A semi-automated hybrid HPLC-MS approach for in-depth characterization of intact non-covalent heterodimer glycoforms of gonadotropin biopharmaceuticals

Marco, F. di; Blümel, G.; Blöchl, C.; Wuhrer, M.; Huber, C.G.

Analytica Chimica Acta 1274 (2023) 341574

Contents lists available at ScienceDirect

Analytica Chimica Acta

journal homepage: www.elsevier.com/locate/aca

A semi-automated hybrid HPLC-MS approach for in-depth characterization of intact non-covalent heterodimer glycoforms of gonadotropin biopharmaceuticals

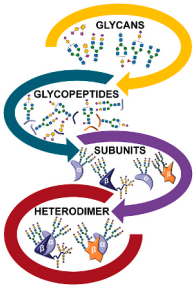
Fiammetta Di Marco [a,1], Gabriele Blümel [a,1], Constantin Blöchl [a,b], Manfred Wuhrer [b], Christian G. Huber [a,*]

[a] Department of Biosciences and Medical Biology, Bioanalytical Research Labs, University of Salzburg, Hellbrunner Straße 34, 5020, Salzburg, Austria
[b] Center for Proteomics and Metabolomics, Leiden University Medical Center, Albinusdreef 2, 2333 ZA, Leiden, the Netherlands

HIGHLIGHTS

- Hybrid HPLC-MS approach for in-depth characterization of the glycosylation of gonadotropin biopharmaceuticals.
- Analysis of glycoproteins at different structural levels.
- Acquisition of gonadotropin intact subunit and dimer mass spectra.
- Annotation of intact glycoforms of gonadotropin subunits and dimers.
- Comparison of urinary and recombinant biopharmaceuticals.

GRAPHICAL ABSTRACT



ARTICLE INFO

Handling Editor: Dr. L. Liang

Keywords:
Biopharmaceutical characterization
Gonadotropin
hCG
hFSH
Glycosylation
Intact glycoforms
Hybrid-MS
HPLC-MS

ABSTRACT

Background: Gonadotropins are a class of heavily glycosylated protein hormones, thus extremely challenging to characterize by mass spectrometry. As biopharmaceuticals, gonadotropins are prescribed for the treatment of infertility and are derived from different sources: either from pooled urine of pregnant women or upon production in genetically modified Chinese Hamster Ovary cells. Human chorionic gonadotropin (hCG) is sold as a biopharmaceutical under the name Pregnyl® (urinary hCG, u-hCG) and Ovitrelle® (recombinant hCG, r-hCG), and recombinant human follicle stimulating hormone (r-hFSH) is marketed as Gonal-f®. Recently, we reported the exhaustive characterization of r-hCG at different structural levels.
Results: We implement size exclusion (SE) HPLC-MS to automatize the acquisition of native mass spectra of r-hCG dimer, but also u-hCG and r-hFSH, comparing the drug products up to intact heterodimer level. A hybrid HPLC-MS approach was employed for the characterization of r-hCG, u-hCG and r-hFSH drug products at different structural levels. Released glycans were analyzed by porous graphitized carbon (PGC)-HPLC-MS/MS, glycopeptides by reversed-phase (RP)-HPLC-MS/MS, subunits by RP-HPLC-MS and finally the intact native

Abbreviations: r-hCG, recombinant human chorionic gonadotropin; u-hCG, urinary human chorionic gonadotropin; r-hFSH, recombinant human follicle stimulating hormone; CHO, Chinese hamster ovary; MS, mass spectrometry; HPLC, high-performance liquid chromatography; PGC, porous graphitized carbon; RP, reversed-phase; SE, size exclusion; EICCs, extracted ion current chromatograms; M_r, molecular mass.

* Corresponding author.
E-mail address: c.huber@plus.ac.at (C.G. Huber).
[1] These authors contributed equally to this work.

https://doi.org/10.1016/j.aca.2023.341574
Received 3 April 2023; Received in revised form 7 June 2023; Accepted 28 June 2023
Available online 29 June 2023

heterodimers by semi-automated online buffer exchange SE-HPLC-MS. The data were integrated using bioinformatic tools, to finally unravel the composition of 1481 co-existing dimeric glycoforms for r-hCG, 1167 glycoforms for u-hCG, and 1440 glycoforms for r-hFSH, and to compare critical quality attributes of the different drug products such as their degree of sialylation and *O*-glycosylation.

Significance and novelty: The strong alliance of bioanalytics and bioinformatics data integration at the different structural levels allowed the identification of more than thousand different glycoforms of r-hCG, u-hCG, and r-hFSH. The results showed that these biopharmaceuticals differ considerably in their glycosylation patterns and highlight the importance of in-depth characterization of biopharmaceuticals for quality control.

1. Introduction

Gonadotropins are a class of glycoprotein hormones playing a central role in reproduction and are used in the biopharmaceutical field for the treatment of infertility [1–4]. In humans, the three gonadotropins consist of the human luteinizing hormone (hLH), the human follicle stimulating hormone (hFSH), and the human chorionic gonadotropin (hCG), whose molecular structure is characterized by a heavily glycosylated non-covalent heterodimer sharing a common alpha subunit and differing in the beta subunit. The alpha subunit contains two *N*-glycosylation sites (N52, N78), and the bare amino acid sequence has a molecular mass (M_r) of ≈10.2 kDa. The beta subunit of hCG is the most complex among the three hormones containing two *N*-glycosylation sites (N13, N30) and up to 9 *O*-glycosylation sites (S118-121, S127, S130, S132, S138, T140) with an amino acid M_r of ≈15.5 kDa [5,6]. The FSH beta subunit is less complex with two *N*-glycosylation sites (N7, N24) and a protein backbone M_r of ≈12.5 kDa [7].

From a biopharmaceutical perspective, the drug products are mainly produced from two sources, by affinity extraction of the hormone from a pool of urine of pregnant women using specific antibodies or biotechnologically by recombinant expression using Chinese Hamster Ovary (CHO) cells in fermentation systems [3,8]. However, it was observed that the urinary hCG product contains more impurities compared to the recombinant product, *e.g.* contaminations related to prion proteins [9] or fragments of the hCG beta subunit – the so-called nicked beta [10].

Even if clinical studies have been conducted to assess the biological activity of urinary in comparison to recombinant biopharmaceuticals [11,12], the knowledge on how protein glycosylation influences the efficacy of the drug product is still limited [2,13,14]. It is known that the presence of *O*-glycans and/or terminally sialylated glycans increases half-life in circulation [13,15–17] and that the folding of the subunit in the non-covalent heterodimer is facilitated by the presence of glycosylation [13]. The receptor-binding affinity and signal transduction are also affected by the glycosylation state of the gonadotropins: gonadotropins deglycosylated at the N52 of the alpha subunits showed an increased affinity for the receptor but a reduced stimulation of the cAMP signaling, resulting in reduced biological activity [14].

Since glycosylation may greatly impact the efficacy of a drug product, the bioanalytical characterization of glycoprotein biopharmaceuticals is extremely important. To comprehensively characterize highly complex glycoproteins, novel high-performance liquid chromatography - mass spectrometry (HPLC-MS) based approaches tend towards the analysis of the protein at the different structural levels of released glycans [18–20], glycopeptides [21–23], intact protein subunits [5,24], and intact complexes [25–30], the so-called hybrid MS approaches [5,23,31–33]. Very recently, size exclusion high-performance liquid chromatography (SE-HPLC) hyphenated to MS was proposed as an alternative method to perform native MS in an automated fashion [34,35]. In fact, using this technique, buffer exchange of the drug products with a solution compatible with native MS can be performed directly in the HPLC system where low M_r drug additives are separated from the protein and a native mass spectrum can be acquired in an automated fashion upon direct hyphenation with MS.

Previous studies aimed to characterize hFSH and hCG at glycopeptide level [5,36–42], however only few studies dealt with the characterization of these complex glycoproteins at intact protein level by MS. Grass et al. reported the characterization of r-hFSH intact subunits acquiring non-isotopically resolved mass spectra [40], Al Matari et al. reported a partial characterization of r-hFSH alpha subunit [43], while we and others reported the full characterization of hCG subunits [5, 43–45]. However, neither isotopically resolved spectra of intact hFSH and u-hCG subunits nor mass spectra of the dimers were reported. Our previous study is the only one dealing with the characterization of r-hCG at intact dimer level and with the annotations of the intact r-hCG glycoforms at subunit and dimer level *via* bioinformatic integration of the data from the different structural levels [5]. The key to unravel the structural complexity of the plethora of co-existing hCG glycoforms was to study the protein at the different structural levels: from released glycans by porous graphitized carbon (PGC)-HPLC-MS/MS and capillary gel electrophoresis with laser induced fluorescence detection to glycopeptides by reversed phase (RP)-HPLC-MS/MS, to the intact subunits using a RP-HPLC-MS approach and finally to the intact non-covalent dimer by conventional direct infusion native MS approach with a static nano-ESI source operated under mild conditions.

In this study we went an important step further. We optimized the previous approach substituting the native MS experiment, comprising a step of sample preparation prior to MS analysis for removal of drug additives, with an automated online buffer exchange SE-HPLC-MS approach and we applied this method to characterize and compare the gonadotropin-based drug products, Ovitrelle® (r-hCG), Pregnyl® (urinary hCG, u-hCG) and Gonal-f® (recombinant hFSH, r-hFSH).

We attempted the comprehensive characterization of u-hCG and r-hCG and r-hFSH up to the intact dimer level and the annotations of their intact glycoforms by the bioinformatic workflow reported in our previous publication [5]. The results were used to compare the drug products and showed that u-hCG and r-hCG heavily differ from a glycosylation perspective and thus are not identical products. Moreover, the r-hFSH alpha subunit also differs in the glycosylation pattern compared to r-hCG, even though they share the same alpha subunit structure and they are both recombinantly produced by the same pharmaceutical company. This study proposes a semi-automated HPLC-MS approach to extensively characterize complex glycoproteins and stresses the importance to implement methods for in-depth characterization of intact glycoforms of heavily glycosylated biopharmaceuticals in the quality control environment and in the regulatory process for biopharmaceuticals approval.

2. Materials and methods

2.1. Materials

The gonadotropin-based biopharmaceuticals Ovitrelle® containing r-hCG (Merck, Lot: BA059433, expiration date 04/2021, BA056714 expiration date 01/2021, Lot: BA065472, expiration date 04/2022), Pregnyl® containing u-hCG (Merck Lot: S024812, expiration date 02/2021), and Gonal-f® containing r-hFSH (Merck, Lot: AU028271,

expiration date 07/2021), were purchased from a local pharmacy. Ovitrelle® had a concentration of r-hCG of 0.5 μg μL^{-1}, lyophilized Pregnyl® was dissolved in MilliQ water to a u-hCG concentration of 0.5 μg μL^{-1} and Gonal-f® had a concentration of 0.044 μg μL^{-1} of r-hFSH. Dithiothreitol (DTT), guanidine hydrochloride (Gnd-HCl), ammonium acetate (AmAc), ammonium bicarbonate, iodoacetamide (IAA), and formic acid (FA) were purchased from Sigma-Aldrich (St. Louis, MO). Trypsin and chymotrypsin were purchased from Promega (Fitchburg, WI). LC-MS grade acetonitrile (ACN) was obtained from VWR (Radnor, PA). Ultrapure H_2O was obtained by in-house purification with an MilliQ Integral 3 system from Merck Millipore or was purchased from VWR.

2.2. Sample preparation and enzymatic treatment

N- and *O*-glycans of r- and u-hCG were released and reduced using an established 96-well plate protocol based on a polyvinylidene difluoride (PVDF) membrane workflow [18]: 20 μg of protein (25 μL) were directly applied and bound to the activated PVDF membranes (Immobilon-P; Millipore, Amsterdam, The Netherlands) followed by protein denaturation applying guanidine HCl (Thermo Fisher Scientific) to 5.8 mol L^{-1} and dithiothreitol (Sigma Aldrich) to 5.0 mmol L^{-1}. Samples were shaken for 15 min and subsequently incubated at 60 °C for 30 min. After removal of unbound material by washing with water, 2.0 U of PNGase F (Roche Diagnostics) were added for *N*-glycan release and incubated at 37 °C overnight. At this point, an internal standard (10 ng maltoheptaose DP7; Elicityl, Crolles, France) was added as quality control for sample preparation as well as system performance. Released *N*-glycans were recovered by centrifugation from the plates in their glycosylamin form and subsequently hydrolyzed in approximately 6 mmol L^{-1} ammonium acetate (pH 5.0; Sigma-Aldrich) for 1.0 h at room temperature. Samples were dried by vacuum centrifugation and transformed into their alditol forms by reduction in 20 μL of 50 mmol L^{-1} KOH (Honeywell Fluka, Thermo Fisher Scientific) supplemented with 1.0 mol L^{-1} $NaBH_4$ (Sigma-Aldrich) and incubation at 60 °C for 3.0 h. *O*-glycans could be released from the same PVDF-immobilized proteins by reductive β-elimination (50 mmol L^{-1} KOH and 500 mmol L^{-1} $NaBH_4$) at 55 °C for 16.0 h. Similar to *N*-glycans, 5 ng maltopentaose DP5 (Elicityl) served as an internal standard for *O*-glycans. Both *N*- and *O*-glycan alditols were desalted on a strong cation exchange resin (Dowex 50 W X8; Merck, Darmstadt, Germany) that was packed into 96-well filter plates (Orochem Technologies, Naperville, IL, USA). Residual boric acid was removed by multiple rounds of co-evaporation with methanol in a vacuum centrifuge. Released glycans were purified on PGC material (Carbograph, Grace Discovery Sciences, Columbia, TN, USA) self-packed into 96-well filter plates. Eventually, samples were evaporated to dryness and dissolved in 10 μL of H_2O prior to analysis.

For glycopeptide analysis, 10 μL of u-hCG and r-hCG (5.0 μg of protein) were diluted with 50 mmol L^{-1} ammonium bicarbonate to a volume of 55 μL. Regarding r-hFSH preparation, 5.0 μL of 2.0 mol L^{-1} ammonium bicarbonate were added to 113.64 μL of Gonal-f® (5.0 μg of protein). The protein was then reduced and denatured by the addition of 10 μL of 40 mmol L^{-1} DTT and 50 μL of 6.0 mol L^{-1} Gnd-HCl, incubating for 1.0 h at 50 °C under shaking (900 rpm). Alkylation was performed by subsequently adding 15 μL of freshly prepared 250 mmol L^{-1} aqueous IAA solution for 1.0 h at 22 °C in the dark under shaking (900 rpm). For *O*-glycopeptide analysis of Ovitrelle® and Pregnyl®, 6.0 μL of 2.0 mol L^{-1} ammonium bicarbonate were added to 60 μL of sample (30 μg of protein). Subsequently, 60 μL of 6.0 mol L^{-1} Gnd-HCl and 12 μL of 40 mmol L^{-1} DTT were added, followed by an incubation for 1.0 h at 50 °C in a thermoshaker. For alkylation, 18 μL of freshly prepared 250 mmol L^{-1} IAA solution was added to the sample and incubated for 1.0 h at 22 °C in the dark while shaking (900 rpm).

Before tryptic digestion, Pierce® C18 tips (100 μL, Thermo Fisher Scientific) were used for glycopeptide sample clean-up. First, C18 tips were conditioned twice with 50:50 ACN:H_2O, then twice with H_2O + 0.10% FA. The samples were then acidified to a final concentration of 1.0% FA and loaded onto the C18 tips by pipetting the solution up and down 10 times until the whole volume was loaded. The C18 tips were then washed twice with 2.5:97.5 ACN:H_2O + 0.10% FA and eluted twice with 99:1 ACN:H_2O + 0.10% FA. After vacuum centrifugation, the samples were reconstituted in 20 μL (for *N*-glycopeptide analysis) or 60 μL (for *O*-glycopeptide analysis) of 50 mmol L^{-1} ammonium bicarbonate. 0.50 μg (or 1.50 μg for *O*-glycopeptides) trypsin (or chymotrypsin for r-hFSHβ *N*-glycopeptides) were added and the samples were incubated at 37 °C overnight in a thermoshaker.

2.3. PGC-HPLC-MS/MS analysis of released hCG N- and O-glycans

Released *N*- and *O*-glycans of r- and u-hCG were analyzed by nano PGC-HPLC-MS/MS using the MS parameters stated in Zhang et al. [18]. Briefly, *N*- and *O*-glycan alditols were chromatographically separated on a Thermo Scientific™ Ultimate 3000 RSLCnano UHPLC system (Thermo Fisher Scientific) coupled to an amaZon ETD speed ion trap mass spectrometer *via* a CaptiveSpray ESI source (Bruker Daltonics, Bremen, Germany) where isopropanol was used as dopant solvent. A self-packed trap column (5 μm particles, 30 mm × 0.32 mm, i.d.) and a self-packed separation column (3 μm particle diameter, 100 mm × 0.075 mm i.d.) were installed. Hypercarb™ KAPPA material (Thermo Fisher Scientific) was used as material for packing the columns. For the separation, column oven temperature was kept at 45 °C and the injection volume was 1.0 μL. Mobile phase A compromised H_2O + 10 mM ammonium bicarbonate and mobile phase B 60% ACN +10 mM ammonium bicarbonate. Glycans were injected into the trap at a flow rate of 6.0 μL min^{-1} and a mobile phase composition of 1.0% B that was held for 5.0 min. N-glycans were separated employing a gradient elution of: 2.0–9.0% B in 1.0 min, 9.0–54.0% B in 100.0 min, 54.0–95.0% B in 9.0 min, 95.0% B for 8 min, 95.0–2.0% B in 5.0 min, and 2.0% B for 17.0 min at a flow rate of 500 nL min^{-1}. *O*-glycans were separated as follows: 1.0–41.0% B in 80.0 min, 41.0–95.0% B in 3.0 min, 95.0% B for 5 min, 95.0–1.0% B in 8.0 min, and 1.0% B for 19.0 min also at a flow rate of 500 nL min^{-1}.

2.4. RP-HPLC-MS/MS analysis of gonadotropin glycopeptides

Glycopeptides were analyzed using a Thermo Scientific™ Ultimate 3000 RSLCnano UHPLC system (Thermo Fisher Scientific) coupled with a Thermo Scientific™ Q Exactive™ Plus Hybrid Quadrupole-Orbitrap™ mass spectrometer (Thermo Fisher Scientific) equipped with a Thermo Scientific™ Nanospray Flex™ Ion Source (Thermo Fisher Scientific) and operated using Chromeleon 7.2.10. Chromatographic separation of *N*-glycopeptides was achieved using a 150 mm × 300 μm i.d. Acclaim Pepmap 100C18 LC column (Thermo Fisher Scientific), particle size 2 μm and 100 Å pore size. The column was operated at 50 °C and a flow rate of 1.2 μL min^{-1}. The injection volume was 5.0 μL. Water with 0.10% FA was used as mobile phase A and ACN with 0.10% FA was utilized as mobile phase B. Gradient elution for u-hCG and r-hCG and r-hFSHα tryptic glycopeptides was carried out. The mobile phase composition was kept at 1.0% for 5.0 min, followed by 1.0–30% B over 30 min, 30–60% B in 5.0 min, 99% B was kept for 5.0 min and equilibration was performed at 1.0% B for 10 min (total run time 55 min). For r-hFSHβ chymotryptic glycopeptides, the mobile phase composition was kept at 1.0% for 5.0 min, followed by 1.0–30% B over 30 min, 30–99% B in 15 min, 99% B was kept for 5.0 min and equilibration was performed at 1.0% B for 20 min (total run time 75 min).

For r- and u-hCG *O*-glycans analysis, 0.90 μL of sample were injected to a 750 mm × 75 μm i.d. nanoPepMap column (Thermo Fisher Scientific) with a particle size of 2 μm and a pore size of 100 Å. The column was operated at a flow rate 200 μL min^{-1} at 50 °C. The same eluents A and B as for *N*-glycans analysis were used. First a gradient from 1.0 to 30% B in 165 min was carried out, followed by 30–60% B in 35 min. Then the column was kept at 99% B for 20 min and equilibrated at 1.0% B for 60 min. For both methods, the instrument was equipped with a

Thermo Scientific™ Nanospray Flex™ Ion Source equipped with a nano spray fused silica emitter with a pulled tip, o.d. 360 μm, i.d. 20 μm, Tip i.d. 10 μm, length 12 cm (TIP1002010-12, CoAnn Technologies, LLC, MS Wil, Aarle-Rixtel, the Netherlands). The mass spectrometer was operated with the following settings: ion-source spray voltage of 1.5 kV, positive polarity, capillary temperature of 250 °C, in-source CID of 0, S-lens RF level of 60, sheath, sweep and auxiliary gas 0. Data were acquired in data dependent MS/MS mode. For MS^1, the scan range was set to *m/z* 400–2000 at a resolution setting of 70,000, with automatic gain control (AGC) target of 3×10^6, the maximum injection time (IT) was 100 ms. For MS/MS, the mass range was set to *m/z* 200–2000 with a resolution of 17,500. The AGC target value was 1×10^5, the maximum IT was set to 50 ms and the normalized collision energy (NCE) was 28. Microscans were not averaged. The dynamic exclusion was 10.0 s. Instrument calibration was performed with a Pierce® LTQ Velos ESI Positive Ion Calibration Solution (Thermo Fisher Scientific). All samples were measured in technical triplicates, separated by a blank run. For quality assurance before and after a sequence, 500 fmol bovine serum albumin (Pierce™ BSA protein digest, Thermo Fisher Scientific) (*N*-glycan analysis) or 100 fmol *(O*-glycan analysis) were injected.

2.5. RP-HPLC-MS analysis of intact gonadotropin subunits

Intact gonadotropin subunits were measured with a Thermo Scientific™ Ultimate 3000 UHPLC system (Thermo Fisher Scientific) coupled to a Thermo Scientific™ Q Exactive™ Hybrid Quadrupole-Orbitrap™ mass spectrometer (Thermo Fisher Scientific), operated with Chromeleon 7.2.10. Chromatographic separation was achieved using a 2.1 mm × 150 mm i.d., 3 μm particle diameter, 300 Å pore size Discovery Wide Pore C18 column (Supelco, Bellefonte, PA). The injection volume was 7.0 μL. Mobile phase A comprised water and 0.10% FA; mobile phase B comprised ACN and 0.10% FA. A gradient from 15 to 30% eluent B was applied over a time span of 67.5 min at a flow rate of 100 μL min^{-1} and a temperature of 60 °C, followed by 99% B for 5 min and equilibration at 15% B for 15 min. The ion-source spray voltage was set to 4.5 kV, the capillary temperature to 300 °C, sheath gas at 10, auxiliary gas flow rate at 5, sweep gas at 0 and probe heater temperature at 150 °C. The Orbitrap mass analyzer mass range was set to 1000–3500 *m/z* with a resolution of 140,000. The S-lens RF level was set to 50 and the AGC target value was 3×10^6 with a maximum IT of 150 ms, 10 microscans were averaged.

2.6. Native SE-HPLC-MS analysis of the intact gonadotropin dimers

Native MS analyses were performed with a Thermo Scientific™ Ultimate 3000 UHPLC system (Thermo Fisher Scientific) equipped with two pumps and a post-column six-port-valve, coupled to a Thermo Scientific™ Q Exactive™ Plus Hybrid Quadrupole-Orbitrap™ mass spectrometer (Thermo Fisher Scientific). 4 μL of the Ovitrelle® and Pregnyl® samples and 20 μL of the Gonal-f® sample were injected onto an Acquity UPLC Protein BEH SEC Guard column 4.6 mm × 30 mm, 1.7 μm particle size, 200 Å pore size to online buffer exchange the drug products directly in the HPLC system. To avoid mass spectrometer contamination by drug additives, the valve position was switched to divert the flow into the waste after protein elution to the mass spectrometer (at 3.0 min for Pregnyl® and Ovitrelle® and 3.6 min for Gonal-f®). As eluents for both pumps, 150 mmol L^{-1} AmAc was used and pumped at a flow rate of 100 μL min^{-1}, leading to a constant flow in the HPLC system and in the tip of the HESI source for both valve positions during the whole chromatographic separation and spectra acquisition. The Q Exactive™ Plus was used with the BioPharma feature activated in Protein Mode, which allows ions to be transferred, trapped, and cooled in the HCD cell resulting in increased signal intensities. The source and MS parameters were set as follows: spray voltage 3.6 kV, capillary temperature 250 °C, in-source CID 50, S-lens RF level 50, polarity positive, resolution 17,500 at *m/z* 200, sheath gas 20, auxiliary gas 5 sweep gas 0 and probe heater temperature 250 °C. The *m/z* range was acquired from *m/z* 1000 to 5000. Trapping gas pressure setting was 1 and spectral averaging 0. The AGC target value was 3×10^6 with a maximum IT of 200 ms. The mass spectrometer was calibrated by using LTQ Velos ESI positive ion calibration solution, while ammonium hexafluorophosphate was used for calibration of the high mass range, in order to cover the *m/z* range from 195.09 to 5559.98.

2.7. MS data processing and peak assignment

N- and *O*-glycan structure assignments are based on MS^1 spectra and retention time matching or MS^2 spectra when unambiguous identification was not possible based on MS^1 and retention time and were evaluated using GlycoMod. EICCs were visualized employing Skyline and the implemented small molecule interface [46].

Glycopeptide identification based on MS^2 spectra was carried out using PMI-Byonic™ (v3.11.3, Protein Metrics Inc., Cupertino, CA, USA). The parameters were set as follows: cleavage residues RK for trypsin and FWLY for chymotrypsin, cleavage side C-terminal, digestion specificity was set to "fully specific" with a maximum of missed cleavages of 2, precursor mass tolerance and fragment mass tolerance were set to 20 ppm. Carbamidomethylation/+57.021464 was set as a fixed modification, deamidation of N/+0.984016 as rare 2 and oxidation/+15.994915 as rare 1. A maximum of 2 common and 2 rare modifications were set. No automatic score cutoff was used and the option "show all *N*-glycopeptides" was selected. *N*- and *O*-glycans were searched separately.

Semi-quantitation of the glycopeptides based on extracted ion current chromatograms (EICCs) of glycopeptide MS^1 ions was performed using Skyline (v20.2.0.343, MacCoss Lab, Department of Genome Science, University of Washington) with the proteomics interface [47]. Glycans identified with Byonic™ were then integrated as peptide modifications in Skyline and validated based on retention time, isotopic pattern and mass error matching. EICC areas of validated glycopeptides were then exported and fractional abundances were calculated using Excel (Microsoft 365).

Isotopically resolved raw mass spectra of gonadotropin subunits were deconvoluted using Thermo Xcalibur 4.2.28.14 (Thermo Fisher Scientific) equipped with a license of the Xtract™ algorithm. Data of intact gonadotropins was deconvoluted with the ReSpect algorithm embedded in Thermo Scientific™ BioPharma Finder™ software v. 3.0 (Thermo Fisher Scientific). The default method was chosen and the parameters were optimized according to the specific spectrum (*m/z* range, selected retention time, model mass range, charge state range, minimum charges).

For glycoform annotations of the intact deconvoluted mass spectra of hCG and FSH subunits and dimers, the annotation tool MoFi [48] was used. The libraries built from glycopeptide or subunit data were integrated in the software and the monosaccharides were considered as variable modifications (see our previous publication [5] for more details). MoFi settings for glycoform annotations were set as follows. For all the subunit searches, mass set was selected as monoisotopic and oxidation (max 2) and Off by ± 1 were considered as variable modification. For u-hCG, r-hCGα and r-hFSHα, disulfide bridges were set at 5, cleavage of the first two *N*-terminal amino acids (AP) were considered as variable modification and a mass tolerance of 20 ppm was selected for the search. For u-hCG and r-hCGβ, disulfide bridges were set at 6 and mass tolerance of 31 ppm. For r-hFSHβ glycoform annotations, disulfide bridges were set at 6, cleavage of the first *N*-terminal amino acid (N) was considered as variable modification and a mass tolerance of 20 ppm was selected for the search. For all the dimer annotations, disulfide bridges were set at 11, mass set average (IUPAC) and a mass tolerance of 5 Da was selected for the search.

3. Results

3.1. Hybrid HPLC-MS approach to characterize gonadotropins at the different structural levels and bioinformatic data integration across the different structural levels

To comprehensively characterize u-hCG, r-hCG and r-hFSH contained in the drug products Pregnyl®, Ovitrelle®, and Gonal-f®, a hybrid HPLC-MS approach was employed involving: *i.* nano PGC HPLC-MS/MS analysis for released glycan analyses of u-hCG and r-hCG to obtain information about the *N*- and *O*-glycan structures, *ii.* micro and nano RP-HPLC-MS/MS for glycopeptide identification and semi-quantitation of u-hCG, r-hCG, and r-hFSH, *iii.* RP-HPLC-MS for the analysis of intact glycoforms of u-hCG and r-hCG, and r-hFSH subunits and *iv.* native SE-HPLC-MS for native non-covalent u-hCG, r-hCG, and r-hFSH heterodimer glycoforms analysis (Fig. S1).

Bioinformatic tools were used for a stepwise integration of the data at the different structural levels. From released glycan data, a qualitative library of glycan structures was built by combining data from u-hCG and r-hCG and by manually adding some glycans that could potentially have not been identified at released glycan level. This library was later also used for r-hFSH glycopeptide identification. The qualitative glycan library was integrated into Byonic® (Protein Metrics) for glycopeptide identification and semi-quantitation. In this way, a site-specific semi-quantitative glycan library was built and used later for u-hCG, r-hCG and r-hFSH subunit glycoform annotation. To annotate deconvoluted spectra of gonadotropins, the annotation tool MoFi [48] was used. The site-specific semi-quantitative glycan library from glycopeptide data was integrated in MoFi and annotation of intact u-hCG and r-hCG and r-hFSH subunit glycoforms was achieved. Finally, libraries of glycoforms of the alpha and the beta subunits of u-hCG and r-hCG, as well as r-hFSH, were built and integrated in MoFi to annotate the deconvoluted mass spectra of u-hCG and r-hCG and r-hFSH dimer glycoforms. This information was then used to finally obtain the composition of the biologically relevant glycoform complexes. More details about the workflow can be found in our previous publication [5].

3.2. Urinary vs recombinant hCG: released glycan analysis by PGC-HPLC-MS/MS and glycopeptide analysis by RP-HPLC-MS/MS

N-glycans of r- and u-hCG were enzymatically released upon treatment with PNGase F and *O*-glycans were chemically released *via* reductive β-elimination and both analyzed by nano PGC-HPLC-MS/MS. Glycan compositions were elucidated based on acquired MS^1 masses, and glycan structures were assessed based on chromatographic retention time and MS^2 spectra obtained by CID fragmentation when available (see Table S1 and supplementary information 2 and 3). Due to limited availability of the batches of Ovitrelle® that were used for released glycan analysis glycopeptide, subunit and dimer data was obtained from another batch. Therefore, to account for potential batch variability, two batches of Ovitrelle® were analyzed at the level of released glycans and showed no differences in the glycans present (Fig. S2 and S3). Since released glycan results were used in this approach to build a broad qualitative library of glycans and the identification and semi-quantitation is based on glycopeptide data, it was considered reasonable to use the data from two other Ovitrelle® batches at this structural level.

Already at released *N*-glycan level, u-hCG showed to be more heterogeneous than r-hCG with 34 *N*-glycans identified compared to 19 *N*-glycans of r-hCG (see Table S1 and supplementary information 2 and 3), of which only three were highly abundant: A2S2F, A2S2 and A2G1S1 (Figs. S2 and S4). To some extent, this heterogeneity is caused by the variability in sialic acid linkages of u-hCG of human origin (α-2,3- and α-2,6-linked sialic acids) compared to CHO produced r-hCG (α-2,3-linked sialic acids only) that can be separated by PGC chromatography [49], see, for example glycans 20a and 20b in the supplementary information. This data was collected in a library of *N*-glycan structures for its later use in glycopeptide identification. The library was extended with other 91 glycan structures comprising all the possible degrees of core-fucosylation, galactosylation and sialylation of the structures found and pentaantennary *N*-glycans, for a total of 129 different *N*-glycans. As an example, since the *N*-glycan A4S4F was found in Pregnyl®, we manually added the corresponding afucosylated structure A4S4 but also all the other possible related tetraantennary structures with different degrees of sialylation (*e.g.* A4G3S1) and galactosylation (*e.g.* A4G2S1), for a total of 32 tetraantennary structures manually added to the library (see Supplement Table S1).

Released *O*-glycan analysis showed the presence of only three different *O*-glycan structures in r-hCG, core 1, core 1 + 1xS and core 1 + 2xS (Fig. S3, see Table S2 for glycan nomenclature and supplementary information 2 and 4). Of note, u-hCG showed also the glycans core 1 + 1xS and core 1 + 2xS as most abundant structures but in addition core 2 *O*-glycans were also found (Fig. S5, supplementary information 2 and 4). From this data, two libraries of *O*-glycan structures were built, one related to r-hCG and one to u-hCG (Table S2).

Tryptic glycopeptides were produced upon reduction, cysteine alkylation and trypsin digestion of u-hCG and r-hCG and analyzed by micro RP-HPLC-MS/MS for *N*-glycopeptides (Figs. S6–9) and nano RP-HPLC-MS/MS for *O*-glycopeptides (Figs. S10–11). The results showed major differences in the *N*-glycan structures attached to different glycosylation sites and in the glycosylation of r- and u-hCG. Of note, the most abundant *N*-glycans for the two glycosylation sites in hCGα (N58 and N72) were monoantennary *N*-glycans (A1S1) in u-hCGα, while complex biantennary *N*-glycans (A2G1S1 and A2S2) prevailed in r-hCGα (Fig. S12). Regarding hCGβ, 39% of site N13 was unmodified in r-hCGβ, while this was observed to a much lower extent for u-hCGβ with only 13% unmodified peptide at site N13. For r-hCGβ, the main glycovariants were carrying biantennary complex *N*-glycans half core-fucosylated in site N13 and completely core-fucosylated in site N30. Interestingly, site N13 of u-hCGβ, showed a predominance of biantennary complex *N*-glycans and paucimannose *N*-glycans (M2, M2F, M3, M3F) while site N30 showed a predominance of biantennary complex *N*-glycans (Fig. S13).

Unambiguous specific assignment of *O*-glycosylation sites was prevented by the close proximity of the serine and threonine residues in the *C*-terminus of hCGβ. Three different glycopeptides were formed upon tryptic digestion, the first (from site 115-122) carrying four serines, the second (from site 123-133) with three serines, and the third (from site 134-145) containing one serine and one threonine that could potentially be *O*-glycosylated. For r-hCG, 2 Core1 + 2xS was the most abundant structure for the peptides from site 115-122 and from site 134-145, while the peptide from site 123-133 carried Core1 + 1xS as main variant, but also 2 Core1 + 3xS was present at a lower abundance (Fig. S14A). Since only Core1 *O*-glycans were identified at released glycan level in r-hCG, the most abundant glycoform at intact level would be the one carrying 5 *O*-glycosylation sites co-modified and a maximum amount of 6 *O*-glycosylation sites may co-exist based on glycopeptide data.

For u-hCG, the peptides from site 115-122 and from site 134-145 were carrying mainly the structure Core1 + 1xS and Core1 + 2xS, respectively, while for the peptide from site 123 to 133 the most abundant structure was 2 x Core1 + 2xS or Core2 + 1G + 2xS (Fig. S14B). When this data is combined at intact level, the glycoforms with highest abundances are expected to carry 3 or 4 *O*-glycosylated sites. It is worth to mention that since also Core2 + 1 G *O*-glycans were found in u-hCG at released glycan level, the prediction of the number of *O*-glycosylation sites co-modified in an intact glycoform is not completely achievable at glycopeptide level.

From r- and u-hCG glycopeptide data, a semi-quantitative site-specific glycan library was built and used for intact subunit glycoform annotations.

3.3. Urinary vs recombinant hCG: Intact hCG subunit glycoforms analysis by RP-HPLC-MS

Intact glycoforms of r- and u-hCG subunits were analyzed by RP-HPLC-MS. The denaturing conditions of the mobile phase composition utilized for the separation, consisting of a gradient of acidified organic modifier and water, combined with the temperature of 60 °C caused the dissociation of the hCG dimer directly in the HPLC system and the separation of the hCGα from the hCGβ subunits (Fig. S15A). This chromatographic method also enabled the partial resolution of sialic acid variants of hCGα and β subunits (Fig. S15A). Isotopically resolved raw mass spectra of r- and u-hCG subunit glycoforms were acquired using this MS-based approach (Fig. S15B). To obtain the masses of the proteoforms, deconvolution to zero-charge mass spectra was performed using the Xtract™ algorithm embedded in the software Xcalibur™. Compared to Ovitrelle®, additional chromatographic peaks could be observed in the total ion current chromatogram (TICC) of Pregnyl® (Fig. S15A). The raw mass spectra associated with the chromatographic peaks of Pregnyl® impurities were deconvoluted resulting in three main patterns of mass peaks (Fig. S16). One pattern of peaks appeared in a mass range from 8.5 to 10.5 kDa and their nature remained unknown, a second pattern of peaks spanned from 12.5 to 15 kDa and was attributed to u-hCGα mono-oxidized, and a third pattern ranged from 20 to 24.5 kDa and corresponded to u-hCGβ glycoforms with an additional mass of 18 Da (Fig. S16).

The deconvoluted mass spectra of intact r- and u-hCGα and β, obtained after deconvolution, are reported in Figs. 1 and 2. The alpha subunit of the recombinant product resulted heavier than the urinary one: the M_r of r-hCGα proteoforms spanned from 13 to 15.5 kDa (Fig. 1A) while u-hCGα reached from 12.5 to 14.8 kDa (Fig. 1B). For both products, sialic acid variants can be observed displaying mass differences of 291 Da. When comparing the two deconvoluted mass spectra of r- and u-hCGβ (Fig. 2A and B), a double pattern of mass peaks can be observed in the r-hCGβ but not in u-hCGβ. This pattern was attributed to the partial lack of *N*-glycosylation at site N13 in r-hCGβ while for u-hCGβ the two sites are mainly *N*-glycosylated in accordance with glycopeptide data (see Fig. S13). Contrary to the alpha subunits, proteoforms of u-hCGβ were heavier (M_r from 15.0 to 21.9 kDa) than r-hCGβ (M_r from 20.5 to 25.0 kDa).

For peak annotation in the deconvoluted mass spectra of u-hCG and r-hCG subunit glycoforms, the four site-specific semi-quantitative glycan libraries built from glycopeptide data were integrated in MoFi (see supplementary files, Site-specific semi-quantitative glycan library_*protein subunit_drug name*.csv). This analysis resulted in hierarchical lists of proteoform compositions for u-hCG and r-hCGα and β. Considering oxidation and other variable modifications, a total of 1672 proteoforms were annotated for r-hCGα, 8111 proteoforms for r-hCG β, 2272 proteoforms for u-hCGα and 17,142 proteoforms for u-hCG β (see supplementary files, Hits MoFi_*protein subunit_drug name_settings*.csv). To retrieve the number of glycoforms unraveled, we summed the fractional abundances of glycoforms with the same glycans carrying different additional modifications (*e.g.* oxidized/non-oxidized or sequence variants). This resulted in an annotation of 283 glycoforms for r-hCGα and 4316 glycoforms for r-hCGβ (see supplementary files Subunit glycoform list with abundances_r-hCG_Ovitrelle.xlsx) and 401 glycoforms for u-hCGα and 9979 glycoforms for u-hCGβ (see supplementary files Subunit glycoform list with abundances_u-hCG_Pregnyl.xlsx). The main glycoform structures annotated for the most abundant mass peaks are reported in Figs. 1 and 2. It is worth to mention that when comparing the fractional abundances of glycoforms (see supplementary files Subunit glycoform list with abundances_*protein_drug name*.xlsx) with the relative intensity of the mass peaks in the deconvoluted mass spectrum, the most intense mass peak not always corresponded to the mass of the most abundant glycoform. For r-hCGα the main glycoforms carried two biantennary complex *N*-glycans with two or three sialic acid residues attached (A2G1S1/A2S2 or A2G1S1/A2G1S1, glycoforms indicated by letter D in Fig. 1A) while the main glycoform of u-hCGα comprised one monoantennary and one biantennary complex *N*-glycan with three sialic acids attached (A1S1/A2S2, glycoform indicated by letter C in Fig. 1B).

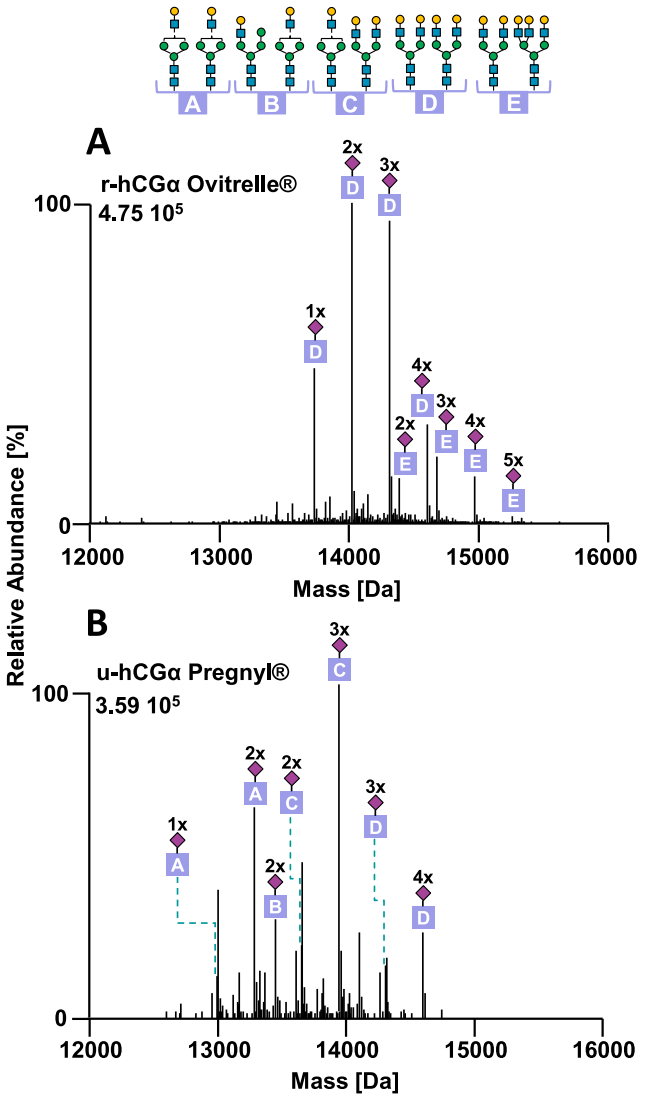


Fig. 1. Deconvoluted mass spectra of proteoforms present in **A)** r-hCGα, Ovitrelle® and **B)** u-hCGα, Pregnyl® obtained by RP-HPLC-MS analysis. The main glycoforms annotated are indicated by letters A-E. The numbers above the peaks refer to the number of sialic acid attached to the *N*-glycan structures. A complete list of annotations can be found in the supplementary files, Subunit glycoform list with abundances_*protein_drug name*.xlsx

When comparing r- and u-hCGβ, the most abundant glycoform for r-hCGβ was "Unmod/A2S2F-2 x Core1 + 2xS/2 x Core1 + 3xS/Core1+2xS" with only site N30 *N*-glycosylated and five sites *O*-glycosylated (glycoform indicated by letter B in Fig. 2A). For u-hCGβ, "A2S2/A2S2F-2 x Core1 + 2xS or Core2+1xG+2xS/Core1+2xS/Core1+2xS" was the most abundant glycoform with both sites *N*-glycosylated and four sites *O*-glycosylated (glycoform indicated by letter E in Fig. 2B).

3.4. Urinary vs recombinant hCG: Intact hCG heterodimer glycoform analysis by native SE-HPLC-MS

To automate the acquisition of native mass spectra of hCG non-covalent heterodimers, an SE-HPLC-MS method was developed to on-line buffer exchange the biopharmaceuticals Pregnyl® and Ovitrelle® directly in the HPLC system to conditions favoring the formation of the

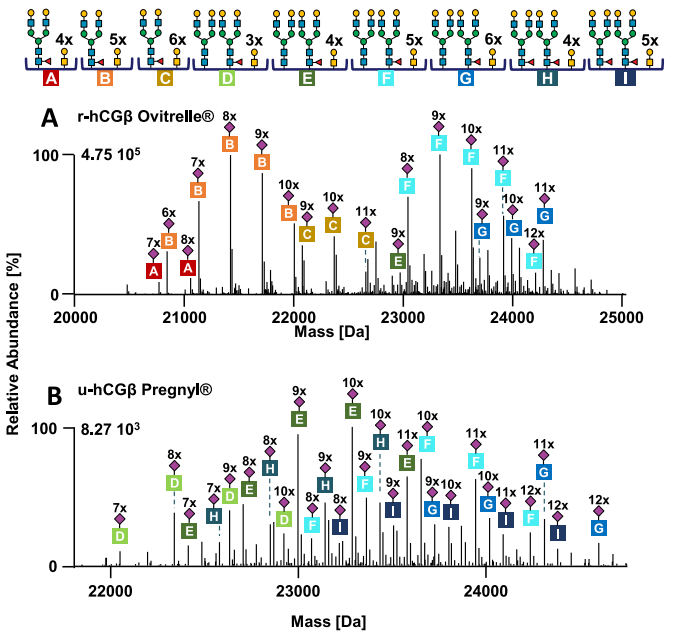


Fig. 2. Deconvoluted mass spectra of proteoforms present in **A)** r-hCGβ, Ovitrelle® and **B)** u-hCGβ, Pregnyl® obtained by RP-HPLC-MS analysis. The main glycoforms annotated are indicated by letters A-I. The numbers above the peaks refer to the number of sialic acids attached to the glycan structures. A complete list of annotations can be found in supplementary files, Subunit glycoform list with abundances_*protein_drug name*.xlsx

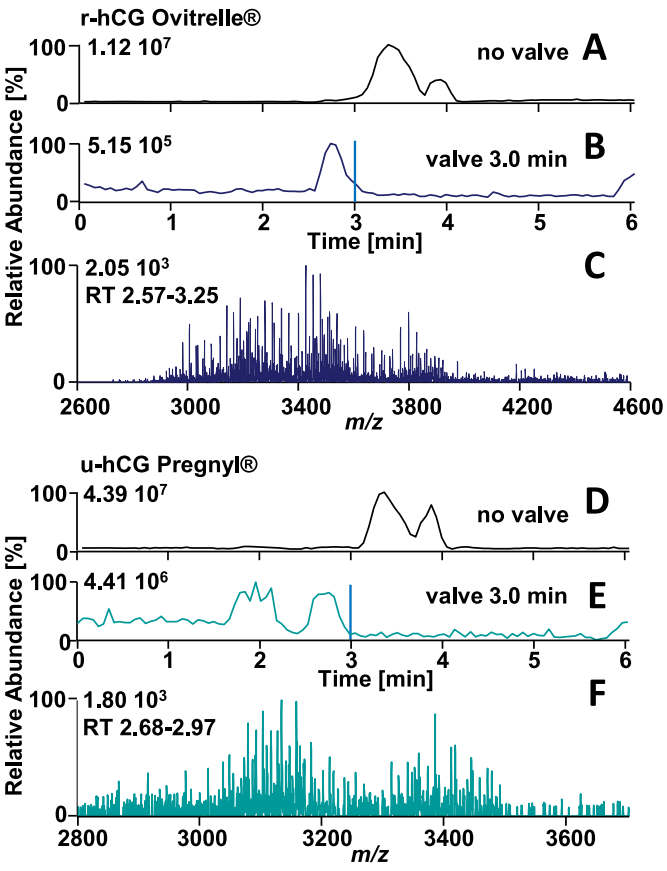


Fig. 3. Results of native SE-HPLC-MS analysis of r- and u-hCG dimer. Chromatogram **A)** without and **B)** with valve switching at 3.0 min of the drug product Ovitrelle®, r-hCG. **C)** raw mass spectrum of r-hCG dimer. Chromatogram **D)** without and **E)** with valve switching at 3.0 min of the drug product Pregnyl®, u-hCG. **F)** raw mass spectrum of u-hCG dimer.

noncovalent heterodimer complexes. The use of a short SE column (3 cm length) allowed the separation of the drug additives (polymers, salts, sugars etc.) from the glycoprotein based on their differences in the molecular mass in a short time (5 min). The glycoprotein eluted in the void volume of the chromatographic column while the additives eluted later. The removal of additives such as salts and polymers is fundamental for the acquisition of native mass spectra to prevent sodium adduct formation and ionization suppression during electrospray ionization. The column was pre-conditioned with 150 mmol L^{-1} AmAc, the standard solution used to perform native MS, therefore the glycoprotein-heterodimer eluted in AmAc and online buffer exchange was achieved directly in the HPLC system. A six-port-valve was installed between the column outlet and the ESI source to divert the HPLC flow to the waste after the elution of the glycoprotein (after 3.0 min) to the ESI source. In this way, contamination of the mass spectrometer by salts and polymers was avoided (Fig. 3A, B, D and E).

Using this method, raw mass spectra of the non-covalent dimer of r-hCG and for the first time of u-hCG were acquired (Fig. 3C and F). The three charge states of r- and u-hCG from 10+ to 12+ in the raw mass spectrum together with the preservation of the non-covalent interaction in the heterodimer indicate the maintenance of the quasi-native conformation of the protein. Raw mass spectra were deconvoluted using Thermo Scientific™ BioPharma Finder™ software. The zero-charge mass spectrum of r-hCG dimer glycoforms spanned a M_r of 35–40 kDa and showed two main peak patterns (from 35 to 37 kDa and from 37 to 40 kDa) (Fig. 4) due to the fact that around 40% of site N13 of the r-hCGβ was unmodified (Fig. S13). The dimeric glycoforms of u-hCG showed a different glycosylation pattern compared to r-hCG, the M_r spanned from 35 to 40 kDa but the double pattern of mass peaks present in r-hCG was not observable for the urinary version (Fig. 4B). This is in accordance with glycopeptide data, because u-hCGβ did not show a high percentage of unmodified N13 site as the recombinant product (Fig. S13). To annotate the dimeric glycoforms of r- and u-hCG, the libraries of u-hCG and r-hCGα and β glycoforms were built from the annotations of intact subunit glycoforms and integrated in MoFi. To avoid the explosion of the combinatorial space of MoFi, a cut-off of 0.3% of fractional abundance of the glycoforms was used (see supplementary files, Subunit glycoform library_*protein_drug name*_0.3% cutoff.xlsx). Taking this cut-off into account, the results showed the presence of 1481 glycoforms of r-hCG dimer and 1167 glycoforms of u-hCG dimer (see supplementary files, Hits MoFi_*protein* Dimer_*drug name*_5 Da_0.3% cutoff.xlsx). The most abundant glycoform for r-hCG dimer carried the glycoform A2G1S1/A2S2 for r-hCGα and the glycoform Unmod/A2S2F/Core1+2xS/2x Core1+3xS/2x Core1+3xS for r-hCGβ with a M_r of 36343.5 Da (glycoform indicated by a green 7-point star in Fig. 4A). For u-hCG dimer, the most abundant glycoform was A1S1/A2S2 for u-hCGα and A2S2/A2S2F/2x Core1+2xS or Core2+1xG+2xS/Core1+2xS/Core1+2xS for u-hCGβ with a M_r of 37237.2 Da (glycoform indicated by an orange triangle in Fig. 4B).

3.5. Comprehensive characterization of r-hFSH by hybrid HPLC-MS

In a next step, the hybrid HPLC-MS approach used to characterize r- and u-hCG, was applied to study the biopharmaceutical Gonal-f® (r-hFSH) at glycopeptide, intact subunit and dimer-levels. First, glycopeptides were analyzed by micro RP-HPLC-MS/MS. The sample was digested with trypsin for r-hFSHα *N*-glycopeptides and with chymotrypsin for r-hFSHβ *N*-glycopeptides, as the use of the latter peptidase enabled the identification of more glycovariants, as previously reported by Grass et al. [40]. *N*-glycopeptides were identified based on the *N*-glycan library compromising 129 different *N*-glycan structures, previously created for r- and u-hCG. The most abundant glycovariants of the two glycosylation sites of r-hFSHα (N52, N78) carried complex biantennary *N*-glycans with a different sialylation degree (A2G1S1 or A2S2)

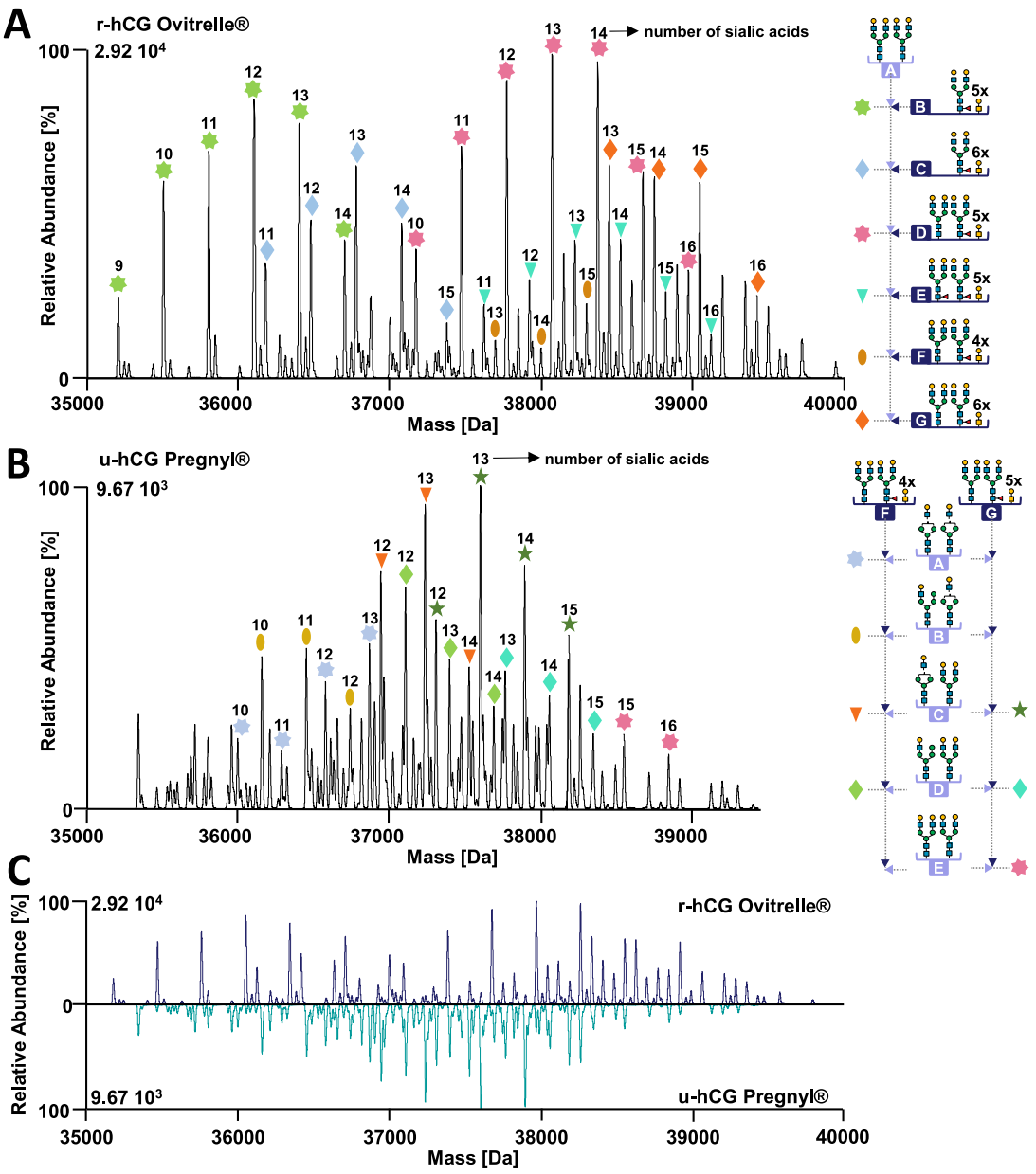


Fig. 4. Deconvoluted mass spectra of proteoforms present in **A)** r-hCG dimer, Ovitrelle® and **B)** u-hCG dimer, Pregnyl® obtained by native SE-HPLC-MS analysis. The main glycoforms annotated are indicated by colored symbols and represent a combination of glycans present on hCGα (light violet) and hCGβ (dark violet). The number above the peaks refer to the number of sialic acids attached to the glycan structures. A complete list of annotations can be found in supplementary files, Hits MoFi_*protein* Dimer_*drug name*_5 Da_0.3% cutoff.xlsx. **C)** Mirror plot of Ovitrelle® and Pregnyl® deconvoluted mass spectra. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

(Fig. S19). Regarding r-hFSHβ, also embedding two *N*-glycosylation sites (N7, N24), 24.5% of N7 and 69.0% of N24 were detected as unmodified. In site N7, the most abundant *N*-glycan structures consisted of core-fucosylated complex tri- or tertraantennary *N*-glycans (Fig. S19). Site N24 of r-hFSHβ compromised complex biantennary *N*-glycans (Fig. S19) as main variants. Although the human gonadotropins share the same α subunit sequence and even if Ovitrelle® and Gonal-f® are both produced recombinantly by the same pharmaceutical company, the glycopeptide profiles identified were vastly different (Figs. S12 and S19). When comparing r-hCGα and r-hFSHα, the latter had a higher sialylation degree and additionally also showed triantennary *N*-glycans on both glycosylation sites.

Intact subunits of r-hFSH were analyzed by RP-HPLC-MS enabling the separation of r-hFSHβ eluting before r-hFSHα (Fig. S20A). Isotopically resolved mass spectra of r-hFSH subunits were acquired for the first time (Fig. S20B), deconvoluted using the Xtract™ algorithm and annotated *via* MoFi integrating the site-specific semi-quantitative glycan libraries (see supplementary files, Site-specific semi-quantitative glycan library_r-FSH*subunit*_Gonal-f.xlsx) as previously described. Both subunits of r-hFSH showed *N*-terminal cleavage which was integrated in the bioinformatic workflow. As a result, two hierarchical lists of r-hFSHα and β proteoforms were obtained (see supplementary files, Hits MoFi_*protein subunit_drug name_settings*.csv). The deconvoluted mass spectrum of r-hFSHα proteoforms (Fig. 5A) showed mass peaks ranging from 13.5 to 16 kDa in M_r. The peaks in the deconvoluted mass spectrum of r-FSHβ spanned a M_r of 15–19 kDa (Fig. 5B). The main variants annotated are reported in Fig. 5. After summing up glycoforms with the same glycan composition but carrying different modifications, a total of 118 glycoforms could be identified for r-hFSHα and 111 glycoforms for r-hFSHβ (see supplementary files, Subunit glycoform list with abundances_r-

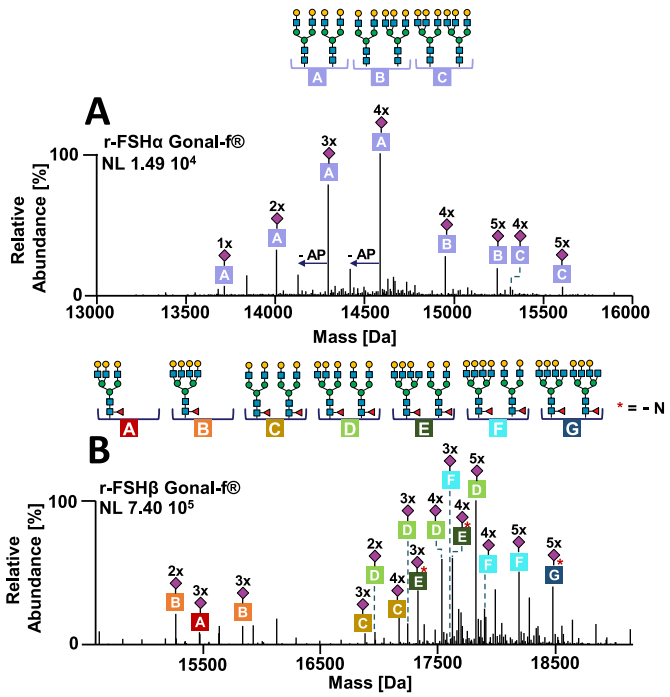


Fig. 5. Deconvoluted mass spectra of proteoforms present in **A)** r-hFSHα and **B)** r-hFSHβ, Gonal-f® obtained by RP-HPLC-MS analysis. The main glycoforms annotated are indicated by letters A-C for r-hFSHα and A-F for r-hFSHβ. The number above the peaks refers to the number of sialic acids attached to the *N*-glycan structures. In panel A, -AP indicates the cleavage of *N*-terminal alanine and proline (-AP) in r-hFSHα. The red asterisk in panel B indicates the cleavage of the *N*-terminal asparagine (-N) in r-hFSHβ. A complete list of annotation can be found in supplementary files, Subunit glycoform list with abundances_r-FSH_Gonal-f.xlsx. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

FSH_Gonal-f.xlsx). Of note, mass peaks 17,167.7, 17,987.9 and 18,479.1 Da in r-hFSHα could only be annotated *via* MoFi when a mass tolerance up to 5 Da was set (see supplementary files, Hits MoFi_r-FSHalpha_Gonal-f_5 Da_ox_AP mod.csv).

Next an intact raw spectrum of r-hFSH dimer was acquired by utilizing native SE-HPLC-MS for online removal of drug additives, as previously described. As the r-hFSH dimer has a lower M_r compared to the hCG dimer, the valve switching time point was delayed to 3.6 min (Fig. 6A and B). For the first time, a raw mass spectrum of r-hFSH dimer was acquired (Fig. 6C). Deconvolution of the charge states 10+ and 11+

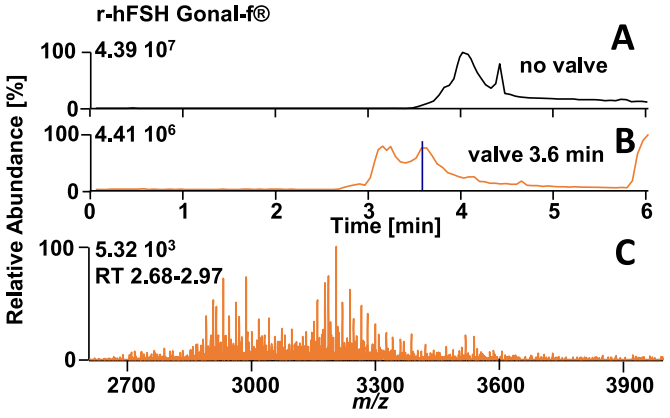


Fig. 6. Results of native SE-HPLC-MS analysis of r-hFSH dimer. Chromatogram **A)** without and **B)** with valve switching at 3.6 min of the drug product Gonal-f®, r-hFSH. **C)** raw mass spectrum of r-hFSH dimer.

was accomplished with the ReSpect™ algorithm embedded in Thermo Scientific™ BioPharma Finder™. The zero-charge mass spectrum of the r-hFSH dimer reached from 29 to 36 kDa, whereas two peak patterns can be observed, a pattern with a low abundance from 29 to 31 kDa and a highly abundant pattern from 31 to 36 kDa (Fig. 7). The pattern with the lower M_r results from unmodified N24 in r-hFSHβ. Surprisingly N24 was unmodified to a large extent (69.0%) at glycopeptide level, however in the intact spectra partially unmodified glycovariants are only present at a low extent. The dimeric glycoforms were annotated as described before based on a library with the previously identified glycoforms for r-hFSHα and for r-hFSHβ with a fractional abundance cut-off of 0.01% to avoid the explosion of the combinatorial space (see supplementary files, Subunit glycoform library_r-FSH_Gonal-f_0.01% cutoff). Using this approach, 1440 glycovariants could be identified in the intact spectra (see supplementary files, Hits MoFi_r-FSH Dimer_Gonal-f_5 Da_N mod.xlsx). Annotation for the most abundant peaks is shown in Fig. 7. The most abundant glycoform for r-hFSH dimer bore the glycoform A2S2/A2S2 for r-hFSHα and the glycoform A3S3F/A2S2F and had a M_r of 32,440 Da.

4. Discussion

We described the in-depth characterization of Ovitrelle®, Pregnyl® and Gonal-f®, three biopharmaceuticals comprising gonadotropins, which are heavily glycosylated proteins. The in-depth characterization of these proteins was fraught with difficulty, because the high heterogeneity of their glycosylation state gives rise to a myriad of different co-existing glycoforms. Moreover, the isobaricity of the different monosaccharides (*e.g.* galactose and mannose, *N*-acetylgalactosamine and *N*-acetylglucosamine) leads to the presence of isobaric glycoforms that are then indistinguishable by MS. To unravel the heterogeneity hidden behind these biopharmaceuticals, the glycoproteins were characterized at the different structural levels of released glycan, glycopeptides and intact subunits and non-covalent heterodimer by employing an alliance of different analytical approaches. This allowed the acquisition of different structural information: glycan from released glycan analysis, site-specific glycan assignment and relative quantitation of glycans from glycopeptide analysis, and masses of intact glycoforms from intact subunits and dimer analysis. This information was then integrated using a bioinformatic workflow that finally allowed the annotation of thousands of different glycoforms for the different drug products. The analysis at the different structural levels is crucial not only to determine the molecular structure of the glycans, but also to reduce the number of possible glycan combinations for annotation at dimeric level. Taking as an example r-hCG, if a direct annotation of the dimer would have been performed using glycopeptide data and not intact subunit data, this would have led to a possible number of intact dimeric glycoforms of 12,768,000, just based on a combinatorial model (multiplication of the number of glycovariants identified per glycosylation site, see Figs. S12–14). On the opposite, when first the glycopeptide data is integrated into the subunits data, we reduced the number of possible combinations at intact dimeric level to 1,221,428 (see supplementary files subunit glycoform list with abundances_r-hCG_Ovitrelle.xlsx). Moreover, using this approach, the annotations at dimeric level are supported by the information retrieved at the different structural levels. Data integration also strengthens our annotation results, because the glycoforms annotated at dimeric levels are first observed as monomers at subunit level.

From an analytical perspective, a semi-automated SE-HPLC-MS approach was demonstrated to acquire native mass spectra of the non-covalent heterodimer of r-hCG, and for the first time u-hCG and r-hFSH. Native mass spectra can only be obtained when ions, like Na^+, forming adducts with the protein, and polymers, that suppress protein ionization by electrospray are removed. Conventionally, protein buffer exchange with solutions compatible with native MS (normally 150 mmol L^{-1} AmAc) is performed "offline" by using preparative size

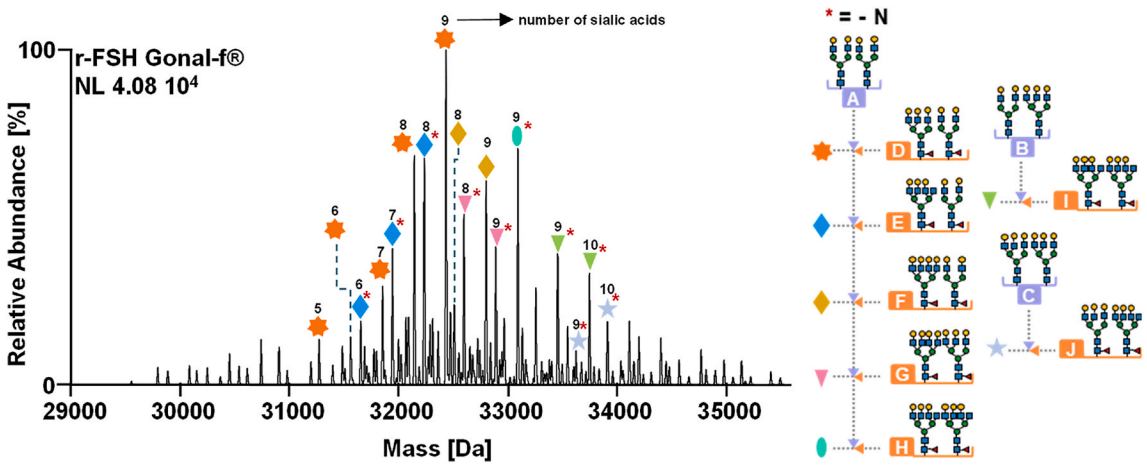


Fig. 7. Deconvoluted mass spectrum of proteoforms present in r-hFSH dimer, Gonal-f® obtained by native SE-HPLC-MS analysis. The main glycoforms annotated are indicated by colored symbols and represent a combination of glycans present on r-hFSHα (light violet) and r-hFSHβ (orange). The numbers above the peaks refer to the number of sialic acids attached to the glycan structures. The red asterisk indicates the cleavage of the *N*-terminal asparagine (-N) in r-hFSHβ. A complete list of annotation can be found in supplementary files, Hits MoFi_r-FSH Dimer_Gonal-f_5 Da_N mod.xlsx. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

exclusion columns or ultrafiltration. The buffer-exchanged protein is then infused into the mass spectrometer *via* a static nano ESI source requiring considerable manual handling, and a native mass spectrum is acquired. Here, we managed to online buffer-exchange proteins in biopharmaceuticals by chromatographically separating low M_r drug additives such as detergents or salts from the protein dimer using a semi-automated, fast SE-HPLC-MS method. Of note, this novel approach preserved the non-covalent interactions between the alpha and the beta subunit of the gonadotropin dimers, allowing the acquisition of their native mass spectra.

The here described hybrid HPLC-MS approach, although being followed by a challenging data processing workflow, is easily adaptable to other biopharmaceuticals. By using our previously published RP-HPLC-MS method [5], isotopically resolved mass spectra of r-hCG and, for the first time, u-hCG and r-hFSH subunits were acquired. This allowed high throughput spectra deconvolution and accurate assessment of the monoisotopic mass of gonadotropin subunits. With the aid of bioinformatic annotation, we were able to identify 283 glycoforms for r-hCGα and 4316 glycoforms for r-hCGβ, 401 glycoforms for u-hCGα and 9979 glycoforms for u-hCGβ and 118 glycoforms for r-hFSHα and 111 glycoforms for r-hFSHβ. Additionally, the use of the here reported SE-HPLC-MS approach to online buffer exchange human gonadotropins enabled us to obtain spectra of the intact non-covalent heterodimer of r-hCG and for the first time of u-hCG and r-hFSH. Intact mass spectra were readily annotated using the annotation tool MoFi, which allowed the annotation of 1481 glycoforms for r-hCG dimer, 1167 glycoforms for u-hCG dimer and 1440 glycoforms for r-hFSH dimer using an abundance cutoff of 0.3% and 0.01%, respectively.

From the glycoform annotations obtained, observations about critical quality attributes of these biopharmaceuticals could be drawn. The analytical characterization showed that Pregnyl® and Ovitrelle® are vastly different from a glycosylation perspective (Figs. 2–4) even if they are analogously prescribed to patients for the treatment of infertility. In particular, Ovitrelle® was more *O*-glycosylated than Pregnyl®, however Pregnyl® showed a higher sialylation degree of 83% compared to 72% for Ovitrelle® at dimer level (for calculation see Fig. S21). It is known that *O*-glycosylation and a higher degree of sialylation increase the serum half-life of the drug products [13,15–17], thus from an *O*-glycosylation perspective it can be assumed that Ovitrelle® would be more stable than Pregnyl® in the blood circulation, but from a sialylation point of view Pregnyl® would have a longer serum half-life time. Moreover, Pregnyl® showed more impurities than Ovitrelle® (Fig. S16), thus pointing out the benefit of the recombinant drug product with regard to purity.

It is interesting to notice that, even if r-hFSH and r-hCG are both recombinantly produced in CHO cells by the same pharmaceutical company and carry the same α subunit, their glycosylation pattern is highly different (Figs. 2A and 5A). In particular, r-hFSHα is characterized by a higher sialylation degree of 79% compared to r-hCGα with 62% (for calculation see Fig. S21) and a higher number of antennae in the *N*-glycan structures. We hypothesize that the differences in the glycosylation patterns could originate from different biotechnological strategies such as different recombinant DNA, expression systems, fermentation conditions (pH, temperature, dissolved O_2, stirring speed, culture medium) and/or downstream processing.

Incongruences between the data of glycopeptides and intact subunits and dimer of r-hFSH were observed. In particular, at r-hFSHβ glycopeptide level, the site N24 was detected as mainly non-glycosylated (Fig. S19), while in the intact spectra of r-hFSHβ and r-hFSH dimer only a small fraction of glycoforms was partially non-glycosylated. This incongruence could either have arisen from glycopeptide or intact data. We hypothesize that the chymotryptic glycopeptide carrying the glycosylation site N24, lacking the nucleophilic amino acids lysine and arginine typical for tryptic peptides, may have suffered from a different electrospray ionization efficiency between the glycosylated and unmodified peptides, leading to a higher abundance of unmodified peptide detected. At intact level, partial glycosylation of r-hFSHβ leads to *m/z* signals of these glycoforms in the mass spectrum appearing at a lower *m/z* range. These signals may potentially overlap with the signals of the adjacent lower charge state in the mass spectra of the r-hFSHβ and the r-hFSH dimer (Figs. S20B and 6C), causing the inability to detect the partially non-glycosylated glycoforms. The impossibility to fully match variant distribution among the different structural level was already discussed with mathematical rigor in a previous publication [50]. However even with some limitations, more than thousands of different glycoforms could be identified for r-hFSH dimer for the first time.

The here suggested semi-automated hybrid HPLC-MS method proposes an approach for characterization of complex glycoproteins and allows studying their heterogeneity which could be used in terms of good manufacturing practice in the process of drug approval.

5. Conclusion

An experimental and computational workflow involving a hybrid

HPLC-MS approach and stepwise data integration across different levels of glycoprotein structure allowed to comprehensively characterize gonadotropin-based biopharmaceuticals. The strong alliance of the hybrid HPLC-MS approach with the data integration at the different structural levels *via* a bioinformatic workflow enabled the annotation of more than one thousand different glycoforms at intact dimer levels of r-hCG, u-hCG and r-hFSH. The method proved to be useful to compare biopharmaceuticals, *e.g.* obtained from different sources (recombinant *vs* urinary) or by different biotechnological strategies (r-hFSHα and r-hCGα). Moreover, the semi-automation of native mass spectra acquisition of gonadotropin non-covalent heterodimer puts the basis for the approach to be implemented in the quality control environment to assess similarity of drug products.

Funding

This work was supported by the Christian Doppler Society, Austria, the Austrian Federal Ministry for Digital and Economic Affairs, the National Foundation of Research, Technology, and Development, a Start-up Grant of the State of Salzburg. Moreover, grants of the Austrian Science Fund (W1213, FG 12 N) financially supported this work.

CRediT authorship contribution statement

Fiammetta Di Marco: Conceptualization, Software, Formal analysis, Investigation, Data curation, Writing – original draft, Writing – review & editing, Visualization. **Gabriele Blümel:** Software, Formal analysis, Investigation, Data curation, Writing – original draft, Writing – review & editing, Visualization. **Constantin Blöchl:** Formal analysis, Investigation, Data curation, Writing – review & editing. **Manfred Wuhrer:** Resources, Supervision, Writing – review & editing. **Christian G. Huber:** Conceptualization, Resources, Writing – review & editing, Supervision, Project administration, Funding acquisition, All authors agreed to the final version of the manuscript.

Declaration of competing interest

The authors declare the following competing financial interests: Thermo Fisher Scientific and Novartis AG provided financial support for the Christian Doppler Laboratory for Innovative Tools for Biosimilar Characterization. The salary of Christian G. Huber was partly funded by the Christian Doppler Laboratory for Biosimilar Characterization. The authors declare no other competing financial interest.

Data availability

Raw files and Byonic search results are available from Zenodo (10.5281/zenodo.7547480).

Acknowledgements

We thank Wolfgang Esser-Skala, Maximilian Lebede, and Therese Wohlschlager for the scientific inputs given to the project. Tao Zhang is acknowledged for his expert assistance on released glycan analysis.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.aca.2023.341574.

References

[1] L.A. Cole, Biological functions of hCG and hCG-related molecules, Reprod. Biol. Endocrinol. 8 (2010) 1–14, https://doi.org/10.1186/1477-7827-8-102.
[2] R.C. Anderson, C.L. Newton, R.A. Anderson, R.P. Millar, Gonadotropins and their analogs: current and potential clinical applications, Endocr. Rev. 39 (2018) 911–937, https://doi.org/10.1210/er.2018-00052.
[3] A. Bachmann, T. Rabe, K. Bühler, G. Griesinger, N. Sänger, C. Albring, C. Gnoth, H. Kentenich, K. König, J. Krüssel, E. Merkle, T. Steck, T. Strowitzki, A. Weghofer, W. Würfel, The gonadotropins FSH and LH and their use in adult women an overviewbr(joint statement by the German society for gynecological endocrinology and reproductive medicine [DGGEF] and the German professional association of gynecologists [BVF]), J. Für Reproduktionsmedizin Und Endokrinol. - J. Reprod. Med. Endocrinol. 12 (2015) 360–376. www.kup.at/repromedizin. (Accessed 26 October 2022).
[4] G.R. Bousfield, J.A. Dias, Synthesis and secretion of gonadotropins including structure-function correlates. https://doi.org/10.1007/s11154-011-9191-3, 2011.
[5] M. Lebede, F. Di Marco, W. Esser-Skala, R. Hennig, T. Wohlschlager, C.G. Huber, Exploring the chemical space of protein glycosylation in noncovalent protein complexes: an expedition along different structural levels of human chorionic gonadotropin by employing mass spectrometry, Anal. Chem. 93 (2021) 10424–10434, https://doi.org/10.1021/ACS.ANALCHEM.1C02199/SUPPL_FILE/AC1C02199_SI_003.ZIP.
[6] A.J. Lapthorn, D.C. Harris, A. Littlejohn, J.W. Lustbader, R.E. Canfield, K. J. Machin, F.J. Morgan, N.W. Isaacs, Crystal structure of human chorionic gonadotropin, Nature 369 (1994) 455–461, https://doi.org/10.1038/369455a0.
[7] K.M. Fox, J.A. Dias, P. Van Roey, Three-dimensional structure of human follicle-stimulating hormone, Mol. Endocrinol. 15 (2001) 378–389, https://doi.org/10.1210/MEND.15.3.0603.
[8] R. de B.F. Leão, S.C. Esteves, Gonadotropin therapy in assisted reproduction: an evolutionary perspective from biologics to biotech, Clinics 69 (2014) 279–293, https://doi.org/10.6061/clinics/2014(04)10.
[9] 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 6 (2011), https://doi.org/10.1371/journal.pone.0017815.
[10] S. Birken, Y. Chen, M.A. Gawinowicz, J.W. Lustbader, S. Pollak, G. Agosto, R. Buck, J. O'Connor, Separation of nicked human chorionic gonadotropin (hCG), intact hCG, and hCG beta fragment from standard reference preparations and raw urine samples, Endocrinology 133 (1993) 1390–1397, https://doi.org/10.1210/ENDO.133.3.7689955.
[11] S. Daya, Updated meta-analysis of recombinant follicle-stimulating hormone (FSH) versus urinary FSH for ovarian stimulation in assisted reproduction, Fertil. Steril. 77 (2002) 711–714, https://doi.org/10.1016/S0015-0282(01)03246-0.
[12] E.G. Papanikolaou, H. Fatemi, M. Camus, D. Kyrou, N.P. Polyzos, P. Humaidan, B. Tarlatzis, P. Devroey, H. Tournaye, Higher birth rate after recombinant hCG triggering compared with urinary-derived hCG in single-blastocyst IVF antagonist cycles: a randomized controlled trial, Fertil. Steril. 94 (2010) 2902–2904, https://doi.org/10.1016/j.fertnstert.2010.04.077.
[13] A. Ulloa-Aguirre, J.A. Dias, G.R. Bousfield, Gonadotropins and the importance of glycosylation, in: Post-translational Modif. Protein Biopharm., Wiley, 2009, pp. 109–147, https://doi.org/10.1002/9783527626601.ch5.
[14] A. Stockell Hartree, A.G.C. Renwick, Molecular structures of glycoprotein hormones and functions of their carbohydrate components, Biochem. J. 287 (1992) 665, https://doi.org/10.1042/BJ2870665.
[15] P.S. Lapolt, K. Nishimori, F.A. Fares, E. Perlas, I. Boime, A.J.W. Hsueh, Enhanced stimulation of follicle maturation and ovulatory potential by long acting follicle-stimulating hormone agonists with extended carboxyl-terminal peptides, Endocrinology 131 (1992) 2514–2520, https://doi.org/10.1210/ENDO.131.6.1446593.
[16] R. Hoermann, K. Kubota, S.M. Amir, Role of subunit sialic acid in hepatic binding, plasma survival rate, and in vivo thyrotropic activity of human chorionic gonadotropin, Thyroid 3 (1993) 41–47, https://doi.org/10.1089/thy.1993.3.41.
[17] L. Wide, K. Eriksson, P.M. Sluss, J.E. Hall, Serum half-life of pituitary gonadotropins is decreased by sulfonation and increased by sialylation in women, J. Clin. Endocrinol. Metab. 94 (2009) 958–964, https://doi.org/10.1210/jc.2008-2070.
[18] T. Zhang, K. Madunić, S. Holst, J. Zhang, C. Jin, P. Ten Dijke, N.G. Karlsson, K. Stavenhagen, M. Wuhrer, Development of a 96-well plate sample preparation method for integrated N- and O-glycomics using porous graphitized carbon liquid chromatography-mass spectrometry, Mol. Omi. 16 (2020) 355–363, https://doi.org/10.1039/C9MO00180H.
[19] P.H. Jensen, N.G. Karlsson, D. Kolarich, N.H. Packer, Structural analysis of N- and O-glycans released from glycoproteins, Nat. Protoc. 7 (2012) 1299–1310, https://doi.org/10.1038/nprot.2012.063.
[20] M. Wuhrer, A.R. De Boer, A.M. Deelder, Structural glycomics using Hydrophilic interaction chromatography (HILIC) with mass spectrometry, Mass Spectrom. Rev. 28 (2009) 192–206, https://doi.org/10.1002/mas.20195.
[21] H. Yin, J. Zhu, Methods for quantification of glycopeptides by liquid separation and mass spectrometry, Mass Spectrom. Rev. (2022) 1–31, https://doi.org/10.1002/MAS.21771.
[22] L.R. Ruhaak, G. Xu, Q. Li, E. Goonatilleke, C.B. Lebrilla, Mass spectrometry approaches to glycomic and glycoproteomic analyses, Chem. Rev. 118 (2018) 7886–7930, https://doi.org/10.1021/acs.chemrev.7b00732.
[23] Y. Yang, V. Franc, A.J.R. Heck, Glycoproteomics: a balance between high-throughput and in-depth analysis, Trends Biotechnol. 35 (2017) 598–609, https://doi.org/10.1016/j.tibtech.2017.04.010.
[24] J. Camperi, A. Combes, J. Guibourdenche, D. Guillarme, V. Pichon, T. Fournier, 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, J. Pharm. Biomed. Anal. 161 (2018) 35–44, https://doi.org/10.1016/j.jpba.2018.07.056.

[25] A.J.R. Heck, R.H.H. Van Den Heuvel, Investigation of intact protein complexes by mass spectrometry, Mass Spectrom. Rev. 23 (2004) 368–389, https://doi.org/10.1002/mas.10081.
[26] R.J. Rose, E. Damoc, E. Denisov, A. Makarov, A.J.R. Heck, High-sensitivity Orbitrap mass analysis of intact macromolecular assemblies, Nat. Methods 9 (2012) 1084–1086, https://doi.org/10.1038/nmeth.2208.
[27] S. Rosati, R.J. Rose, N.J. Thompson, E. Van Duijn, E. Damoc, E. Denisov, A. Makarov, A.J.R. Heck, Exploring an orbitrap analyzer for the characterization of intact antibodies by native mass spectrometry, Angew. Chem. Int. Ed. 51 (2012) 12992–12996, https://doi.org/10.1002/anie.201206745.
[28] N.J. Thompson, S. Rosati, A.J.R. Heck, Performing native mass spectrometry analysis on therapeutic antibodies, Methods 65 (2014) 11–17, https://doi.org/10.1016/j.ymeth.2013.05.003.
[29] T. Wohlschlager, K. Scheffler, I.C. Forstenlehner, W. Skala, S. Senn, E. Damoc, J. Holzmann, C.G. Huber, Native mass spectrometry combined with enzymatic dissection unravels glycoform heterogeneity of biopharmaceuticals, Nat. Commun. 9 (2018), https://doi.org/10.1038/s41467-018-04061-7.
[30] S. Tamara, M.A. Den Boer, A.J.R. Heck, High-resolution native mass spectrometry, Chem. Rev. 122 (2022) 7269–7326, https://doi.org/10.1021/ACS.CHEMREV.1C00212/ASSET/IMAGES/MEDIUM/CR1C00212_0029.GIF.
[31] Y. Yang, F. Liu, V. Franc, L.A. Halim, H. Schellekens, A.J.R. Heck, Hybrid mass spectrometry approaches in glycoprotein analysis and their usage in scoring biosimilarity, Nat. Commun. 7 (2016) 1–10, https://doi.org/10.1038/ncomms13397.
[32] S. Tamara, V. Franc, A.J.R. Heck, A wealth of genotype-specific proteoforms fine-tunes hemoglobin scavenging by haptoglobin, Proc. Natl. Acad. Sci. U.S.A. 117 (2020) 15554–15564, https://doi.org/10.1073/pnas.2002483117.
[33] Y.H. Lin, J. Zhu, S. Meijer, V. Franc, A.J.R. Heck, Glycoproteogenomics: a frequent gene polymorphism affects the glycosylation pattern of the human serum fetuin/α-2-HS-Glycoprotein, Mol. Cell. Proteomics 18 (2019) 1479–1490, https://doi.org/10.1074/mcp.RA119.001411.
[34] E. Deslignière, M. Ley, M. Bourguet, A. Ehkirch, T. Botzanowski, S. Erb, O. Hernandez-Alba, S. Cianférani, Pushing the limits of native MS: online SEC-native MS for structural biology applications, Int. J. Mass Spectrom. 461 (2021), https://doi.org/10.1016/j.ijms.2020.116502.
[35] Z.L. VanAernum, F. Busch, B.J. Jones, M. Jia, Z. Chen, S.E. Boyken, A. Sahasrabuddhe, D. Baker, V.H. Wysocki, Rapid online buffer exchange for screening of proteins, protein complexes and cell lysates by native mass spectrometry, Nat. Protoc. 15 (2020) 1132–1157, https://doi.org/10.1038/s41596-019-0281-0.
[36] A. Gervais, Y.A. Hammel, S. Pelloux, P. Lepage, G. Baer, N. Carte, O. Sorokine, J. M. Strub, R. Koerner, E. Leize, A. Van Dorsselaer, Glycosylation of human recombinant gonadotrophins: characterization and batch-to-batch consistency, Glycobiology 13 (2003) 179–189, https://doi.org/10.1093/GLYCOB/CWG020.
[37] L. Valmu, H. Alfthan, K. Hotakainen, S. Birken, U.-H. Stenman, Site-specific glycan analysis of human chorionic gonadotropin b-subunit from malignancies and pregnancy by liquid chromatography-electrospray mass spectrometry, Glycobiology 16 (2006) 1207–1218, https://doi.org/10.1093/glycob/cwl034.
[38] X. Bai, D. Li, J. Zhu, Y. Guan, Q. Zhang, L. Chi, From individual proteins to proteomic samples: characterization of O-glycosylation sites in human chorionic gonadotropin and human-plasma proteins, Anal. Bioanal. Chem. 407 (2015) 1857–1869, https://doi.org/10.1007/s00216-014-8439-7.
[39] H. Zhu, C. Qiu, A.C. Ruth, D.A. Keire, H. Ye, A LC-MS all-in-one workflow for site-specific location, identification and quantification of N-/O- glycosylation in human chorionic gonadotropin drug products, AAPS J. 19 (2017) 846–855, https://doi.org/10.1208/s12248-017-0062-z.
[40] J. Grass, M. Pabst, M. Chang, M. Wozny, F. Altmann, Analysis of recombinant human follicle-stimulating hormone (FSH) by mass spectrometric approaches, Anal. Bioanal. Chem. 400 (2011) 2427–2438, https://doi.org/10.1007/s00216-011-4923-5.
[41] H. Wang, X. Chen, X. Zhang, W. Zhang, Y. Li, H. Yin, H. Shao, G. Chen, Comparative assessment of glycosylation of a recombinant human FSH and a highly purified FSH extracted from human urine, J. Proteome Res. 15 (2016) 14, https://doi.org/10.1021/acs.jproteome.5b00921.
[42] R. Mastrangeli, A. Satwekar, F. Cutillo, C. Ciampolillo, W. Palinsky, S. Longobardi, In-vivo biological activity and glycosylation analysis of a biosimilar recombinant human follicle-stimulating hormone product (Bemfola) compared with its reference medicinal product (GONAL-f), PLoS One 12 (2017), https://doi.org/10.1371/JOURNAL.PONE.0184139.
[43] 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. 412 (2020) 1–13, https://doi.org/10.1007/s00216-020-02794-3.
[44] J. Camperi, A. Combès, T. Fournier, V. Pichon, N. Delaunay, Analysis of the human chorionic gonadotropin protein at the intact level by HILIC-MS and comparison with RPLC-MS, Anal. Bioanal. Chem. (2020), https://doi.org/10.1007/s00216-020-02684-8.
[45] H. Toll, P. Berger, A. Hofmann, A. Hildebrandt, H. Oberacher, H.P. Lenhof, C. G. Huber, Glycosylation patterns of human chorionic gonadotropin revealed by liquid chromatography-mass spectrometry and bioinformatics, Electrophoresis 27 (2006) 2734–2746, https://doi.org/10.1002/elps.200600022.
[46] C.A. Cooper, E. Gasteiger, N. Packer, GlycoMod - a software tool for determining glycosylation compositions from mass spectrometric data, Proteomics 1 (2001) 340–349, https://doi.org/10.1002/1615-9861(200102)1:2<340::AID-PROT340>3.0.CO;2-B.
[47] B. MacLean, D.M. Tomazela, N. Shulman, M. Chambers, G.L. Finney, B. Frewen, R. Kern, D.L. Tabb, D.C. Liebler, M.J. MacCoss, Skyline: an open source document editor for creating and analyzing targeted proteomics experiments, Bioinformatics 26 (2010) 966, https://doi.org/10.1093/BIOINFORMATICS/BTQ054.
[48] W. Skala, T. Wohlschlager, S. Senn, G.E. Huber, C.G. Huber, MoFi: a software tool for annotating glycoprotein mass spectra by integrating hybrid data from the intact protein and glycopeptide level, Anal. Chem. 90 (2018) 5728–5736, https://doi.org/10.1021/acs.analchem.8b00019.
[49] J.L. Abrahams, M.P. Campbell, N.H. Packer, Building a PGC-LC-MS N-glycan retention library and elution mapping resource, Glycoconj. J. 35 (2018) 15–29, https://doi.org/10.1007/s10719-017-9793-4.
[50] P.D. Compton, N.L. Kelleher, J. Gunawardena, Estimating the distribution of protein post-translational modification states by mass spectrometry, J. Proteome Res. 17 (2018) 2727–2734, https://doi.org/10.1021/acs.jproteome.8b00150.