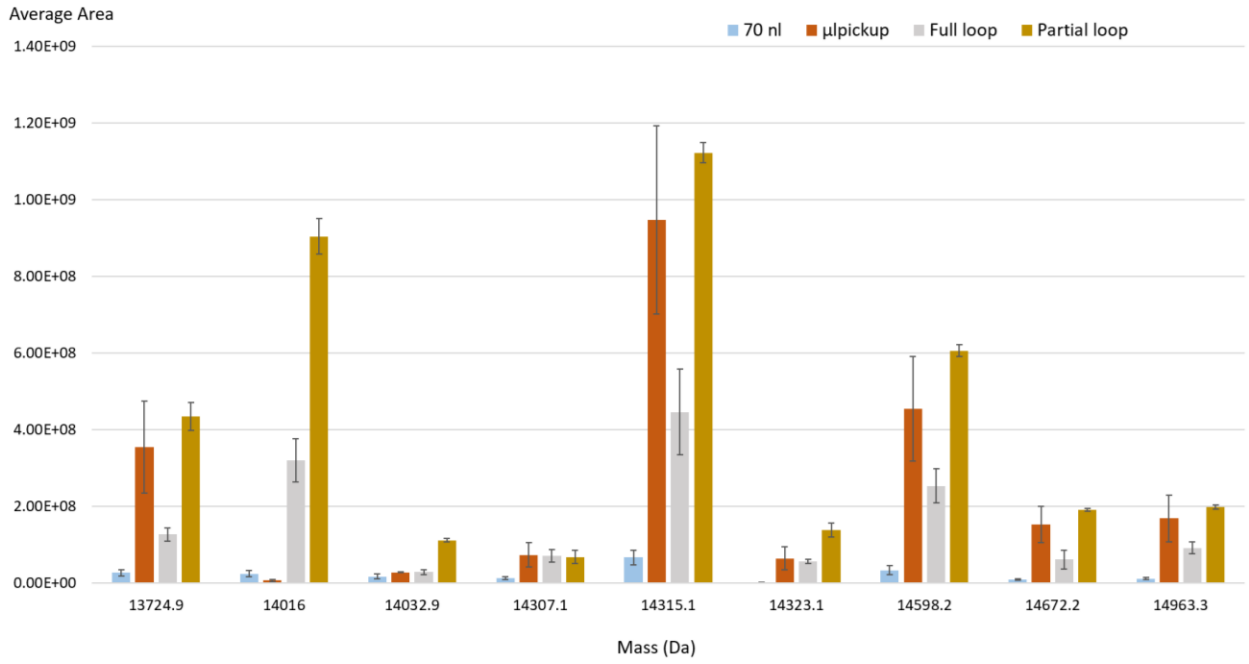


Figure 26. Mean value of peak areas (and standard deviations (n=3)) of 9 hCGα glycoforms analyzed in nano-LC-HRMS by injecting 1 µl of rhCG at 35 µg ml-1 with preconcentration on the trapping precolumn using three different injection modes . Mobile phase: 2% Formic acid in H_2O/ACN mixture (v/v). Gradient: from 10 to 30% ACN with a slope of 0.66% ACN/min. Other conditions: see section 2.3.

4.3.2. Mobile phase gradient optimization

Despite the better sensitivity achieved with the trapping precolumn and optimized injection conditions (volume and mode), it was still impossible to detect any hCGβ glycoforms, even when analyzing a higher quantity of rhCG (100 ng *vs* 35 ng). Another optimization was next investigated on the basis of an improvement of the separation to try to favor the ESI ionization process and thus perhaps allow the detection of hCGβ glycoforms not yet detected. For this, different gradient slopes were tested (see **Table S5** for their description), keeping constant the initial ACN content at 10% and using the same batch of rhCG. It is important to note that the end of each gradient was different, in order to keep an acceptable analysis time and to finish with mobile phase of a high elution strength (high content of ACN) to ensure the elution of all the compounds present in the rhCG-based drug. Nevertheless, the retention times of the hCGα and hCGβ glycoforms were always inferior to the times of the end changes of the gradients. This is

why only the slope of the tested gradient was here relevant on the hCG glycoform separation. The corresponding Base Peak Chromatograms (BPCs) are presented in **Figure S4** in the supplementary data. As expected, when the gradient slope decreased the retention time and thus the apparent retention factor, of hCG glycoforms increased. The resolution between two successive glycoforms may be affected by the residence time of the glycoforms in a media at high temperature, 60°C here. Moreover, in addition to the quality of the separation, i.e. the co-elution of glycoforms, the ESI ionization process efficiency may also be affected by the composition of the mobile phase solubilizing a given glycoform. **Table 8** presents the number of hCGα and hCGβ glycoforms detected with each mobile phase gradient. Despite the difficulty to explain all the observed results, the main observation is that the modification of the gradient slope allowed in some cases the detection of hCGβ glycoforms for the first time in nanoLC-(Orbitrap)MS. The best gradient corresponds to the slope of 1.02% ACN/min, which enabled the detection of the highest number of both hCGα and hCGβ glycoforms, 37 and 19, respectively. Nevertheless, it is also worthwhile to notice that, as expected, there is a strong difference in intensity between hCGα and hCGβ glycoforms. Indeed, the intensity of the signals for hCGβ glycoforms is similar to those of hCGα glycoforms giving the low intensity signal. Indeed, a ratio of 212 was obtained between the highest signal intensity hCGα and lowest signal intensity hCGβ glycoforms while a ratio of 200 was obtained between the highest and smallest signal intensity hCGα glycoforms. As hCG is a dimeric protein composed of one hCGα and one hCGβ, it means that hCGβ glycoforms are less efficiently ionized than hCGα glycoforms, may be due to their higher number of glycosylation sites (6 *vs* 2, respectively).

Table 8 Number of detected hCGα and hCGβ glycoforms analyzed by nano-LC-HRMS with five different gradient slopes. Mobile phase: 2% Formic acid in H2O/ACN mixture (v/v). rhCG at 100 µg ml-1. Injection volume 1 µl in partial loop mode. Other conditions: see section 2.3.

Gradient slope of ACN/min	Number of detected hCGα glycoforms	Number of detected hCGβ glycoforms
1.49%	21	0
1.02%	37	19
0.86%	23	5
0.67%	33	0
0.56%	20	4

4.3.3. Effect of the nature and content of acidic additive introduced in the mobile phase

The previous results have shown that improving the glycoform separations was a good approach to increase the number of hCGα and hCGβ glycoforms detected, certainly favoring their ionization. This is why, the nature and content of the mobile phase additive was also varied. The mobile phase additive FA (2%, v/v) has proven previously its ability to favor the multiple ionization of hCG glycoforms better than other percentages [15]. It was here replaced by TFA, everything else being unchanged. TFA was selected because we previously observed that it has a strong positive effect on the separation of the hCG glycoforms in HILIC due to its ion pairing properties [79]. Very low contents (0.05 or 0.1 %v/v) were first tested as TFA has unfortunately also a strong ion suppression effect during electrospray process, as observed also for hCG [82]. The typical BPCs obtained using 2% FA and 0.05% or 0.1% TFA for the analysis of rhCG at 100 µg ml^{-1} are presented in **Figure 27.** A dramatic change in the separation was observed as the co-elution of hCGα and hCGβ glycoforms observed with FA disappeared with TFA, thus demonstrating that the ion pairing properties of TFA also had a positive effect on the separation of hCG glycoforms in RPLC. However, TFA induced also a huge decrease in the intensity even

with a very low content (for instance a decrease of about a factor 50 with 0.1% TFA in comparison with 0.05% TFA was observed). **Figure 28** shows a comparison of the number of identified hCGα and hCGβ glycoforms identified when applying the conditions of Figure 3. In total, 42 hCGα and 33 hCGβ glycoforms were detected after the addition of 0.05% TFA in the mobile phase, whereas it was only 37 and 19 with 2% FA, respectively. An increase in the TFA content from 0.05 to 0.1% induced a strong drop in the number of glycoforms detected (22 and 5, respectively). In the latter case, the gain in separation power is no longer as interesting because of the strong decrease of the signal intensities. Thus, the addition of 0.05% TFA in the mobile phase was the best compromise, promoting significantly the separation with only a reduced loss in sensitivity and allowing thus the detection of a high number of hCGα and hCGβ glycoforms.

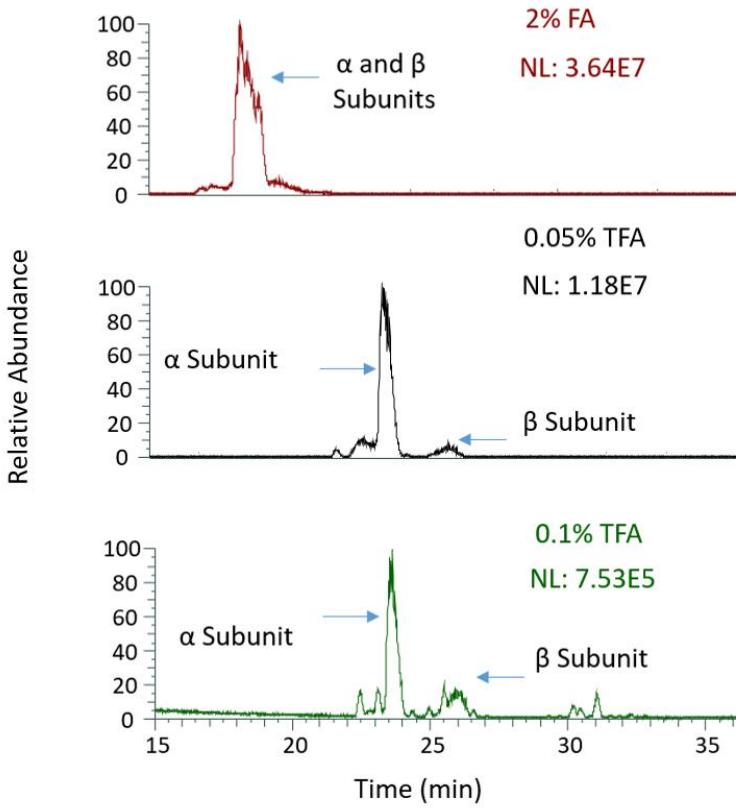


Figure 27. BPCs of rhCG analyzed by nano-LC-HRMS with either FA or TFA as the mobile phase acidifier. Mobile phase: 2% FA, 0.05% TFA, or 0.1% TFA in H2O/ACN mixture (v/v). Gradient slope: 1.02% ACN/min. rhCG at 100 µg ml-1. Injection volume: 1 µl in partial loop mode. Other conditions: see section 2.3.

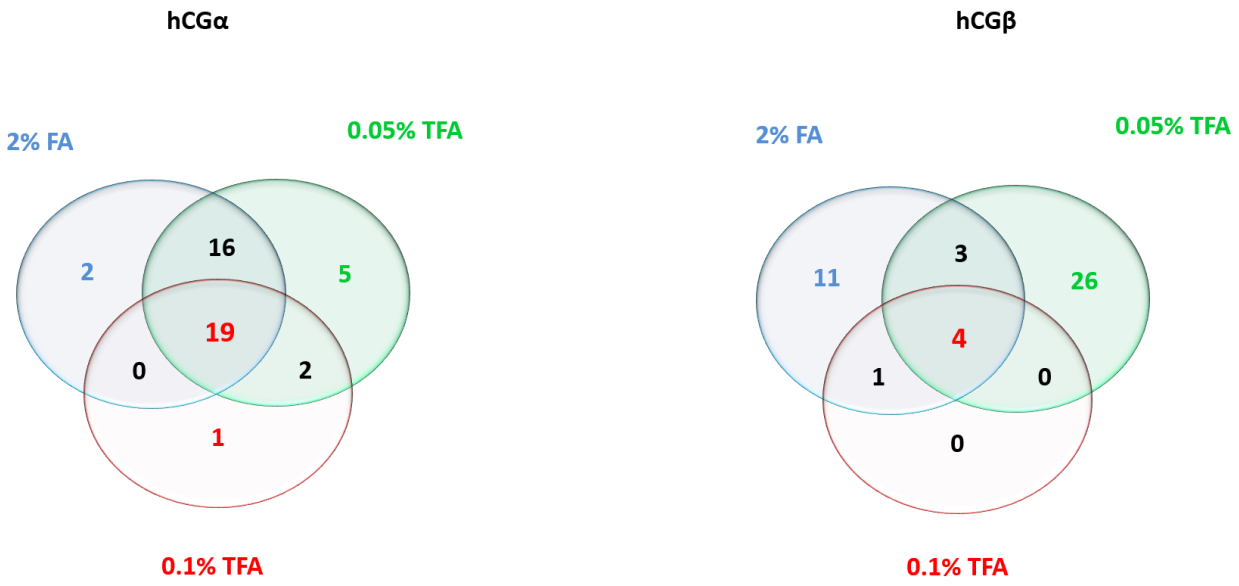


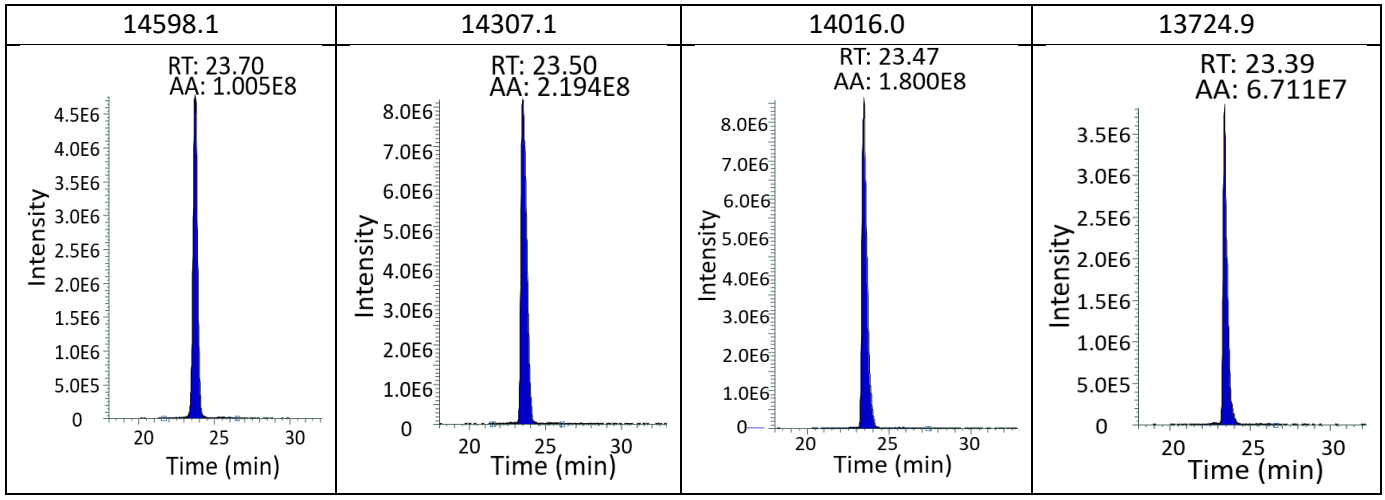
Figure 28. Venn diagram comparing the number of hCGα and hCGβ glycoforms detected after the analysis by nanoLC-HRMS of rhCG according to the nature and content of acidic additive in the mobile phase. Analytical conditions: see Figure 3.

4.3.4. Figures of merit of the nanoLC-HRMS method

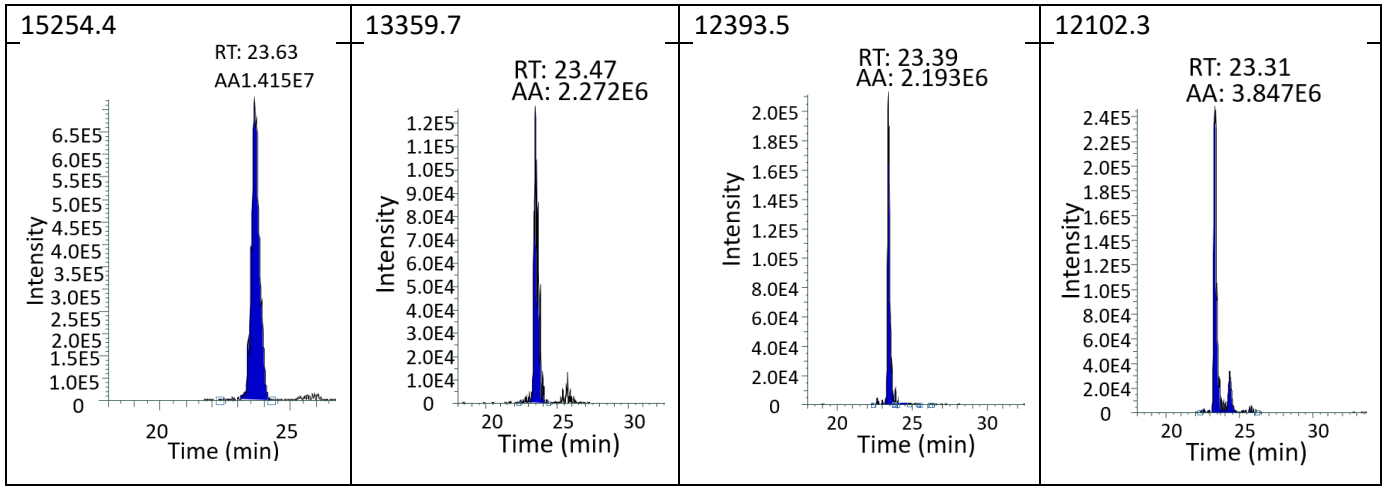
The repeatability of this method that consists in injecting 1 µl of rhCG using partial loop injection mode, a trapping precolumn and a gradient slope of 1.02% ACN/min with 0.05% of TFA in the mobile phase was investigated. For this, 8 rhCG samples were prepared at various concentrations and analyzed, each in triplicate, starting from 500 µg mL^{-1} and ending with 6 µg mL^{-1}. For each analysis, a list of all detected hCGα and hCGβ glycoforms was obtained and for each glycoform the XIC was studied to determine its corresponding retention time (t_R) and peak area. The peak area of each hCGα glycoform was next normalized by the area of the hCGα glycoform leading to the highest area value measured on its own XIC (hCGα having a MW of 14,307.1 Da). Similarly, for each hCGβ glycoform, its area was normalized by the one having the highest area, i.e. a MW of 21,723.8 Da. The obtained relative area was then used for the relative semi-quantification calculation. As an example, **Figure 29** presents the XICs obtained for 4 high intensity signal hCGα glycoforms (i.e. having a relative area fixed arbitrarily above

15%), 4 low intensity signal hCGα glycoforms (i.e. having a relative area fixed arbitrarily below 15%) and 4 hCGβ glycoforms. It is important to mention that a ratio of 84 was obtained between the highest and the lowest intensities of all the detected glycoforms at a given rhCG concentration with the optimized analytical conditions. All detected hCGα and hCGβ glycoforms and their corresponding average retention time and average relative area and their RDS values at 100 µg mL^{-1} are compiled in **Tables S6 and S7** in the supplementary data.

High intensity hCGα glycoforms



Low intensity hCGα glycoforms



hCGβ glycoforms

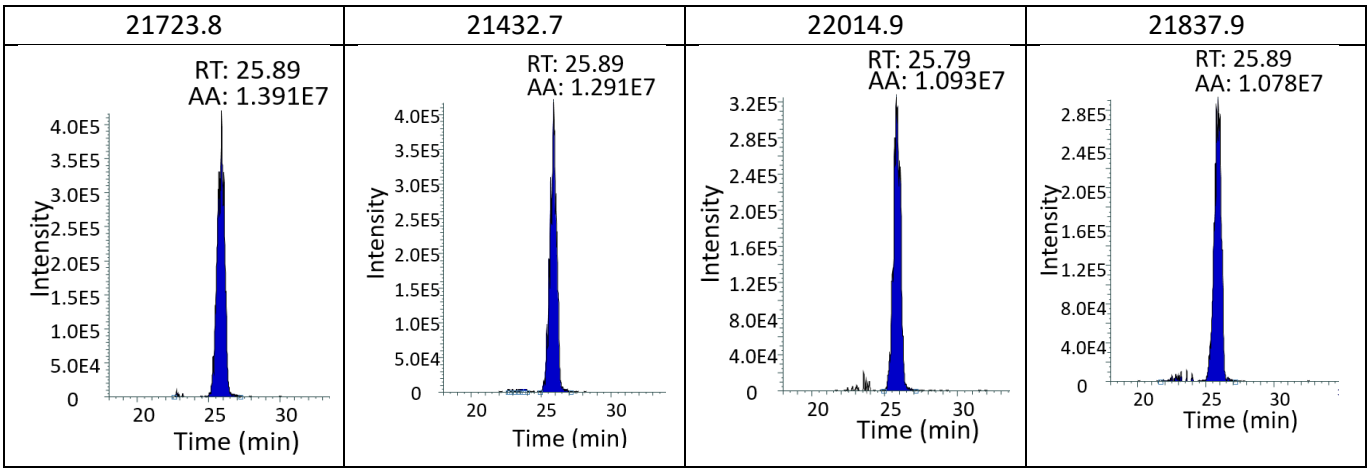


Figure 29. XIC of 4 high intensity hCGα glycoforms, 4 low intensity hCGα and 4 hCGβ glycoforms analyzed by nano-LC-HRMS with all the optimized conditions. Mobile phase: 0.05% TFA in H2O/ACN mixture (v/v). Gradient: from 10 to 65% ACN with a slope of 1.02% ACN/min. rhCG at 100 µg ml^{-1}. Injection volume: 1 µl in partial loop mode.

The resulting RSD values of the retention times are compiled in **Table 9** for hCGα glycoforms and hCGβ glycoforms for the different tested concentrations. It appears that the RSD values are all inferior to 1.1 %. The RSD ranges of relative areas and average for each tested concentration are also reported in **Table 9** for hCGα and hCGβ glycoforms. In general, the average RSD values are below 20% and there is no correlation between RSD values and injected concentrations, even for the lowest ones.

Table 9 RSD range and its average value for the retention times and relative peak areas of the hCGα and hCGβ glycoforms detected after the analysis by nano-LC-HRMS with the optimized conditions of rhCG at 8 different concentrations, in triplicate for each. Mobile phase: 0.05% TFA in H_2O/ACN mixture (v/v). Gradient: from 10 to 65% ACN with a slope of 1.02% ACN/min. rhCG at 100 µg ml^{-1}. Injection volume: 1 µl in partial loop mode. The hCGα glycoforms are said of high intensity is their relative area is above 15% of the relative area of the hCGα glycoform giving the highest peak area. Similarly, the hCGα glycoforms are said of low intensity is their relative area is below 15%.

	The 42 hCGα glycoforms	**The 7 hCGα glycoforms of high intensity**	**The 35 hCGα glycoforms of low intensity**	**The 33 hCGβ glycoforms**	
rhCG concentration (µg/ml)	**RSD range (average) of t_R (%)**	**RSD range (average) of relative area (%)**	**RSD range (average) of relative area (%)**	**RSD range (average) of t_R (%)**	**RSD range (average) of relative area (%)**
500	0.17-0.85 (0.50)	4.7-19.7 (12.7)	4.0- 66.5 (19.6)	0.07-1.08 (0.38)	0.5-11.7 (6.6)
250	0.37-0.95 (0.59)	1.6-6.8 (3.3)	0.5-19.6 (7.0)	0.11-0.61 (0.29)	2.0-14.7 (7.2)
125	0.04-0.61 (0.24)	0.8-7.9 (4.4)	1.6-49.8 (11.2)	0.02-0.49 (0.22)	0.9-17.1 (7.6)
100	0.04-0.39 (0.18)	1.8-10.7 (3.9)	2.0-40.5 (12.8)	0.01-0.21 (0.10)	1.9-17.7 (7.7)
50	0.02-0.61 (0.23)	6.4-14.7 (8.7)	4.4-44.2 (17.9)	0.03-0.47 (0.17)	1.2-26.2 (12.1)
25	0.02-0.48 (0.22)	1.0-17.0 (9.1)	2.0-27.4 (11.1)	0.05-0.37 (0.22)	3.0-30.6 (14.3)
12.5	0.33-0.72 (0.53)	1.3-9.0 (4.4)	2.3-51.7 (12.4)	-	-
6	0.40-0.50 (0.46)	4.0-11.3 (6.0)	-	-	-

The linearity of this method was next studied for the hCGα glycoforms giving the highest peak areas for 8 concentrations of rhCG ranging from 6 to 500 µg mL^{-1}. All the correlation coefficient (R^2) values and the equations of regression of each glycoform are gathered in **Table S8** in the supplementary data. The linearity of the method was better achieved between 6 and 250 µg mL^{-1} of rhCG with R^2 values between 0.93 and 0.98, than between 6 and 500 µg mL^{-1} as illustrated in **Figure S5** (supplementary data). It means that a concentration of 500 µg mL^{-1} was definitely too high and induced ion suppression phenomena during ionization. Finally, the limit of quantification (LOQ) of the method could be estimated, even if the LOQ value was determined as a function of a global concentration of rhCG because the exact concentration of each

glycoform present in the sample is unknown. The 7 high intensity signal hCGα glycoforms (fixed arbitrarily above 15% for their relative areas) are all detected down to 6 µg mL^{-1} of the global hCG content with a signal to noise ratio measured on the XIC of each glycoform above 10. The low intensity hCGα glycoforms are all detected injecting 12.5 µg mL^{-1} of rhCG, while injecting less led to the total loss of the signals indicating that the LOQ of the low intensity glycoforms was 12.5 µg mL^{-1}. Following the same reasoning the LOQ for hCGβ glycoforms is 25 µg mL^{-1}.

4.3.5. Identification of the hCGα and hCGβ glycoforms

For the identification of hCGα glycoforms, the MWs of the detected glycoforms were compared to those of the primary amino acid sequence of hCGα plus the MW of all the possible combinations of two potential N-glycans already described in the literature for hCGα [16], [32], [55]and to the ones that we recently identified with a bottom–up approach [14]. The identifications that were possible are reported in **Table 10**. It was sometimes not possible to elucidate the glycoform structure that may be a result of a combination of glycosylation and other PTMs. In addition, for several identified glycoforms a difference of 16 or 32 Da was detected, which could be the indication of an oxidation that leads to the addition of one or two oxygen atoms to a susceptible amino acid side-chain.

For hCGβ glycoforms, the identification of the exact structures is much more complicated due to the presence of the additional 4 O-glycosylation sites that lack the common core glycan structure of the N-glycans. This multiplies the number of possible combinations making the exact identification very difficult. Nevertheless, an important observation was achieved when selecting the whole retention time range corresponding to hCGβ glycoforms (retention time between 25 and 27 min) at once. The resulting average mass spectrum with the determined charge states is

presented in **Figure S6** where we can notice the presence of two m/z values 2116.39 and 2336.86 that corresponds to the MWs 21,141.74 and 23,346.51 Da, respectively. These MWs correspond to two glycoforms that differs by the mass of one N-glycan bearing two sialylated antennas (HexNAc(4)Hex(5)NeuAc(2), with Hex, hexose; HexNAc, N-acetylglucosamine; and NeuAc, N-acetylneuraminic), which proves the ability to differentiate between hCGβ glycoforms with two or only one occupied N-glycosylation sites. Therefore, the newly developed nano-LC-HRMS method allows the rapid and reproducible evaluation of a qualitative and semi-quantitative mapping of not only the hCGα but also of the hCGβ glycoforms that are present in a sample, which could be helpful to compare more accurately hCG-based samples.

Table 10 MW of the identified hCGα glycoforms detected after nano-LC-HRMS analysis of rhCG in triplicate and use of the new optimized method with glycan structure suggestions for some in comparison with bottom-up data (marked with *) or with GlyGen (marked with**), while those confirmed by both are marked with ***. The notation 1st and 2nd glycan is not correlated to their localization on the 2 N-glycosylation sites of hCGα. The glycoforms were highlighted in dark gray when addition of two oxygen atoms was suspected while those highlighted in light gray were suspected to have the addition of only one oxygen atom.

Mass (Da)	1st N-Glycan Proposed structure	2nd N-Glycan Proposed structure	Structure identification method
14963.3	Hex5 HexNAc4 NeuAc1	Hex6 HexNAc5 NeuAc3	*
14672.1	Hex5 HexNAc4 NeuAc2	Hex6 HexNAc5 NeuAc1	**
14598.1	Hex5 HexNAc4 NeuAc2	Hex5 HexNAc4 NeuAc2	**
14437.2	Hex4 HexNAc4 dHex1 NeuAc1	Hex5 HexNAc4 dHex1 NeuAc2	*
	Hex4 HexNAc5 dHex1	Hex6 HexNAc5 NeuAc1	**
14307.1	Hex5 HexNAc4 dHex1 NeuAc2	Hex5 HexNAc4 dHex1	**
14146.1	Hex4 HexNAc4 dHex1 NeuAc1	Hex5 HexNAc4 dHex1 NeuAc1	*
14104.1	Hex5 HexNAc4 NeuAc2	Hex5 HexNAc3 NeuAc1	***
14090.0	Hex5 HexNAc4	Hex6 HexNAc5 NeuAc1	***
14055.9	Hex4 HexNAc5 dHex1 NeuAc1	Hex4 HexNAc4 dHex1	**
14048.9	Hex6 HexNAc3	Hex6 HexNAc5 NeuAc1	***
14016.0	Hex5 HexNAc4 NeuAc1	Hex5 HexNAc4 NeuAc1	**
13942.9	Hex4 HexNAc5	Hex3 HexNAc6 dHex1	**
13812.8	Hex5 HexNAc4 NeuAc1	Hex5 HexNAc3 NeuAc1	***
13741.9	Hex6 HexNAc4	Hex5 HexNAc4 dHex1	*

13724.9	Hex5 HexNAc4	Hex5 HexNAc4 NeuAc1	***
13683.8	Hex6 HexNAc3 dHex1	Hex5 HexNAc4	*
13650.8	Hex5 HexNAc4 NeuAc1	Hex4 HexNAc3 NeuAc1	***
13609.8	Hex5 HexNAc3 NeuAc1	Hex5 HexNAc3 NeuAc1	***
13521.7	Hex5 HexNAc4	Hex5 HexNAc3 NeuAc1	***
13359.7	Hex5 HexNAc4 NeuAc1	Hex3 HexNAc4	*
	Hex6 HexNAc3	Hex3 HexNAc4 NeuAc1	**

4.4. Conclusion

In this study a new nano-LC-HRMS method implementing an Orbitrap analyzer was developed and applied to the detection of rhCG glycoforms. After optimizing various key parameters such as injection volume, injection mode, the gradient slope and mobile phase additive, this method allowed the detection of numerous hCGα but also for the first time hCGβ glycoforms with this approach. The good figures of merit of the method enable its use for the semi-quantification of these glycoforms. This is why this method could be implemented in the future for fast mapping of hCG samples and should be evaluated also with other gonadotropins.

Acknowledgements

This work has received the support of the Conseil Régional d'Île-de-France with the DIM Analytics, and Horizon 2020, Actions Marie-Sklolowska Curie, Cofund UPtoParis, ESPCI Paris PSL.

Chapter 5: Analysis of different recombinant and urinary gonadotropin proteins in nanoLC-(Orbitrap)MS at the intact level

The successful optimization of the nanoLC-(Orbitrap)MS method that enabled the detection of a high number of hCGα and hCGβ glycoforms present in a recombinant hCG-based drug was the trigger to this part that is dedicated to test the potentials of the application of this method for the analysis of two urinary hCG-based drugs, but also other recombinant and urinary gonadotropin-based drugs. Indeed, the similarity between the selected gonadotropins lies in the fact that they share the same α subunit, but each is bound to a unique β subunit. The selected samples were rLH, rFSH, a mixture of menopausal LH and FSH, in addition to two commercial purified urinary hCG samples. Keeping the previously optimized injection conditions that led to an increase in the injected volume and the mobile phase gradient, all the samples were analyzed with a mobile phase containing either 2% FA or 0.05% TFA to determine which acidic additive led to the best compromise between selectivity that favors the glycoform ionization and sensitivity as a function of the gonadotropin nature. This chapter is in the form of a published article in Journal of Chromatography B: A. Al Matari, A. Combès, T. Fournier, V. Pichon, and N. Delaunay, “Analysis of different recombinant and urinary gonadotropin proteins in nanoLC-(Orbitrap)MS at the intact level”

Analysis of different recombinant and urinary gonadotropin proteins in nanoLC-(Orbitrap)MS at the intact level

Amira Al Matari[1], Audrey Combès[1], Thierry Fournier[2], Valérie Pichon[1, 3], Nathalie Delaunay[1]

[1]Laboratory of Analytical, Bioanalytical Sciences and Miniaturization, Chemistry, Biology and Innovation (CBI) UMR 8231, ESPCI Paris PSL, CNRS, PSL Research University, Paris, France

[2]Université de Paris, INSERM, UMR-S1139, «Pathophysiology & Pharmacotoxicology of the Human Placenta, pre & postnatal Microbiota», 3PHM, Paris, France.

[3]Sorbonne Université, France

Corresponding author: Nathalie Delaunay, UMR CBI 8231, ESPCI Paris PSL, 10 rue Vauquelin, 75231 Paris cedex 05, France. Tel: (+33)1 4079 4651. nathalie.delaunay@espci.fr. ORCID: 0000-0003-3008-0286

Compliance with Ethical Standards

a) The authors declare that there are no conflicts of interest.

b) Research involving Human Participants and/or Animals: no.

Keywords: Glycoprotein, gonadotropins, intact protein, mass spectrometry, nano-liquid chromatography.

Abbreviations:

AA, amino acid; BPC, Base Peak Chromatogram; ACN, acetonitrile; CHO, Chinese hamster ovary; FA, formic acid; FSH, follicle-stimulating hormone; hCG, human chorionic gonadotropin; hCGα, human chorionic gonadotropin alpha subunit; HRMS, high-resolution mass spectrometry; LH, luteinizing hormone; LC, high performance liquid chromatography; m/z, mass-to-charge; MS, mass spectrometry; MW, molecular weight; PTM, post-translational modification; rhCG, a recombinant hCG-based drug (Ovitrelle); RSD, relative standard deviation; TFA: trifluoroacetic acid.

Abstract

The human chorionic gonadotropin hormone (hCG), follicle-stimulating hormone (FSH), and luteinizing hormone (LH) belong to the gonadotropin family and regulate normal growth, sexual development, and reproductive functions. They are heterodimeric glycoproteins made of two subunits, α and β. They share the same α-subunit structure, but each is bound to a unique β-subunit resulting in differences in their physiological roles and signaling. These hormones are considered highly glycosylated proteins due to the presence of 2 to 8 glycosylation sites, according to the gonadotropin nature, leading to numerous glycoforms. The analysis at the intact level is the only way to correlate a glycoform with its different glycans and determine the total number of glycoforms and their relative abundances present in a given sample. Implementing nano-reversed phase liquid chromatography coupled to mass spectrometry (MS) with an Orbitrap analyzer, two mobile phase compositions (2% formic Acid (FA) or 0.05% trifluoroacetic acid (TFA) as additives) were previously optimized using a recombinant hCG-based drug, Ovitrelle®, as standard. The potential of these two methods for the characterization of the gonadotropin glycoforms was here evaluated for the first time to three different kinds of gonadotropin: two formulations of filtered urinary hCG called CG10 and Orb80731, two recombinant drugs Puregon® and Luveris® containing FSH and LH, respectively, and a highly purified human Menopausal Gonadotropin preparation containing both FSH and LH, called Merional®. The differences in gonadotropin's MWs and number of N- and O-glycosylation sites between their beta subunits favored the use of only one or two compositions of mobile phase for the characterization of their glycoforms at the intact level. For the analysis of hCG the use of TFA alone is preferred, while the complementarity of both additives was demonstrated for the other gonadotropins.

5.1. Introduction

Glycosylation is the attachment of sugar moieties to certain proteins. It is the most common and complex form of post translational modifications [108]. It is estimated that 70% of human proteins are glycosylate [109]. The added carbohydrates are essential for various biological processes, such as molecular recognition, inter- and intracellular signaling, fertilization, embryonic development, immune defense, inflammation, cell adhesion and division processes, viral replication and parasitic infections. Moreover, glycan chains can alter protein conformation by which it may consequently modify the functional activity of a protein as well as protein/protein interactions [110]. Several well-known glycoproteins are being used as pharmaceutical drugs such as antibodies [111] and gonadotropin hormones [53].

The gonadotropins belong to a family of homologous heterodimeric proteins secreted by the pituitary gland of all vertebrates in addition to placental cells of primates. These hormones include the following pituitary hormones: follitropin, also called follicle-stimulating hormone (FSH), lutropin (or luteinizing hormone, LH) and thyrotropin (thyroid-stimulating hormone, TSH), in addition to human chorionic gonadotropin (hCG) [112]. They share the same α subunit that possesses both Asn52 and Asn78 N-glycosylation sites. They are also composed of a β subunit, which is specific to each gonadotropin. As an example, FSHβ has two N-glycosylation sites [113], [114], whereas LHβ has only one N-glycosylation site [115], and hCGβ has two N-glycosylation and four O-glycosylation sites, leading to a very high number of different glycoforms [9], with a molecular weight (MW) approximately ranges between 22-24 kDa [28].

The glycosylation state of these gonadotropins affects their stability, solubility, and clearance. Therefore the characterization of these glycosylation patterns is essential during the development of the pharmaceutical preparations in each of the following steps: discovery, preclinical, and

clinical steps in addition to the monitoring of their production, storage, and delivery [116]. However, the analysis of these glycoforms has to be down at the intact level to correlate a glycoform with its different glycans and determine the total number of glycoforms and their relative abundances present in a given sample. This is a highly challenging objective requiring the use of high performance separation and detection methods [81], [110], [117]. This is why, recently, we have optimized two methods based on nano-reversed phase liquid chromatography coupled to mass spectrometry (nanoLC-MS) with an Orbitrap analyzer for the analysis of hCG glycoforms at the intact level using a recombinant hCG-based drug, Ovitrelle® (rhCG), as standard [107], [118]. One of the two main differences between the two developed methods is in one case the implementation of a preconcentration step with a trapping precolumn between sample injection and glycoform separation and detection, which allowed an increase in the injected volume from 70 nl to 1 µl and thus in sensitivity, without losing any glycoforms [118]. The second difference is the nature and content of the acidic additive in the mobile phase, 2% formic acid (FA) or 0.05% trifluoroacetic acid (TFA), an ion pairing agent promoting selectivity but inducing signal suppression in MS [118].In the case of the analysis of rhCG, 2% FA led to a higher sensitivity but to a lower number of detected hCGα and hCGβ glycoforms whereas 0.05% TFA allowed the detection and semi-quantification of 42 hCGα glycoforms and 33 hCGβ glycoforms at intact level after having optimized the mobile phase gradient slope. In the present work, the idea was to go further in the comparison of the nature and content of the acidic additive in the mobile phase, keeping the trapping precolumn in the set-up, for the analysis in nanoLC-MS of different categories of gonadotropins, with the objective to determine the best approach for each kind of gonadotropin that enables the detection of the highest number of glycoforms. First, two commercial urinary-based hCG samples, called CG10 and Orb80731,

were tested in comparison with the recombinant hCG-based drug previously used. Then two recombinant FSH- and LH-based drugs, Puregon® and Luveris®, respectively, were analyzed. Finally, a highly purified Human Menopausal Gonadotropin (hMG) preparation containing FSH and LH, called Merional®, was tested.

5.2. Experimental Part

5.2.1. Materials and Reagents

HPLC-grade acetonitrile (ACN), FA, TFA and LC-MS-grade water were supplied by Carlo Erba (Val de Reuil, France). Puregon® (Organon, Oss, The Netherlands) is a drug made of recombinant FSH from Chinese hamster ovary (CHO) cell line and contains 900 IU of FSH in a 1.08 mL aqueous solution. It was converted to its corresponding µg content using [119], which is equivalent to 83.3 µg mL^{-1}. Luveris® (Merck Serono, Lyon, France) is a drug made of recombinant LH from CHO cell line and contains 75 IU in the form of lyophilized powder dissolved in 1ml aqueous solution which is equivalent to 3 µg mL^{-1} [120]. Merional® (IBSA, Lugano, Switzerland) is a highly purified Human Menopausal Gonadotropin preparation containing 75 IU of FSH and 75 IU of LH in the form of lyophilized powder dissolved in 1ml aqueous solution. The conversion of IU to µg was done following the reference [121], which corresponds to 90 µg mL^{-1} of LH and 23 µg mL^{-1} of FSH are present in this sample. CG10 (Sigma Aldrich, Saint Louis, USA) is a lyophilized powder of filtered urinary hCG containing 10,000 IU (46 µg) hCG per vial reconstituted in water to obtain a final concentration of 200 µg mL^{-1}. Orb80731 (Biorbyt, Saint Louis, USA) is a lyophilized powder (5 mg) of filtered urinary hCG with an activity of 56,251 IU mg^{-1}.

5.2.2. Sample preparation

All drug's stock, either in solutions or reconstituted powders, were desalted and filtered for 10 min at 4°C with ultra-centrifugal units (Microcon, Merck, Darmstadt, Germany) with a MW cut-off of 10 kDa. After this step that aims to remove additives and excipients from the formulation, the residues were washed with water and finally the samples were recovered from the inverted membrane by centrifugation for 4 min at 14,000 g (≈ 11,000 rpm) and the volumes were adjusted with water to obtain a final concentration of: 200 µg mL^{-1} of FSH, 30 µg mL^{-1} of LH , 230 µg mL^{-1} of FSH and 900 µg mL^{-1} of LH (Merional), 200 µg mL^{-1} of hCG (CG10 and orb80731) .

5.2.3. NanoLC-MS system

The analyses were carried out using a nanoLC system (NC 3500 RS, ThermoFisher Scientific, Courtaboeuf, France) coupled to an Orbitrap mass analyzer (Exactive plus, ThermoFisher Scientific). A precolumn C18 PepMap (0.3 x 5 mm internal diameter, 300 Å) and packed with 5 µm particles (ThermoFisher Scientific) and an EASY-Spray PepMap RSLC C18 (150 x 0.075 mm internal diameter, 300 Å) column packed with 5 µm particles (ThermoFisher Scientific) were used. The mobile phase was composed of (A) H_2O/ACN (98/2; v/v) and (B) H_2O/ACN (1/9; v/v) acidified with 2% FA (v/v) or 0.05% TFA (v/v). The column temperature was set at 60°C and the injection volume was 1 µl in partial loop injection mode. The flow-rate was set at 300 nL min^{-1} and the gradient was 10 to 50% ACN in 39 min, followed by a rapid increase to 90% of ACN within 5 min, then a 5 min plateau, and finally a return to the initial 10% of ACN concentration and kept for 15 min for equilibration. The EASY-spray source parameters were as follows: capillary voltage, 1.7 kV; drying temperature, 280°C; and sheath,

auxiliary, and sweep gas flows were set at 0. The focusing voltages were tuned weekly by the diffusion of positive and negative ion calibration solutions (ThermoFisher Scientific). The mass spectrometer functioned in positive full MS mode and using a scan range from 1000 to 2500 mass-to-charge (m/z) with a scan rate of 1 spectrum/s. The mass resolution was set to 140,000 at 200 m/z. The AGC target was 3e6, and the maximum injection time was 200 ms. Finally the deconvolution analyses were processed using Freestyle 1.3 SP2 (ThermoFisher Scientific) with the same deconvolution parameters as the previous optimized method (ref article 1) that uses a mass tolerance for m/z of 30 mDa, and a minimal charge state of 5, with at least 3 different charge states to identify a compound and the accuracy mass tolerance for deconvoluted MW was set at 300 mDa.

5.2.4. Nano-LC-HRMS data treatment for the identification of glycoforms

In this study the spectral deconvolution and data treatment method used was the same as the one previously optimized *[107]*. Briefly each chromatogram was cut in successive time slices of 30 s and next 50 s. This was followed by collecting the average deconvoluted mass spectrum for each slice resulting in the achievement of a list of MWs corresponding to the detected glycoforms present in each sample. Next, the results of both types of time slicing divisions were combined to reduce the risk of data loss at the boundaries of the performed cuts. The resulted list of MWs are compared to the theoretical mass of the amino acid (AA) sequence of each gonadotropin subunit (α or β) in addition to their potential glycans.

5.3. Results and discussion

5.3.1. TFA vs FA as mobile phase additive for different samples of hCG

As mentioned in the introduction, two methods based on nanoLC-MS with an Orbitrap analyzer were developed for the analysis of hCG glycoforms at the intact level that differs by the addition

of a trapping column and the nature and content of acidic additive in the mobile phase [107], [118]. The use of 2% FA in the mobile phase for the analysis of intact hCG led to a higher sensitivity but to a lower number of hCGα and hCGβ glycoforms in comparison with 0.05% TFA (37 hCGα and 19 hCGβ glycoforms with 2% FA and 42 hCGα and 33 hCGβ glycoforms with 0.05% TFA, as reminded in **Table 11**). Nevertheless, if most of the hCGα glycoforms were detected with the two acidic additives, it was not the case for the hCGβ glycoforms as illustrated in **Figure 30**. Indeed, among the 33 hCGβ glycoforms detected with 0.05% TFA, only 7 were also detected with 2% FA plus 12 additional ones. In this case, both methods could be complementary.

Table 11 Gathering of the number of detected glycoforms and their corresponding MW and retention ranges for rhCG, CG10 and Orb80731 hCG-based samples, Luveris (LH based-sample), Puregon (FSH-based sample), and Merional® (LH- and FSH-based sample), analyzed by nanoLC-MS with either 2% FA or 0.05% TFA as mobile phase additive.

	FA			**TFA**		
Sample	**Number of Glycoforms**	**MW range (kDa)**	**t_R range (min)**	**Number of Glycoforms**	**MW range (kDa)**	**t_R range (min)**
rhCG	37 hCGα	12.10-15.25	17.8-19.1	42 hCGα	12.10-15.25	23.3-23.8
	19 hCGβ	21.14-24.29	18.4-20.6	33 hCGβ	21.14-23.93	25.7-25.9
CG10	68 hCGα	12.45-14.61	14.5-20.1	92 hCGα	12.67-15.82	20.4-24.3
	9 hCGβ	21.96-23.35	22.0-23.3	17 hCGβ	22.04-23.78	26.3-26.5
Orb80731	47 hCGα	13.29-14.76	18.8-20.9	55 hCGα	13.29-14.79	18.7-21.7
	0 hCGβ	-	-	11 hCGβ	22.98-24.04	25.8-26.0
Luveris® (LH)	71	14.02-15.98	17.5-19.4	58	13.72-16.30	24.2-31.3
Puregon® (FSH)	64	13.27-16.48	16.9-20.8	39	13.56-15.70	17.3-17.6
	16	18.35-19.30	24.4-28.3	-	-	-
Merional® (FSH + LH)	54	14.61-16.16	14.4-20.5	36	15.0-16.04	20.6-24.3

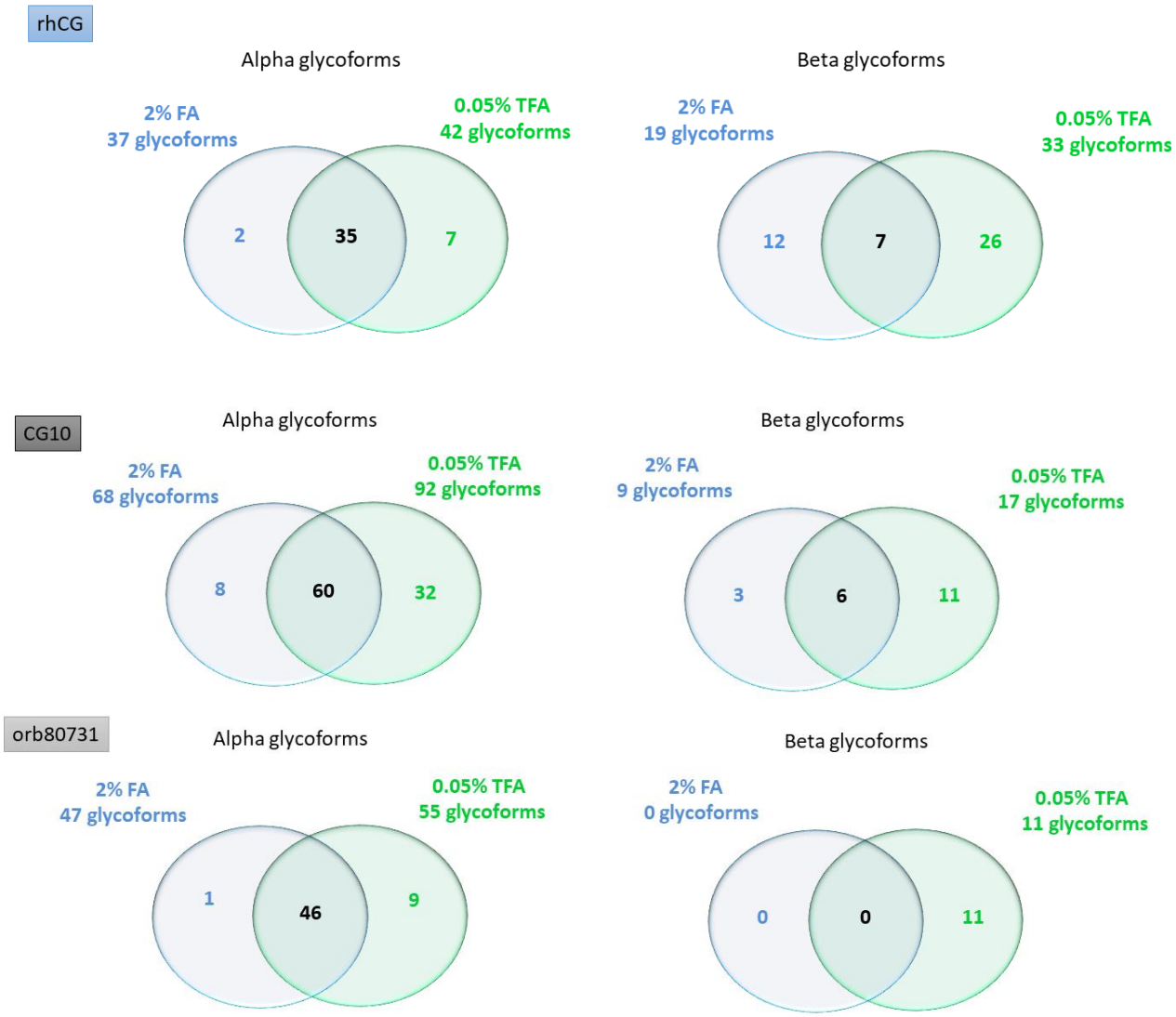


Figure 30. Venn diagrams comparing the number of hCGα and hCGβ glycoforms detected after the analysis of CG10, orb80731, and rhCG by nanoLC-MS according to the nature and content of acidic additive in the mobile phase. Other conditions: see section 2.3.

In the present work, the idea was first to confirm or not this tendency for hCG and determine if one additive is definitely more appropriate for the analysis of this gonadotropin glycoforms, or if both additives led to complementary results, which means that it would rather be necessary to

analyze one given hCG sample with the two LC mobile phase compositions, which would be time-consuming and more expensive. Therefore, nanoLC-MS analysis using the trapping precolumn in the set-up with either 0.05% TFA or 2% FA in the mobile phase was carried out with two other hCG samples, CG10 and orb80731. The typical resulting base peak chromatograms (BPCs) are represented in **Figure 31 (a, b, c, g, h and i)**. First of all, significant different BPCs were obtained even keeping constant the additive nature, showing dramatic differences between the two urinary hCG-based samples and also with rhCG, which means that they are composed of different glycoform mixtures. A large massif of peaks can be observed in all cases, since each sample contains a high number of glycoforms that cannot be baseline resolved by the nanoLC column. However, the separation power of the MS analyzer can partially overcome it. Nevertheless, as already seen for rhCG, the TFA ion pairing agent improved the separation between the alpha and beta glycoforms, leading to the detection of a significant higher number of hCGα and hCGβ glycoforms, as presented in **Table 11**. Indeed, the analysis of hCG named CG10 led to the detection of 92 hCGα and 17 hCGβ glycoforms with TFA *vs* 68 and 9 with FA, respectively. All the detected hCGα and hCGβ glycoforms, characterized by their MW and their corresponding average unique retention time and average relative area and their respective RDS values, are compiled in the supplementary data **(Table S9** and **S10**, respectively). It is important to note that for several peaks on the mass spectra corresponding to detected glycoforms, a second, and sometimes even a third, less intense peaks were observed with a mass difference of 16 Da which could result from (i) an oxidation process that took place during LC separation or ionization process or (ii) the replacement of a sialic acid by a glycolylneuraminic acid on the glycan chain.

The same analysis carried out with orb80731 led to the detection of 55 hCGα and 11 hCGβ

glycoforms with TFA *vs* only 47 hCGα and no detection of hCGβ glycoforms with FA (Table 11). All the detected hCGα and hCGβ glycoform MWs and their corresponding average retention time and average relative area and their RDS values are compiled in **Tables S11** and **S12**, respectively.

If the MW range of the detected glycoforms did not vary with the nature of the acidic additive in the mobile phase, this is not the case for the retention time range: TFA induced an increase in the retention on the apolar stationary phase, and thus glycoforms elute with a mobile phase containing a higher content of ACN. It seems also that an improvement of the hCGα and hCGβ glycoforms retention and chromatographic resolution was obtained, as their corresponding elution time range gap increased with TFA. Both effects contribute to an improvement of the glycoform ionization and may explain why a higher number of glycoforms are detected with TFA than with FA. Regarding the nature of the detected glycoforms, **Figure 30** presents the Venn diagrams comparing the number of hCGα and hCGβ glycoforms detected after the analysis of CG10 and orb80371 hCG-based samples with FA or TFA. It appears that contrary to what was observed with rhCG, even if FA allows the detection of some new glycoforms, their number remains very low: 8 compared to 92 in one case and 1 compared to 56 in the other for hCGα glycoforms, and 3 compared to 17 in one case and 0 compared to 11 in the other for hCGβ glycoforms. One can wonder if this could be linked to a problem in sensitivity when analyzing with TFA rather than FA, but depending of the sample and of the subunit nature of the glycoforms, the non-common glycoforms are of either high or low intensity. It seems thus that there is no correlation between the missed glycoforms and their intensity. It is also worthwhile to notice that (i) the two glycoforms giving the highest signal are the same using FA or TFA and (ii) the relative areas of common glycoforms are not significantly different in most cases (>80%)

using FA or TFA. This means that the nanoLC-MS method using 2% FA as additive did not provide a dramatic improvement of the glycoform detection in hCG samples. Therefore, TFA should be always preferred to FA for the analysis of hCG glycoforms at the intact level if only one method can be performed.

Finally, a comparison between the detected glycoforms in the three hCG samples was done and the results are represented in **Figure 32**. It can be noticed that the recombinant (rhCG) and urinary preparations (CG10 and orb80731) of hCG share a low number of hCGα and hCGβ glycoforms, 16 and 1, respectively. The three samples are significantly different from each other, so these biological activities too.

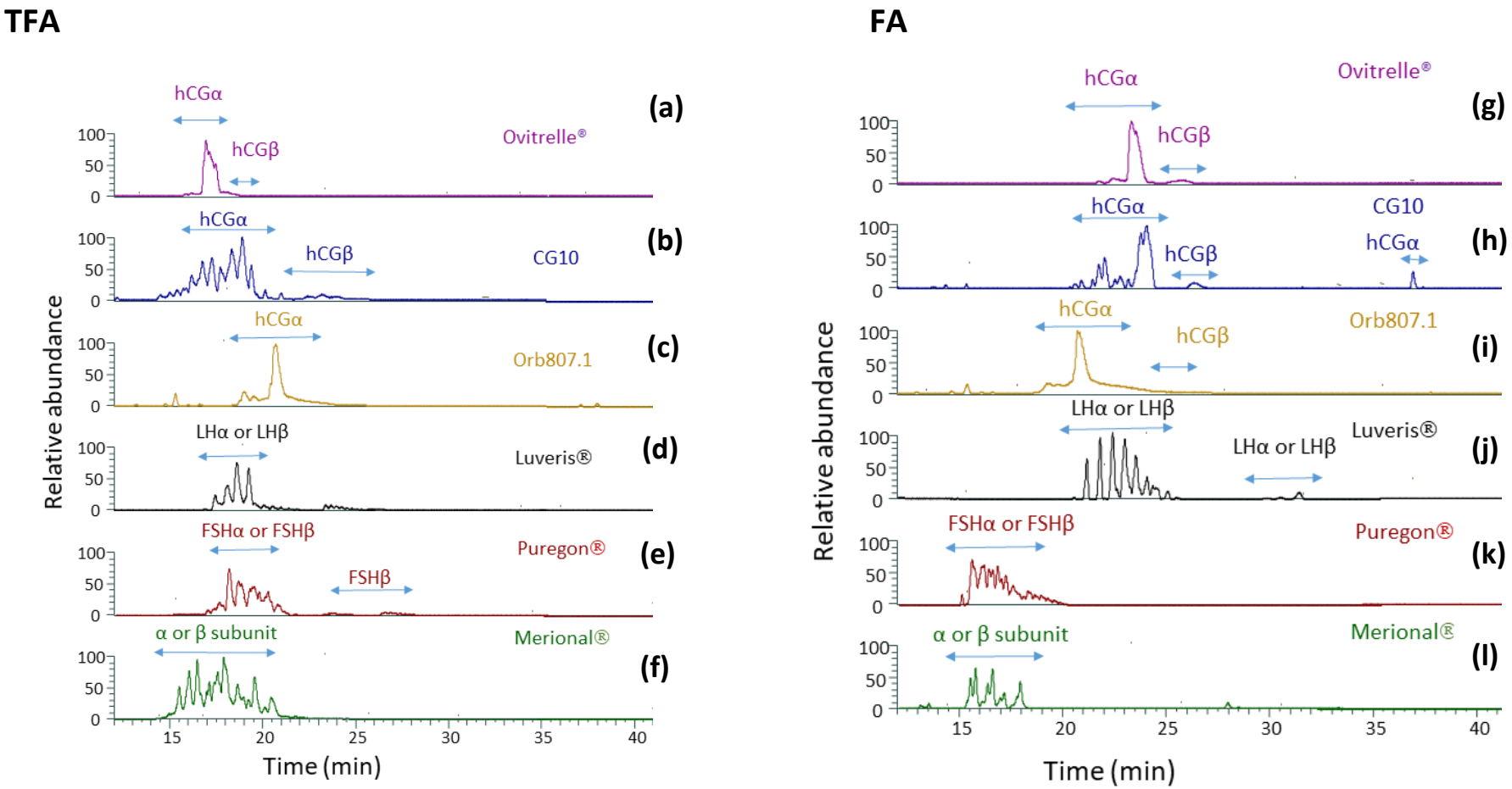


Figure 31. BPCs of Ovitrelle® (rhCG), CG10, Orb80731, Luveris®, Puregon®, and Merional®, analyzed by nanoLC-MS with 0.05% TFA or 2% FA as the mobile phase acidifier. Other conditions: see section 2.3.

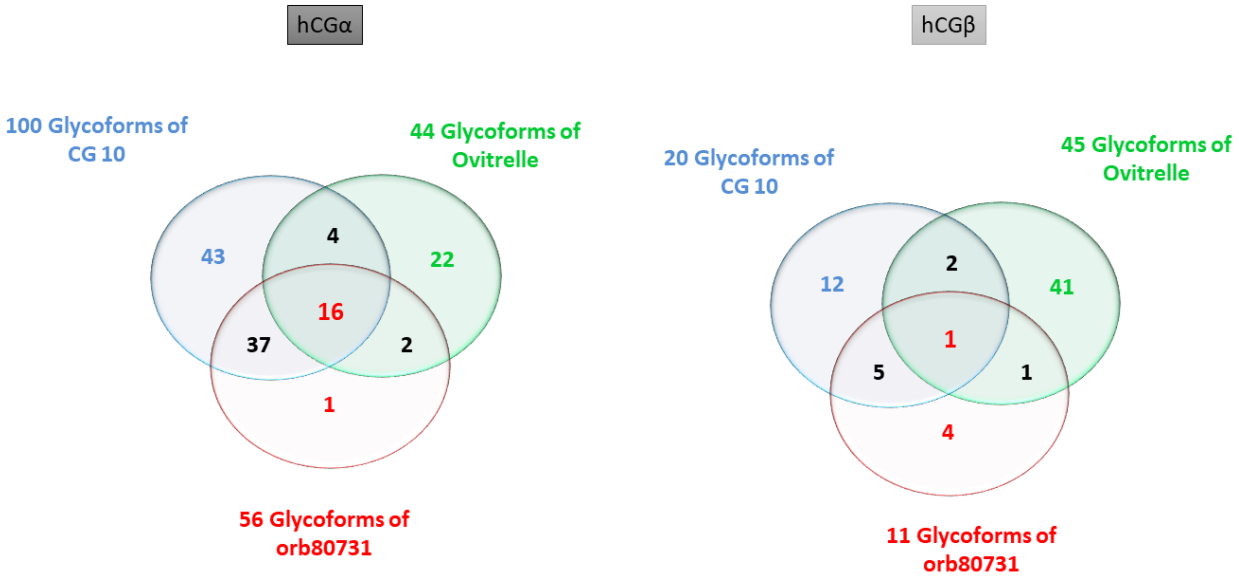


Figure 32. Venn diagrams comparing the total number of hCGα and hCGβ glycoforms detected after the analysis of 3 hCG-based samples (CG10, orb80731 and Ovitrelle® (rhCG)) by nanoLC-MS with both methods using 2% FA or 0.05% TFA in the mobile phase. Other conditions: see section 2.3.

5.3.2. TFA vs FA as mobile phase additive for gonadotropins other than hCG

The potential of the first method that we developed for the analysis of hCGα glycoforms by nanoLC-MS using 2% FA as acidic modifier in the mobile phase and without trapping precolumn **[107]** was evaluated for another gonadotropin, FSH, and allowed the detection of FSHα glycoforms. This was not surprising as all the gonadotropins have similar alpha subunit, with the same 92 amino acids-based sequence and 2 N-glycosylation sites (MW without glycan of 10,188 Da). However, they differ by their beta subunit, in terms of amino acid sequence and number of N- and O-glycosylation sites. Indeed, if hCGβ has 145 AAs (MW without glycan of 15,509 Da), 2 N- and 4 O-gycosylation sites, FSHβ has 111 AAs and 2 N-glycosylation sites (MW without glycan of 12485.1 Da), and LHβ 120 AAs and 1 N-glycosylation site (MW without glycan of 13202.5 Da). It means that FSHα and FSHβ glycoforms have MWs that are

less different than the glycoforms of the two hCG subunits, and it is also the case for LHα and LHβ glycoforms. Therefore, for these two gonadotropins, the potential of the nanoLC-MS methods involving the trapping precolumn and either 2% FA or 0.05% TFA in the mobile phase gradient previously optimized for hCG glycoforms [118] was for the first time evaluated for the characterization of the FSH and LH glycoforms.

For this, two recombinant drugs, Puregon® and Luveris®, derived from CHO cell lines as Ovitrelle® and based on FSH and LH, respectively, were analyzed. The typical resulting BPCs are represented in **Figure 31 (d, e, j and k)**. The profiles are again very different from one sample to another one, keeping constant the analytical conditions. This confirmed the strong differences in terms of glycoform mixture.

All detected LH glycoforms and their corresponding average retention time and average relative area and their RDS values are compiled in **Table S13** in the supplementary data and **Table 11** presents the number of detected glycoforms and their MW and retention time ranges. First of all, the MW range starts with a MW value equal to 13,724.8 Da, which could correspond either to an LHα or LHβ glycoform. Indeed, contrary to hCG, the MWs of the two subunits are close and cannot be distinguished according to this single criterion. This is why, in this case, only LH glycoforms are discussed.

The analyses of Luveris® led to the detection of 58 and 71 glycoforms with TFA and FA, respectively, with a total of 108 different glycoforms, with only 21 were in common, as illustrated in **Figure 33**. Thereby, both methods are here complementary for the analysis of this LH-based drug. As previously observed, TFA induced an increase in retention time, and also in MW and retention time ranges. Moreover, with TFA, two elution zones appeared: a first one between 24.2 and 24.4 min with glycoforms between 13,724.8 and 15,254.3 Da, and a second

one between 29.6 and 31.3 min with glycoforms between 14,982.2 and 16,302.6 Da, thus with a higher average MW. This may correspond to the separation between LHα and LHβ glycoforms, as observed for hCG. Indeed, even if LHα and LHβ glycoforms have about the same MWs, they should be quite different from a physicochemical point of view as LHα has 2 N-glycosylation sites but only 92 AAs and LHβ only 1 N-glycosylation site but 120 AAs.

With regards to the analysis of Puregon®, the FSH-based drug, all the detected glycoforms and their corresponding average retention times and average relative areas and their RDS values are compiled in **Table S14** in the supplementary data. **Table 11** gives the number of detected glycoforms with the corresponding MW and retention time ranges using 2% FA or 0.05% TFA. Surprisingly, with this FSH sample, results that are the opposite of what had been observed so far: TFA led to shorter retention times and to only one elution area (between 17.3 and 17.6 min and 13,562.8-15,703.6 Da MWs) whereas FA led to higher retention and two distinct elution areas ((i) between 16.9 and 20.8 min with 13,271.7-16,475.1 Da MWs and (ii) between 24.4 and 28.3 min with 18,351.8-19,301.1 Da MWs). Regarding the MW ranges and in comparison with the MW range observed for hCGα (12.1-15.8 kDa), these two elution zones may correspond to FSHα and FSHβ glycoforms, respectively, as the two subunits have both 2 N-glycosylation sites with either 92 or 111 AAs. With this hypothesis, the analyses of Puregon® led to the detection of 39 and 64 FSHα glycoforms with TFA and FA, respectively, with most of them (34) being common (see **Figure 33**). Moreover, with FA, 16 additional FSHβ glycoforms were detected. Thereby, these results led to the conclusion that FA should be preferred to TFA for the analysis of FSH glycoforms at the intact level if only one method can be performed. Of course, this should be confirmed by the analysis of other FSH samples.

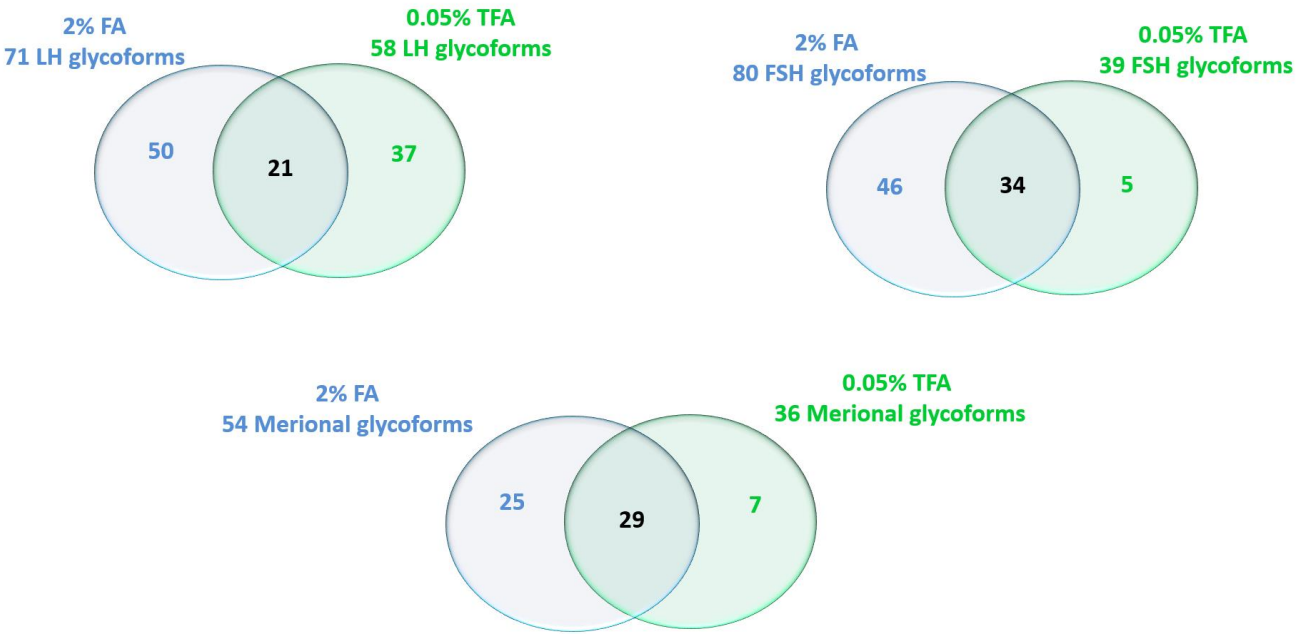


Figure 33. Venn diagrams comparing the number of glycoforms detected after the analysis of Luveris® (LH-based sample), Puregon® (FSH-based sample), and Merional® (LH- and FSH-based sample) by nanoLC-MS according to the nature and content of acidic additive in the mobile phase. Other conditions: see section 2.3.

Next, the analysis of Merional®, a highly purified human menopausal gonadotropin preparation containing both FSH and LH, took place. The resulting BPC is presented **Figure 31 (f and l)**. A profile different from the ones obtained with the other samples was again obtained, keeping constant the analytical conditions. One more time, this indicates that the samples are composed of different glycoform mixtures, even if they contain all gonadotropins sharing the same alpha subunit structure.

As for the other drugs, all detected glycoforms and their detils are compiled in **Table S15** in the supplementary data. The number of detected glycoform and their MW and retention time ranges are given in **Table 11**. The MW range between 14,598.1 and 16,160.7 Da may correspond to alpha glycoforms only. It is important to note that separation between FSH and LH glycoforms in this sample is impossible as they share the same alpha subunit. It is just possible to say that the analysis with TFA and FA led to the detection of 36 and 54 glycoforms, respectively, with a total

number of 61 since 25 and 7 additional glycoforms were detected using FA or TFA, respectively (see **Figure 33**). Both methods are therefore complementary.

In summary, if for hCG TFA is highly recommended to favor the separation and thus the ionization of the highly glycosylated hCGβ glycoforms from the hCGα ones, for other gonadotropins less glycosylated and having their two subunits closer in MW and glycosylation state, it seems that the use of both FA and TFA is recommended to increase the number of detected glycoforms as much as possible. Even if the implementation of two mobile phases in nanoLC-MS for one given sample is time-consuming, as it is done on the same set-up and at the intact level that allows a fast and simple sample handling, it still allows a rapid characterization of the glycoforms present in the sample.

5.3.3. Comparison between gonadotropin glycoforms

With the analysis of the different samples, a high number of glycoforms were detected in total: 89 in rhCG, 120 in CG10, 68 in orb80731, 85 in Puregon®, 108 in Luveris®, and 61 in Merional® with both FA and TFA as a mobile phase additive are represented in **Figures 30 and 33**. When comparing all the results of the analyzed samples, it was noticed that only one glycoform was detected in all samples and in both methods with FA and TFA in the mobile phase, the one having a MW of 14,614.1 Da, which confirmed the first impression given by the very different BPC profiles (**Figure 31**). It is quite surprising when it is claimed that the gonadotropins share the same alpha subunit and may be explained by the different production sources and processes of each gonadotropin.

5.4. Conclusion

In this study, it was demonstrated for the first time that according to the nature of the gonadotropin, especially its potential differences in MW and glycosylation state between their alpha and beta subunits, only one or two compositions of mobile phase is required for the characterization of the glycoforms by their analysis at the intact level in nanoLC-MS with an Orbitrap analyzer. The application of this fast and simple method, requiring virtually no sample handling except a filtration-concentration step, should constitute a powerful tool available to the analytical chemists for the characterization of these complex glycoproteins in purified samples. However, its application for the analysis of more complex biological samples should require the development of a selective extraction step, which is the prospect of this work.

Acknowledgements

This work has received the support of PSL University, the Conseil Régional d'Île-de-France with the DIM Analytics, and Horizon 2020, Actions Marie-Sklolowska Curie, Cofund UPtoParis, ESPCI Paris PSL.

Chapter 6: synthesis and first characterization of immunsorbents dedicated to the selective extraction of hCG glycoforms

Immunosorbent synthesis and first evaluations for the extraction of hCG glycoforms

In all analytical procedures, the direct analysis of samples is the ideal situation without any pre-treatment step. However, biological samples require the addition of pre-treatment techniques due to the high complexity of their matrix, which can affect the sensitivity of the measurements. Therefore, various pretreatment methodologies are being utilized, mainly carrying out solid phase extraction (SPE) [122]. Various SPE configurations using conventional sorbents are being adapted, such as C18-bonded silica or non-conventional sorbents such as the polymeric (hydrophilic, hydrophobic or mixed mode) and monolithic materials and carbon-based nanomaterials. However, these sorbents can suffer from a lack of selectivity and lead to the co-extraction of numerous interfering compounds. This is why more selective sorbents can also be used such as immunosorbents (ISs) and molecularly imprinted polymers (MIPs) [122].

Immunosorbents, also called immunoaffinity (IA) sorbents, are based on biological tools, antibodies, and permit the specific molecular recognition of the targeted compounds. Indeed, the interaction between an antibody (Ab) and its antigen (Ag) has a high affinity and is very specific. Therefore, the use of ISs enables the selective extraction and concentration of individual compounds or classes of compounds existing in liquid matrices, in addition to the sample purification of extracts from solid matrices. As selective antibodies for large molecules such as proteins, peptides, antibodies, and viruses are available, immunoextraction is very popular in biological field [123]. In a similar way, numerous antibodies have been commercialized for hCG, allowing the implementation of hCG immunoextraction.

In this chapter, after some generalities about ISs, the state-of-the-art of their use for hCG is presented. Their synthesis and first evaluations carried out to extract all the hCG glycoforms from a sample are also discussed.

6.1.Characteristics of ISs

IS could be characterized by various variables such as the used Abs, solid support, antibody immobilization procedure, bonding density, capacity, specificity, cross-reactivity, and recovery. The preparation of an IS requires the linkage of antibodies to a solid support that is packed in a SPE cartridge and then stored in a wet media, usually phosphate-buffered saline (PBS) [122]. Both monoclonal (mAbs) and polyclonal (pAbs) antibodies can be immobilized. Recently mAbs are being used dominantly regardless of their higher cost due to the fact that they guarantee a long-term availability of reproducible antibodies [23], [124].

Another important parameter is the solid support selected for the immobilization of the antibodies. Several properties of the support must be available such as the chemical, biochemical, and mechanical stability, uniformity in the large particle's size, activation, and hydrophilicity to limit next the non-specific interactions. In addition the immobilization procedure should retain the bio-specific activity of antibodies. The most commonly applied approach for immobilizing antibodies is their covalent bonding by the reaction of the free amino group of the antibody with a support containing a reactive epoxide or aldehyde groups at the surface or groups that could be activated using carbonyldiimidazole, cyanogen bromide or *N*-hydroxysuccinimide. Traditional supports are made of silica, cellulose, agarose, and synthetic polymers. Out of these agarose and silica are the most commonly used sorbents for IS synthesis. Silica based-ISs are preferred for their pressure resistance and the ability to operate at high flow rates. They have been used predominantly for the on-line coupling of immunoextraction to chromatographic separation techniques [123]. The immobilization of antibodies to activated agarose- and silica-based supports is considered random because it includes covalent coupling via lysine amino groups, which are encountered throughout antibody molecules, which leads to several different antibody orientations [122].

One of the key parameters when synthesizing an IS is the bonding density, which is the number of antibodies linked to the surface per volume or mass of the sorbent and thus expressed in mg/ml or mg/g of sorbent. It can be measured experimentally, as an example by measuring the concentration of the antibodies in the binding solution by spectrophotometry, and determines the future antigen-binding capacity [23], [122]. It depends on the specific surface area of the solid support accessible for the immobilization of the antibodies. Supports could have small pore sizes with a high surface area, but this leads to low accessibility of the large antibodies, or they could

have large pore sizes with good accessibility but low surface area. The choice of the type of support to be used must be a compromise between the two factors and this were discussed in a study by Clarke et al in 2000 [125] that showed 30-50 nm pores were the best for the immobilization of immunoglobulin G (IgG) or Fab fragments due to the fact that the diameter of an antibody is 8-10 nm expecting that this could be applied to other proteins that are to be immobilized to HPLC supports. The average bonding density is about 5 mg/ml for agarose gel and 20 mg/g of silica independent whether mAbs or pAbs are being used [123].

Another key parameter of an IS is its capacity, which corresponds to the total number of accessible immobilized antibodies which cannot be calculated directly due to the random orientation and steric hindrance that might block the access of the analyte to the specific complementary-determining region of the antibody. In general capacities are always in the range of the hundreds of ng to some mg for 1 g of sorbent [122], [123].

Another parameter is the IS cross reactivity that can be defined by the ability of the antibodies to bind one or more analytes with a structure similar to the analyte that has induced the immune response. It is usually considered as a drawback but, fortunately it is avoided by the separation step that follows the immunoextraction allowing an individual quantification of the trapped analytes during the detection step [123].

Finally, depending on the previously discussed parameters and on the extraction protocol, extraction recoveries of the targeted compounds can be determined targeting a 100% value as long as the capacity but also the breakthrough volume has not been exceeded. The breakthrough volume depends on the affinity between the analytes and the antibody.

6.2.Immunoextraction procedure

An immunoextraction procedure is divided into 3 main steps summarized in **figure 34**. The first percolation step aims for the introduction of the sample that includes the analytes of interest in addition to other compounds present in the sample mixture. Next a washing step is crucial for the removal of interfering compounds. The last step is elution of the analytes. The ideal solution to be used for the elution step must disrupt the analyte-antibody interactions without adversely affecting the immobilized Abs. The most common elution strategies are the use of chaotropic agents, the reduction of the pH to 2-3 and/or the use of a water-miscible organic solvent mixture. The IS regeneration and reuse is possible by percolation of PBS (with sodium azide) and storage at 4 °C for a couple of days. Neverthelees, a decrease in capacity is often observed with time and number of uses [123].

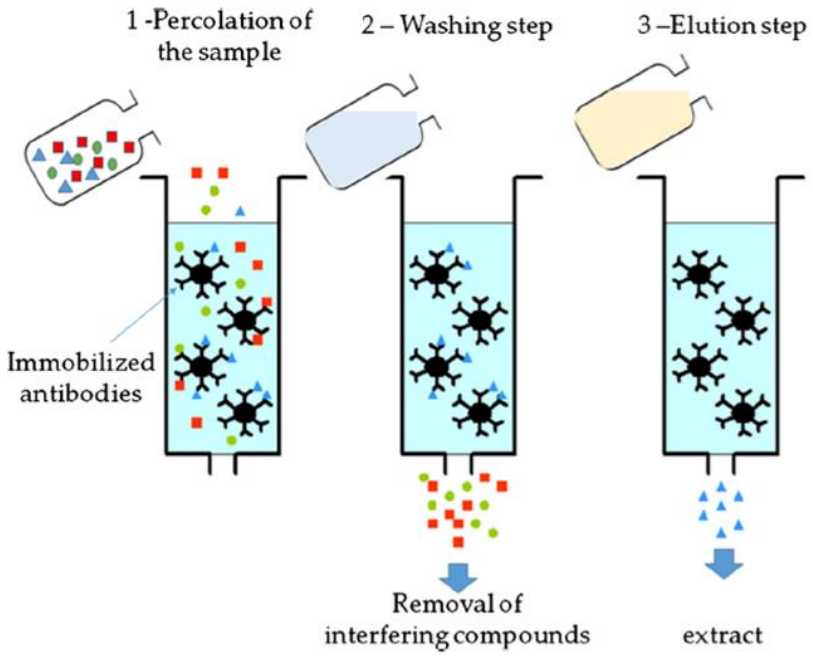


Figure 34. Immunoextraction procedure. Modified from [126].

6.3.Immunoextraction of hCG

Various studies have been previously conducted aiming for the implementation of immunoextraction technique for the analysis of hCG in complex biological samples and are described in **Table 12**. These studies mainly detailed the developed analytical method. Two solid supports grafted with antibodies were mainly use, CNBr-activated Sepharose (an agarose gel) packed between frits in a disposable cartridge and magnetic beads dispersed in samples. Immunoextraction using Ab-magnetic beads was adopted by the majority of the studies aiming to extract hCG. These studies applied very similar immunoextraction protocol represented in **figure 35** which shows dual immunoextraction protocol that was selected to extract first intact hCG and next hCGβ variant.

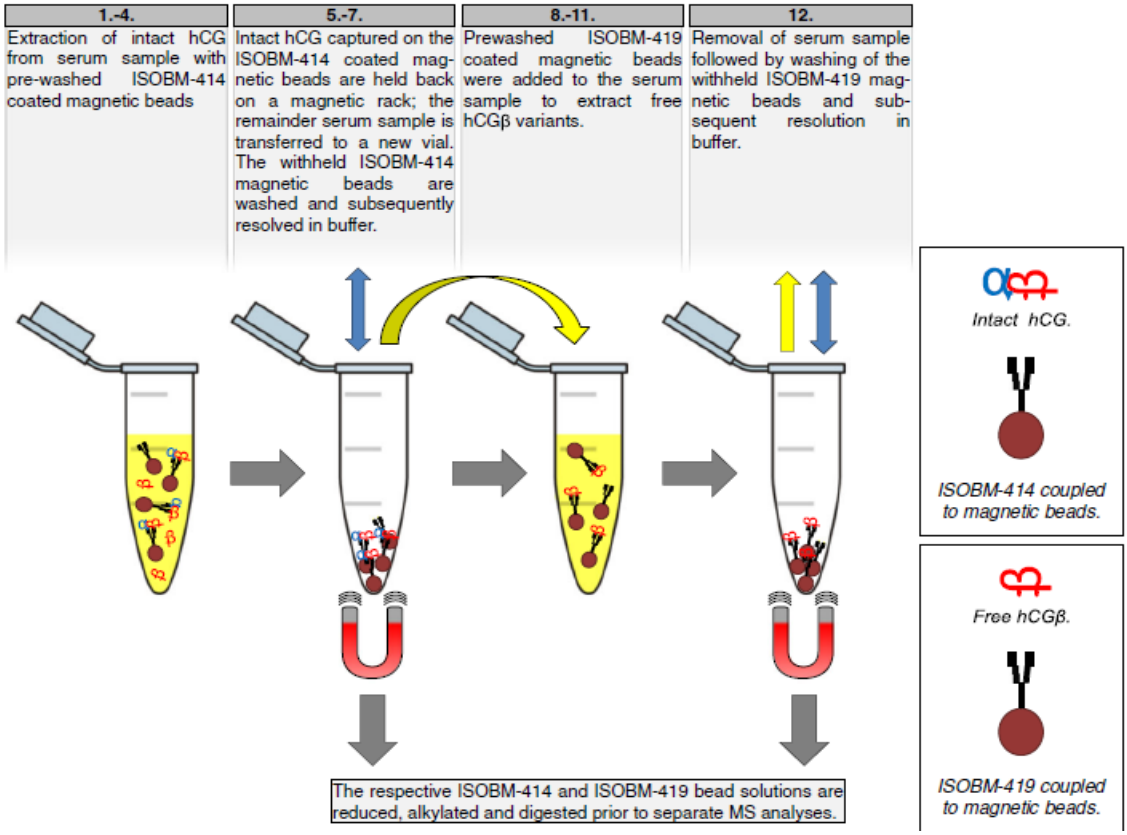


Figure 35. Dual-immunoextraction method applied by Egelant et al 2016 **[65]**.

It is important to note that these studies succeeded in the extraction of hCG and some variants: hCGβ, hCGn 47/48, hCGn 44/45, free hCGβn 44/45, free hCGβn 47/48, hCGβ core fragment (hCGβcf). They were determined with the bottom-up approach after tryptic digestion, which was the common analytical method applied by all. In some cases, hCG was digested directly in the epindorf tubes, which means that the ISs are no longer usable because the antibodies are also digested and therefore antibody peptides are also generated, which makes the analysis of hCG more complex. Therefore, none of these studies were aiming for the analysis of hCG glycoforms present in the tested samples and none targeted the analysis of hCG at the intact level. In addition, the use of CNBr-activated Sepharose based immunosorbents packed in SPE cartridge was not clearly studied and optimized for immunoextraction of hCG despite their interest, by carrying out step elutions, to evaluate the affinity of different antigen for the antibodies. For all of the mentioned reasons in this manuscript, we were interested in the development and optimization of an immunoextraction procedure based on the use of anti-hCG Abs, grafted on Sepharose in order to extract all the hCG glycoforms, before their analysis at the intact level using the previously optimized nanoLC-(Orbitrap)MS method. Sepharose based immunosorbents were selected due to the fact that they could be coupled online in the future as along term perspective for the analysis of hCG. Therfore CNBr-activated Sepharose based immunosorbents were preferred over the use of magnetic beads.

Different types of anti-hCG antibodies were used in the different studies, but no indication was given about their ability to extract all the hCG glycoforms. Therefore, we decided to select two antibodies, an anti-hCGα and an anti-hCGβ, and compare their potentials in selectively extracting the hCG glycoforms.

Table 12 Summary of the various studies implementing the immunoextraction of hCG.

Sample	Ab used	Solid sorbent	Known IS details	IS extraction protocol	Recovery	Analytical process	Objective results	Ref
Urine spiked with pregnyl®	mAb B108	Ab1 or Sepharose-CNBr beads	Ca: 200 IU/ml of beads K: 0.132 nM	100 µl of beads in10 ml of urine, 2 h at room temp Settling and washing with 1 ml of PBS Elution with 0.4 ml of 6 M guanidine HC1	98-99%	IS-dSPE + reduction-alkylation + tryptic digestion + LC-MS/MS	LOQ: 25 IU/l (0.07 pM) of intact hCG	[54]
urine from cancer patients and pregnant woman, medium from a choriocarcinoma cell line, pregnyl®	mAb 6G5, mAb B152, mAb 9C11-IS	Sepharose-CNBr	-	1 ml of urine or 5 ml of serum on 1 ml of mAb6G5-IS Washing with 40 mL of 0.1 M ammonium acetate (pH 4.5) Elution with 15 mL of 3 M acetic acid. free hCGβ from urine on mAb 9C11-affinity column hCGh by immunoaffinity chromatography with MAb B152	-	IS extraction or immunoaffinity extraction + reduction-alkylation + tryptic digestion digestion + LC-MS/MS	3 different processes of purifying hCG, differences in the site-specific glycan structure of hCGβ from pregnancy or tumors	[73]
Pregnyl® spiked serum from cancer patients and urine from	mAb E27	96-well plates	0.5 μg of Abs/well Ca: 10 IU/mL (1.1 ng/l)	200 μL sample per well, 1 h washing with PBS pH 7.4 + 0.05% Tween 20, 2x with PBS, and 10	-	Immunocapture + in-well tryptic digestion + LC-MS/MS	E27 able to extract hCG (intact, hCGh, free), hCGβcf, hCGβn 44/45, hCGβn 47/48 in	[59]

pregnant women				mM Tris-HCl pH 7.4 Elution with ACN + 0.1% TFA			serum and urine LOD: 100 IU/l (0.29 nM)	
Serum of cancer patients, serum and urine of pregnant women, serum and urine of one subject after injection of pregnyl®	mAb E27	Magnetic beads	15 µg of Ab/mg of beads	20 µl of beads in 1 mL of sample washing with 500 µL of PBS + 0.05% Tween 20; 500 µL of PBS; 500 µL of 10 mM Tris−HCl pH 7.4	-	IS-dSPE + tryptic digestion on beads + SPE + LC-MS/MS	quantitation of total hCG LOD: 2 and 5 IU/L (5.8-14.5 pM) for urine and serum respectively	[23]
urine and serum after injection of pregnyl® or ovitrelle® to men	mAb E27	Magnetic beads	15 µg of Ab/mg of beads	20 µl of beads in 1 mL of sample washing with 500 µL of PBS + 0.05% Tween 20; 500 µL of PBS; 500 µL of 10 mM Tris−HCl pH 7.4	-	IS-dSPE + tryptic digestion on beads + SPE + LC-MS/MS	E27 able to extract hCG intact, free hCGβ, hCGn 47/48, hCGn 44/45, hCGβcf, free hCGβn 44/45, free hCGβn 47/48, hCGβ core fragment	[60]
six 1st International Reference Reagents (IRRs) for hCG	30 mAbs	Tosyl-activated magnetic beads	15 µg of Ab/mg of beads	20 µl of beads in 1 mL of sample washing with PBS + 0.005% Tween 20 (500 µL); PBS (500 µL); 10 mM Tris–HCl pH 7.4 (500 µL)	-	IS-dSPE + tryptic digestion on beads + LC-MS/MS	Characterization of 30 mAbs for recognition of intact hCG, hCGn,hCGβ, hCGβcf, hCGβn, and hCGα	[61]

Male urine after the administration of Novarel® and Ovidrel®	β7 (INN-hCG-68) mAb, β1 (INN-hCG-2) mAb	Carboxylate-activated magnetic beads	12 µg of Ab/mg of beads	0.5 mg of β7 - beads in 1 ml of sample recognizing hCGβ and hCGβcf (intact hCG in supernatant) washing with 1 mL of 10 mM Tris-HCl pH 7.4 elution with 15 µl of 0.1 M glycine + 1 mML EDTA pH 2.55 the supernatant in β1-beads followed by intact hCG elution	91-109%	IS-dSPE + tryptic digestion + LC-MS/MS	Quantification and differentiation between intact hCG and free hCGβ LOQ: 0.2 IU/L (0.58 pM)	[62]
Spiked serum with pregnyl® or ovitrelle® of cancer patients	9 mAbs (ISOBM :387, 389, 411, 414, 419, 425, 436, 446, 447)	Magnetic beads	15 µg of Ab/mg of beads Ca: 6 nM in serum	40 µl of ISOMB-414-beads in 1 ml of sample Washing with 500 µL PBS + 0.05 % Tween 20, 500 µL PBS, 500 µL of 10 mM Tris-HCl pH 7.4 Addition of 50 mM ABC + 0.001 % PEG and 20 µL of ISOMB-419-beads Washing with 500 µL PBS + 0.05 % Tween 20, 500 µL PBS, 500 µL of 10 mM Tris-HCl	34%	Dual-IS-dSPE + tryptic digestion + LC-MS/MS	Differentiation between hCG and free subunits Differentiation between hCG variants (not glycoforms)	[66]

				pH 7.4 Addition of 50 mM ABC + 0.001 % PEG				
Urine after injection of Ovidrel® to men	β7 (INN-hCG-68) mAb, β1 (INN-hCG-2) mAb	Magnetic beads	-	0.5 mg of β7-beads in 1 ml of sample recognizing hCGβ and hCGβcf (intact hCG in supernatant) washing with 1 mL of 10 mM Tris-HCl pH 7.4 Elution of hCGβ and hCGβcf with 15 µl of 0.1 M glycine + 1 mML EDTA pH 2.55 the supernatant in β1-beads followed by intact hCG elution	-	Dual-IS-dSPE + tryptic digestion + LC-MS/MS	A proof that total hCG immunoassays significantly overestimate hCG concentrations and can produce false positive results.	[19]
Ovitrelle® spiked serum	mAb E146, mAb E27	Magnetic beads	-	0.5 mg of Ab-beads in 1 ml of sample Several washes were followed by elution with 15 µL of either 2% FA or 0.1% TFA for the samples to be digested after elution. While 2 µL of FA (100%) or TFA (5%) was added to the samples after	80%	IS-dSPE + tryptic digestion (prior or after elution) + nanoLC-MS/MS	sequence coverage and number of identified peptides of both subunits similar or higher after on bead digestion No clear answer to whether or not the protein should be eluted from the antibody prior to	[66]

				ended on-bead digestion			digestion	
Dried spots of Pregnyl® - hCG spiked blood serum, plasma and urine	E27	Magnetic beads	-	0.5 mg of E27-beads in 1 ml of sample washing with PBS + Tween 20, PBS and Tris-HCL	-	IS-dSPE + Immunoaffinity cleanup + tryptic digestion+ nanoLC-MS/MS	Determination of total hCG content LOD: 2,1-10,5 IU/l (=6,1-30,5 pM)	[67]
Urine	β7 (INN-hCG-68) mAb, β1 (INN-hCG-2) mAb	magnetic beads	-	0.5 mg of β7-beads in 1 ml of sample recognizing hCGβ and hCGβcf (intact hCG in supernatant) washing with 1 mL of 10 mM Tris-HCl pH 7.4 elution with 15 µl of 0.1 M glycine + 1 mM EDTA pH 2.55 the supernatant in β1-beads followed by intact hCG elution	-	Dual-IS-dSPE + tryptic digestion + LC-MS/MS	Mean concentrations of intact hCG, hCGβ and hCGβcf: 0.04 IU/L, 0.47 pmol/L and 0.16 pmol/L, upper reference limits: 0.21 IU/L, 0.40 pmol/L, and 1.86 pmol/L, respectively	[11]

ABC, ammonium bicarbonate; Abs, antibodies; ACN, acetonitrile; Ca, capacity; CNBr, Cyanogen bromide; EDTA, Ethylenediaminetetraacetic acid; hCGβcf, human chorionic gonadotropin beta core fragment; hCG*β*n human chorionic gonadotropin beta nicked; HCl, Hydrochloric acid; hCG, human chorionic gonadotropin;hCGh, hyperglycosylated human chorionic gonadotropin; IS, immunosorbent; IS-dSPE, immunosorbent-dispersive solid phase extraction; K, affinity constant; PBS, phosphate buffer saline; PEG, polyethylene glycolhCGh hyperglycosylated; LOD, limit of detection; LOQ, limit of quantification; LC-MS/MS, liquid chromatography tandem mass spectrometry; mAb, monoclonal antibody; SPE, solid phase extraction; TFA, Trifluoroacetic acid.

6.4.Complementary Results

As discussed before the use of immunoextraction prior to the analytical separation and detection method of hCG is a promising tool for the identification of hCG glycoforms present in complex biological samples. Therefore, in this last part, first experiments related to the preparation of anti-hCG ISs and their use for the trapping of hCG glycoforms followed by nanoLC-MS analysis are presented.

6.4.1. Materials and reagents

HPLC-grade acetonitrile (ACN), trifluoroacetic acid (TFA) and LC-MS-grade water were supplied by Carlo Erba (Val de Reuil, France). Cyanogen bromide-activated-Sepharose®, Sodium acetate, disodium hydrogenophosphate, dipotassium hydrogenophosphate, Trizma hydrochloride, Trizma base, sodium azide, hydrochloric acid (2 M), and sodium chloride were purchased from Sigma. Potassium dihydrogenophosphate, sodium bicarbonate and formic acid (FA) (99%) were purchased from VWR (Fontenay-sous-Bois, France). A0231 anti-hCGβ pAbs were supplied from Dako (Denmark) and HT13.3 anti-hCGα mAbs were kindly provided by Yves Combarnou (Director of: Physiologie de la Reproduction et des Comportements). Ovitrelle® (Organon, Oss, The Netherlands) is a drug made of recombinant hCG from a CHO cell line. rhCG contains 250 µg of rhCG suspended in 500 µL of suspension liquid containing: mannitol, methionine, Poloxamer 188, phosphoric acid (for pH adjustment), sodium hydroxide (for pH adjustment) and water for injections.

6.4.2. Sample preparation

The rhCG drug stock solution was desalted and filtered for 10 min at 4°C with ultra-centrifugal units (Microcon) (Merck, Darmstadt, Germany) with a MW cut-off of 10 kDa. This step aims for the removal of additives and excipients from the formulation. Next, the residues were washed with water and finally the samples were recovered from the inverted membrane by centrifugation for 4 min at 14,000 g (≈ 11,000 rpm) and the volumes were adjusted with water to obtain a final concentration of 500 µg mL^{-1} of hCG.

6.4.3. Immunosorbent synthesis

The immobilization of the antibodies on Sepharose-CNBr was adapted from the procedure already used in our laboratory [127]. Briefly, 35 mg of Sepharose-CNBr were swollen in 1 mL of an acid solution (HCl, 1 mM) during 15 min and washed two times with 1 mL of a mixture of $NaHCO_3$ (0.1 M) and NaCl (0.5 M) at pH 8.3. 100 µg of anti-hCG antibodies contained in 100 µL of PBS were then added. Antibodies were incubated during 24 h. The resulting sorbent was packed into 1 ml disposable cartridge. This process was applied two times and two ISs were prepared, one using HT13.3 mAbs and the second using A0231 pAbs. The same process was done for the preparation of two new immunosorbents, one using HT13.3 mAbs and the second using A0231 pAbs. The only difference was doubling the amount of Sepharose-CNBr and Abs used.

6.4.4. Quantification of antibodies grafted on Sepharose

The incubation of Abs with Sepharose resulted in the solid phase to which Abs are bound and the supernatant containing unbound Abs. A bicinchoninic acid assay (BCA) was performed to

quantify the antibodies remaining in the supernatant to evaluate the amount of antibodies immobilized on Sepharose. Three calibration standards of rhCG in water were prepared at three concentrations (1, 0.5, and 0.25 mg/ml). Next, 10 µL of each standard solution and the supernatant were added in triplicate (3 wells) and 100 µL of working reagent (mix of bicinchoninic acid and Cu^{2+}) was added in each microwell. The microplate was incubated at 37°C for 30 min with an agitation at 300 rpm. Absorbance was measured at 562 nm with a spectrophotometer (SpectraMax M2, Molecular Devices, St Gregoire, France).

6.4.5. Imunoextraction procedure

The percolation of 10 µg in the first trial and 5 µg in the second trial of hCG in water after a conditioning step (4 ml PBS) took place. The washing of the IS with 1 ml PBS and the elution with 1 ml FA 0.5 M pH 2.2 were performed. Finally, the ISs were stored in a wet medium (PBS). To prevent microbial contamination, sodium azide, a bacteriostatic agent, was added in the storage buffer. Each fraction, percolation, wash, elution and regeneration, was collected, desalted and concentrated using ultra-centrifugal units (Microcon, Merck, Darmstadt, Germany) with a MW cut-off of 10 kDa.

6.4.6. NanoLC-MS analysis

The analyses of each collected fraction were carried out using a nano-LC system (NC 3500 RS, ThermoFisher Scientific, Courtaboeuf, France) coupled to an Orbitrap mass analyzer (Exactive plus, ThermoFisher Scientific). A C18 PepMap precolum (0.3 x 5 mm internal diameter, 300 Å) packed with 5 µm particles (ThermoFisher Scientific) was used, while the column was an EASY-Spray PepMap RSLC C18 (150 x 0.075 mm internal diameter, 300 Å) packed with 5 µm particles (ThermoFisher Scientific). The mobile phase was composed of (A) H_2O/ACN (98/2;

v/v) and (B) H_2O/ACN (1/9; v/v) acidified with 0.05 % TFA v/v. The column temperature was set at 60°C. The injection volume was 1 µl in partial loop mode. The flow-rate was set at 300 nL min^{-1}. The gradient was 10 to 50% ACN in 39 min, followed by a rapid increase to 90% of ACN within 5 min, then a 5 min plateau, and finally a return to the initial 10% of ACN concentration and kept for 15 min for equilibration. The EASY-spray source parameters are as follows: capillary voltage, 1.7 kV; drying temperature, 280°C; and sheath, auxiliary, and sweep gas flows were set at 0. The mass spectrometer was operated in the positive full MS mode and using a scan range from 1000 to 2500 m/z with a scan rate of 1 spectrum/s. The focusing voltages were tuned on weekly bases by the transmission of positive and negative ion calibration solutions (ThermoFisher Scientific). The mass resolution was similarly set to 140,000 at 200 m/z, the AGC target was 3e6, and the maximum injection time was 200 ms. The deconvolution analyses were processed using Freestyle 1.3 SP2 (ThermoFisher Scientific) with deconvolution parameters that uses a mass tolerance for m/z of 30 mDa, and a minimal charge state of 5, with at least 3 different charge states to identify a compound, accuracy mass tolerance for deconvoluted MW was set at 300 mDa.

6.5.Results and discussion

6.5.1. Preparation of ISs

The choice of the immobilization support is a key parameter for antibody grafting. In our case, CNBr-activated Sepharose, which has been widely used for antibody immobilization, was selected for its low hydrophobic properties and the efficiency of the coupling of biomolecules it offers. Two anti-hCG antibodies were chosen to compare their potential for the selective extraction of all the hCG glycoforms. Therefore, the first selected Ab was A0231, an anti-hCGβ pAb, and the second was HT13.3, an anti-hCGα mAb. An important aspect when selecting the

antibodies was their ability to interact with all hCG glycoforms which requires that the antigenic determinents of these antibodies to reside far from the oligosaccharide part of hCG. This was proven for HT13.3 indicating that its antigenic determinants do not reside on the oligosaccharide part of the a-subunit [128].

The IS synthesis consisted in the covalent grafting of 100 and 200 µg of polyclonal or monoclonal anti-hCG antibodies to 35 and 70 mg of CNBr-activated Sepharose, respectively, that was next packed between frits into 1 ml SPE cartridge. The procedure conventionally used in the LSABM was selected, as it was observed that these conditions led to good grafting yields [126], [127], [129]. Indeed, excellent grafting yields (percentage of grfted/bounded antibodies) were obtained for all the anti-hCG ISs synthesized. The measured grafting yields were 86% and 98% for the two anti-hCGβ pAb-based ISs, and 96% and 97% for the two anti-hCGα mAb-based ISs.

The theoretical capacities of each IS were calculated by considering that 100% of the grafted mAbs are accessible and specific, but that it is the case for only 10% for the grafted pAbs. This proportion is commonly used for pAbs because pAbs correspondond to a mixture of antibodies that are not all specific of hCG [123] . This means that the capacity was 0.06 and 0.07 µg of hCG per mg of sorbent for the two anti-hCGβ pAb-based ISs, leading to a capacity value of 2.0 µg and 4.6 µg of hCG per anti-hCGβ SPE cartridge. The capacity values were of 0.64 and 0.65 of hCG per mg of sorbent for for the anti-hCGα ISs, leading to a capacity value of 22.4 µg and 45.3 µg of hCG per anti-hCGα SPE cartridge. These theorical capacity values were very low compared to the minimal hCG amount necessary to test the ISs, due to the sensitivity of the nanoLC-MS method for the quantification of all the hCG glycoforms, especially for those giving low signal intensities, i.e. some of the glycoforms hCGα and all the hCGβ. Indeed, even if each

fraction obtained during the immunoextraction process would be concentrated before its injection in the nanoLC-MS system, about 5-10 µg of hCG were needed to be percolated. This lack of capacity was essentially true for the pAbs-based ISs, less for the mAbs-based ISs. Unfortunately, the amounts of available antibodies were too low at that time and prevented us from additional IS synthesis to try to increase the grafting density. Keeping this limitation in mind, first evaluations of IS potential to extract hCG glycoforms were done.

6.5.2. First evaluations of the IS potential to extract hCG glycoforms

The potential of the ISs for the extraction of hCG glycoforms was first evaluated by using a conventional immunoextraction protocol, involving the percolation of rhCG in water, followed by a washing with PBS, and an elution with a strongly acidic solution (0.5 M FA, pH 2.2). For each experiment, percolation, washing, elution, fractions were collected, filtered and concentrated by a factor 4 using a ultra-centrifugal units (Microcon). The resulting concentrated extract was analyzed by the nanoLC-(Orbitrap)MS method developed before [118]. In parallel, rhCG at 50 µg ml^{-1} was analyzed with the same method to act as a reference standard to be used for the quantification of the extraction recovery.

Using 10 µg of hCG and 35 mg of grafted ISs, all the hCG glycoforms (α and β) detected in our previous work [118] were detected in the percolation, wash and elution fractions of anti-hCGβ pAbs-based IS while all hCG glycoforms were detected in the wash and elution fraction of anti-hCGα mAbs-based IS. These results were expected since they agree with the calculated theoretical capacity values, with an experimental capacity definitely too low for the pAbs-based IS. A comparison of the relative area of the high intensity hCGα glycoforms, normalized by the area of the glycoform revealing the highest area value, in the mentioned fractions is represented

in **figures 36 and 37** for anti-hCGβ pAbs- and anti-hCGα mAbs-based ISs, respectively. In both cases, the relative proportions of the glycoforms in the washing and elution fractions differ and could reflect differences in affinity between the glycoforms and the grafted antibody.

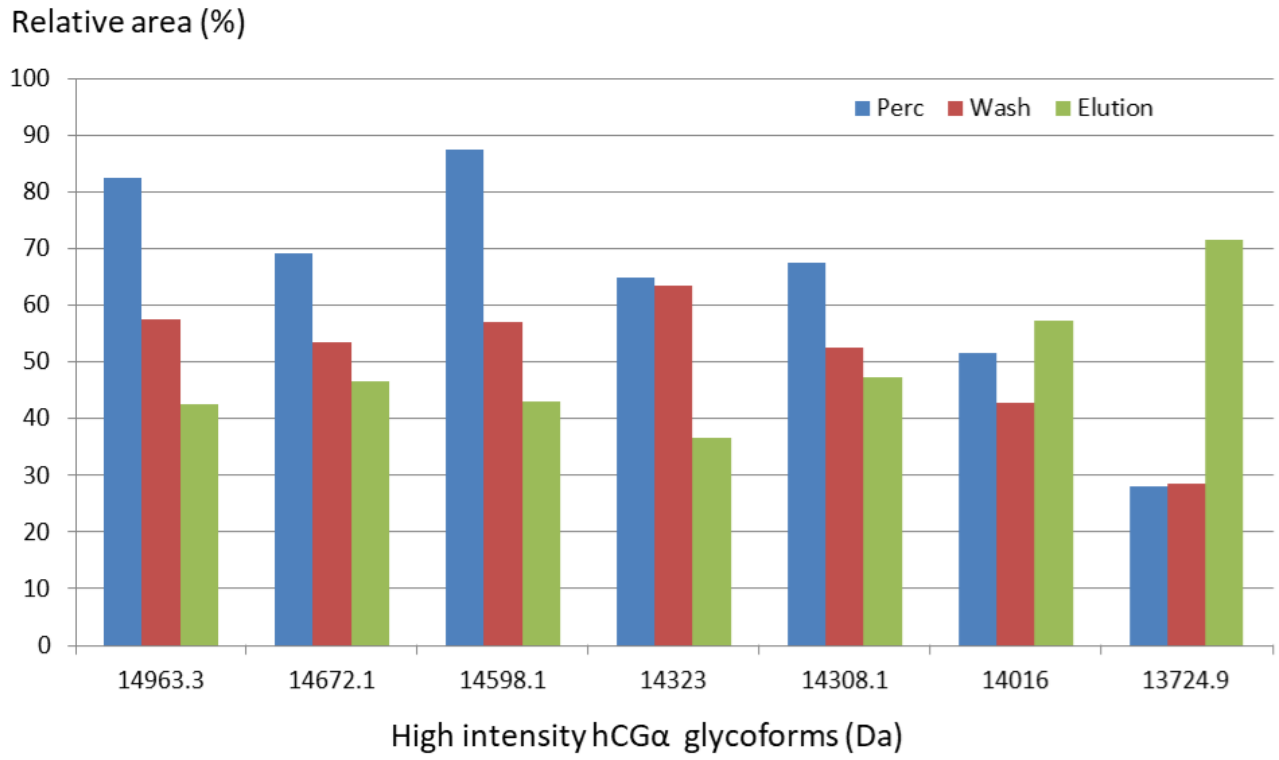


Figure 36. Relative areas of the high intensity hCGα glycoforms in the immunoextracted fractions obtained with the anti-hCGβ pAbs-based IS. Percolation of 10 µg of rhCG on 35 mg of grafted sorbent.

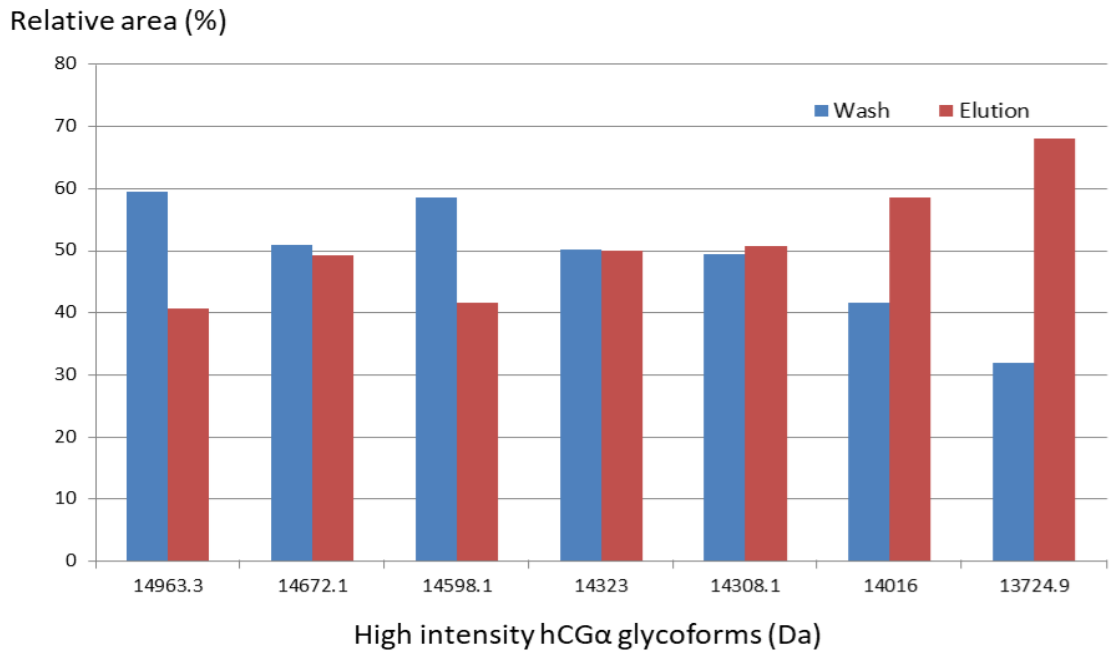


Figure 37. Relative areas of the high intensity hCGα glycoforms in the immunoextracted fractions obtained with the anti-hCGα mAbs-based IS. Percolation of 10 µg of rhCG on 35 mg of grafted sorbent.

In a second immunoextration experiment, the amount of the grafted sorbent was doubled from 35 to 70 mg on one hand, and on the other the amount of hCG was reduced from 10 to 5 µg, allowing both to try to prevent from capacity issues. In this case, the analysis of the percolation, washing, and regeneration fractions showed no presence of hCG glycoforms, while all the hCG glycoforms detected in our previous work [118] were detected in the elution fraction of both immunosorbents. This confirmed that capacity was previously the issue when percolating 10 µg of only 35 mg of grafted sorbent.

The extraction recovery was calculated for each glycoform with its peak area values measured on the XICs obtained after injecting the concentrated elution fraction and rhCG at 50 µg ml^{-1} and taking into account a correction factor. For the elution fraction obtained with the anti-hCGβ pAbs-based IS, the extraction recovery of hCGα glycoforms has an average of 53% [24-92% range], while that of hCGβ glycoforms has an average of 43% [35-59% range]. For the anti-

hCGα mAbs-based IS, the extraction recovery of hCGα glycoforms has an average of 30% [10-71 range], while that of hCGβ glycoforms has an average of 21% [15-35% range]. Checking high intensity hCGα glycoforms separately showed that with the anti-hCGβ pAbs-based IS, the extraction recovery of these glycoforms has an average of 29% [16-53] while with the anti-hCGα mAbs-based IS, the extraction recovery was an average of 51% [36-73]. These extraction recoveries are represented in **figures 38** and **37** showing the extraction recoveries of all hCGα and hCGβ glycoforms respectively.

Extraction recovery determined in the elution fraction (%)

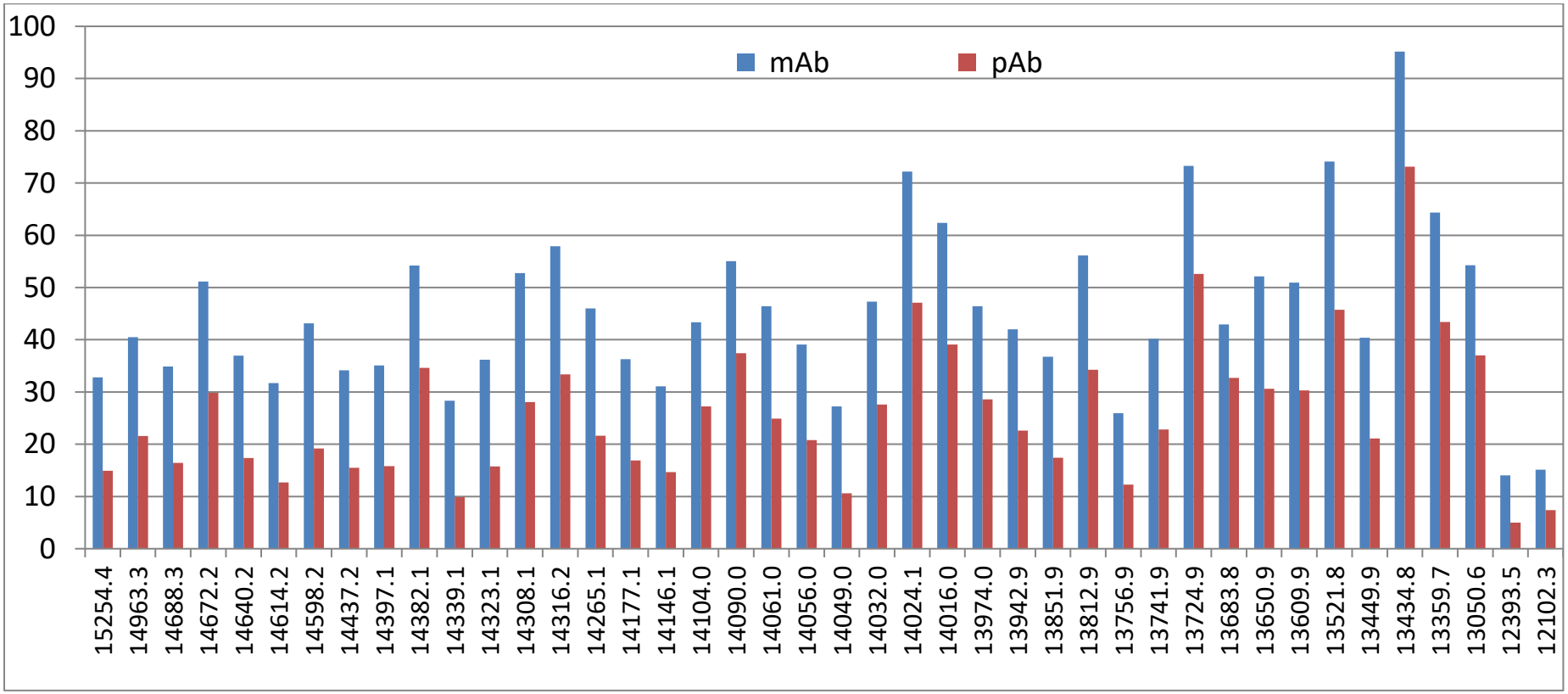


Figure 38. Extraction recoveries of each hCGα glycoform determined in the elution fraction (%) obtained with both mAbs and pAbs-based ISs. Percolation of 5 µg of rhCG on 70 mg of grafted sorbent.

Extraction recovery determined in the elution fraction (%)

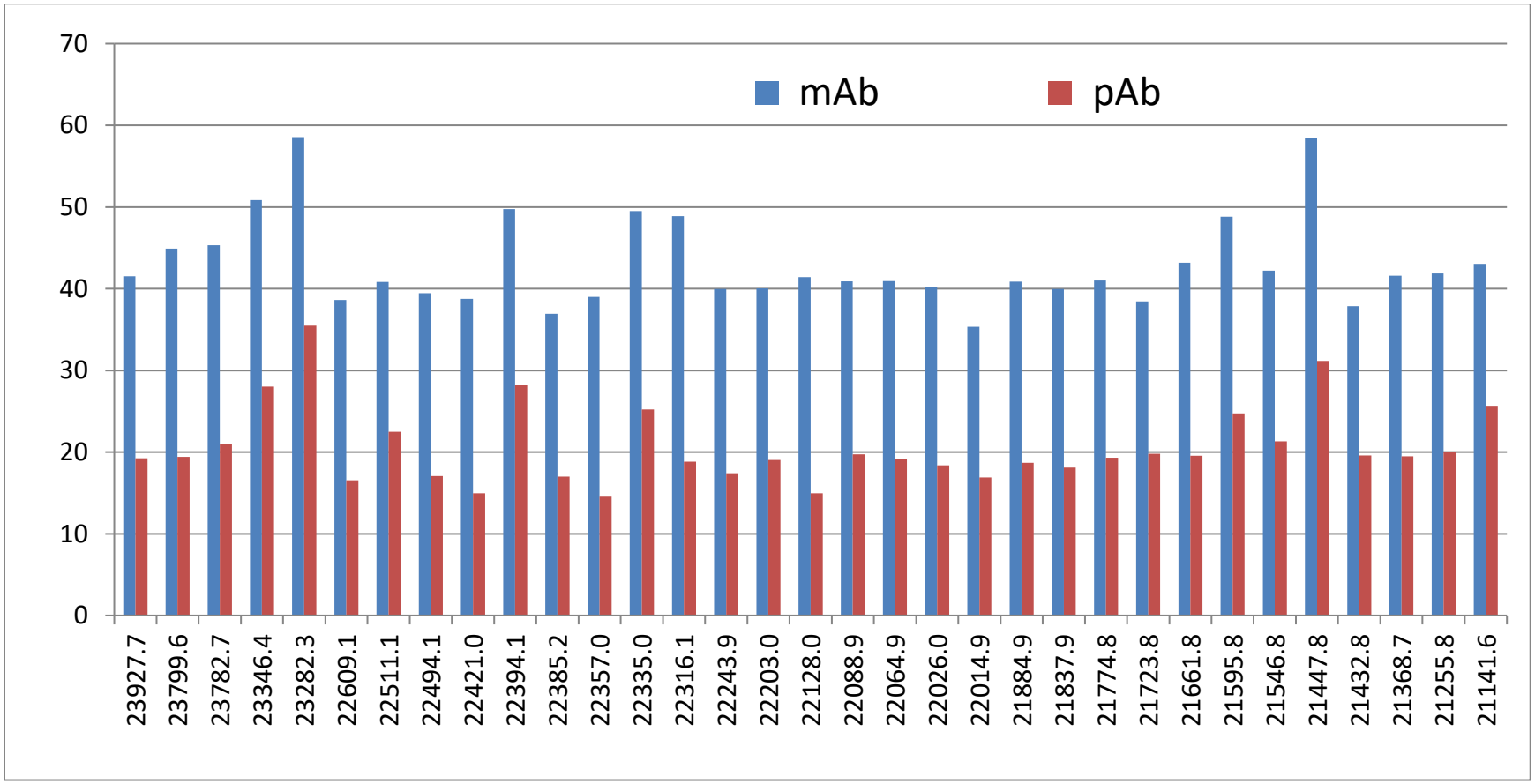


Figure 39. Extraction recoveries of each hCGβ glycoform determined in the elution fraction (%) obtained with both mAbs and pAbs-based ISs. Percolation of 5 µg of rhCG on 70 mg of grafted sorbent.

It is important to note that the LOQ the high intensity signal hCGα glycoforms is 6 µg mL^{-1} of the global hCG content and that of the low intensity hCGα glycoforms is 12.5 µg mL^{-1} of rhCG, while the LOQ for hCGβ glycoforms is 25 µg mL^{-1}. For instance the anaysed fractions showing no presence of hCG glycoforms could be a reason of quantification resulting from the presence of hCG below 12.5-25 µg mL^{-1} in general which corresponds to an extraction recovery of 9.6-40%. Therefore, the values far from 100% raise questions. Are they due to a lack of sensitivity, which makes impossible the hCG glycoform detection in percolation and washing fractions? Are they due to glycoforms loss during the different immunoextraction and reconcentration process? Is there some precipitation phenomenon in the low pH elution fraction? Additional experiments are required to optimize these iextraction recoveries, but the results are promising as all the hCG glycoforms were extracted with the both tested ISs. Additional experiments are required to optimize these extraction recoveries, but the results are promising as all the hCG glycoforms were extracted with the both tested ISs, paving the way for future determination of hCG glycoforms present in complex biological samples, such as pregnant women sera.

General conclusion and perspectives

Human Chorionic Gonadotropin, the hormone of pregnancy, is a highly glycosylated heterodimeric protein. The presence of certain glycoforms of hCG is expected to be biomarkers for some pregnancy outcomes. This was the trigger of our aim to develop an analytical method able to determine the different hCG glycoforms in biological samples. The first step consisted in the development of a nanoLC-(Orbitrap)MS method for the qualification detection and semi-quantification of the hCGα glycoforms, using a recombinant hCG-based drug as sample. This method was accompanied by testing a new data treatment method that resulted in detecting a significant higher number of hCGα glycoforms, even if it is non-automated and time-consuming. The good figures of merit of the method demonstrated its great potential for fast glycoprofiling of glycoproteins, as it was illustrated with another similar gonadotropin-based drug, containing FSH that shares the same α-subunit with hCG.

However, the inability of the method to detect the hCGβ glycoforms was the trigger to its optimization. Various parameters were then modified such as the addition of a trapping precolumn upstream of the nanoLC separation allowing an increase in the injection volume. The injection mode, gradient slope and mobile phase additive were next varied. This modified method allowed the detection of numerous hCGα but also for the first time hCGβ glycoforms. Good figures of merit of the method were also obtained and proved its ability to be used for the semi-quantification of all these hCG glycoforms. Next, the potential of this nanoLC-HRMS method to analyze different gonadotropins from recombinant or urinary source were tested. According to the nature and content of the mobile phase acidic additive (2% FA or 0.05% TFA),

it was observed that even if gonadotropins have all the same α-subunit, their very different β-subunit in terms of MW and glycosylation state induce significant chromatographic retention differences, and thus MS ability ionization and detection. A generic approach for the analysis of these different gonadotropins is thus not straightforward.

The last aim of this work was the characterization of the hCG glycoforms in real biological samples, which required the development of a sample handling step. This is why the development of an immunoextraction step able to selectively extract and concentrate all the hCG glycoforms present in a given complex sample was investigated. Two immunosorbents based on either anti-hCGβ pAbs or anti-hCGα mAbs were successfully synthesized. First immunoextraction experiments were carried out using the recombinant hCG-based drug as sample. The primary results were promising and require further investigations for their use to immunoextract hCG glycoforms present in biological samples. Therefore optimization of the immunoextractio protocol is required followed by its evaluation for the extraction of more complex samples such as uhCG-based preparations and other gonadotropins to check their cross-reactivity. This would complete the tool box to carry out a fast and simple hCG glycoprofiling in numerous sera of pregnant women in order to try to identify one or set of glycoforms as biomarkers of some pregnancy pathologies.

Supplementary Data

Table S1: Compilation of all the detected glycoforms with their average retention time (<t_R>) and average relative area (<relative area>) measured on the XIC with their corresponding RSD values obtained with the analysis by nanoLC-HRMS of 3 batches of rhCG in triplicate. The grey line corresponds to the glycoform leading to the most abundant signal in MS.

	Batch 1				**Batch 2**				**Batch 3**			
MW (Da)	**<t_R> (min)**	**RSD(t_R) (%)**	**<relative area> (%)**	**RSD(relative area) (%)**	**<t_R> (min)**	**RSD(t_R) (%)**	**<relative area> (%)**	**RSD(relative area) (%)**	**<t_R> (min)**	**RSD(t_R) (%)**	**<relative area> (%)**	**RSD(relative area) (%)**
13756.3	x	x	x	x	17.65	0.67	3	26	17.53	0.49	2	28
14048.9	x	x	x	x	17.76	0.57	3	28	17.65	0.28	3	10
14688.2	x	x	x	x	17.97	0.48	5	26	17.96	0.51	4	18
14339.0	x	x	x	x	17.92	0.16	5	12	17.94	0.22	3	18
13852.9	x	x	x	x	18.50	0.14	2	6	18.41	0.67	5	17
13521.8	x	x	x	x	x	x	x	x	18.43	0.19	4	6
14032.0	18.01	1.79	22	17	17.83	0.61	28	13	17.8	0.24	26	7
13434.8	18.10	0.14	5	13	18.39	0.68	5	11	18.33	0.14	8	9
14323.0	18.13	0.22	15	31	18.06	0.08	20	15	18.03	0.03	28	20
13724.9	18.14	0.40	31	21	18.42	0.57	34	3	18.39	0.19	9	11

14089.0	18.15	0.30	2	22	x	x	x	x	18.36	0.20	3	26
13847.8	18.15	0.24	3	30	18.44	0.88	5	13	18.38	0.16	8	5
14381.0	18.22	0.37	9	17	18.52	0.49	9	23	18.50	0.27	7	33
14139.0	18.28	0.43	4	16	18.53	0.60	4	21	18.50	0.24	5	17
14607.2	18.29	6.08	48	10	19.18	0.34	7.2	0	19.19	0.33	50	1
14016.0	18.31	0.40	78	7	18.60	0.43	44	3	18.60	0.13	78	9
14148.0	18.31	0.40	2	23	18.87	0.15	1	10	18.67	0.77	1	38
14024.1	18.32	0.45	30	15	x	x	x	x	18.66	0.60	43	25
13974.0	18.34	0.43	4	9	18.60	0.41	2	31	18.58	0.16	4	2
13609.8	18.37	0.41	2	11	18.64	0.30	2	37	18.63	0.30	2	13
13812.9	18.38	0.30	7	13	18.66	0.20	7	16	18.66	0.22	1	17
14614.2	18.39	0.40	6	37	18.53	0.64	7	19	18.40	0.30	6	4
13050.6	18.41	0.41	7	14	18.71	0.60	11	6	18.71	0.21	11	13
14672.2	18.42	0.29	15	23	18.72	0.66	16	26	18.68	0.22	16	22
13650.9	18.45	0.64	3	19	18.73	0.60	3	23	18.66	0.27	3	20

14307.1	18.50	0.49	100	0	18.72	0.26	100	0	18.78	0.05	100	0
14316.1	18.50	0.46	40	18	x	x	x	x	18.79	0.13	61	3
14265.1	18.54	0.52	4	12	18.77	0.51	2	6	18.74	0.24	1	7
14104.0	18.57	0.53	8	6	18.79	0.43	8	5	18.82	0.11	7	9
14963.3	18.58	0.61	18	7	18.83	0.25	18	14	18.87	0.22	19	18
13942.9	18.58	0.82	3	13	18.87	0.26	3	31	18.89	0.37	3	7
14598.2	18.96	0.41	48	10	19.18	0.35	49	13	19.20	0.17	50	11
14640.2	18.97	0.43	3	7	19.18	0.56	3	12	x	x	x	x
15254.4	18.99	0.43	6	35	19.21	0.47	6	1	19.28	0.15	6	3

Table S2: Study of the linearity with the corresponding determined R^2 values and regression equations of the hCGα glycoforms detected with a high intensity by nanoLC-HRMS analysis (batch 1 of rhCG). The concentration range was from 7.5 to 250 µg ml^{-1}.

Mass (Da)	R^2 Value	Equation of regression
13724.9	0.995	y=2545.6x-17160
14016.0	0.994	y=5739.4x-1214.3
14024.1	0.992	y=2383.5x-2488.8
14032.0	0.992	y=4375.4x-24447
14598.1	0.997	y=5481.2x+42806
14307.1	0.993	y=2973.9x+16723
14316.2	0.997	y=2530.4x+706.12
14323.0	0.991	y=1354.7x+19025
14607.2	0.993	y=2973.9x+16723
14672.2	0.991	y=965.56x+9304.9
14963.3	0.995	y=1080.4x-831.49

Table S3: MW of each FSHα glycoform detected after nanoLC-HRMS analysis of rFSH (3 batches, 3 runs each) and use of the second data treatment approach with glycan structure suggestions for some of them in comparison with bottom-up data (marked with *) or with GlyGen (marked with **), while when those confirmed by both are marked with ***. x: no structure to propose. The notation 1st and 2nd glycan is not correlated to their localization on the 2 N-glycosylation sites of FSHα.

Mass (Da)	Number of times the glycoforms was detected among the three runs			1st N-Glycan Proposed structure	2nd N-Glycan Proposed structure	Structure identification method
	Batch 1	Batch 2	Batch 3			
13562.8	3/3	3/3	3/3	x	x	
13688.8	0/3	0/3	3/3	x	x	
13724.8	3/3	3/3	3/3	HexNAc(4)Hex(5) or HexNAc(3)Hex(4) or HexNAc(5)Hex(6) or HexNAc(6)Hex(4)NeuAc (1)	HexNAc(4)Hex(5)NeuAc(1) or HexNAc(5)Hex(6)NeuAc(1) or HexNAc(3)Hex(4)NeuAc(1) or HexNAc(4)Hex(3)	* ** ** **
13848.9	3/3	3/3	3/3	x	x	
13852.90	3/3	3/3	3/3	HexNAc(4)Hex(5)NeuAc(2) or HexNAc(5)Hex(4)NeuAc(1)	HexNAc(4)Hex(4) or HexNAc(3)Hex(5)NeuAc(1)	* **
14016.0	3/3	3/3	3/3	HexNAc(4)Hex(5)NeuAc(1)	HexNAc(4)Hex(5)NeuAc(1)	***
14032.0	3/3	03/3	3/3	HexNAc(5)Hex(4)NeuAc(1) or HexNAc(3)Hex(5)NeuAc(1) or HexNAc(4)Hex(4)Fuc(1)	HexNAc(5)Hex(4)Fuc(1) or HexNAc(5)Hex(6)Fuc(1) or HexNAc(6)Hex(4)NeuAc(1)	**
14307.1	3/3	3/3	3/3	HexNAc(4)Hex(5)NeuAc(1) or HexNAc(3)Hex(5)NeuAc(1)	HexNAc(3)Hex(6) or HexNAc(5)Hex(5)NeuAc(2)	* **
14323.1	3/3	3/3	3/3	HexNAc(5)Hex(5)NeuAc(2) or HexNAc(5)Hex(4)	HexNAc(5)Hex(3)Fuc(1) or HexNAc(5)Hex(4)NeuAc(2)Fuc(1)	**
14349.0	3/3	3/3	3/3	x	x	
14381.0	3/3	3/3	3/3	HexNAc(4)Hex(5)NeuAc(1) or	HexNAc(3)Hex(6) or	***

				HexNAc(3)Hex(5) or HexNAc(5)Hex(4) or HexNAc(5)Hex(4)NeuAc(1)	HexNAc(6)Hex(6)NeuAc(1) or HexNAc(6)Hex(4)NeuAc(2) or HexNAc(6)Hex(4)NeuAc(1)	** ** **
14391.1	2/3	3/3	3/3	x	x	
14397.1	3/3	3/3	3/3	x	x	
14431.0	3/3	3/3	3/3	x	x	
14440.0	3/3	3/3	3/3	x	x	
14472.0	3/3	3/3	3/3	x	x	
14512.0	3/3	3/3	3/3	x	x	
14598.1	3/3	3/3	3/3	HexNAc(4)Hex(5)NeuAc(2)	HexNAc(4)Hex(5)NeuAc(2)	***
14614.2	3/3	3/3	3/3	x	x	
14640.1	3/3	3/3	3/3	x	x	
14672.2	3/3	3/3	3/3	HexNAc(4)Hex(5)NeuAc(2)	HexNAc(5)Hex(6)NeuAc(1)	*
14680.2	3/3	3/3	3/3	x	x	
14682.2	3/3	3/3	3/3	x	x	
14963.3	3/3	3/3	3/3	HexNAc(4)Hex(5)NeuAc(2)	HexNAc(5)Hex(6)NeuAc(2)	*
14979.3	3/3	3/3	3/3	x	x	
15254.3	3/3	3/3	3/3	HexNAc(4)Hex(5)Fuc(1)NeuAc(2)	HexNAc(5)Hex(6)NeuAc(3)	*
15328.3	3/3	3/3	3/3	x	x	
15338.4	3/3	3/3	3/3	x	x	
15619.4	3/3	3/3	3/3	x	x	
15910.7	3/3	3/3	3/3	x	x	

Table S4: Compilation of the detected glycoforms with their average retention time (<t_R>) and average relative area (<relative area>) measured on the XIC with their corresponding RSD values obtained with the analysis by nanoLC-MS of three batches of rFSH in triplicate. The grey line corresponds to the most intense glycoform.

	Batch 1				Batch 2				Batch 3			
Mass (Da)	<t_R> (min)	RSD(t_R) (%)	<relative area> (%)	RSD(relative area) (%)	<t_R> (min)	RSD(t_R) (%)	<relative area> (%)	RSD(relative area) (%)	<t_R> (min)	RSD(t_R (%)	<relative area> (%)	RSD(relative area) (%)
13848.9	18.67	0.43	13	14	18.31	0.14	13	14	18.63	1.11	5	9
13688.8	x	x	x	x	x	x	x	x	18.66	0.36	1	20
13724.8	18.67	0.34	17	9	18.33	0.19	18	35	18.83	1.21	1	26
14307.1	19.01	0.29	20	36	18.83	0.66	15	19	18.81	0.74	9	8
14032.0	18.81	0.44	2	12	x	x	x	x	18.60	0.86	7	14
14016.0	18.86	0.29	100	0	18.50	0.38	100	0	18.87	0.70	100	0
14381.0	18.80	0.29	16	10	18.46	0.34	19	29	18.88	0.59	2	23
14323.1	18.89	0.25	21	32	18.13	0.21	25	14	18.79	0.80	9	18
14349.0	19.29	0.43	23	28	18.84	0.77	26	31	18.82	0.24	6	37

14431.0	19.12	0.23	9	9	18.83	0.24	9	11	19.01	0.63	2	23
14397.1	19.28	0.77	8	21	18.85	0.68	14	15	18.80	0.24	2	12
14391.1	19.26	0.47	22	34	18.77	0.69	31	32	18.84	0.46	7	21
14672.2	19.09	0.13	35	1	18.65	0.37	39	11	18.84	0.75	35	31
14440.0	19.00	0.47	6	17	18.79	0.20	6	34	18.95	0.22	6	31
14963.3	19.28	0.29	86	5	19.02	0.13	87	17	19.08	0.40	37	34
14472.0	19.18	0.13	8	12	18.94	0.17	6	17	19.04	0.68	6	20
14979.3	18.91	1.13	24	16	18.74	1.90	20	16	19.03	0.65	8	19
13562.8	19.26	0.14	30	17	19.09	0.21	23	6	18.92	0.89	21	34
14640.1	19.37	0.21	8	6	19.02	0.30	10	20	19.27	0.33	5	39
13852.9	19.34	0.19	4	6	19.03	0.17	3	5	19.01	0.85	2	20
14512.0	19.26	0.65	13	26	18.87	0.32	14	35	19.11	0.82	5	21
15328.3	19.16	0.33	17	23	18.91	0.14	26	16	19.01	0.77	9	30

14614.2	19.28	0.21	48	4	19.05	0.33	42	13	19.24	0.35	13	26
14598.1	19.35	0.19	33	2	19.05	0.15	42	16	19.29	0.61	19	15
14682.2	19.43	0.24	60	4	19.09	0.00	60	8	19.37	0.75	18	11
15254.3	19.48	0.15	69	5	19.24	0.10	72	13	19.74	0.53	32	22
14680.2	19.41	0.34	8	8	19.09	0.16	9	27	19.39	0.60	4	37
15619.4	19.40	0.39	26	6	19.18	0.22	35	7	19.56	0.75	14	28
15338.4	19.53	0.21	22	2	19.28	0.33	24	11	19.87	0.38	12	37
15910.7	19.52	0.20	15	2	19.26	0.29	18	14	20.18	0.31	6	31

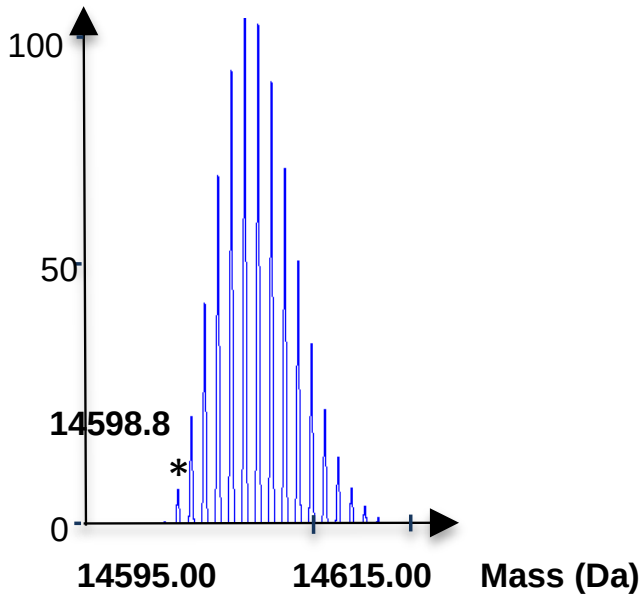


Figure S1: Theoretical mass distribution corresponding to the most commonly described hCGα glycoforms (hCGα with 2 HexNAc(4)Hex(5)NeuAc(2) N-glycans).

Table S5: Five tested gradients for the analysis of rhCG at the intact level by nanoLC-HRMS.

Gradient slope	**Gradient**	**Equilibration**
1.49%/min	10 to 80% ACN in 47 min, 90% of ACN within 5 min, 5 min plateau	return to 10% of ACN and kept for 15 min
1.02%/min	10 to 65% ACN in 54 min, 90% of ACN within 5 min 5 min plateau	return to 10% of ACN and kept for 15 min
0.86%/min	10 to 65% ACN in 64 min, 90% of ACN within 5 min 5 min plateau	return to 10% of ACN and kept for 15 min
0.67%/min	10 to 30% ACN in 30 min, 62% of ACN within 5 min 5 min plateau	return to 10% of ACN and kept for 15 min
0.56%/min	10 to 40% ACN in 54 min, 90% of ACN within 5 min 5 min plateau	return to 10% of ACN and kept for 15 min

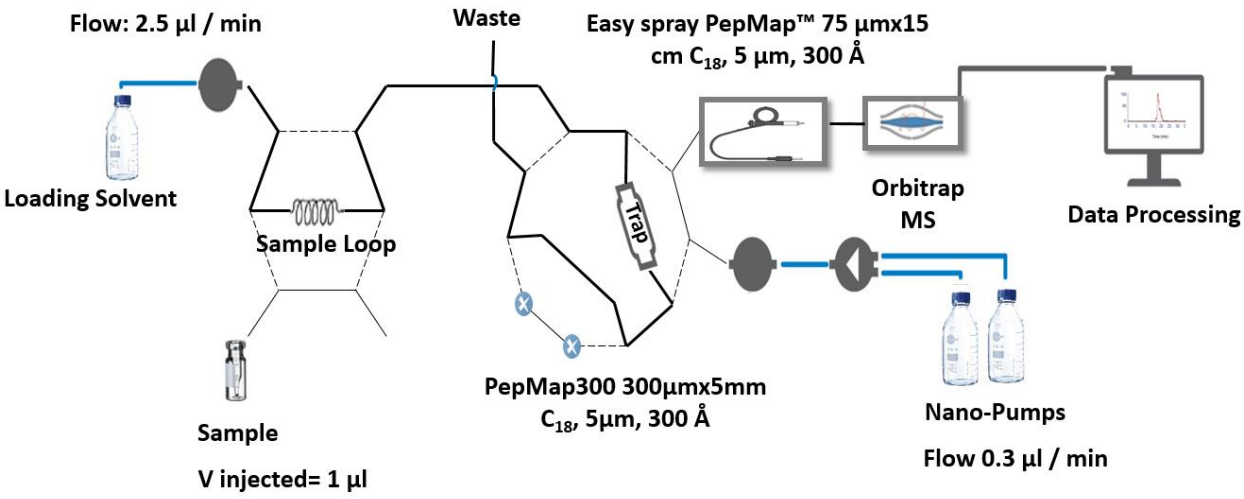


Figure S2: Experimental setup used for the nanoLC-HRMS analysis with the addition of a trapping precolumn.

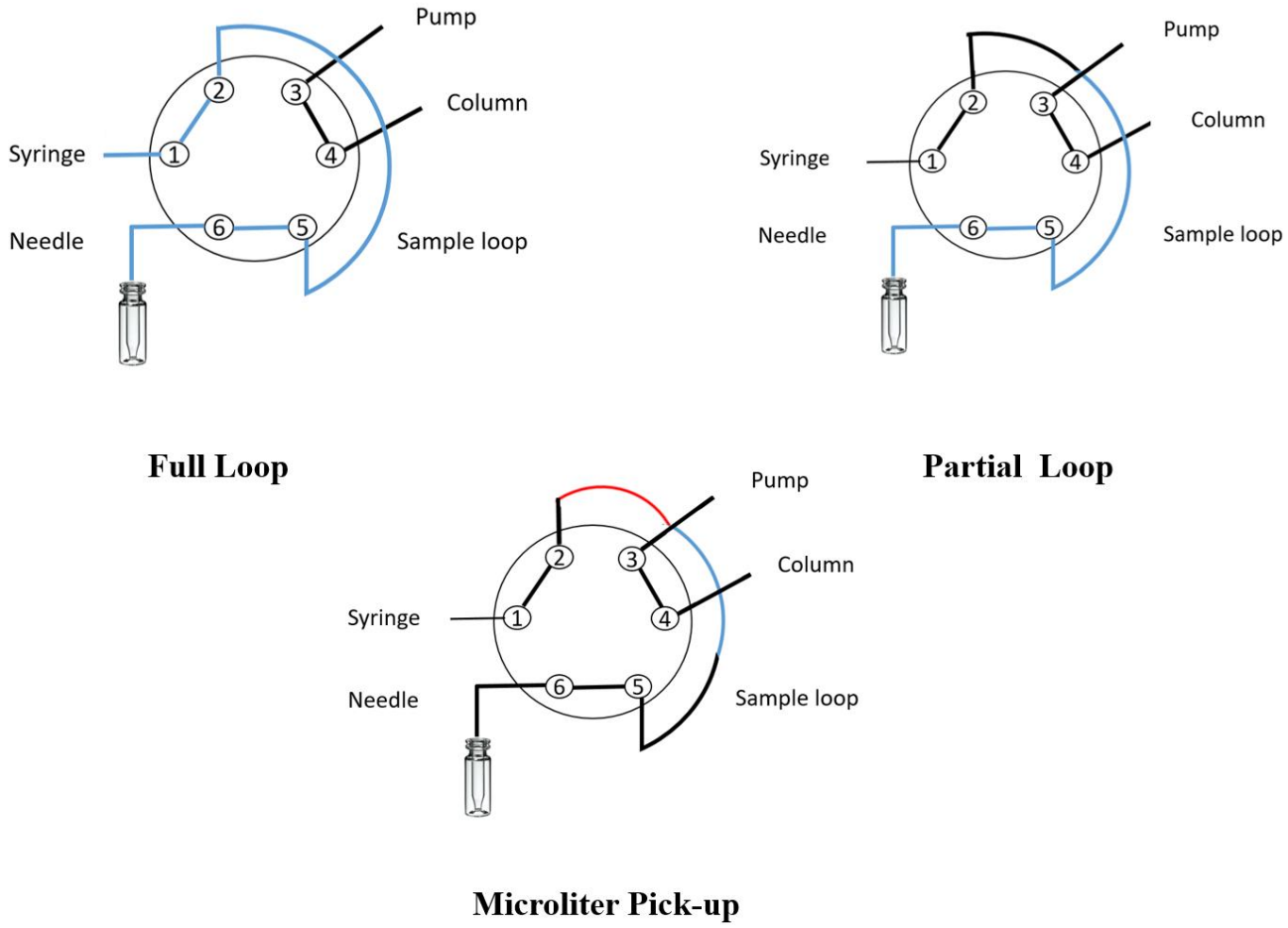


Figure S3: Schematic representations of the three tested injection modes based on full loop, partial loop or microliter pick-up (blue: sample; black: Mobile phase; red: water).

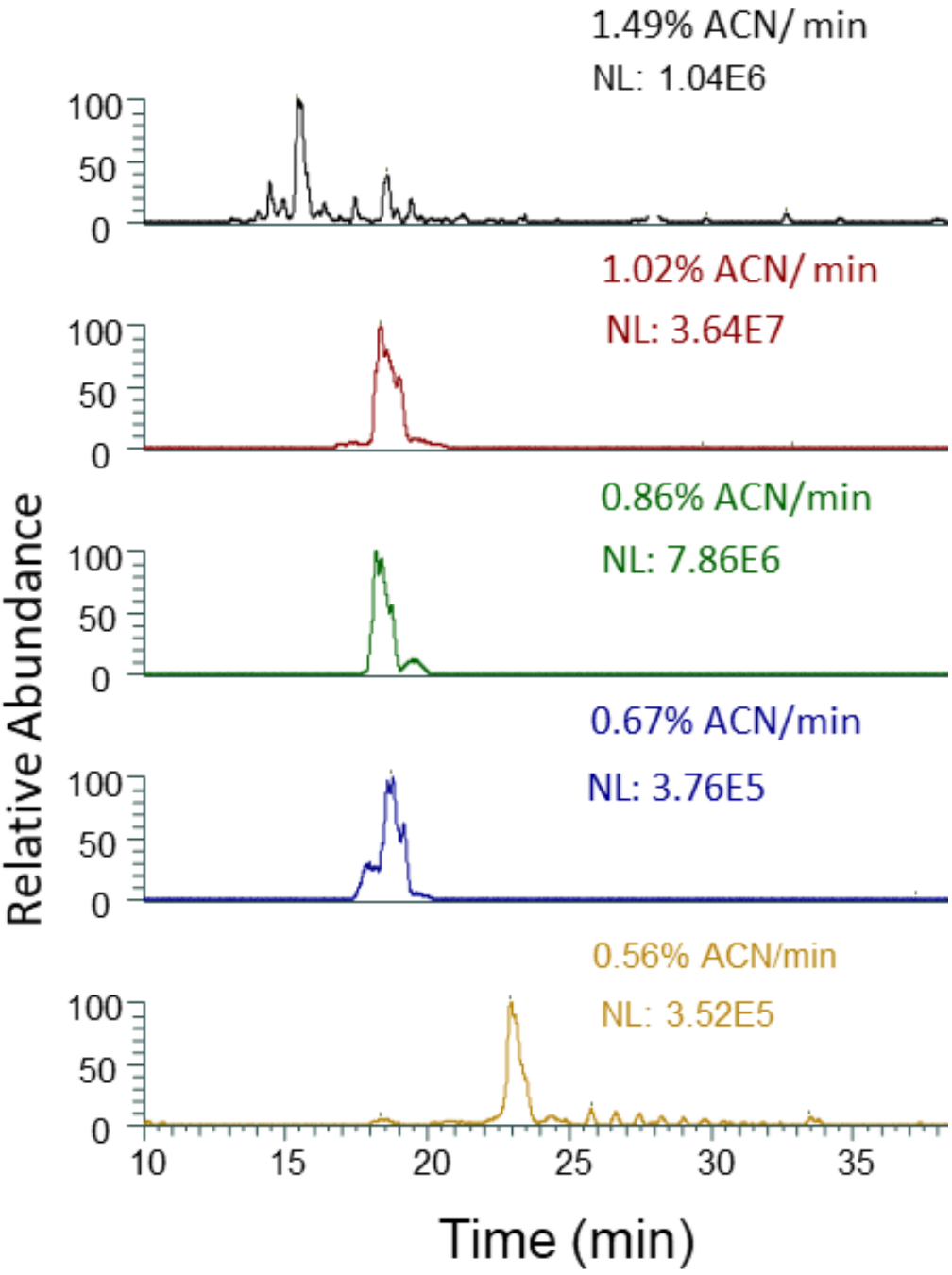


Figure S4: BPC of rhCG analyzed by nanoLC-HRMS with 5 different gradients (described in Table S1). Mobile phase: 2% FA in H_2O/ACN mixture (v/v). hCG at 100 µg ml^{-1}. Injection volume: 1 µl with partial loop mode. Other conditions: see section 2.3.

Table S6: Compilation of all the detected hCGα glycoforms with their average retention time (<t_R>) and average relative area (<relative area>) measured on the XIC with their corresponding RSD values obtained with the analysis by nano-LC-HRMS of rhCG at 100 µg ml^{-1} in triplicate with the optimized conditions. The grey line corresponds to the glycoform leading to the most abundant signal in MS.

	hCGα			
MW (Da)	**<t_R> (min)**	**RSD(t_R) (%)**	**<relative area> (%)**	**RSD(relative area) (%)**
15254.4	23.67	0.19	6.8	4.3
14963.3	23.46	0.04	21.3	4.7
14688.2	23.42	0.15	6.8	7.9
14672.1	23.41	0.15	21.3	2.2
14640.2	23.80	0.21	3.1	8.1
14614.1	23.67	0.26	13.3	10.7
14598.1	23.72	0.07	47.5	1.8
14437.2	23.43	0.09	8.1	6.2
14397.1	23.36	0.28	4.2	6.9
14381.1	23.37	0.09	11.4	2.1
14339.0	23.37	0.11	5.9	15.1
14323.0	23.40	0.17	28.6	10.7
14315.2	23.52	0.24	8.7	3.5
14307.1	23.52	0.11	100	0.0
14265.1	23.51	0.24	3.9	8.9
14177.1	23.43	0.17	1.5	21.0
14146.1	23.38	0.39	8.9	6.1
14104.1	23.54	0.07	6.6	5.9
14090.0	23.40	0.34	3.9	8.3
14061.0	23.52	0.06	2.2	12.0
14056.1	23.52	0.29	2.5	14.5
14048.9	22.55	0.30	1.3	30.3
14032.0	23.31	0.13	1.9	9.8
14024.1	23.43	0.13	6.3	3.9
14016.0	23.44	0.16	78.2	2.1
13974.1	23.51	0.40	3.0	40.5
13942.9	23.61	0.10	3.4	12.5
13852.8	23.38	0.12	5.2	5.3
13812.9	23.48	0.37	4.8	38.8
13756.8	23.36	0.25	1.7	14.8
13741.9	23.32	0.13	8.4	11.7
13724.9	23.39	0.22	29.0	2.1
13683.8	23.42	0.16	1.7	14.7

13650.86	23.52	0.20	3.1	19.3
13609.88	23.54	0.27	1.4	8.4
13521.77	23.44	0.26	1.8	25.5
13449.9	23.69	0.08	2.8	7.6
13434.82	23.42	0.13	3.9	7.9
13359.75	23.50	0.35	0.9	18.7
13050.62	23.42	0.17	1.8	15.1
12393.51	23.41	0.11	1.1	10.9
12102.31	23.32	0.07	1.4	11.9

Table S7: Compilation of all the detected hCGβ glycoforms with their average retention time (<t_R>) and average relative area (<relative area>) measured on the XIC with their corresponding RSD values obtained with the analysis by nanoLC-HRMS of rhCG at 100 µg ml^{-1} in triplicate with the optimized conditions. The grey line corresponds to the glycoform leading to the most abundant signal in MS.

	hCGβ			
MW (Da)	**<t_R> (min)**	**RSD(t_R) (%)**	**<relative area> (%)**	**RSD(relative area) (%)**
23927.7	25.73	0.06	31.8	5.3
23799.6	25.81	0.20	22.5	10.0
23782.7	25.66	0.18	14.2	17.7
23346.4	25.76	0.08	41.9	9.5
23282.3	25.76	0.20	39.5	1.9
22609.1	25.70	0.17	31.8	6.9
22511.1	25.71	0.04	24.8	9.3
22494.1	25.70	0.09	49.3	9.4
22421.0	25.76	0.07	26.2	8.0
22394.1	25.70	0.06	32.1	5.8
22385.2	25.69	0.21	38.5	9.9
22357.0	25.71	0.14	41.4	10.6
22335.0	25.65	0.05	28.7	5.3
22316.1	25.71	0.10	42.3	10.8
22243.9	25.80	0.05	52.7	6.5
22203.0	25.77	0.09	48.2	2.7
22128.0	25.72	0.08	55.8	8.0
22088.9	25.72	0.01	40.6	7.2
22064.9	25.75	0.16	55.1	5.7
22026.0	25.71	0.03	45.3	15.1
22014.9	25.74	0.08	77.5	3.1
21884.9	25.74	0.09	36.3	11.6
21837.9	25.74	0.06	77.0	3.7
21774.8	25.80	0.07	51.3	8.9
21723.8	25.78	0.12	100	0
21661.8	25.86	0.08	59.8	7.3
21595.8	25.77	0.08	39.2	7.9
21546.8	25.79	0.18	73.7	5.5
21447.8	25.73	0.08	33.0	5.8
21432.8	25.82	0.10	95.8	2.8
21368.7	25.84	0.12	41.5	5.1
21255.8	25.81	0.07	52.7	4.6
21141.6	25.84	0.05	24.8	15.5

Table S8: Study of the linearity with the corresponding determined R^2 values and regression equations of the hCGα glycoforms detected with the highest intensity by nanoLC-HRMS analysis with all the optimized conditions. The concentration range was from 6 to either 250 or 500 µg ml^{-1}.

	Concentration range 6-500 µg ml^{-1}		Concentration range 6-250 µg ml^{-1}	
Mass (Da)	R^2 value	Equation of Regression	R^2 value	Equation of Regression
14963.33	0.92	y=456076.8x+6781230.2	0.97	y= 501894x+4E+06
14672.18	0.89	y=422718.4x+7049749.7	0.98	y= 475827.1x+3667679.4
14598.16	0.92	y=980328.4x+13644799.7	0.98	y= 1081773.9x+7281143.1
14323.09	0.84	y=323755.1x+18568084.9	0.93	y= 741634.6x+1686055.5
14307.11	0.84	y=323755.1x+18568084.9	0.96	y=2566797.8x+3400889.4
14016.03	0.91	y=1511427.8x+24115292.6	0.96	y= 1624890.4x+16163090.3
13724.91	0.91	y=592251.9x+5329270.9	0.96	y= 588886.7x+4907328.8

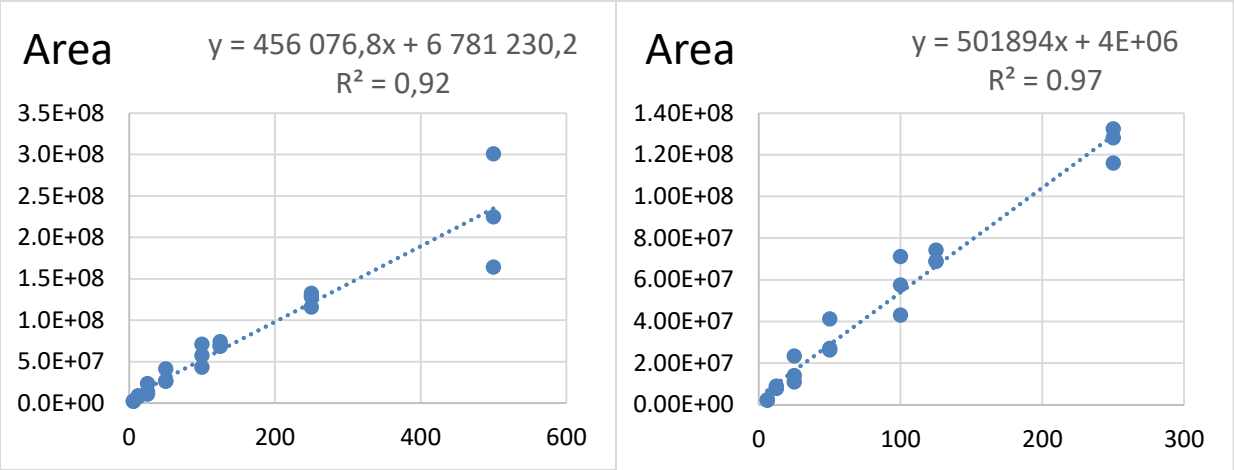


Figure S5: Calibration curves for the hCGα glycoform having a MW of 14963.3 Da injected between (Left) 6 and 500 µg ml^{-1} or (Right) 6 and 250 µg ml^{-1} (n=3 for each concentration).

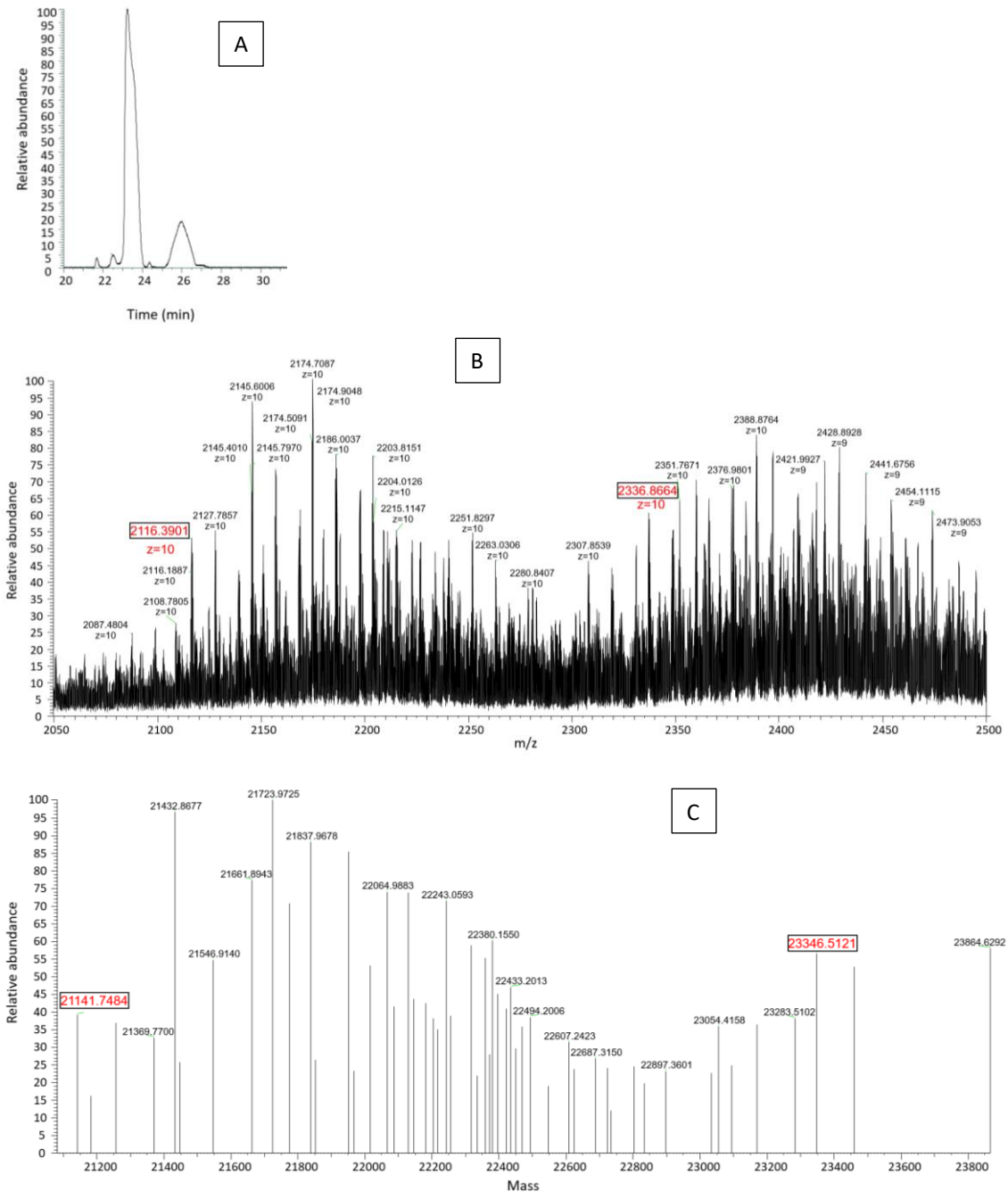


Figure S6: (A) TIC of rhCG analyzed by nanoLC-HRMS. B: Average mass spectrum corresponding to the totality of the chromatographic peak observed (retention time range: 25-26 min). (C) Deconvoluted mass spectrum corresponding to the totality of the chromatographic peak observed (retention time range: 25-26 min). Mobile phase: 0.05% TFA in H_2O/ACN mixture (v/v). Gradient: from 10 to 65% ACN with a slope of 1.02% ACN/min. rhCG at 100 µg ml^{-1}. Injection volume: 1 µl with partial loop mode. Other conditions: see section 2.3.

Table S9: Compilation of all the detected hCGα glycoform MWs analyzing in triplicate CG 10 by nano-LC-MS either with FA or TFA as additive giving their average retention time (<t_R>) and average relative area (<relative area>) measured on the XIC with their corresponding RSD values. CG 10 at 200 µg ml^{-1}. The underlined glycoforms correspond to the ones leading to the most abundant signal in MS with FA or TFA mobile phase acidifier. The glycoforms were highlighted in dark gray when addition of two oxygen atoms was suspected while those highlighted in light gray were suspected to have the addition of only one oxygen atom.

	CG 10α-FA				CG 10α-TFA			
MW (Da)	<t_R> (min)	RSD(t_R) (%)	<relative area> (%)	RSD(relative area) (%)	<t_R> (min)	RSD(t_R) (%)	<relative area> (%)	RSD(relative area) (%)
15822.9	-	-	-	-	22.10	0.08	7.6	6.3
14712.2	-	-	-	-	24.26	0.06	9.6	4.3
14614.2	17.67	0.14	27.0	1.0	22.23	0.08	10.4	9.8
14598.1	19.46	0.42	31.7	6.7	24.25	0.07	39.5	6.4
14550.1					21.06	0.07	2.4	5.2
14535.1	19.68	0.13	3.2	16.7	23.22	0.07	7.8	3.5
14494.0	-	-	-	-	24.01	0.05	23.0	19.0
14381.0	-	-	-	-	24.00	0.06	18.6	4.6
14315.1	18.74	0.54	16.8	6.1	23.91	0.10	22.0	8.8
14299.1	16.81	0.38	30.3	2.1	21.74	0.07	8.0	10.6
14282.0	17.28	0.69	15.5	12.3	22.07	0.07	23.5	3.5
14265.0	18.86	0.52	12.1	14.6	24.00	0.06	69.2	3.2
14235.0	-	-	-	-	22.10	0.06	6.8	10.9
14218.0	-	-	-	-	24.07	0.05	20.9	10.8
14170.0	-	-	-	-	24.13	0.04	34.1	11.8
14153.1	-	-	-	-	20.64	0.06	3.9	14.0
14136.0	15.73	0.29	11.6	15.0	21.76	0.06	8.6	7.7
14122.0	-	-	-	-	23.71	0.05	15.5	8.6

14119.0	17.28	0.65	65.8	1.8	22.10	0.07	25.0	7.7
14104.9	18.94	0.48	65.2	3.3	24.04	0.07	74.5	4.8
14088.0	18.96	0.36	8.6	5.6	24.02	0.06	25.0	14.3
14073.0	-	-	-	-	22.15	0.10	7.9	17.0
14056.9	-	-	-	-	24.14	0.09	24.5	5.3
14032.0	14.89	0.21	1.9	18.1	20.40	0.06	1.0	7.6
14016.0	16.19	0.75	31.7	2.6	21.50	0.08	4.1	11.1
13999.0	16.62	0.70	7.3	15.3	21.80	0.08	14.6	11.2
13990.0	-	-	-	-	20.67	0.07	5.1	12.7
13983.0	18.16	0.41	31.9	1.3	23.74	0.07	42.4	5.0
13974.0	15.78	0.22	15.5	15.3	21.80	0.09	14.3	9.5
13957.9	17.35	0.32	14.1	14.8	22.16	0.07	29.8	6.3
13942.9	18.97	0.52	88.2	2.81	24.12	0.06	100.0	0.0
13909.9	15.24	0.23	2.6	16.6	-	-	-	-
13893.9	-	-	-	-	21.07	0.06	2.1	10.5
13884.9	-	-	-	-	23.94	0.08	7.9	8.7
13877.9	19.04	0.15	2.6	10.1	23.24	0.09	7.1	0.5
13869.0	14.89	0.28	4.3	4.4	20.41	0.06	1.2	8.3
13852.0	16.26	0.77	13.4	4.2	21.53	0.06	6.0	3.5
13836.9	16.65	0.73	14.4	16.2	21.82	0.08	15.4	6.1
13820.9	18.20	0.62	37.6	4.1	23.76	0.07	46.6	5.1
13789.8	16.87	0.63	21.0	3.4	21.88	0.08	10.5	8.2
13772.9	18.40	0.81	22.8	1.4	23.85	0.04	27.6	4.3
13756.8	-	-	-	-	21.76	0.08	3.2	8.0
13748.8	16.23	0.44	1.9	21.3	-	-	-	-
13741.8	-	-	-	-	22.10	0.06	7.8	7.2
13732.8	17.76	0.59	3.4	12.5	-	-	-	-

13724.9	-	-	-	-	24.07	0.11	23.7	3.7
13707.9	14.93	0.27	5.7	8.4	21.77	0.06	4.4	12.6
13690.9	16.24	0.47	13.8	16.3	21.55	0.07	7.4	6.3
13674.8	16.69	0.67	58.7	2.6	21.85	0.08	21.0	6.6
13667.8	-	-	-	-	22.11	0.08	6.9	3.9
13657.9	18.27	0.65	21.0	14.0	23.83	0.07	55.9	3.3
13641.8	16.45	0.80	20.1	15.0	21.75	0.07	10.3	9.8
13634.7	20.14	0.08	6.7	1.5	24.20	0.06	33.8	6.6
13625.8	16.87	0.60	89.1	2.5	22.09	0.07	32.5	7.1
13609.8	18.42	0.36	100.0	0.0	24.09	0.05	86.1	1.0
13577.8	-	-	-	-	22.11	0.07	4.8	9.7
13562.8	-	-	-	-	24.14	0.05	13.5	3.5
13529.7	17.79	0.47	5.7	9.5	23.79	0.09	11.4	6.4
13513.7	-	-	-	-	24.24	0.08	21.2	4.6
13503.7	16.29	0.83	3.5	10.6	-	-	-	-
13495.8	15.10	0.28	3.1	12.8	20.64	0.06	4.2	7.6
13479.8	16.45	0.50	13.7	11.5	21.77	0.07	7.6	12.6
13463.8	16.91	0.51	21.4	13.8	22.14	0.05	19.4	6.1
13449.8	18.45	0.36	58.6	0.3	24.15	0.06	60.5	2.8
13440.7	-	-	-	-	23.84	0.07	21.8	6.2
13399.7	15.95	0.43	1.0	19.9	24.22	0.08	11.4	8.9
13383.7	16.26	0.77	2.5	15.7	21.87	0.06	8.9	17.2
13374.8	14.73	0.61	5.6	7.3	20.41	0.06	1.7	6.8
13367.7	17.80	0.54	19.1	2.6	23.88	0.11	15.7	5.1
13359.8	15.98	0.47	34.5	3.5	21.53	0.08	6.6	6.5
13342.7	16.30	0.83	61.0	1.1	21.84	0.07	22.1	7.1
13325.7	17.87	0.89	56.0	5.4	23.83	0.07	59.3	4.0

13317.7	16.53	0.68	15.1	13.1	21.01	0.06	3.3	13.8
13301.7	16.99	0.79	54.6	1.4	22.20	0.09	16.2	5.5
13285.7	18.53	0.40	52.7	5.2	24.23	0.06	51.2	4.9
13278.7	-	-	-	-	23.91	0.06	41.6	14.3
13213.7	14.76	0.72	3.5	7.6	20.41	0.06	1.0	9.1
13197.7	16.00	0.57	10.2	12.7	21.56	0.07	4.1	10.6
13180.7	16.31	0.74	17.5	15.9	21.88	0.07	13.8	7.0
13165.6	17.88	0.74	16.8	16.5	23.88	0.08	34.9	4.0
13117.6	-	-	-	-	23.95	0.07	12.8	9.4
13107.5	14.85	0.19	0.6	17.6	-	-	-	-
13067.5	15.87	0.87	0.9	11.7	24.07	0.06	19.9	6.3
13050.7	16.09	0.88	4.4	18.2	21.84	0.09	11.3	5.1
13035.7	16.04	0.85	3.9	16.0	23.80	0.08	24.4	9.8
13018.6	16.41	0.85	14.8	15.0	21.92	0.07	11.7	7.5
13003.5	17.91	0.55	7.1	14.3	23.94	0.07	31.9	5.4
12985.6	16.10	0.70	2.9	5.5	-	-	-	-
12969.6	16.43	0.58	10.3	16.8	22.07	0.07	6.9	2.0
12953.5	-	-	-	-	24.04	0.06	21.8	10.3
12686.5	16.11	0.90	2.8	13.9	21.83	0.07	4.6	8.3
12670.5	17.62	0.63	2.7	10.3	23.82	0.08	14.3	4.5
12613.3	14.50	0.46	0.2	12.1	-	-	-	-
12451.3	14.53	0.33	0.3	13.2	-	-	-	-
11472.6	-	-	-	-	36.91	0.04	0.8	8.6
11390.7	-	-	-	-	36.94	0.04	1.6	16.5
11374.7	-	-	-	-	23.94	0.08	23.6	6.3
11358.6	-	-	-	-	36.92	0.03	18.2	10.3
11230.6	-	-	-	-	36.99	0.04	1.4	10.5

10942.5	-	-	-	-	37.39	0.03	1.3	11.7

Table S10: Compilation of all the detected hCGβ glycoform MWs analyzing in triplicate CG 10 by nano-LC-MS either with FA or TFA as additive giving their average retention time (<t_R>) and average relative area (<relative area>) measured on the XIC with their corresponding RSD values. CG 10 at 200 µg ml^{-1}. The underlined glycoforms correspond to the ones leading to the most abundant signal in MS with FA or TFA mobile phase acidifier. The glycoform highlighted in dark gray is suspected to have addition of two oxygen atoms.

	CG 10β-FA				CG 10β-TFA			
MW (Da)	**<t_R> (min)**	**RSD(t_R) (%)**	**<relative area> (%)**	**RSD(relative area) (%)**	**<t_R> (min)**	**RSD(t_R) (%)**	**<relative area> (%)**	**RSD(relative area) (%)**
23784.5	-	-	-	-	26.39	0.23	63.6	11.5
23752.6	-	-	-	-	26.41	0.61	52.0	6.9
23500.5	-	-	-	-	26.44	0.75	45.4	10.97
23459.5	-	-	-	-	26.36	0.43	49.1	111.0
23346.4	23.06	0.39	56.7	6.1	26.31	0.27	56.3	9.5
23272.4	-	-	-	-	26.31	0.56	63.7	8.7
23240.4	-	-	-	-	26.32	0.32	68.2	5.5
23210.4	-	-	-	-	26.41	0.51	100.0	0.0
23126.4	-	-	-	-	26.32	0.25	45.9	10.6
23095.3	-	-	-	-	26.47	0.39	57.0	5.9
22983.5	23.24	0.38	79.5	8.3	26.32	0.22	67.8	5.8
22835.3	22.47	0.33	66.6	6.0	26.45	0.22	62.0	7.5
22804.3	-	-	-	-	26.43	0.51	40.1	7.0
22690.2	22.46	0.36	61.5	7.9	26.29	0.25	50.6	9.1
22470.2	22.84	0.65	29.3	12.6	-	-	-	-
22439.3	-	-	-	-	26.34	0.08	37.3	2.9
22398.2	22.04	0.70	27.4	7.8	-	-	-	-
22325.3	22.65	0.33	100.0	0.0	26.45	0.10	64.0	5.0
22035.2	22.17	0.38	24.8	3.6	26.43	0.39	24.4	2.4

21959.9	23.32	0.19	14.4	10.9	-	-	-	-

Table S11: Compilation of all the detected hCGα glycoform MWs analyzing in triplicate orb80731 by nano-LC-MS either with FA or TFA as additive giving their average retention time (<t_R>) and average relative area (<relative area>) measured on the XIC with their corresponding RSD values. orb80731 at 200 µg ml^{-1}. The underlined glycoforms correspond to the ones leading to the most abundant signal in MS with FA or TFA mobile phase acidifier. The glycoforms were highlighted in dark gray when addition of two oxygen atoms was suspected while those highlighted in light gray were suspected to have the addition of only one oxygen atom.

	orb80731α-FA				orb80731α-TFA			
MW (Da)	**<t_R> (min)**	**RSD(t_R) (%)**	**<relative area> (%)**	**RSD(relative area) (%)**	**<t_R> (min)**	**RSD(t_R) (%)**	**<relative area> (%)**	**RSD(relative area) (%)**
14793.3	-	-	-	-	19.25	0.74	0.7	10.7
14777.3	-	-	-	-	19.21	0.48	1.6	3.9
14761.2	20.81	0.35	6.5	2.6	20.95	0.26	9.8	3.9
14744.2	20.80	0.39	10.7	6.1	20.92	0.34	9.4	3.0
14728.2	-	-	-	-	20.98	0.24	5.6	4.8
14712.2	20.85	0.44	4.7	11.8	21.02	0.24	3.6	6.8
14647.2	18.75	0.48	1.0	10.5	19.24	0.62	0.8	10.9
14630.2	18.77	0.42	6.0	4.0	20.74	0.41	3.3	9.6
14614.2	20.84	0.38	12.1	2.7	19.25	0.57	5.7	2.5
14598.2	20.84	0.40	31.3	3.3	20.99	0.38	28.9	4.5
14511.1	20.59	0.22	4.7	0.9	20.80	0.46	4.3	6.7
14469.1	20.60	0.49	6.4	11.3	20.78	0.45	4.1	9.5
14453.1	20.55	0.40	7.2	5.1	20.76	0.43	4.7	9.1
14430.1	-	-	-	-	21.07	0.24	4.8	9.2
14397.1	20.64	0.45	9.6	12.5	20.81	0.56	7.4	5.9
14381.1	20.57	0.32	12.5	5.3	20.84	0.67	7.9	12.0
14307.1	20.60	0.56	12.9	5.4	20.83	0.63	12.6	3.0
14299.1	18.83	0.25	18.7	2.9	19.38	0.70	16.2	3.7
14282.1	20.40	0.33	72.9	2.6	20.68	0.63	63.4	0.3

14265.1	20.58	0.42	90.2	3.6	20.81	0.43	82.3	4.2
14235.0	20.62	0.40	11.5	12.4	20.81	0.58	8.7	2.8
14219.0	20.62	0.26	13.8	4.5	20.87	0.43	13.1	4.8
14153.0	18.77	0.55	6.1	4.0	18.70	0.44	1.3	15.5
14137.0	18.78	0.19	16.7	3.0	19.44	0.49	16.1	2.4
14120.0	20.40	0.33	70.9	4.0	20.68	0.46	64.7	2.1
14104.0	20.64	0.49	89.2	3.5	20.86	0.47	86.0	3.2
14088.0	20.66	0.32	18.8	9.7	20.82	0.29	10.4	6.5
14073.0	20.65	0.52	10.5	7.6	20.89	0.46	6.9	6.5
14056.0	20.72	0.31	16.2	2.2	20.91	0.40	10.4	12.1
14016.0	-	-	-	-	20.84	0.53	10.8	4.2
13991.0	-	-	-	-	18.74	0.67	1.4	14.7
13974.0	18.77	0.37	20.9	1.5	19.54	0.35	2.3	2.1
13958.0	19.00	0.46	18.9	3.2	20.73	0.64	66.6	4.2
13942.9	20.69	0.25	100.0	0.0	20.91	0.36	92.1	7.5
13852.9	18.83	0.32	1.2	17.6	20.90	0.10	8.3	4.63
13836.9	20.46	0.55	2.6	9.1	20.66	0.59	5.0	4.7
13820.9	20.48	0.34	6.4	11.4	21.55	0.78	15.3	1.9
13789.9	20.47	0.27	4.5	5.2	20.74	0.34	8.4	8.9
13772.9	20.76	0.29	15.1	17.4	20.97	0.31	8.2	2.2
13756.9	-	-	-	-	20.88	0.50	8.6	4.7
13741.9	20.45	0.33	8.6	6.9	20.63	0.76	3.7	5.0
13723.9	-	-	-	-	21.52	0.26	8.7	10.1
13674.9	20.38	0.42	4.9	12.6	20.66	0.10	13.6	6.2
13657.9	18.79	0.70	3.9	12.6	19.16	0.51	2.3	3.7
13641.8	18.84	0.52	8.8	8.0	20.68	0.20	19.5	3.2
13625.8	20.41	0.39	75.8	5.9	20.71	0.32	77.1	0.9

13609.8	20.65	0.74	80.6	6.5	21.34	0.63	100.0	0.0
13495.8	18.89	0.55	2.3	8.1	19.16	0.68	1.8	9.5
13479.8	18.84	0.45	5.4	5.4	20.60	0.23	17.4	3.2
13463.8	20.43	0.51	25.0	11.4	20.60	0.15	23.4	2.9
13449.8	21.11	0.33	29.8	6.2	21.49	0.84	60.9	1.2
13342.8	20.38	0.12	4.1	12.6	-	-	-	-
13333.7	-	-	-	-	19.19	0.78	0.9	15.8
13317.7	18.86	0.34	6.6	14.9	20.25	0.32	9.2	4.4
13301.7	20.40	0.48	25.4	4.5	20.59	0.41	16.3	5.4
13285.7	21.24	0.23	29.8	2.0	21.70	0.48	43.8	2.5

Table S12: Compilation of all the detected hCGβ glycoform MWs analyzing in triplicate orb80731 by nano-LC-MS with TFA as additive giving their average retention time (<t_R>) and average relative area (<relative area>) measured on the XIC with their corresponding RSD values.orb80731 at 200 µg ml^{-1}. The underlined glycoform corresponds to the ones leading to the most abundant signal in MS with TFA mobile phase acidifier. The glycoform highlighted in dark gray is suspected to have addition of two oxygen atoms.

MW (Da)	<tR> (min)	RSD (tR) (%)	<relative area> (%)	RSD (relative area) (%)
24041.7	25.84	0.51	60.0	7.4
23864.7	25.91	0.47	86.2	8.2
23752.6	25.92	0.43	100.0	0.0
23637.6	25.89	0.54	92.0	6.1
23574.6	25.89	0.77	74.6	4.7
23532.5	25.93	0.50	53.1	10.0
23500.5	25.92	0.62	80.5	7.0
23459.5	25.88	0.49	74.0	2.4
23385.5	25.98	0.63	82.1	1.6
23272.4	25.91	0.35	80.8	4.4
22983.4	25.94	0.45	54.2	6.5

Table S13: Compilation of all the detected glycoform MWs analyzing in triplicate Luveris®, a LH-based drug, by nano-LC-MS either with FA or TFA as additive giving their average retention time (<t_R>) and average relative area (<relative area>) measured on the XIC with their corresponding RSD values. LH at 30 µg ml^{-1}. The underlined glycoforms correspond to the ones leading to the most abundant signal in MS with FA or TFA mobile phase acidifier. The glycoforms were highlighted in dark gray when addition of two oxygen atoms was suspected while those highlighted in light gray were suspected to have the addition of only one oxygen atom.

	FA				TFA			
MW (Da)	**<t_R> (min)**	**RSD(t_R) (%)**	**<relative area> (%)**	**RSD(relative area) (%)**	**<t_R> (min)**	**RSD(t_R) (%)**	**<relative area> (%)**	**RSD(relative area) (%)**
16302.6	-	-	-	-	31.30	0.32	10.3	9.1
16239.5	-	-	-	-	31.22	0.21	4.8	2.2
16188.6	-	-	-	-	31.27	0.23	19.6	10.5
16126.5	-	-	-	-	31.25	0.22	8.7	3.8
16011.5	-	-	-	-	31.23	0.27	12.8	3.3
15984.6	19.25	0.68	1.3	8.1	-	-	-	-
15897.5	-	-	-	-	31.23	0.13	19.5	7.8
15874.4	-	-	-	-	31.26	0.27	21.1	5.7
15808.4	-	-	-	-	31.29	0.29	6.2	17.7
15784.4	-	-	-	-	30.34	0.32	6.7	10.5
15760.4	19.28	0.60	2.7	4.5	31.26	0.14	42.1	3.4
15703.5	19.34	0.65	1.9	7.3	-	-	-	-
15696.4	-	-	-	-	31.23	0.18	12.0	7.8
15663.5	-	-	-	-	31.26	0.09	21.6	4.9
15661.5	19.32	0.63	1.6	2.5	-	-	-	-
15647.4	-	-	-	-	31.25	0.16	66.4	2.6
15619.5	19.30	0.60	13.9	3.7	-	-	-	-
15615.4	-	-	-	-	31.25	0.46	12.2	4.0
15583.4	-	-	-	-	31.25	0.30	24.9	4.2

15580.5	-	-	-	-	31.29	0.37	20.3	9.3
15549.4	-	-	-	-	31.25	0.16	33.35	1.6
15533.4	-	-	-	-	31.27	0.13	100.00	0.00
15517.4	-	-	-	-	31.28	0.39	10.5	2.5
15500.4	-	-	-	-	29.76	0.29	15.2	8.1
15485.4	19.03	3.61	1.2	19.4	31.25	0.49	13.5	6.7
15474.4	18.58	0.46	2.8	3.0	-	-	-	-
15470.3	-	-	-	-	31.24	0.23	34.9	5.6
15435.4	-	-	-	-	30.41	0.13	15.9	12.2
15420.3	-	-	-	-	30.43	0.18	39.0	1.3
15403.4	-	-	-	-	31.22	0.32	11.9	1.3
15400.4	19.32	0.71	6.2	5.7	-	-	-	-
15386.4	-	-	-	-	29.66	0.03	17.0	10.5
15372.3	-	-	-	-	31.24	0.31	12.9	12.3
15355.3	-	-	-	-	31.26	0.18	43.7	7.7
15338.4	19.40	0.68	8.8	10.6	-	-	-	-
15328.4	18.61	0.35	14.2	2.7	-	-	-	-
15322.3	-	-	-	-	31.24	0.12	8.3	14.9
15305.4	18.68	0.11	6.3	6.5	-	-	-	-
15295.3	19.39	0.72	7.9	10.7	-	-	-	-
15288.4	18.75	0.84	5.1	22.2	31.20	0.21	14.5	6.5
15273.3	19.33	0.72	24.0	13.7	29.71	0.48	28.5	7.2
15258.3	-	-	-	-	31.21	0.08	17.6	7.5
15254.3	19.36	0.69	56.6	10.5	24.32	0.27	21.2	26.6
15241.3	-	-	-	-	31.23	0.19	53.1	7.2
15209.2	-	-	-	-	29.64	0.15	9.5	14.1
15193.3	18.63	0.39	1.6	4.5	-	-	-	-
15183.3	18.04	0.40	1.2	7.3	-	-	-	-
15160.3	-	-	-	-	29.62	0.29	5.6	11.2
15144.3	-	-	-	-	29.77	0.34	9.1	13.6
15129.3	-	-	-	-	30.35	0.31	21.7	7.3

15125.3	18.59	0.57	5.8	15.1	-	-	-	-
15110.3	18.60	0.35	9.2	2.8	-	-	-	-
15096.2	-	-	-	-	29.69	0.06	11.1	13.2
15077.3	18.68	0.46	5.9	4.2	-	-	-	-
15063.3	18.57	0.62	5.9	13.2	-	-	-	-
15047.2	18.67	0.45	13.2	1.0	29.61	0.34	4.5	6.9
15037.3	18.09	0.53	5.9	5.0	-	-	-	-
15021.3	18.63	0.17	6.5	11.4	-	-	-	-
15014.3	18.11	0.36	2.8	9.6	-	-	-	-
15005.3	18.70	0.32	10.6	2.6	-	-	-	-
14995.3	18.10	0.47	7.6	21.7	-	-	-	-
14982.2	-	-	-	-	29.74	0.20	15.2	8.9
14979.4	18.57	0.81	43.8	16.7	24.23	0.26	11.3	25.3
14963.2	18.68	0.33	100.00	0.00	24.31	0.08	37.7	23.0
14940.2	18.69	0.50	3.9	7.6	-	-	-	-
14920.2	18.68	0.30	3.6	5.8	-	-	-	-
14872.2	18.73	0.93	2.3	4.2	-	-	-	-
14834.2	18.06	0.31	2.7	15.0	-	-	-	-
14818.2	18.08	0.47	4.9	3.0	-	-	-	-
14796.2	18.70	0.29	4.1	6.0	-	-	-	-
14786.1	18.14	0.60	4.5	6.0	24.25	0.25	3.9	8.6
14759.2	18.19	0.38	6.9	5.0	-	-	-	-
14755.2	-	-	-	-	24.32	0.19	3.0	25.6
14745.2	18.65	0.28	7.3	8.8	24.33	0.35	2.4	28.0
14729.2	18.74	0.28	4.5	14.3	-	-	-	-
14712.2	18.72	0.48	7.3	4.3	24.35	0.23	3.9	30.5
14704.2	17.63	0.68	5.1	19.5	-	-	-	-
14688.3	17.98	0.35	25.8	18.3	24.30	0.23	8.7	27.7
14672.1	18.14	0.31	51.1	4.0	24.28	0.36	26.9	24.4

14650.1	18.17	0.28	4.9	10.0	-	-	-	-
14640.1	18.73	0.47	7.6	1.2	-	-	-	-
14630.1	18.15	0.62	5.9	17.4	-	-	-	-
14614.2	18.67	0.28	27.2	16.5	24.35	0.22	6.3	37.5
14598.1	18.70	0.36	61.3	2.1	24.38	0.25	17.4	19.6
14525.1	18.68	0.57	3.3	14.8	-	-	-	-
14508.1	18.80	0.46	4.5	0.8	-	-	-	-
14485.1	18.01	0.94	1.9	16.6	-	-	-	-
14469.1	18.05	0.58	3.1	9.6	-	-	-	-
14453.1	18.11	0.58	4.4	2.0	-	-	-	-
14438.1	18.09	0.83	1.5	18.2	-	-	-	-
14430.0	18.74	0.51	1.7	4.5	-	-	-	-
14421.0	18.16	0.15	3.4	7.4	24.34	0.19	3.2	28.1
14397.1	17.68	0.93	2.5	26.8	-	-	-	-
14391.1	18.20	0.29	4.7	5.4	-	-	-	-
14381.0	17.87	0.76	6.9	19.9	24.25	0.33	8.5	14.6
14365.1	18.09	0.42	3.3	16.7	-	-	-	-
14349.1	18.21	0.58	4.4	2.3	-	-	-	-
14341.0	17.70	0.49	3.9	20.7	-	-	-	-
14339.0	17.67	0.41	4.5	25.4	-	-	-	-
14323.1	17.72	0.46	27.8	16.0	24.23	0.27	7.4	23.7
14307.0	18.19	0.31	55.8	3.1	24.33	0.24	22.3	21.9
14179.0	17.75	0.64	0.6	29.3	-	-	-	-
14162.0	17.81	0.99	1.3	13.9	-	-	-	-
14143.0	18.21	0.30	1.9	2.0	-	-	-	-
14047.9	17.54	0.92	0.7	23.2	-	-	-	-
14032.0	17.70	0.85	5.4	25.7	24.26	0.07	2.9	20.3

14016.9	17.87	0.75	11.9	15.9	24.28	0.36	10.8	5.3
13724.8	-	-	-	-	24.29	0.23	3.0	12.6

Table S14: Compilation of all the detected glycoform MWs analyzing in triplicate Puregon®, a FSH-based drug, by nano-LC-MS either with FA or TFA as additive giving their average retention time (<t_R>) and average relative area (<relative area>) measured on the XIC with their corresponding RSD values. FSH at 200 µg ml^{-1}. The underlined glycoforms correspond to the ones leading to the most abundant signal in MS with FA or TFA mobile phase acidifier. The glycoforms were highlighted in dark gray when addition of two oxygen atoms was suspected while those highlighted in light gray were suspected to have the addition of only one oxygen atom.

	FSH-FA				FSH-TFA			
MW (Da)	**<t_R> (min)**	**RSD(t_R) (%)**	**<relative area> (%)**	**RSD(relative area) (%)**	**<t_R> (min)**	**RSD(t_R) (%)**	**<relative area> (%)**	**RSD(relative area) (%)**
19301.1	27.04	0.52	2.8	0.6	-	-	-	-
19219.1	27.06	0.69	3.0	2.8	-	-	-	-
19135.0	27.12	0.70	2.5	4.0	-	-	-	-
19060.0	28.14	0.82	1.8	3.5	-	-	-	-
19018.0	28.27	0.84	2.3	2.6	-	-	-	-
18975.0	28.00	0.47	1.9	3.3	-	-	-	-
18934.0	28.04	0.52	3.3	4.0	-	-	-	-
18887.0	26.27	0.49	1.7	4.8	-	-	-	-
18843.9	25.85	0.86	3.5	4.7	-	-	-	-
18742.9	26.70	0.45	2.7	2.4	-	-	-	-
18726.9	26.44	0.59	4.5	0.7	-	-	-	-
18698.9	26.41	0.79	3.1	3.7	-	-	-	-
18642.9	26.59	0.68	4.4	2.1	-	-	-	-
18552.8	24.41	0.74	1.9	3.4	-	-	-	-
18496.8	27.20	0.77	1.8	2.3	-	-	-	-
18351.8	25.63	0.68	2.4	1.7	-	-	-	-
16475.1	20.09	0.50	1.81	9.73	-	-	-	-
16275.7	20.69	0.56	4.66	1.38	-	-	-	-
16036.7	20.83	0.74	4.16	3.92	-	-	-	-

16012.7	20.81	0.70	2.84	3.93	-	-	-	-
15994.6	20.81	0.60	9.10	4.65	-	-	-	-
15984.6	19.82	0.68	6.74	2.03	-	-	-	-
15952.6	20.81	0.62	8.38	1.70	-	-	-	-
15910.6	20.78	0.70	21.09	0.51	-	-	-	-
15703.6	19.94	0.40	10.87	2.77	17.55	0.38	6.40	14.07
15661.5	19.92	0.49	11.62	4.19	-	-	-	-
15636.5	19.90	0.58	11.26	1.78	-	-	-	-
15619.5	19.89	0.60	36.03	2.73	17.54	0.21	18.12	1.37
15422.4	20.27	0.53	7.65	4.04	-	-	-	-
15396.4	19.15	0.67	5.48	8.04	-	-	-	-
15380.4	20.26	0.45	8.87	1.71	-	-	-	-
15369.4	-	-	-	-	17.58	0.24	9.92	6.92
15338.4	20.20	0.74	21.99	3.57	17.61	0.49	11.16	0.87
15328.4	19.12	0.74	25.27	6.10	17.36	0.12	18.43	3.71
15295.4	20.17	0.55	20.74	1.73	17.60	0.22	8.67	6.45
15271.4	20.00	0.63	17.98	2.21	-	-	-	-
15254.4	20.03	0.51	61.78	2.01	17.60	0.35	31.76	2.50
15169.3	20.34	0.50	5.19	6.85	-	-	-	-
15087.3	20.27	0.49	11.83	4.91	17.50	0.31	9.90	6.28
15047.3	19.28	0.80	16.03	3.72	17.52	0.17	15.09	6.08
15037.3	18.47	0.83	15.39	10.02	17.30	0.18	10.63	4.65
15005.3	19.22	0.66	16.99	1.99	17.46	0.66	13.59	7.65
14979.3	17.58	0.76	8.43	3.86	17.36	0.46	14.38	4.59
14963.3	19.18	0.58	73.85	1.97	17.45	0.34	58.62	1.24
14795.2	19.64	0.56	14.48	1.08	17.32	0.18	15.08	1.86
14765.2	-	-	-	-	17.58	0.23	6.88	12.28

14756.2	18.72	0.43	8.53	4.07	17.45	0.59	11.29	4.42
14745.2	17.94	0.78	6.23	9.83	-	-	-	-
14725.2	19.60	0.74	7.35	4.91	-	-	-	-
14712.2	18.72	0.35	9.57	5.28	17.50	0.27	23.45	1.92
14698.2	17.64	0.79	3.87	2.71	-	-	-	-
14682.2	19.49	0.75	24.64	2.01	17.56	0.21	24.80	4.04
14672.1	18.68	0.65	63.03	0.96	17.31	0.18	44.06	1.92
14640.1	19.43	0.80	24.22	2.21	17.55	0.21	21.26	6.02
14614.2	17.60	0.72	13.19	2.19	17.51	0.21	22.68	6.49
14598.1	19.32	0.45	100.00	0.00	17.51	0.24	88.91	2.87
14512.1	19.63	0.64	6.23	2.68	17.47	0.23	14.28	4.85
14509.1	18.72	0.83	17.04	4.48	-	-	-	-
14472.1	19.63	0.59	5.06	7.04	-	-	-	-
14440.1	17.99	0.56	7.11	1.47	-	-	-	-
14431.1	19.50	0.49	15.87	4.34	17.37	0.25	18.57	1.88
14421.1	-	-	-	-	17.31	0.21	22.05	5.26
14391.1	18.81	0.82	20.86	3.64	17.47	0.43	19.93	3.22
14381.1	17.95	0.79	10.59	9.19	17.27	0.23	16.71	6.60
14349.0	18.68	0.45	28.32	3.43	17.39	0.64	17.04	3.56
14325.1	18.52	0.56	31.14	1.50	17.30	0.17	18.15	5.71
14323.1	17.26	0.77	5.42	18.34	-	-	-	-
14307.0	18.54	0.67	21.21	1.23	17.33	0.17	100.00	0.00
14219.0	18.15	0.45	15.27	2.22	17.35	0.09	9.81	2.44
14145.2	18.82	0.84	33.12	3.61	17.56	0.28	18.49	6.91
14139.0	-	-	-	-	17.28	0.26	16.12	11.70
14130.0	-	-	-	-	17.30	0.22	12.81	4.32
14100.0	18.11	0.72	10.36	9.54	-	-	-	-

14058.0	18.13	0.81	13.11	5.12	-	-	-	-
14032.1	16.92	0.74	5.78	9.14	17.27	0.23	9.33	6.75
14016.1	18.05	0.88	33.79	6.56	17.29	0.19	57.92	5.86
13852.9	18.10	0.54	11.53	6.55	17.27	0.30	20.57	10.16
13724.9	17.63	0.84	10.48	20.21	17.27	0.22	11.65	9.17
13688.8	18.87	0.72	5.78	4.47	-	-	-	-
13650.8	18.16	0.67	4.86	5.52	-	-	-	-
13578.8	16.73	0.80	1.24	17.14	-	-	-	-
13562.8	17.71	0.75	10.98	19.57	17.33	0.15	12.71	4.46
13398.7	17.70	0.74	5.15	12.41	-	-	-	-
13359.7	17.71	0.69	1.30	20.24	-	-	-	-
13271.7	17.47	0.77	1.60	11.72	-	-	-	-

Table S15: Compilation of all the detected glycoform MWs analyzing in triplicate Merional®, a FSH- and LH-based drug, by nano-LC-MS either with FA or TFA as additive giving their average retention time (<t_R>) and average relative area (<relative area>) measured on the XIC with their corresponding RSD values. FSH at 250 µg ml^{-1} and LH at 750 µg ml^{-1}. The underlined glycoforms correspond to the ones leading to the most abundant signal in MS with FA or TFA mobile phase acidifier. The glycoforms were highlighted in dark gray when addition of two oxygen atoms was suspected while those highlighted in light gray were suspected to have the addition of only one oxygen atom.

	Merional-FA				Merional-TFA			
MW (Da)	**<t_R> (min)**	**RSD(t_R) (%)**	**<relative area> (%)**	**RSD(relative area) (%)**	**<t_R> (min)**	**RSD(t_R) (%)**	**<relative area> (%)**	**RSD(relative area) (%)**
16160.7	16.65	0.61	11.4	2.5	-	-	-	-
16144.7	17.11	0.65	50.5	0.5	-	-	-	-
16041.8	-	-	-	-	22.28	0.40	38.3	2.8
15958.7	16.62	0.60	18.3	3.4	-	-	-	-
15942.7	17.09	0.66	88.7	0.7	21.00	0.26	49.9	7.0
15926.6	18.66	0.43	75.3	1.3	22.29	0.25	62.9	9.2
15910.7	20.45	0.34	21.2	1.4	24.21	0.33	47.5	5.8
15870.7	15.95	0.71	16.7	13.8	20.55	0.16	43.1	8.5
15854.7	16.33	0.71	11.8	4.2	20.93	0.36	56.7	9.5
15838.6	17.94	0.33	17.9	13.5	22.28	0.43	65.7	5.3
15797.6	17.79	0.52	15.5	4.7	-	-	-	-
15710.6	15.96	0.74	14.1	3.1	-	-	-	-
15693.6	16.44	0.68	48.4	7.1	-	-	-	-
15677.5	17.92	0.34	13.0	1.3	-	-	-	-
15667.5	15.93	0.97	11.6	5.8	-	-	-	-
15659.5	19.59	0.36	37.0	2.4	-	-	-	-
15650.5	16.38	0.57	12.4	11.5	-	-	-	-
15635.5	17.87	0.70	28.0	1.0	22.27	0.34	73.1	4.4

15619.6	19.51	0.30	9.0	8.7	-	-	-	-
15578.5	19.37	0.34	8.6	4.2	-	-	-	-
15572.5	18.43	0.49	14.3	10.0	-	-	-	-
15545.5	15.99	0.69	8.4	11.1	-	-	-	-
15530.5	16.44	0.62	54.7	1.4	-	-	-	-
15514.5	17.93	0.51	18.6	1.7	-	-	-	-
15505.5	16.00	0.84	19.2	3.7	20.63	0.23	58.6	4.5
15497.5	19.56	0.35	10.4	6.7	-	-	-	-
15489.5	16.42	0.51	85.3	2.4	21.02	0.26	69.7	4.6
15474.5	17.94	0.45	42.2	0.9	22.33	0.36	72.2	4.5
15457.5	19.57	0.41	24.1	3.6	24.28	0.37	50.6	3.3
15417.5	-	-	-	-	20.60	0.24	43.7	6.8
15400.5	-	-	-	-	20.98	0.28	34.8	8.7
15369.4	18.44	0.28	31.1	4.5	24.25	0.44	40.0	3.0
15360.4	15.74	0.75	13.7	5.0	-	-	-	-
15343.4	15.97	0.87	17.5	4.5	20.61	0.16	35.7	13.1
15328.4	16.42	0.63	65.3	1.9	-	-	-	-
15302.4	15.98	0.87	23.6	6.9	20.60	0.25	76.7	6.6
15295.4	19.55	0.33	6.3	6.5	24.25	0.39	29.8	4.3
15287.4	16.42	0.67	100.0	0.0	21.01	0.34	89.4	3.2
15270.4	17.93	0.49	30.3	2.7	22.30	0.31	100.0	0.0
15254.4	19.57	0.25	15.1	6.1	-	-	-	-
15215.4	15.45	0.81	12.6	4.2	-	-	-	-
15198.4	15.83	0.75	43.7	1.0	20.94	0.26	55.6	8.9
15182.4	17.35	0.84	14.7	5.2	-	-	-	-
15166.4	18.95	0.37	7.4	10.5	24.23	0.46	45.9	6.6
15052.3	15.42	0.89	13.8	4.2	20.59	0.32	45.1	6.6

15037.3	15.86	0.89	41.1	4.9	20.96	0.30	48.7	8.5
15020.3	17.35	0.64	16.6	8.9	22.28	0.27	41.2	6.8
15011.3	15.47	0.72	13.9	5.3	20.58	0.20	46.5	3.6
14996.3	15.83	0.61	42.2	1.4	20.94	0.24	56.7	13.9
14979.3	17.30	0.72	21.3	6.8	-	-	-	-
14963.4	18.90	0.28	9.5	9.1	24.20	0.30	37.3	5.7
14874.3	15.84	0.87	6.6	12.2	-	-	-	-
14849.3	15.49	0.99	11.1	4.7	20.61	0.18	37.5	8.8
14833.2	15.87	0.81	48.6	2.0	-	-	-	-
14760.3	-	-	-	-	20.59	0.17	24.4	14.0
14745.3	-	-	-	-	20.98	0.28	33.3	3.2
14672.3	-	-	-	-	21.00	0.26	26.3	0.2
14646.2	15.45	0.94	18.5	8.4	20.60	0.24	47.2	13.1
14630.2	15.86	0.75	11.2	8.6	20.98	0.31	48.8	3.7
14614.3	17.36	0.51	25.1	1.7	22.28	0.31	52.9	2.0
14598.1	-	-	-	-	24.25	0.37	29.5	11.8

List of tables and figures

List of figures

Figure 1. Human placental structure. (A) Structure and orientation of fetus and placenta in uterus. (B) Enlargement of box in panel A. [26] 20

Figure 2. Different sources of hCG production: syncytiotrophoblast (ST in green) and invasive extravillous cytotrophoblast (EVT in red). Modified from [9] 21

Figure 3. hCG structure with its two subunits, α in red and β in blue. Modified from [33] 23

Figure 4. Scheme of hCG subunits assembly process that occurs in vivo. Modified from [34] 24

Figure 5. Three dimensional structure of deglycosylated hCG shown by X-ray defraction .*O* and *N* mark *O*- and *N*-linked oligosaccharide sites. Grey is hCGα and black is hCGß [22] 25

Figure 6. Representation of the N-glycan family and the common core structure. 26

Figure 7. A: Representation of the O-glycan family. B: Frequently occurring core structures of mucin-type O-linked glycans. 27

Figure 8. IFSS nomenclature of hCG and its closely related molecules. S-S represents disulfide bonds, CTP shows carboxyl-terminal peptide. N-linked carbohydrates are annotated by  while O-linked carbohydrates are annotated by  [39] 28

Figure 9. Representation that summarizes the structures of hCG produced by placental cells and related molecules in the placenta, urine and blood [44] 30

Figure 10. Cartoon Illustrating the Overall Fold of FSH [48]. 32

Figure 11. Representation of a timeline of major events in the development of gonadotropin-based drugs. [53] 34

Figure 12. Structures and relative intensities of the glycans attached to Asn-13 of hCGβ. Panel A shows oligosaccharides from the urine of a normally pregnant woman at 35 weeks of pregnancy, panel B from

the urine of a patient suffering from invasive mole, and panel C from the urine of a patient with testicular cancer. The proposed sugar structures have been calculated by the mass-matching approach Symbols: , GlcNAc; , Man; , Gal; , Fuc; , NeuAc. Fuc is attached to the first GlcNAc within the structures when present. [73] 44

Figure 13. Separation of the N-glycans of rFSH by HPAEC-MALDI-TOF-MS. (A) rFSH batch with a high biospecific activity; (B) rFSH batch with a low biospecific activity. Triangles: sialic acid; closed circles: galactose; closed squares: N-acetylgalactosamine; open circles: mannose. The attribution of the glycan structures to the peaks is deduced from MALDI-TOF mass spectra [21]. 49

Figure 14. Effect of various concentrations of diaminopropane on the electrophoretic migration and CZE separation of hCG 2292 (4 mg/ml). The DAP concentrations used were 0 (A), 1.0 (B), 2.5 (C) and 5.0 mM (D). UV detection. [76] 55

Figure 15. CZE-MS analysis of (A) rhCG and (B) uhCG [17] 56

Figure 16. TIC (a) and the averaged mass spectra (b-d) extracted at the positions indicated in (a) showing the separation of hCGα and hCGβafter the RPLC-MS [32] 58

Figure 17. Deconvoluted spectrum of (A) rhCG and (B) uhCG between, with the different possible N-glycan structures obtained by RPLC-HRMS [16] 59

Figure 18. (A) TIC of hCG analyzed by nanoLC-HRMS. (B) Average mass spectrum corresponding to the totality of the chromatographic peak observed (retention time range: 15.5-21 min). Conditions: see section 2.3. rhCG, batch 1, run 1. 79

Figure 19. Zoom on average mass spectrum on the mass range [1825.6 – 1828.0 Da] showing the experimental isotopic profile (m/z = 1825.78) corresponding to the most commonly described hCGα glycoforms (hCGα with 2 HexNAc(4)Hex(5)NeuAc(2) N-glycans). 81

Figure 20. Deconvoluted average mass spectrum, determined over the entire chromatographic peak observed in Figure 1, representing the 12 detected hCGα glycoforms and the N-glycan structures of the 9 glycoforms that were identified among the 12. 82

Figure 21. Comparison of the relative intensity of the glycoforms of hCGα present in 3 batches of rhCG analyzed by nanoLC-HRMS (n=3 for each). The comparison is based on the average relative area of each glycoform characterized by its MW and ordered by their retention times. 87

Figure 22. A: TIC of FSH obtained by nanoLC-HRMS. B: average deconvoluted mass spectrum for the 17.5-20 min range. rFSH, batch 1, run 1. Conditions: see section 2.3. 91

Figure 23. Comparison of the glycoforms of FSHα present in 3 batches of rFSH analyzed by nanoLC-HRMS (n=3 for each). The comparison is based on the average relative area of each glycoform characterized by its mass and ordered by their retention time. 93

Figure 24. A: Venn diagram comparing the number of the α-subunit glycoforms identified by the analysis of rhCG and rFSH by nanoLC-HRMS. B: Pie charts representing the average relative area of each detected glycoform in (Top) rhCG and (Bottom) rFSH. For each drug, 3 batches were analyzed 3 times each (n=9). 95

Figure 25. Mean value of peak areas (and standard deviation, n=3 injections) of 6 hCGα glycoforms obtained after injecting 35 ng of rhCG using a 70 nl loop or after the preconcentration of 1, 1.5 or 2 µl of sample on the trapping precolumn. Mobile phase: 2% formic acid in H_2O/ACN mixture (v/v). Gradient: from 10 to 30% ACN with a slope of 0.66% ACN/min. Injection mode: µl pickup. Other conditions: see section 2.3 109

Figure 26. Mean value of peak areas (and standard deviations (n=3)) of 9 hCGα glycoforms analyzed in nano-LC-HRMS by injecting 1 µl of rhCG at 35 µg ml-1 with preconcentration on the trapping precolumn using three different injection modes . Mobile phase: 2% Formic acid in H_2O/ACN mixture (v/v). Gradient: from 10 to 30% ACN with a slope of 0.66% ACN/min. Other conditions: see section 2.3. 111

Figure 27. BPCs of rhCG analyzed by nano-LC-HRMS with either FA or TFA as the mobile phase acidifier. Mobile phase: 2% FA, 0.05% TFA, or 0.1% TFA in H2O/ACN mixture (v/v). Gradient slope: 1.02% ACN/min. rhCG at 100 µg ml-1. Injection volume: 1 µl in partial loop mode. Other conditions: see section 2.3. 114

Figure 28. Venn diagram comparing the number of hCGα and hCGβ glycoforms detected after the analysis by nanoLC-HRMS of rhCG according to the nature and content of acidic additive in the mobile phase. Analytical conditions: see Figure 3. 115

Figure 29. XIC of 4 high intensity hCGα glycoforms, 4 low intensity hCGα and 4 hCGβ glycoforms analyzed by nano-LC-HRMS with all the optimized conditions. Mobile phase: 0.05% TFA in H2O/ACN mixture (v/v). Gradient: from 10 to 65% ACN with a slope of 1.02% ACN/min. rhCG at 100 µg ml^{-1}. Injection volume: 1 µl in partial loop mode. 117

Figure 30. Venn diagrams comparing the number of hCGα and hCGβ glycoforms detected after the analysis of CG10, orb80731, and rhCG by nanoLC-MS according to the nature and content of acidic additive in the mobile phase. Other conditions: see section 2.3. 133

Figure 31. BPCs of Ovitrelle® (rhCG), CG10, Orb80731, Luveris®, Puregon®, and Merional®, analyzed by nanoLC-MS with 0.05% TFA or 2% FA as the mobile phase acidifier. Other conditions: see section 2.3. 137

Figure 32. Venn diagrams comparing the total number of hCGα and hCGβ glycoforms detected after the analysis of 3 hCG-based samples (CG10, orb80731 and Ovitrelle® (rhCG)) by nanoLC-MS with both methods using 2% FA or 0.05% TFA in the mobile phase. Other conditions: see section 2.3. 138

Figure 33. Venn diagrams comparing the number of glycoforms detected after the analysis of Luveris® (LH-based sample), Puregon® (FSH-based sample), and Merional® (LH- and FSH-based sample) by nanoLC-MS according to the nature and content of acidic additive in the mobile phase. Other conditions: see section 2.3. 141

Figure 34. Immunoextraction procedure. Modified from [126]..... 150

Figure 35. Dual-immunoextraction method applied by Egelant et al 2016 [65]. 151

Figure 36. Relative areas of the high intensity hCGα glycoforms in the immunoextracted fractions obtained with the anti-hCGβ pAbs-based IS. Percolation of 10 µg of rhCG on 35 mg of grafted sorbent. 164

Figure 37. Relative areas of the high intensity hCGα glycoforms in the immunoextracted fractions obtained with the anti-hCGα mAbs-based IS. Percolation of 10 µg of rhCG on 35 mg of grafted sorbent. 165

Figure 38. Extraction recoveries of each hCGα glycoform determined in the elution fraction (%) obtained with both mAbs and pAbs-based ISs. Percolation of 5 µg of rhCG on 70 mg of grafted sorbent. 167

Figure 39. Extraction recoveries of each hCGβ glycoform determined in the elution fraction (%) obtained with both mAbs and pAbs-based ISs. Percolation of 5 µg of rhCG on 70 mg of grafted sorbent. 168

List of tabes

Table 1: The biological functions of hCG during pregnancy. Adapted from [22] 22

Table 2 Compiling of the various analytical studies of gonadotropins using bottom-up approach..... 38

Table 3 Compiling several studies of gonadotropins using released-glycans method. 46

Table 4 Compiling of studies involving separation methods to analyze gonadotropins at the intact level. 52

Table 5 Identified O-glycan structures reported in gonadotropin literature, their monoisotopic masses, composition, and method of detection. 61

Table 6 Identified N-glycan structures reported in gonadotropin literature, their monoisotopic masses, composition, and method of detection. 62

Table 7 MW of each hCGα glycoform detected after nanoLC-HRMS analysis of rhCG (3 batches, 3 runs each) and use of the second data treatment approach with glycan structure suggestions for some of them in comparison with bottom-up data (marked with *) or with GlyGen database (marked with **), while when those confirmed by both are marked with ***. x: no structure to propose. The notation 1st and 2nd glycan is not correlated to their localization on the 2 N-glycosylation sites of hCGα. 84

Table 8 Number of detected hCGα and hCGβ glycoforms analyzed by nano-LC-HRMS with five different gradient slopes. Mobile phase: 2% Formic acid in H2O/ACN mixture (v/v). rhCG at 100 µg ml-1. Injection volume 1 µl in partial loop mode. Other conditions: see section 2.3. 113

Table 9 RSD range and its average value for the retention times and relative peak areas of the hCGα and hCGβ glycoforms detected after the analysis by nano-LC-HRMS with the optimized conditions of rhCG at 8 different concentrations, in triplicate for each. Mobile phase: 0.05% TFA in H_2O/ACN mixture (v/v). Gradient: from 10 to 65% ACN with a slope of 1.02% ACN/min. rhCG at 100 µg ml^{-1}. Injection volume: 1 µl in partial loop mode. The hCGα glycoforms are said of high intensity is their relative area is above 15%

of the relative area of the hCGα glycoform giving the highest peak area. Similarly, the hCGα glycoforms are said of low intensity is their relative area is below 15%. 118

Table 10 MW of the identified hCGα glycoforms detected after nano-LC-HRMS analysis of rhCG in triplicate and use of the new optimized method with glycan structure suggestions for some in comparison with bottom-up data (marked with *) or with GlyGen (marked with**), while those confirmed by both are marked with ***. The notation 1st and 2nd glycan is not correlated to their localization on the 2 N-glycosylation sites of hCGα. The glycoforms were highlighted in dark gray when addition of two oxygen atoms was suspected while those highlighted in light gray were suspected to have the addition of only one oxygen atom. 120

Table 11 Gathering of the number of detected glycoforms and their corresponding MW and retention ranges for rhCG, CG10 and Orb80731 hCG-based samples, Luveris (LH based-sample), Puregon (FSH-based sample), and Merional® (LH- and FSH-based sample), analyzed by nanoLC-MS with either 2% FA or 0.05% TFA as mobile phase additive. 132

Table 12 Summary of the various studies implementing the immunoextraction of hCG. 153

References

[1] V. Wittmann, Glycopeptides and Glycoproteins Synthesis, Structure, and Application. Springer Berlin Heidelberg New York: Springer, 2007.

[2] Y. Mechref and D. C. Muddiman, “Recent advances in glycomics, glycoproteomics and allied topics,” Anal Bioanal Chem, vol. 409, no. 2, pp. 355–357, Jan. 2017, doi: 10.1007/s00216-016-0093-9.

[3] L. Valmu, H. Alfthan, K. Hotakainen, S. Birken, and U.-H. Stenman, “Site-specific glycan analysis of human chorionic gonadotropin -subunit from malignancies and pregnancy by liquid chromatography--electrospray mass spectrometry,” Glycobiology, vol. 16, no. 12, pp. 1207–1218, Jul. 2006, doi: 10.1093/glycob/cwl034.

[4] M. K. Sethi and J. Zaia, “Extracellular matrix proteomics in schizophrenia and Alzheimer’s disease,” Anal Bioanal Chem, vol. 409, no. 2, pp. 379–394, Jan. 2017, doi: 10.1007/s00216-016-9900-6.

[5] R. Goulabchand, T. Vincent, F. Batteux, J. Eliaou, and P. Guilpain, “Impact of autoantibody glycosylation in autoimmune diseases,” Autoimmunity Reviews, vol. 13, no. 7, pp. 742–750, Jul. 2014, doi: 10.1016/j.autrev.2014.02.005.

[6] M. J. Kailemia, D. Park, and C. B. Lebrilla, “Glycans and glycoproteins as specific biomarkers for cancer,” Anal Bioanal Chem, vol. 409, no. 2, pp. 395–410, Jan. 2017, doi: 10.1007/s00216-016-9880-6.

[7] M. Shimojo, R. Sakakibara, and M. Ishiguro, “Immature Human Chorionic Gonadotropin (hCG) in First Trimester Placental Cells Is Bound to an ATP-Binding Protein Forming High-Molecular-Weight hCG,” The Journal of Biochemistry, vol. 114, no. 1, pp. 83–87, Jul. 1993, doi: 10.1093/oxfordjournals.jbchem.a124144.

[8] P. Laidler, R. C. Hider, D. A. Cowan, A. T. Kicman, and A. Keane, “Tryptic mapping of human chorionic gonadotropin by matrix-assisted laser desorption/ionization mass spectrometry,” Rapid Commun. Mass Spectrom., vol. 9, no. 11, pp. 1021–1026, 1995, doi: 10.1002/rcm.1290091110.

[9] T. Fournier, J. Guibourdenche, and D. Evain-Brion, “Review: hCGs: Different sources of production, different glycoforms and functions,” Placenta, vol. 36, pp. S60–S65, Apr. 2015, doi: 10.1016/j.placenta.2015.02.002.

[10] E. Keikkala, P. Vuorela, H. Laivuori, J. Romppanen, S. Heinonen, and U.-H. Stenman, “First trimester hyperglycosylated human chorionic gonadotrophin in serum – A marker of early-onset preeclampsia,” Placenta, vol. 34, no. 11, pp. 1059–1065, Nov. 2013, doi: 10.1016/j.placenta.2013.08.006.

[11] A. W. Butch, B. D. Ahrens, and N. K. Avliyakulov, “Urine reference intervals for human chorionic gonadotropin (hCG) isoforms by immunoextraction-tandem mass spectrometry to detect hCG use,” Drug Test Anal, vol. 10, no. 6, pp. 956–960, Jun. 2018, doi: 10.1002/dta.2333.

[12] C. M. G. Thomas and F. J. L. Reljnders, “Human Choriogonadotropin (hCG): Comparisons between Determinations of Intact hCG, Free hCG /3-Subunit, and ‘Total’ hCG + f3 in Serum during the First Half of High-Risk Pregnancy,” p. 5.

[13] R. Thennati, S. K. Singh, N. Nage, Y. Patel, S. K. Bose, V. Burade, and R. S. Ranbhor, “Analytical characterization of recombinant hCG and comparative studies with reference product,” BTT, vol. Volume 12, pp. 23–35, Jan. 2018, doi: 10.2147/BTT.S141203.

[14] S. Perchepied, N. Eskenazi, C. Giangrande, J. Camperi, T. Fournier, J, Vinh, N. Delaunay, and V. Pichon, “Development of Immobilized Enzyme Reactors for the characterization of the glycosylation heterogeneity of a protein,” Talanta, vol. 206, p. 120171, Jan. 2020, doi: 10.1016/j.talanta.2019.120171.

[15] J. Camperi, “Development of separation methods for the characterization of a glycoprotein at the intact scale: application to the human chorionic hormone gonadotropin,” Thesis manuscript, PSL Research University, Laboratoire Sciences Analytiques, Bioanalytiques et Miniaturisation, 2018.
[16] J. Camperi, A. Combes, J. Guibourdenche, D. Guillarme, V. Pichon, T. Fournier, and N. Delaunay, “An attempt to characterize the human Chorionic Gonadotropin protein by reversed phase liquid chromatography coupled with high-resolution mass spectrometry at the intact level,” Journal of Pharmaceutical and Biomedical Analysis, vol. 161, pp. 35–44, Nov. 2018, doi: 10.1016/j.jpba.2018.07.056.
[17] J. Camperi, B. D. Cock, V. Pichon, A. Combes, J. Guibourdenche, T. Fournier, Y. V. Heyden, D. Mangelings, and N. Delaunay, “First characterizations by capillary electrophoresis of human Chorionic Gonadotropin at the intact level,” Talanta, vol. 193, pp. 77–86, Feb. 2019, doi: 10.1016/j.talanta.2018.09.095.
[18] M. A. Chlenov, E. I. Kandyba, L. V. Nagornaya, I. L. Orlova, and Y. V. Volgin, “High-performance liquid chromatography of human glycoprotein hormones,” Journal of Chromatography A, vol. 631, no. 1–2, pp. 261–267, Feb. 1993, doi: 10.1016/0021-9673(93)80531-C.
[19] A. W. Butch and G. A. Woldemariam, “Urinary human chorionic gonadotropin isoform concentrations in doping control samples: Urinary human chorionic gonadotropin isoform concencentrations in doping control samples,” Drug Test. Analysis, vol. 8, no. 11–12, pp. 1147–1151, Nov. 2016, doi: 10.1002/dta.2061.
[20] A. Stockell Hartree and A. G. C. Renwick, “Molecular structures of glycoprotein hormones and functions of their carbohydrate components,” Biochemical Journal, vol. 287, no. 3, pp. 665–679, Nov. 1992, doi: 10.1042/bj2870665.
[21] A. Gervais, “Glycosylation of human recombinant gonadotrophins: characterization and batch-to-batch consistency,” Glycobiology, vol. 13, no. 3, pp. 179–189, Mar. 2003, doi: 10.1093/glycob/cwg020.
[22] L. A. Cole, “hCG, the wonder of today’s science,” Reprod Biol Endocrinol, vol. 10, no. 1, p. 24, 2012, doi: 10.1186/1477-7827-10-24.
[23] H. Lund, K. Løvsletten, E. Paus, T. G. Halvorsen, and L. Reubsaet, “Immuno–MS Based Targeted Proteomics: Highly Specific, Sensitive, and Reproducible Human Chorionic Gonadotropin Determination for Clinical Diagnostics and Doping Analysis,” Anal. Chem., vol. 84, no. 18, pp. 7926–7932, Sep. 2012, doi: 10.1021/ac301418f.
[24] I. Carrasco-Wong, A. Moller, F. R. Giachini, V. V. Lima, F. Toledo, J. Stojanova, L. Sobrevia, S. San Martín, “Placental structure in gestational diabetes mellitus,” Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease, p. 165535, Aug. 2019, doi: 10.1016/j.bbadis.2019.165535.
[25] R. L. Pemathilaka, D. E. Reynolds, and N. N. Hashemi, “Drug transport across the human placenta: review of placenta-on-a-chip and previous approaches,” Interface Focus, vol. 9, no. 5, p. 20190031, Oct. 2019, doi: 10.1098/rsfs.2019.0031.
[26] V. B. Zeldovich, J. R. Robbins, M. Kapidzic, P. Lauer, and A. I. Bakardjiev, “Invasive Extravillous Trophoblasts Restrict Intracellular Growth and Spread of Listeria monocytogenes,” PLoS Pathog, vol. 7, no. 3, p. e1002005, Mar. 2011, doi: 10.1371/journal.ppat.1002005.
[27] A. T. Kicman, R. V. Brooks, and D. A. Cowan, “Human chorionic gonadotrophin and sport.,” British Journal of Sports Medicine, vol. 25, no. 2, pp. 73–80, Jun. 1991, doi: 10.1136/bjsm.25.2.73.
[28] P. Berger and A. J. Lapthorn, “The molecular relationship between antigenic domains and epitopes on hCG,” Molecular Immunology, vol. 76, pp. 134–145, Aug. 2016, doi: 10.1016/j.molimm.2016.06.015.
[29] A. W. Butch and G. A. Woldemariam, “Urinary human chorionic gonadotropin isoform concentrations in doping control samples: Urinary human chorionic gonadotropin isoform

concencentrations in doping control samples," Drug Test. Analysis, vol. 8, no. 11–12, pp. 1147–1151, Nov. 2016, doi: 10.1002/dta.2061.
[30] L. A. Cole, "Biological functions of hCG and hCG-related molecules," Reprod Biol Endocrinol, vol. 8, no. 1, p. 102, 2010, doi: 10.1186/1477-7827-8-102.
[31] R. Thennati, S. K. Singh, N. Nage, Y. Patel, S. K. Bose, V. Burade, and R. S. Ranbhor, "Analytical characterization of recombinant hCG and comparative studies with reference product," BTT, vol. Volume 12, pp. 23–35, Jan. 2018, doi: 10.2147/BTT.S141203.
[32] H. Toll, P. Berger, A. Hofmann, A. Hildebrandt, H. Oberacher, H. P. Lenhof, and C. G. Huber, "Glycosylation patterns of human chorionic gonadotropin revealed by liquid chromatography-mass spectrometry and bioinformatics," Electrophoresis, vol. 27, no. 13, pp. 2734–2746, Jul. 2006, doi: 10.1002/elps.200600022.
[33] R. Leao and S. Esteves, "Gonadotropin therapy in assisted reproduction: an evolutionary perspective from biologics to biotech," Clinics, vol. 69, no. 4, pp. 279–293, Apr. 2014, doi: 10.6061/clinics/2014(04)10.
[34] M. P. Bernard, W. Lin, V. Kholodovych, and W. R. Moyle, "Human Lutropin (hLH) and Choriogonadotropin (CG) Are Assembled by Different Pathways: A MODEL OF hLH ASSEMBLY," J. Biol. Chem., vol. 289, no. 20, pp. 14360–14369, May 2014, doi: 10.1074/jbc.M113.535609.
[35] U.-H. Stenman and H. Alfthan, "Determination of human chorionic gonadotropin," Best Practice & Research Clinical Endocrinology & Metabolism, vol. 27, no. 6, pp. 783–793, Dec. 2013, doi: 10.1016/j.beem.2013.10.005.
[36] J. B. L. DAMM, H. VOSHOL, and J. VLIEGENTHART, "The β -Subunit of Human Chorionic Gonadotropin Contains N-Glycosidic Trisialo Tri- and Tri'-antennary Carbohydrate Chains," Glycoconjuga te J, vol. 5, pp. 221–233, 1988.
[37] Q. Zhang, Z. Li, Y. Wang, Q. Zheng, and J. Li, "Mass spectrometry for protein sialoglycosylation," Mass Spec Rev, vol. 37, no. 5, pp. 652–680, Sep. 2018, doi: 10.1002/mas.21555.
[38] H. Geyer and R. Geyer, "Strategies for analysis of glycoprotein glycosylation," Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics, vol. 1764, no. 12, pp. 1853–1869, Dec. 2006, doi: 10.1016/j.bbapap.2006.10.007.
[39] A. Bristow, "Establishment, Value Assignment, and Characterization of New WHO Reference Reagents for Six Molecular Forms of Human Chorionic Gonadotropin," Clinical Chemistry, vol. 51, no. 1, pp. 177–182, Nov. 2004, doi: 10.1373/clinchem.2004.038679.
[40] S. Birken, G. Kovalevskaya, and J. O'Connor, "Metabolism of hCG and hLH to multiple urinary forms," Molecular and Cellular Endocrinology, vol. 125, no. 1–2, pp. 121–131, Dec. 1996, doi: 10.1016/S0303-7207(96)03942-1.
[41] U.-H. Stenman, A. Tiitinen, H. Alfthan, and L. Valmu, "The classification, functions and clinical use of different isoforms of HCG," Human Reproduction Update, vol. 12, no. 6, pp. 769–784, Dec. 2006, doi: 10.1093/humupd/dml029.
[42] S. F. de Medeiros and R. J. Norman, "Human choriogonadotrophin protein core and sugar branches heterogeneity: basic and clinical insights," Human Reproduction Update, vol. 15, no. 1, pp. 69–95, Oct. 2008, doi: 10.1093/humupd/dmn036.
[43] B. C. Nisula, D. L. Blithe, A. Akar, G. Lefort, and R. E. Wehmann, "Metabolic fate of human choriogonadotropin," Journal of Steroid Biochemistry, vol. 33, no. 4, pp. 733–737, Oct. 1989, doi: 10.1016/0022-4731(89)90485-8.
[44] L. A. Cole, "Immunoassay of human chorionic gonadotropin, its free subunits, and metabolites," Clinical Chemistry, no. 12, p. 11, 1997.
[45] J. Smitz, C. Wolfenson, S. Chappel, and J. Ruman, "Follicle-Stimulating Hormone: A Review of Form and Function in the Treatment of Infertility," Reprod Sci, vol. 23, no. 6, pp. 706–716, Jun. 2016, doi: 10.1177/1933719115607992.

[46] R. F. Loureiro, J. E. de Oliveira, P. A. Torjesen, P. Bartolini, and M. T. C. P. Ribela, “Analysis of intact human follicle-stimulating hormone preparations by reversed-phase high-performance liquid chromatography,” Journal of Chromatography A, vol. 1136, no. 1, pp. 10–18, Dec. 2006, doi: 10.1016/j.chroma.2006.09.037.
[47] A. Conforti, A. Vaiarelli, D. Cimadomo, F. Bagnulo, S. Peluso, L. Carbone, F.Di Rella, G. De Placido, F. Ubaldi, I. Huhtaniemi, and C. Alviggi , “Pharmacogenetics of FSH Action in the Female,” Front. Endocrinol., vol. 10, p. 398, Jun. 2019, doi: 10.3389/fendo.2019.00398.
[48] K. M. Fox, J. A. Dias, and P. V. Roey, “Three-Dimensional Structure of Human Follicle-Stimulating Hormone,” vol. 15, no. 3, p. 12, 2018.
[49] X. Jiang, J. A. Dias, and X. He, “Structural biology of glycoprotein hormones and their receptors: Insights to signaling,” Molecular and Cellular Endocrinology, vol. 382, no. 1, pp. 424–451, Jan. 2014, doi: 10.1016/j.mce.2013.08.021.
[50] P. R. Manna, “Synthesis, purification and structural and functional characterization of recombinant form of a common genetic variant of human luteinizing hormone,” Human Molecular Genetics, vol. 11, no. 3, pp. 301–315, Feb. 2002, doi: 10.1093/hmg/11.3.301.
[51] C. M. Carvalho, J.E. Oliveira, B.E. Almeida, E.K.M. Ueda, P.A. Torjesen, P. Bartolini, M.T.C.P. Ribela, “Efficient isolation of the subunits of recombinant and pituitary glycoprotein hormones,” Journal of Chromatography A, vol. 1216, no. 9, pp. 1431–1438, Feb. 2009, doi: 10.1016/j.chroma.2008.12.096.
[52] B. E. Almeida, J. E. Oliveira, C. M. Carvalho, S. L. Dalmora, P. Bartolini, and M. T. C. P. Ribela, “Analysis of human luteinizing hormone and human chorionic gonadotropin preparations of different origins by reversed-phase high-performance liquid chromatography,” Journal of Pharmaceutical and Biomedical Analysis, vol. 53, no. 1, pp. 90–97, Sep. 2010, doi: 10.1016/j.jpba.2010.03.013.
[53] B. Lunenfeld, W. Bilger, S. Longobardi, V. Alam, T. D’Hooghe, and S. K. Sunkara, “The Development of Gonadotropins for Clinical Use in the Treatment of Infertility,” Front. Endocrinol., vol. 10, p. 429, Jul. 2019, doi: 10.3389/fendo.2019.00429.
[54] C.-L. Liu and L. D. Bowers, “Immunoaffinity trapping of urinary human chorionic gonadotropin and its high-performance liquid chromatographic-mass spectrometric confirmation,” Journal of Chromatography B: Biomedical Sciences and Applications, vol. 687, no. 1, pp. 213–220, Dec. 1996, doi: 10.1016/S0378-4347(96)00302-7.
[55] C. Liu and L. D. Bowers, “Mass spectrometric characterization of nicked fragments of the $^{N}_{L}$-subunit of human chorionic gonadotropin,” Clinical Chemistry, no. 7, p. 10, 1997.
[56] C. Liu and L. D. Bowers, “Mass Spectrometric Characterization of the b- Subunit of Human Chorionic Gonadotropin,” p. 10.
[57] M.-T. Vu-Hai, J.C. Huet, K. Echasserieau, J.M. Bidart, C. Floiras, J.C. Pernollet, and E. Milgrom, “Posttranslational Modifications of the Lutropin Receptor: Mass Spectrometric Analysis †,” Biochemistry, vol. 39, no. 18, pp. 5509–5517, May 2000, doi: 10.1021/bi992913f.
[58] L.-H. Gam, S.-Y. Tham, and A. Latiff, “Immunoaffinity extraction and tandem mass spectrometric analysis of human chorionic gonadotropin in doping analysis,” Journal of Chromatography B, vol. 792, no. 2, pp. 187–196, Jul. 2003, doi: 10.1016/S1570-0232(03)00264-2.
[59] H. Lund, S. B. Torsetnes, E. Paus, K. Nustad, L. Reubsaet, and T. G. Halvorsen, “Exploring the Complementary Selectivity of Immunocapture and MS Detection for the Differentiation between hCG Isoforms in Clinically Relevant Samples,” J. Proteome Res., vol. 8, no. 11, pp. 5241–5252, Nov. 2009, doi: 10.1021/pr900580n.
[60] H. Lund, A. H. Snilsberg, E. Paus, T. G. Halvorsen, P. Hemmersbach, and L. Reubsaet, “Sports drug testing using immuno-MS: clinical study comprising administration of human chorionic

gonadotropin to males," Anal Bioanal Chem, vol. 405, no. 5, pp. 1569–1576, Feb. 2013, doi: 10.1007/s00216-012-6566-6.
[61] H. Lund, E. Paus, P. Berger, U-H. Stenman, T. Torcellini, T. G. Halvorsen, L. Reubsaet, "Epitope analysis and detection of human chorionic gonadotropin (hCG) variants by monoclonal antibodies and mass spectrometry," Tumor Biol., vol. 35, no. 2, pp. 1013–1022, Feb. 2014, doi: 10.1007/s13277-013-1135-y.
[62] G. A. Woldemariam and A. W. Butch, "Immunoextraction–Tandem Mass Spectrometry Method for Measuring Intact Human Chorionic Gonadotropin, Free β-Subunit, and β-Subunit Core Fragment in Urine," Clinical Chemistry, vol. 60, no. 8, pp. 1089–1097, Aug. 2014, doi: 10.1373/clinchem.2014.222703.
[63] X. Bai, D. Li, J. Zhu, Y. Guan, Q. Zhang, and L. Chi, "From individual proteins to proteomic samples: characterization of O-glycosylation sites in human chorionic gonadotropin and human-plasma proteins," Anal Bioanal Chem, vol. 407, no. 7, pp. 1857–1869, Mar. 2015, doi: 10.1007/s00216-014-8439-7.
[64] S. V. Egeland, L. Reubsaet, and T. G. Halvorsen, "The pros and cons of increased trypsin-to-protein ratio in targeted protein analysis," Journal of Pharmaceutical and Biomedical Analysis, vol. 123, pp. 155–161, May 2016, doi: 10.1016/j.jpba.2016.02.011.
[65] S. V. Egeland, L. Reubsaet, E. Paus, and T. G. Halvorsen, "Dual-immuno-MS technique for improved differentiation power in heterodimeric protein biomarker analysis: determination and differentiation of human chorionic gonadotropin variants in serum," Anal Bioanal Chem, vol. 408, no. 26, pp. 7379–7391, Oct. 2016, doi: 10.1007/s00216-016-9818-z.
[66] M. C. S. Levernæs, M. N. Broughton, L. Reubsaet, and T. G. Halvorsen, "To elute or not to elute in immunocapture bottom-up LC–MS," Journal of Chromatography B, vol. 1055–1056, pp. 51–60, Jun. 2017, doi: 10.1016/j.jchromb.2017.03.044.
[67] C. Rosting, E. V. Tran, A. Gjelstad, and T. G. Halvorsen, "Determination of the low-abundant protein biomarker hCG from dried matrix spots using immunocapture and nano liquid chromatography mass spectrometry," Journal of Chromatography B, vol. 1077–1078, pp. 44–51, Mar. 2018, doi: 10.1016/j.jchromb.2018.01.016.
[68] V. Borromeo, A. Berrini, F. De Grandi, F. Cremonesi, N. Fiandanese, P. Pocar , C. Secchi, "A novel monoclonal antibody-based enzyme-linked immunosorbent assay to determine luteinizing hormone in bovine plasma," Domestic Animal Endocrinology, vol. 48, pp. 145–157, Jul. 2014, doi: 10.1016/j.domaniend.2014.03.004.
[69] G. R. Bousfield, V. Y. Butnev, J.-M. Bidart, D. Dalpathado, J. Irungu, and H. Desaire, "Chromatofocusing Fails To Separate hFSH Isoforms on the Basis of Glycan Structure †," Biochemistry, vol. 47, no. 6, pp. 1708–1720, Feb. 2008, doi: 10.1021/bi701764w.
[70] H. Wang, X. Chen, X. Zhang, W. Zhang, Y. Li, H. Yin, H. Shao, and G. Chen, "Comparative Assessment of Glycosylation of a Recombinant Human FSH and a Highly Purified FSH Extracted from Human Urine," J. Proteome Res., vol. 15, no. 3, pp. 923–932, Mar. 2016, doi: 10.1021/acs.jproteome.5b00921.
[71] A. Van Dorsselaer, C. Carapito, F. Delalande, C. Schaeffer-Reiss, D.Thierse, H. Diemer, D. S. McNair, D. Krewski, N. R. Cashman, "Detection of Prion Protein in Urine-Derived Injectable Fertility Products by a Targeted Proteomic Approach," PLoS ONE, vol. 6, no. 3, p. e17815, Mar. 2011, doi: 10.1371/journal.pone.0017815.
[72] V. Dotz , R. Haselberg , A. Shubhakar , R. P. Kozak , D. Falck , Y. Rombouts , D. Reusch , G. W. Somsen , D. L. Fernandes , and M. Wuhrer, "Mass spectrometry for glycosylation analysis of biopharmaceuticals," TrAC Trends in Analytical Chemistry, vol. 73, pp. 1–9, Nov. 2015, doi: 10.1016/j.trac.2015.04.024.

[73] L. Valmu, H. Alfthan, K. Hotakainen, S. Birken, and U.-H. Stenman, “Site-specific glycan analysis of human chorionic gonadotropin -subunit from malignancies and pregnancy by liquid chromatography--electrospray mass spectrometry,” Glycobiology, vol. 16, no. 12, pp. 1207–1218, Jul. 2006, doi: 10.1093/glycob/cwl034.

[74] K Hård, J B Damm, M P Spruijt, A A Bergwerff, J P Kamerling, G W Van Dedem, and J F Vliegenthart, “The carbohydrate chains of the beta subunit of human chorionic gonadotropin produced by the choriocarcinoma cell line BeWo. Novel O-linked and novel bisecting-GlcNAc-containing N-linked carbohydrates,” Eur J Biochem, vol. 205, no. 2, pp. 785–798, Apr. 1992, doi: 10.1111/j.1432-1033.1992.tb16843.x.

[75] M. Pabst and F. Altmann, “Glycan analysis by modern instrumental methods,” Proteomics, vol. 11, no. 4, pp. 631–643, Feb. 2011, doi: 10.1002/pmic.201000517.

[76] D. E. Morbeck, B. J. Madden, and D. J. McCormick, “Analysis of the microheterogeneity of the glycoprotein chorionic gonadotropin with high-performance capillary electrophoresis,” Journal of Chromatography A, vol. 680, no. 1, pp. 217–224, Sep. 1994, doi: 10.1016/0021-9673(94)80070-7.

[77] R. Bassett, C. De Bellis, L. Chiacchiarini, D. Mendola, E. Micangeli, K. Minari, L. Grimaldi, M. Mancinelli, R. Mastrangeli and R. Bucci, “Comparative characterisation of a commercial human chorionic gonadotrophin extracted from human urine with a commercial recombinant human chorionic gonadotrophin,” Current Medical Research and Opinion, vol. 21, no. 12, pp. 1969–1976, Dec. 2005, doi: 10.1185/030079905X75005.

[78] M. Terenghi, L. Elviri, M. Careri, A. Mangia, and R. Lobinski, “Multiplexed Determination of Protein Biomarkers Using Metal-Tagged Antibodies and Size Exclusion Chromatography−Inductively Coupled Plasma Mass Spectrometry,” Anal. Chem., vol. 81, no. 22, pp. 9440–9448, Nov. 2009, doi: 10.1021/ac901853g.

[79] J. Camperi, V. Pichon, T. Fournier, and N. Delaunay, “First profiling in hydrophilic interaction liquid chromatography of intact human chorionic gonadotropin isoforms,” Journal of Pharmaceutical and Biomedical Analysis, vol. 174, pp. 495–499, Sep. 2019, doi: 10.1016/j.jpba.2019.06.014.

[80] B. E. Almeida, J. E. Oliveira, R. Damiani, S. L. Dalmora, P. Bartolini, and M. T. C. P. Ribela, “A pilot study on potency determination of human follicle-stimulating hormone: A comparison between reversed-phase high-performance liquid chromatography method and the in vivo bioassay,” Journal of Pharmaceutical and Biomedical Analysis, vol. 54, no. 4, pp. 681–686, Mar. 2011, doi: 10.1016/j.jpba.2010.10.018.

[81] J. Camperi, V. Pichon, and N. Delaunay, “Separation methods hyphenated to mass spectrometry for the characterization of the protein glycosylation at the intact level,” Journal of Pharmaceutical and Biomedical Analysis, vol. 178, p. 112921, Jan. 2020, doi: 10.1016/j.jpba.2019.112921.

[82] J. Camperi, A. Combès, T. Fournier, V. Pichon, and N. Delaunay, “Analysis of the human chorionic gonadotropin protein at the intact level by HILIC-MS and comparison with RPLC-MS,” Anal Bioanal Chem, vol. 412, no. 18, pp. 4423–4432, Jul. 2020, doi: 10.1007/s00216-020-02684-8.

[83] G. Weisshaar, J. H. Hiyama, and A. G. C. Renwick, “Site-specific N-glycosylation of human chorionic gonadotrophin— structural analysis of glycopeptides by one- and two-dimensional H NMR spectroscopy,” Glycobiology, vol. Vol. 1, no. 4, pp. 393-404, 1991.

[84] M. M. Elliott, A. Kardana, J. W. Lustbader, and L. A. Cole’, “Carbohydrate and Peptide Structure of the and 13-Subunitsof Human Chorionic Gonadotropin from Normal and Aberrant Pregnancy and Choriocarcinoma,” vol. 7, no. 1, p. 18.

[85] M. Kessler, T. Mise, R. Ghai, and O. Bahl, “Structure and location of the O-glycosydic carbohydrate units of Human Chorionic Gonadotropin,” vol. 254, no. August 25, pp. 7909–7914, 1979.

[86] C.-L. Liu and L. D. Bowers, “Immunoaffinity trapping of urinary human chorionic gonadotropin and its high-performance liquid chromatographic-mass spectrometric confirmation,” Journal of Chromatography B: Biomedical Sciences and Applications, vol. 687, no. 1, pp. 213–220, Dec. 1996, doi: 10.1016/S0378-4347(96)00302-7.

[87] T de Beer, C W van Zuylen, K Hård, R Boelens, R Kaptein, J P Kamerling, and J F Vliegenthart, “Rapid and simple approach for the NMR resonance assignment of the carbohydrate chains of an intact glycoprotein Application of gradient-enhanced natural abundance 1 H- 13 C HSQC and HSQC-TOCSY to the α-subunit of human chorionic gonadotropin,” FEBS Letters, vol. 348, no. 1, pp. 1–6, Jul. 1994, doi: 10.1016/0014-5793(94)00547-8.

[88] J. G. Pierce and T. F. Parsons, “Glycoprotein Hormones: Structure and Function,” Annual Review of Biochemistry, vol. 50, no. 1, pp. 465–495, 1981, doi: 10.1146/annurev.bi.50.070181.002341.

[89] V. Kolli, K. N. Schumacher, and E. D. Dodds, “Engaging challenges in glycoproteomics: recent advances in MS-based glycopeptide analysis,” Bioanalysis, vol. 7, no. 1, pp. 113–131, Jan. 2015, doi: 10.4155/bio.14.272.

[90] R. Zhu, L. Zacharias, K. M. Wooding, W. Peng, and Y. Mechref, “Glycoprotein Enrichment Analytical Techniques,” in Methods in Enzymology, vol. 585, Elsevier, 2017, pp. 397–429.

[91] E. Largy, F. Cantais, G. Van Vyncht, A. Beck, and A. Delobel, “Orthogonal liquid chromatography–mass spectrometry methods for the comprehensive characterization of therapeutic glycoproteins, from released glycans to intact protein level,” Journal of Chromatography A, vol. 1498, pp. 128–146, May 2017, doi: 10.1016/j.chroma.2017.02.072.

[92] D. Thakur, T. Rejtar, B. L. Karger, N. Washburn, C.J. Bosques, N.S. Gunay, Z. Shriver, and G. Venkataraman, “Profiling the Glycoforms of the Intact α Subunit of Recombinant Human Chorionic Gonadotropin by High-Resolution Capillary Electrophoresis−Mass Spectrometry,” Anal. Chem., vol. 81, no. 21, pp. 8900–8907, Nov. 2009, doi: 10.1021/ac901506p.

[93] P. L. Storring, R. E. Gaines Das, J. W. M. Mulders, and M. Halder, “Physicochemical Methods for Predicting the Biological Potency of Recombinant Follicle Stimulating Hormone: An International Collaborative Study of Isoelectric Focusing and Capillary Zone Electrophoresis,” Biologicals, vol. 30, no. 3, pp. 217–234, Sep. 2002, doi: 10.1006/biol.2002.0333.

[94] A. Lombardi, C. Andreozzi, V. Pavone, V. Triglione, L. Angiolini, and P. Caccia, “Evaluation of the oligosaccharide composition of commercial follicle stimulating hormone preparations: Liquid Phase Separations,” ELECTROPHORESIS, vol. 34, no. 16, pp. 2394–2406, Aug. 2013, doi: 10.1002/elps.201300045.

[95] S. R. Wison, T. Vehus, and H. Sjaanes Berg, “Nano-LC in proteomics: recent advances and approaches.” Bioanalysis, 14), 1799–1815 2015.

[96] M. van Scherpenzeel, G. Steenbergen, E. Morava, R. A. Wevers, and D. J. Lefeber, “High-resolution mass spectrometry glycoprofiling of intact transferrin for diagnosis and subtype identification in the congenital disorders of glycosylation,” Translational Research, vol. 166, no. 6, pp. 639-649.e1, Dec. 2015, doi: 10.1016/j.trsl.2015.07.005.

[97] J. F. M. Jacobs, “Fast, robust and high-resolution glycosylation profiling of intact monoclonal IgG antibodies using nanoLC-chip-QTOF,” p. 27.

[98] G. Xu, J. Stupak, L. Yang, L. Hu, B. Guo, and J. Li, “Deconvolution in mass spectrometry based proteomics,” Rapid Commun Mass Spectrom, vol. 32, no. 10, pp. 763–774, May 2018, doi: 10.1002/rcm.8103.

[99] V. Blanchard, R. A. Gadkari, G. J. Gerwig, B. R. Leeflang, R. R. Dighe, and J. P. Kamerling, “Characterization of the N-linked oligosaccharides from human chorionic gonadotropin expressed in the methylotrophic yeast Pichia pastoris,” Glycoconj J, vol. 24, no. 1, pp. 33–47, Dec. 2006, doi: 10.1007/s10719-006-9010-3.

[100] “Isotope Distribution Calculator, Mass Spec Plotter, Isotope Abundance Graphs.” Isotope Distribution Calculator, Mass Spec Plotter, Isotope Abundance Gr https://www.sisweb.com/mstools/isotope.htm (accessed Mar. 09, 2020).
[101] GlyGen database. www. glygen.org accessed on Thursday, December 5,2019. .
[102] D. Liu, X. Han, Xi. Liu, M. Cheng, M. He, G. Rainer, H. Gao, and X. Zhang, “Measurement of ultra-trace level of intact oxytocin in plasma using SALLE combined with nano-LC–MS,” Journal of Pharmaceutical and Biomedical Analysis, vol. 173, pp. 62–67, Sep. 2019, doi: 10.1016/j.jpba.2019.04.023.
[103] Y. Mechref and D. C. Muddiman, “Recent advances in glycomics, glycoproteomics and allied topics,” Anal Bioanal Chem, vol. 409, no. 2, pp. 355–357, Jan. 2017, doi: 10.1007/s00216-016-0093-9.
[104] J. Guibourdenche, K. Handschuh, V. Tsatsaris, P. Gerbaud, M. C. Leguy, F. Muller, D. Evain Brion, and T. Fournier, “Hyperglycosylated hCG Is a Marker of Early Human Trophoblast Invasion,” J Clin Endocrinol Metab, p. 5, 2010.
[105] S. Berndt, S. Blacher, C. Munaut, J. Detilleux, S. Perrier d’Hauterive, I. Huhtaniemi, D. Evain-Brion, A. Noël, T. Fournier, and J-M. Foidart, “Hyperglycosylated human chorionic gonadotropin stimulates angiogenesis through TGF-β receptor activation,” FASEB j., vol. 27, no. 4, pp. 1309–1321, Apr. 2013, doi: 10.1096/fj.12-213686.
[106] I. Sanchez-De Melo, P. Grassi, F. Ochoa, J. Bolivar, F. J. García-Cózar, and M. C. Durán-Ruiz, “N-glycosylation profile analysis of Trastuzumab biosimilar candidates by Normal Phase Liquid Chromatography and MALDI-TOF MS approaches,” Journal of Proteomics, vol. 127, pp. 225–233, Sep. 2015, doi: 10.1016/j.jprot.2015.04.012.
[107] A. Al Matari, A. Combès, J. Camperi, T. Fournier, V. Pichon, and N. Delaunay, “Identification and semi-relative quantification of intact glycoforms by nano-LC–(Orbitrap)MS: application to the α-subunit of human chorionic gonadotropin and follicle-stimulating hormone,” Anal Bioanal Chem, Jul. 2020, doi: 10.1007/s00216-020-02794-3.
[108] X. Bai, D. Li, J. Zhu, Y. Guan, Q. Zhang, and L. Chi, “From individual proteins to proteomic samples: characterization of O-glycosylation sites in human chorionic gonadotropin and human-plasma proteins,” Anal Bioanal Chem, vol. 407, no. 7, pp. 1857–1869, Mar. 2015, doi: 10.1007/s00216-014-8439-7.
[109] Y. Mechref and D. C. Muddiman, “Recent advances in glycomics, glycoproteomics and allied topics,” Anal Bioanal Chem, vol. 409, no. 2, pp. 355–357, Jan. 2017, doi: 10.1007/s00216-016-0093-9.
[110] H. Geyer and R. Geyer, “Strategies for analysis of glycoprotein glycosylation,” Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics, vol. 1764, no. 12, pp. 1853–1869, Dec. 2006, doi: 10.1016/j.bbapap.2006.10.007.
[111] J. Gaston, N. Maestrali, G. Lalle, M. Gagnaire , A. Masiero, B. Dumas, T. Dabdoubi, K. Radošević, and P-F. Berne, “Intracellular delivery of therapeutic antibodies into specific cells using antibody-peptide fusions,” Sci Rep, vol. 9, no. 1, p. 18688, Dec. 2019, doi: 10.1038/s41598-019-55091-0.
[112] A. Stockell Hartree and A. G. C. Renwick, “Molecular structures of glycoprotein hormones and functions of their carbohydrate components,” Biochemical Journal, vol. 287, no. 3, pp. 665–679, Nov. 1992, doi: 10.1042/bj2870665.
[113] G. R. Bousfield, “Comparison of Follicle-Stimulating Hormone Glycosylation Microheterogenity by Quantitative Negative Mode Nano-Electrospray Mass Spectrometry of Peptide-N-Glycanase-Released Oligosaccharides,” J Glycomics Lipidomics, vol. 05, no. 01, 2015, doi: 10.4172/2153-0637.1000129.
[114] B. R. Meher, A. Dixit, G. R. Bousfield, and G. H. Lushington, “Glycosylation Effects on FSH-FSHR Interaction Dynamics: A Case Study of Different FSH Glycoforms by Molecular Dynamics

Simulations," PLoS ONE, vol. 10, no. 9, p. e0137897, Sep. 2015, doi: 10.1371/journal.pone.0137897.
[115] L. Casarini, D. Santi, G. Brigante, and M. Simoni, "Two Hormones for One Receptor: Evolution, Biochemistry, Actions, and Pathophysiology of LH and hCG," Endocrine Reviews, vol. 39, no. 5, pp. 549–592, Oct. 2018, doi: 10.1210/er.2018-00065.
[116] M. C. Patterson, "Metabolic Mimics: The Disorders of N-Linked Glycosylation," Seminars in Pediatric Neurology, vol. 12, no. 3, pp. 144–151, Sep. 2005, doi: 10.1016/j.spen.2005.10.002.
[117] R. Haselberg, G. J. de Jong, and G. W. Somsen, "Capillary electrophoresis-mass spectrometry for the analysis of intact proteins 2007-2010," ELECTROPHORESIS, vol. 32, no. 1, pp. 66–82, Jan. 2011, doi: 10.1002/elps.201000364.
[118] A. Al Matari, Goumenou, Anastasia, N. Delaunay, A. Combes, and V. Pichon, "Identification and semi-relative quantification of intact glycoforms of human chorionic gonadotropin alpha and beta subunits by nanoLC-(Orbitrap)MS," Submitted.
[119] "FSH IU.pdf," https://www.ema.europa.eu/en/documents/product-information/puregon-epar-product-information_en.pdf, Nov. 30, 2020. .
[120] "IU LH.pdf," Nov. 30, 2020. http://www.e-lactancia.org/media/papers/Lutropin_Alpha_Luveris-DS-EMD2012.pdf.
[121] C. Alviggi, A. Revelli, P. Anserini, A. Ranieri, L. Fedele, I. Strina, M. Massobrio, N. Ragni and G. De Placido, "A prospective, randomised, controlled clinical study on the assessment of tolerability and of clinical efficacy of Merional (hMG-IBSA) administered subcutaneously versus Merional administered intramuscularly in women undergoing multifollicular ovarian stimulation in an ART programme (IVF)," Reproductive Biology and Endocrinology, p. 8, 2007.
[122] M. Cruz-Vera, R. Lucena, S. Cárdenas, and M. Valcárcel, "Highly selective and non-conventional sorbents for the determination of biomarkers in urine by liquid chromatography," Anal Bioanal Chem, vol. 397, no. 3, pp. 1029–1038, Jun. 2010, doi: 10.1007/s00216-010-3476-3.
[123] M.-C. Hennion and V. Pichon, "Immuno-based sample preparation for trace analysis," Journal of Chromatography A, vol. 1000, no. 1–2, pp. 29–52, Jun. 2003, doi: 10.1016/S0021-9673(03)00529-6.
[124] N. Delaunay-Bertoncini, V. Pichon, and M.-C. Hennion, "Comparison of immunoextraction sorbents prepared from monoclonal and polyclonal anti-isoproturon antibodies and optimization of the appropriate monoclonal antibody-based sorbent for environmental and biological applications," Chromatographia, vol. 53, no. S1, pp. S224–S230, Jan. 2001, doi: 10.1007/BF02490332.
[125] W. Clarke, J. D. Beckwith, A. Jackson, B. Reynolds, E. M. Karle, and D. S. Hage, "Antibody immobilization to high-performance liquid chromatography supports Characterization of maximum loading capacity for intact immunoglobulin G and Fab fragments," J. Chromatogr. A, p. 10, 2000.
[126] V. Pichon and A. Combès, "Selective tools for the solid-phase extraction of Ochratoxin A from various complex samples: immunosorbents, oligosorbents, and molecularly imprinted polymers," Anal Bioanal Chem, vol. 408, no. 25, pp. 6983–6999, Oct. 2016, doi: 10.1007/s00216-016-9886-0.
[127] A. Cingöz, F. Hugon-Chapuis, and V. Pichon, "Total on-line analysis of a target protein from plasma by immunoextraction, digestion and liquid chromatography–mass spectrometry," Journal of Chromatography B, vol. 878, no. 2, pp. 213–221, Jan. 2010, doi: 10.1016/j.jchromb.2009.07.032.
[128] J. M. Bidart, F. Troalen, G. R. Bousfield, S. Birken, and D. H. Bellet, "Antigenic determinants on human choriogonadotropin alpha-subunit. I. Characterization of topographic sites recognized by monoclonal antibodies.," Journal of Biological Chemistry, vol. 263, no. 21, pp. 10364–10369, Jul. 1988, doi: 10.1016/S0021-9258(19)81525-9.

[129] M. Bonichon, V. Valbi, A. Combès, C. Desoubries, A. Bossée, and V. Pichon, “Online coupling of immunoextraction, digestion, and microliquid chromatography-tandem mass spectrometry for the analysis of sarin and soman-butyrylcholinesterase adducts in human plasma,” Anal Bioanal Chem, vol. 410, no. 3, pp. 1039–1051, Jan. 2018, doi: 10.1007/s00216-017-0640-z.

RÉSUMÉ

La glycosylation est l'addition de glycanes sur une protéine, appelée alors glycoprotéine. C'est l'une des modifications post-traductionnelles les plus courantes et elle a un impact sur l'activité et la stabilité de la protéine. L'hormone chorionique gonadotrope humaine (hCG), appelée communément hormone de grossesse, est une protéine hautement glycosylée, composée de deux sous-unités, hCGα et hCGβ, ayant au total quatre sites de N- et quatre sites de O-glycosylation, conduisant à un nombre très élevé de glycoformes. Or, l'état de glycosylation de cette hormone essentielle de la grossesse a été associé à certaines pathologies de celle-ci. C'est pourquoi il est nécessaire de pouvoir analyser les glycoformes de la hCG présentes dans des échantillons biologiques. Dans ces travaux, de nouvelles méthodes ont été développées pour l'analyse à l'échelle intacte des glycoformes de la hCG par nanochromatographie en phase liquide (nanoLC) couplée à la spectrométrie de masse (MS). Le format miniaturisé de la séparation a permis une augmentation de la sensibilité de la méthode et son couplage avec un analyseur de haute précision et haute résolution en masse, un Orbitrap, a permis de détecter et de semi-quantifier plusieurs glycoformes de la sous-unité α dans des médicaments à base de hCG, mais aussi d'une autre gonadotrophine partageant la même sous-unité α, l'hormone folliculo-stimulante (FSH). L'optimisation d'une étape de préconcentration en amont de la nanoLC-MS et du gradient de phase mobile et de sa nature ont permis ensuite de détecter des glycoformes de la hCGβ pour la première fois et d'un plus grand nombre de glycoformes de la hCGα. Puis, cette méthode analytique utilisant soit de l'acide formique soit de l'acide trifluoroacétique comme additif de la phase mobile a été évaluée pour la détection des glycoformes des sous-unités α et β de différents médicaments à base de gonadotrophines d'origine recombinante ou urinaire. Enfin, dans le but de déterminer les glycoformes de la hCG dans le sérum de femmes enceintes, c'est-à-dire dans des échantillons biologiques complexes plutôt que dans des médicaments concentrés et relativement purs à base de gonadotrophines, il a été nécessaire de développer une étape d'extraction sélective des glycoformes de la hCG avant leur analyse par nanoLC-MS. Ainsi, des immunoadsorbants ont été synthétisés en immobilisant deux anticorps anti-hCG différents sur un support solide. Les premières expériences d'immunoextraction de la hCG ont donné des résultats prometteurs.

MOTS CLÉS

Glycoprotéine, Hormone chorionique gonadotrope humaine, Protéine intacte, Spectrométrie de masse, Nano-liquide chromatographie

ABSTRACT

Glycosylation is the attachment of glycans to a protein, so-called a glycoprotein. It is one of the most common post-translational modifications and it impacts protein activity and stability. Human chorionic gonadotropin (hCG) is a highly glycosylated protein, composed of two subunits, hCGα and hCGβ, having in total four N- and four O-glycosylation sites, leading to a very high number of glycoforms. However, the glycosylation state of this hormone essential for pregnancy was associated with some pregnancy pathologies. For this reason, it is necessary to be able to analyze the hCG glycoforms present in biological samples. In this work, new analytical methods for the analysis at the intact level of hCG glycoforms by nano liquid chromatography (nanoLC) hyphenated with mass spectrometry (MS) were developed. The miniaturization of the separation allowed an increase in sensitivity and its coupling with a high mass accuracy and high mass resolution analyzer, an Orbitrap, allowed the detection and semi-quantification of several α-subunit glycoforms in hCG-based drugs, but also of another gonadotropin sharing the same α-subunit, the follicle-stimulating hormone (FSH). The optimization of both a preconcentration step in the nanoLC-MS set-up and of the mobile phase gradient and nature allowed the detection of hCGβ glycoforms for the first time and also of a higher number of hCGα glycoforms. Next this analytical method using either formic acid or trifluoroacetic acid as mobile phase additive was evaluated for the detection of the α- and β-subunits of different recombinant or urinary gonadotropin-based drugs. Finally, with the aim to determine hCG glycoforms in serum of pregnant women, i.e complex biological samples, rather than in concentrated and fairly pure gonadotropin-based drugs, it was necessary to develop a selective extraction step of hCG glycoforms before their analysis in nanoLC-MS. This is why, immunosorbents were synthesized by immobilizing two different anti-hCG antibodies on a solid sorbent. The primary experiments of hCG immunoextraction showed promising results.

KEYWORDS

Glycoprotein, Human chorionic gonadotropin, Intact protein, Mass spectrometry, Nano-liquid chromatography