scientific reports

Check for updates

OPEN

Comparison of pregnant mare serum gonadotropin products with surprising differences in protein content

Martina Lösle[1], Chia-Wei Lin[2,3], Jane Beil-Wagner[1], Markus Aebi[3] & Thorsten Buch[1]✉

Equine *c*horionic *g*onadotropin (eCG) is a widely used hormone that synchronizes the female cycle and induces estrus in livestock. eCG is a heterodimeric glycoprotein composed of non-covalently linked α- and β-chains whose glycosylation profiles determine the *in vivo* activity of the hormone. The commercially available eCG products are crudely purified from the serum of pregnant mares, hence called *p*regnant *m*are *s*erum *g*onadotropin (PMSG). Appropriate glycosylation of the protein is crucial for the correct binding to the receptor, receptor activation, and its half-life. The exact protein composition of the various commercial PMSG products and their specific glycosylation pattern have not been characterized so far. Therefore, we used proteomic approaches to analyse and compare four commercial PMSG products. Here we show that the examined PMSGs share a surprisingly low level of commonalities (5.5%) in protein composition among the four tested products, with serum proteins as the primary variable. Analysing the site-specific N-glycosylation, we confirmed the presence of O-acetylation of sialic acids at the structure of the glycans of eCG, which we could not find in significant amounts on human CG, suggesting that this modification is species-specific. It remains to be tested whether the O-acetylation plays an important role in the function of PMSG. However, this modification shall be considered while recombinant eCG are produced.

Keywords Equine chorionic gonadotropin (eCG), Lutropin/Luteinizing hormone (LH), O-acetylation, Pregnant mare serum gonadotropin (PMSG), Sulfation

Equine chorionic gonadotropin (eCG) is a hormone frequently used in livestock production and rodent transgenic units. Its applications encompass the synchronization and induction of ovulation of pigs, cattle, and goats and the superovulation of mice. It is currently isolated from blood exclusively from live pregnant horses and is hence also known as *p*regnant *m*are *s*erum *g*onadotropin (PMSG; we use the term PMSG when referring to respective commercial products derived from animals/*ex vivo*). Chorionic girdle cells, a cell lineage of the trophoblast, begin with eCG production ~ 30d during mare pregnancy. Later, the cells migrate into the endometrium to form endometrial cups[1], roughly around day 35 and produce up until ~ 120d of gestation. The hormone eCG has minimal FSH activity in horses; instead, it shows primarily luteinizing hormone (LH) activity[2], facilitating hormonal support for early gestation[3], possibly through induction of supportive ovulation[4,5]. However, in other mammals, it shows dual LH and follicle-stimulating hormone (FSH) activity, making it an important, frequently used hormone.

As a member of the gonadotropin hormone family, eCG is a heterodimeric protein consisting of glycoprotein hormones alpha (α) chain and lutropin/choriogonadotropin beta (β) subunits. It undergoes intricate glycosylation modifications that are crucial for its biological function.

Gonadotropins are evolutionarily conserved regarding structure and glycosylation sites[1]. Yet, chorionic gonadotropins from humans/primates and equids differ from other gonadotropins by expressing a β-chain that is approximately thirty amino acids longer (the carboxy-terminal peptide), carrying additional O-glycosylation sites, that make it 11% more glycosylated than the respective hLHβ. Another peculiarity of the equids is that eCG and eLH use the same peptide for the β-chain. The main difference between eCG and eLH is in the glycosylation itself, where the eCG shows sialic acid and the eLH shows to be sulfated. O-linked sulfated glycans in LH are recognized in the liver and therefore lead to higher and faster clearance rates compared to eCG [6–8].

[1]Institute of Laboratory Animal Science, University of Zurich, Schlieren 8952, Switzerland. [2]Functional Genomics Center Zürich, ETH Zürich, Zürich 8057, Switzerland. [3]Institute of Microbiology, Department of Biology, ETH Zürich, Zürich 8093, Switzerland. ✉email: thorsten.buch@uzh.ch

Glycosylation constitutes up to 45% of the total molecular weight of eCG. In detail, it contains di- and triantenna Asn-linked oligosaccharides terminating with Siaα2,3 or 6Galβ1,4GlcNAc. In contrast, the amino acid sequence-identical eLH presents with >72% of the Asn-linked oligosaccharides with 1 or 2 branches terminating with the sequence $S0_4$-4-GalNAcβ,4GlcNAc[9]. Most likely, this difference is also responsible for the reported long half-life and efficacy of eCG[1,10,11].

Commercially available PMSG products are obtained from the serum of pregnant mares between days 40 and 120 of gestation[12,13]. The hormone is purified and enriched by a simple two-step standard purification scheme. The main contaminants are removed by precipitation using pH fractionation with metaphosphoric acid and alcohol and gel-filtration via Sephadex. It appears as if these two purification steps can be followed by a suitable form of chromatography[14]. The assessed PMSG products were specifically selected as they represent standard products in veterinary medicine and come from different parts of the world (telephone statements prior to order).

There have been numerous discussions regarding PMSG isoforms and their potential impact on glycan patterns. As early as 1941, Cole and Erway[15]speculated on the role of isoforms and de-sialylation. A study conducted in 2008 explored whether significant differences in isoform composition exist among various commercial eCG preparations. This research also compared these findings with the isoform composition found in plasma samples. The study concluded that most immunoactivity (92%) in commercial preparations was observed in the acidic fractions after liquid-phase isoelectric focusing. This finding contrasted sharply with isoform profiles in pregnant mare plasma samples, which displayed a broader range across pH levels. The conclusion was that the isolation process in commercial PMSG production tends to favor the acidic isoforms of PMSG[16]. Furthermore, Cole and Erway noted in their description from 1941 that the removal of roughly 80% of the sialic acid from PMSG resulted in the removal of the hormone from the blood in 1 h. Since then, several papers highlight the importance of the glycosylation (structure and position) for the stability of the protein[7,17].

Because the process for the production of different PMSG products is not standardized, in this study, we aimed to understand the differences among commercial products at the protein level, assessed by mass spectrometry (MS)-based proteomics. Given that we analysed commercial PMSG products, we assume that the preparation process would similarly affect all the samples tested.

In addition, while the glycosylation pattern of purified eCG, constituting up to 45% of its total molecular weight, has been characterized in the past[18,19], a comparison of the N-glycans on PMSG from different vendors has, so far, not been reported.

Results

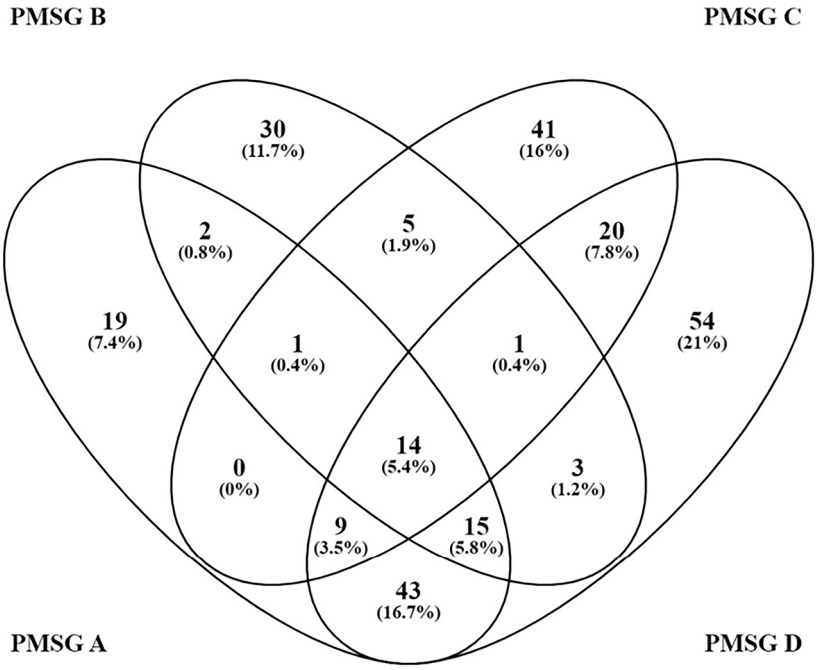
Proteome of PMSG products from different sources

Commercially available PMSG products are known to be only crudely purified from the serum of pregnant mares[20,21]. Nevertheless, the exact protein composition of this frequently used drug has not been assessed in a comparative study of different PMSG products from different vendors and producers. Also, whether glycosylation differs between products has never been addressed before. Therefore, we first aimed to understand the protein composition of four commercial PMSG products, sold by Apollo Scientific, Aviva Systems Biology Corporation, MSD Animal Health, and Genway Biotech INC. (in the following referred to as PMSG A, B, C, and D, respectively). A LC-MS/MS-based proteomic approach was applied to four PMSG products and we confirmed the presence of α- and β-chains of luteinizing hormone/choriogonadotropin (called henceforth α and β-chain) in all samples (Table 1). In addition to the actual pharmaceutically relevant substance, we observed, however, many further serum proteins in the samples (Supplementary Table SI). Surprisingly, 94.5% of proteins were unique to each product and only 5.5% (14 proteins) were common to all samples (Fig. 1). These common serum proteins included e.g. alpha1-microglobulin, maltase-glucoamylase, apolipoprotein, Serpin family (alpha-antitrypsin), transcobalamin, and transthyretin. Three proteins were uncharacterized during the time of manuscript writing. Together, the proteomic data demonstrated that the complexity of protein composition in each product and the discrepancies between the products were astonishingly high.

		lutropin/gonadotropin subunit beta (β-chain)		glycoprotein hormones alpha chain (α-chain)	
		sp|P08751|LSHB_Horse		sp|P01220|GLHA_Horse	
Company	Observed proteins[1]	Best Score (ranking)	Observed peptides	Score/ ranking	Observed peptides
A	103[2]	187.9 (95/103)	1[3]	91.9 (93/103)	1[3]
B	70[2]	556.7 (41/70)	1[3]	103.5 (30/40)[3,4]	1[3,4]
C	91	798.1 (6/91)	3	388.1 (84/91)	2
D	159	732.6 (67/159)	3	494.4 (118/159)	3

Table 1. Summary of the proteomic results of equine lutropin/gonadotropin subunit β and glycoprotein hormones α-chain in each sample. [1]The number of observed proteins. The proteomic data were filtered based on following criteria: 1% false discovery rate FDR and there were at least two unique peptides assigned to the same protein. [2]In this case, only one peptide from α- and β-chains was observed. The proteomic data were filtered based on following criteria: 1% false discovery rate FDR and one peptide assigned to either alpha or β-chains. [3]Only one peptide was observed in this sample. [4]These peptides were only observed in GluC-endopeptidase digested sample.

A



B

Accession number	Gene name	Protein full name
tr|F6TVT0|F6TVT0_HORSE	ABI3BP	ABI family member 3 binding protein
tr|F6UZH0|F6UZH0_HORSE	AMBP	Alpha-1-microglobulin/bikunin precursor
tr|F6XM13|F6XM13_HORSE	APOD	Apolipoprotein D
sp|P01220|GLHA_HORSE	CGA	Glycoprotein hormones alpha chain
sp|P08751|LSHB_HORSE	LHB	Lutropin/choriogonadotropin subunit beta
tr|F6PUW3|F6PUW3_HORSE	MGAM	Maltase-glucoamylase
tr|F6ZLR1|F6ZLR1_HORSE	SERPINA3	Serpin family A member 3
tr|F7CYR1|F7CYR1_HORSE	SERPINC1	Serpin family C member 1
tr|F7CZW9|F7CZW9_HORSE	SERPING1	Serpin family G member 1
tr|F6T7K7|F6T7K7_HORSE	TCN1	Transcobalamin 1
tr|F6UL68|F6UL68_HORSE	TTR	Transthyretin
tr|F7CSL8|F7CSL8_HORSE	SPI2	Alpha-1-antitrypsin
tr|F6XBQ7|F6XBQ7_HORSE	CD55	CD55 molecule (Cromer blood group)
tr|H9GZU9|H9GZU9_HORSE		Uncharacterized protein

Fig. 1. Comparison of the protein composition of commercially available PMSG products (A). A-D represent the proteomes of commercial PMSG products from Apollo, Avivasysbio, MSD animal health and Genway. The numbers in the Venn diagram indicate uniquely observed proteins from each sample; (B) List of common proteins observed in all four products. A complete list of identified proteins is shown in supplementary Table SI.

Site-specific N-glycosylation analysis of PMSG and luteinizing hormone (LH) from horse

Two N-glycosylation sites on the α-chain and one on the β chain of eCG have been described in the past[7]. To understand the glycosylation profile of different PMSG products, we first performed nanoUPLC-MS/MS analysis for digested (glyco)peptides of PMSG D. We found complex bi- and triantenna N-glycan structures on peptides of both α and β-chains. The results of tryptic peptides containing N80 from the alpha chain (Swissprot accession#: P01220) are shown in Fig. 2. The precursor ions of three MS/MS spectra had 42 Da difference (Fig. 2B to D), indicating the presence of acetylation on these glycopeptides. The acetylation could either occur on the side chain of the peptide or the glycans. Apart from the commonly known oxonium ions, $HexNAc^+$, $HexNAcHex^+$, $NeuAc\text{-}H_2O^+$, $NeuAc^+$ and $HexNAcHexNeuAc^+$ at m/z 204.08, 366.12, 274.09, 292.10 and 657.23, the presence of 334.11 and 699.24, corresponding to $acetyl\text{-}NeuAc^+$ and $HexNAcHexNeuAc^+$, indicated that the acetylation occurred on sialic acids (Fig. 2C and D). In addition, the Y1 fragment at m/z 1572.69 with HexNAc was identical among all glycoforms, which was an additional confirmation of an acetyl group on the sialic acid, not on the peptide backbone. The extracted *i*on *c*hromatography (XIC) of each glycoform showed that the main glycoform at N80 was sialylated biantenna N-glycan and was further modified with 1–2 acetyl groups as a minor fraction (Fig. 2A). The same pattern was observed on N106 with core fucosylation (Supplementary Figure S1).

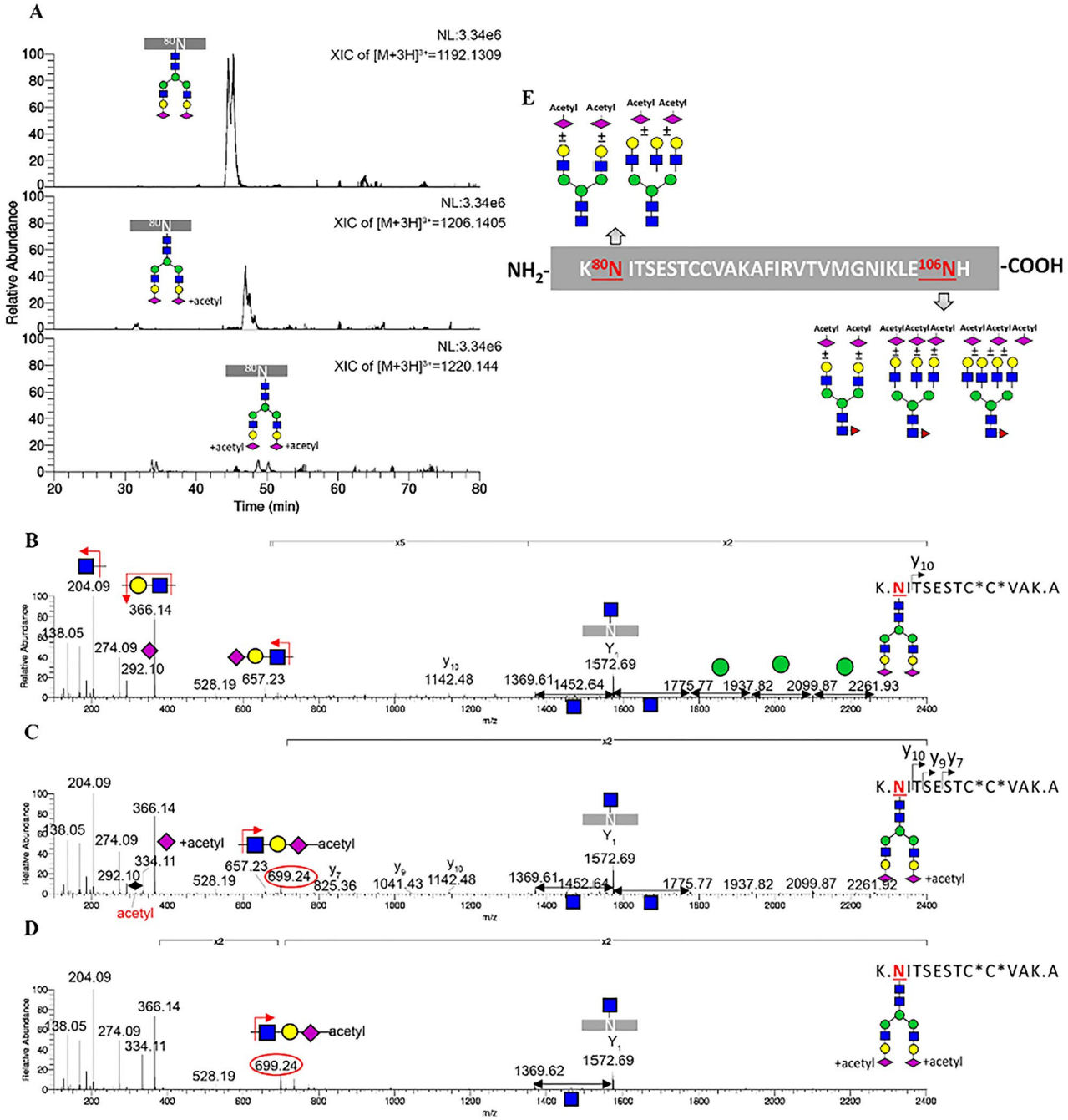


Fig. 2. LC-MS/MS results show acetylated sialic acids as a modification of the α-chain of PMSG. (A) The extracted ion chromatography (XIC) of glycopeptides, K.^{80}NITSESTC*C*VAK, with biantenna with two sialic acids (top) with one O-acetyl group (middle) and two O-acetyl group (bottom). The m/z for each XIC was shown in the figure with the mass tolerance of 10 ppm. NL indicated normalized intensity. The MS/MS spectra of three glycoforms were shown in (B) to (D). The nomenclature of peptide fragment ions and glycan fragmentation ions was described previously[43–45]. The N marked in red and bold was N-glycosylation site. The underlined C represented carbamidomethylated cysteines. (E) Positions and structures of carbohydrate chains of the α-chain.

A summary of glycoforms on the α-chain is shown in Fig. 2E. For the β-chain (Swissprot accession#: P08751), we also observed biantenna N-glycans structures with sialic acid (Supplementary Figure S2). The O-acetylation was also observed on the β-chain but only based on the mass matching (the quality of MS/MS spectra was below our quality criteria). The commonly observed glycopeptides from the four PMSG products are summarized in Table 2 (The observed glycopeptides in each product are listed in Supplementary Table SII a and SII b). In this analysis, we did not find significant differences between the glycosylation profiles of both chains among all four products (Table 2).

Site	Observed m/z[2]	[M + H]+	Sequences[3]	Glycan structures
α-chain				
N80	1192.131^{3+}	3574.393	K.**N**ITSESTC*C*VAK	Core(HexNAcHex)2NeuAc2
	1206.141^{3+}	3616.423		Core(HexNAcHex)2NeuAc2 + acetyl
	1220.144^{3+}	3658.432		Core(HexNAcHex)2NeuAc2 + 2acetyl
	1313.844^{3+}	3939.532		Core(HexNAcHex)3NeuAc2
	1327.850^{3+}	3981.550		Core(HexNAcHex)3NeuAc2 + acetyl
	1341.849^{3+}	4023.547		Core(HexNAcHex)3NeuAc2 + 2acetyl
	1410.871^{3+}	4230.613		Core(HexNAcHex)3NeuAc3
	1424.872^{3+}	4272.616		Core(HexNAcHex)3NeuAc3 + acetyl
	1438.874^{3+}	4314.622		Core(HexNAcHex)3NeuAc3 + 2acetyl
N106	1122.930^{4+}	4488.720	K.LE**N**HTQCYC*STC*YHHK	CoreFuc(HexNAcHex)2NeuAc2
	1133.435^{4+}	4530.740		CoreFuc(HexNAcHex)2NeuAc2 + acetyl
	915.348^{5+}	4572.741		CoreFuc(HexNAcHex)2NeuAc2 + 2acetyl
	1214.211^{4+}	4853.844		CoreFuc(HexNAcHex)3NeuAc2
	1224.718^{4+}	4895.872		CoreFuc(HexNAcHex)3NeuAc2 + acetyl
	1286.991^{4+}	5144.964		CoreFuc(HexNAcHex)3NeuAc3
β-chain				
N33	1126.812^{3+}	3574.393	R.PI**N**ATLAAEK.E	Core(HexNAcHex)2NeuAc2
	1140.815^{3+}	3420.446[2]		Core(HexNAcHex)2NeuAc2 + acetyl
	1154.818^{3+}	3462.456[2]		Core(HexNAcHex)2NeuAc2 + 2acetyl

Table 2. The summary of commonly observed glycopeptides from α- and β -chain from four PMSG products[1]. [1]The results were the summary of data from both trypsin and GluC endopeptidase digestion. The identified glycopeptides in each sample were provided as supplementary Tables. [2]The corresponding assignment was based on observed m/z from MS1 spectra although the quality of MS/MS spectra could not confirm the assignment. [3]C* represents carbamidomethylated cysteines.

Taken together, while the presence of O-acetylated sialic acid for the β-chain of PMSG/eCG has been previously shown [16], we now confirm this observation and extend it by describing the presence of this modification also for the α-chain. This O-acetylation is independent of the PMSG product.

Since the acetylation was observed on all PMSG products, we wanted to assess whether this could also be found on luteinizing hormone (LH) of the horse, which has the identical peptide sequence to eCG. We hence examined the glycosylation profile of luteinizing hormone (LH) from a commercial source (Sigma-Aldrich). Unlike PMSG, the glycosylation of the α-chain from LH showed the unique LacdiNAc with sulfation, but not acetylated sialic acid. The precursor ions of three MS/MS spectra had a 79.996 Da difference (Fig. 3B to D), corresponding to a sulfate or phosphate group. In addition to the commonly known oxonium ions, the presence of 407.16 and 569.22 corresponded to HexNAc2$^+$ and HexNAc2Hex$^+$. Although no diagnostic ion of sulfated LacdiNAc was observed, sulfation has been reported in a previous study[9]. Therefore, the structures were assigned as biantenna with LacdiNAc and sulfate(s). Based on the intensity of XIC, the main glycoforms of the α-chain were the complex N-glycan with the LacdiNAc + 1–2 sulfate group(s) (Fig. 3A). In addition to LacdiNAc + sulfation, the normal complex type N-glycans with sialylation were also observed (Table 3 and Supplementary Figure S3 and Table SIII). These results were consistent with previous studies[9]. Observed positions and structures of the carbohydrate chains within the α-chain are shown in Fig. 3E .

Comparison of equine and human CG N-glycosylation profiles

As we had identified a type of O-acetylation on both α- and β-chain of equine PMSG, we wanted to examine whether this O-acetylation might be a general characteristic of CGs. Therefore, we performed an identical MS analysis for human chorionic gonadotropin (CG). Here, the spectral counts of two sets of diagnostic ions from LC runs were used to demonstrate the expression of O-acetylated sialic acids in human CG and PMSG (Fig. 4A). The MS/MS spectra, containing all four diagnostic ions 204, 274, 292 and 366, were considered N-glycopeptides with all types of sialic acids. The MS/MS spectra containing all five diagnostic ions 204, 274, 292, 334, and 366, were considered N-glycopeptides with purely O-acetylated sialic acid. In PMSG products A and D, the spectral counts of acetyl-NeuAc/NeuAc oxonium ion were 8767/9406 and 4918/5761, respectively. This result indicated that around 85–90% of sialylated N-glycans were acetylated. The same pattern was observed in the products from company B and C (Supplementary Figure S5). However, in human CG, the fraction of acetylated N-glycans was found to be lower than 0.03%. Hence, our data shows that O-acetylated NeuAc was ubiquitously present in equine CG products from different vendors but not in human CG.

For site-specific analysis, the MS profile and the MS / MS spectrum showed that the site N31 from human choriogonadotropin beta subunit (Swissprot accession#: Q6NT52) was glycosylated by complex structures, including fucosylation and sialylation (Fig. 4B and C, Supplementary Table S IVa and Table S IVb). For the α-chain (Swissprot accession#: P01215), glycopeptides were observed when the sample was digested by GluC endopeptidase. As on the β-chain, the major N-glycan structures on the α-chain were bi-antennary N-glycans

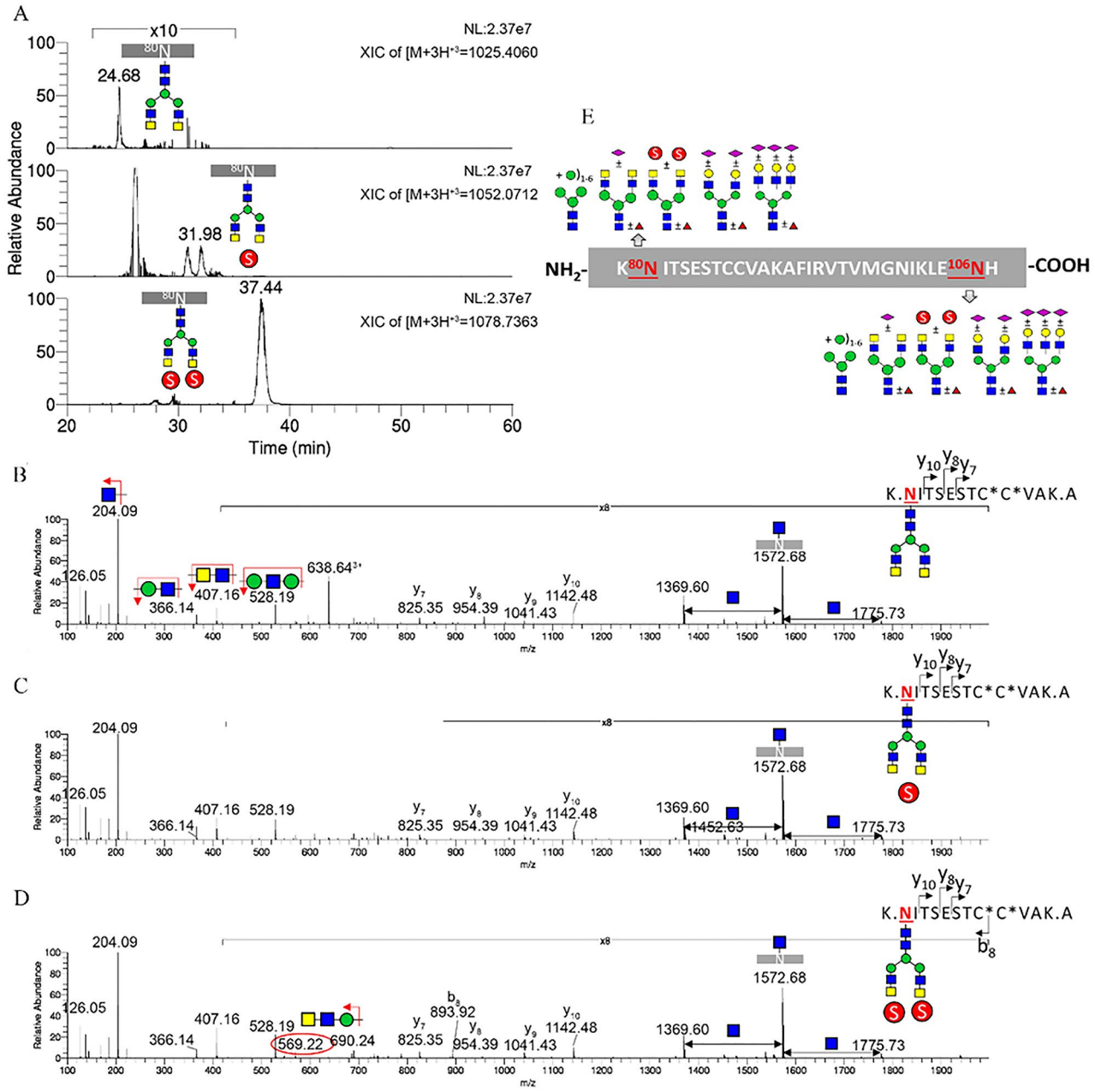


Fig. 3. LC-MS/MS analysis shows that the α-chain of luteinizing hormone is modified by sulfation. (A) The extracted ion chromatography (XIC) of glycopeptides, K.^{80}NITSESTC*C*VAK, with biantenna with LacdiNAc (top) with one O-sulfate group (middle) and two O-sulfate groups (bottom). C*: indicated that cysteines were carbamidomethylated. The m/z for each XIC was shown in the figure with the mass tolerance of 10 ppm. The MS/MS spectra of three glycoforms were shown in (B) to (D). (E) Positions and structures of carbohydrate chains of α-chain.

with core-fucosylation and sialylation and a small amount of Man 5 structures (Supplementary Figure S4). The N-glycan structures observed in this study are summarized in Fig. 5, which also highlights the species specificity of the O-acetylated sialic acid.

Discussion

Commercially distributed eCG is made available as a crude extract from the serum of pregnant mares, commonly called PMSG. Previously there have been attempts to characterize and describe PMSG via HPLC/RP-HPLC and SDS-PAGE to analysis. In fact, Alvarez and colleagues described, using HPLC and four different commercial PMSG, that this method can be used for predicting biological activity[22].

In addition, SDS-PAGE after HPLC was shown by Rodriguez and colleagues, which allowed them to characterize and compare reCG with natural PMSG products and conclude that recombinant molecules and

Site	Observed m/z[1]	$[M+H]^+$	Sequences[2]	Glycan structures
α-chain				
N80	803.332^{3+}	2407.995	K.**N**ITSESTC*C*VAK	CoreHexNAc
	871.025^{3+}	2611.076	K.**N**ITSESTC*C*VAK	CoreHexNAc2
	925.043^{3+}	2773.129	K.**N**ITSESTC*C*VAK	CoreHexNAc2Hex
	938.718^{3+}	2814.153	K.**N**ITSESTC*C*VAK	CoreHexNAc3
	965.371^{3+}	2894.117	K.**N**ITSESTC*C*VAK	CoreHexNAc3 + sulfate
	992.738^{3+}	2976.214	K.**N**ITSESTC*C*VAK	CoreHexNAc3Hex
	1022.074^{3+}	3064.228	K.**N**ITSESTC*C*VAK	CoreHexNAc2HexNeuAc
	1035.747^{3+}	3105.242	K.**N**ITSESTC*C*VAK	CoreHexNAc3NeuAc
	1143.786^{3+}	3429.358	K.**N**ITSESTC*C*VAK	Core(HexNAcHex)2HexNAcNeuAc1
	1203.134^{3+}	3607.401	K.**N**ITSESTC*C*VAK	Core(HexNAcHex)2NeuAc2
	1240.815^{3+}	3720.445	K.**N**ITSESTC*C*VAK	Core(HexNAcHex)2HexNAcNeuAc2
N106	790.346^{5+}	3947.702	K.AFIRVTVMGNIKLE**N**HTQC*YC*STC*.Y[3]	Man5
	822.757^{5+}	4109.758	K.AFIRVTVMGNIKLE**N**HTQC*YC*STC*.Y	Man6
	835.5306+	5008.142	K.AFIRVTVMGNIKLENHTQC*YC*STC*YHHKI.-	Man7
	862.5376+	5170.186	K.AFIRVTVMGNIKLENHTQC*YC*STC*YHHKI.-	Man8
	754.7368^{5+}	3769.655	K.AFIRVTVMGNIKLE**N**HTQC*YC*STC*.Y	CoreFuc
	795.352^{5+}	5144.964	K.AFIRVTVMGNIKLE**N**HTQC*YC*STC*.Y	CoreHexNacFuc
	827.764^{5+}	4134.789	K.AFIRVTVMGNIKLE**N**HTQC*YC*STC*.Y	CoreHexNacHexFuc
	835.968^{5+}	4175.809	K.AFIRVTVMGNIKLE**N**HTQC*YC*STC*.Y	CoreHexNac2Fuc
	868.380^{5+}	4337.871	K.AFIRVTVMGNIKLE**N**HTQC*YC*STC*.Y	CoreHexNac2HexFuc
	750.826^{6+}	4499.916	K.AFIRVTVMGNIKLE**N**HTQC*YC*STC*.Y	Core(HexNacHex)2Fuc
	799.343^{6+}	4791.017	K.AFIRVTVMGNIKLE**N**HTQC*YC*STC*.Y	Core(HexNacHex)2FucNeuAc
	1211.975^{4+}	4844.879	R.VTVMGNIKLE**N**HTQC*YC*STC*YHHKI.-	CoreHexNAc4 + sulfate
β-chain				
N33	688.1373^{5+}	3436.657	A.SRGPLRPLC*RPI**N**ATLAAEK.E[4]	Man5
	900.4333^{4+}	3598.712	A.SRGPLRPLC*RPI**N**ATLAAEK.E	Man6
	752.9591^{5+}	3760.767	A.SRGPLRPLC*RPI**N**ATLAAEK.E	Man7
	785.370^{5+}	3922.820	A.SRGPLRPLC*RPI**N**ATLAAEK.E	Man8
	1021.973^{4+}	4084.869	A.SRGPLRPLC*RPI**N**ATLAAEK.E	Man9
	778.894^{4+}	3112.555	A.SRGPLRPLC*RPI**N**ATLAAEK.E	Core
	652.528^{5+}	3258.612	A.SRGPLRPLC*RPI**N**ATLAAEK.E	CoreFuc
	693.144^{5+}	3461.692	A.SRGPLRPLC*RPI**N**ATLAAEK.E	CoreHexNacFuc
	725.555^{5+}	3623.743	A.SRGPLRPLC*RPI**N**ATLAAEK.E	CoreHexNacHexFuc
	783.773^{5+}	3914.835	A.SRGPLRPLC*RPI**N**ATLAAEK.E	CoreHexNacHexFucNeuAc
	733.760^{5+}	3664.773	A.SRGPLRPLC*RPI**N**ATLAAEK.E	CoreHexNac2Fuc
	766.170^{5+}	3826.823	A.SRGPLRPLC*RPI**N**ATLAAEK.E	CoreHexNac2HexFuc
	798.583^{5+}	3988.884	A.SRGPLRPLC*RPI**N**ATLAAEK.E	Core(HexNacHex)2Fuc
	856.796^{5+}	4279.980	A.SRGPLRPLC*RPINATLAAEK.E	Core(HexNacHex)2FucNeuAc
	915.0151^{5+}	4571.076	A.SRGPLRPLC*RPINATLAAEK.E	Core(HexNacHex)2FucNeuAc2
	971.192^{4+}	3881.746	A.SRGPLRPLC*RPI**N**ATLAAEK.E	CoreHexNac3 + 2Sulfate
	774.376^{5+}	3867.851	A.SRGPLRPLC*RPI**N**ATLAAEK.E	CoreHexNac3Fuc
	987.706^{4+}	3947.803	A.SRGPLRPLC*RPI**N**ATLAAEK.E	CoreHexNac3Fuc-sulfate
	832.592^{5+}	4158.933	A.SRGPLRPLC*RPI**N**ATLAAEK.E	CoreHexNac3FucNeuAc
	806.787^{5+}	4029.905	A.SRGPLRPLC*RPI**N**ATLAAEK.E	CoreHexNac3HexFuc
	839.197^{5+}	4191.955	A.SRGPLRPLC*RPI**N**ATLAAEK.E	CoreHexNac3Hex2Fuc
	897.415^{5+}	4483.047	A.SRGPLRPLC*RPI**N**ATLAAEK.E	CoreHexNac3Hex2FucNeuAc
	955.633^{5+}	4774.134	A.SRGPLRPLC*RPI**N**ATLAAEK.E	CoreHexNac3Hex2FucNeuAc2
	847.403^{5+}	4232.984	A.SRGPLRPLC*RPI**N**ATLAAEK.E	CoreHexNac4HexFuc

Table 3. The summary of observed glycopeptides from α-chain from luteinizing hormone (LH). [1]Observed m/z from MS1 spectrum. [2]C* represents carbamidomethylated cysteines. [3]The MS/MS spectra matched the peptide with non-specific cleavage, meaning that C-terminal of α-chain was truncated. [4]The MS/MS spectra matched the peptide starting from Serine due to the signal peptide.

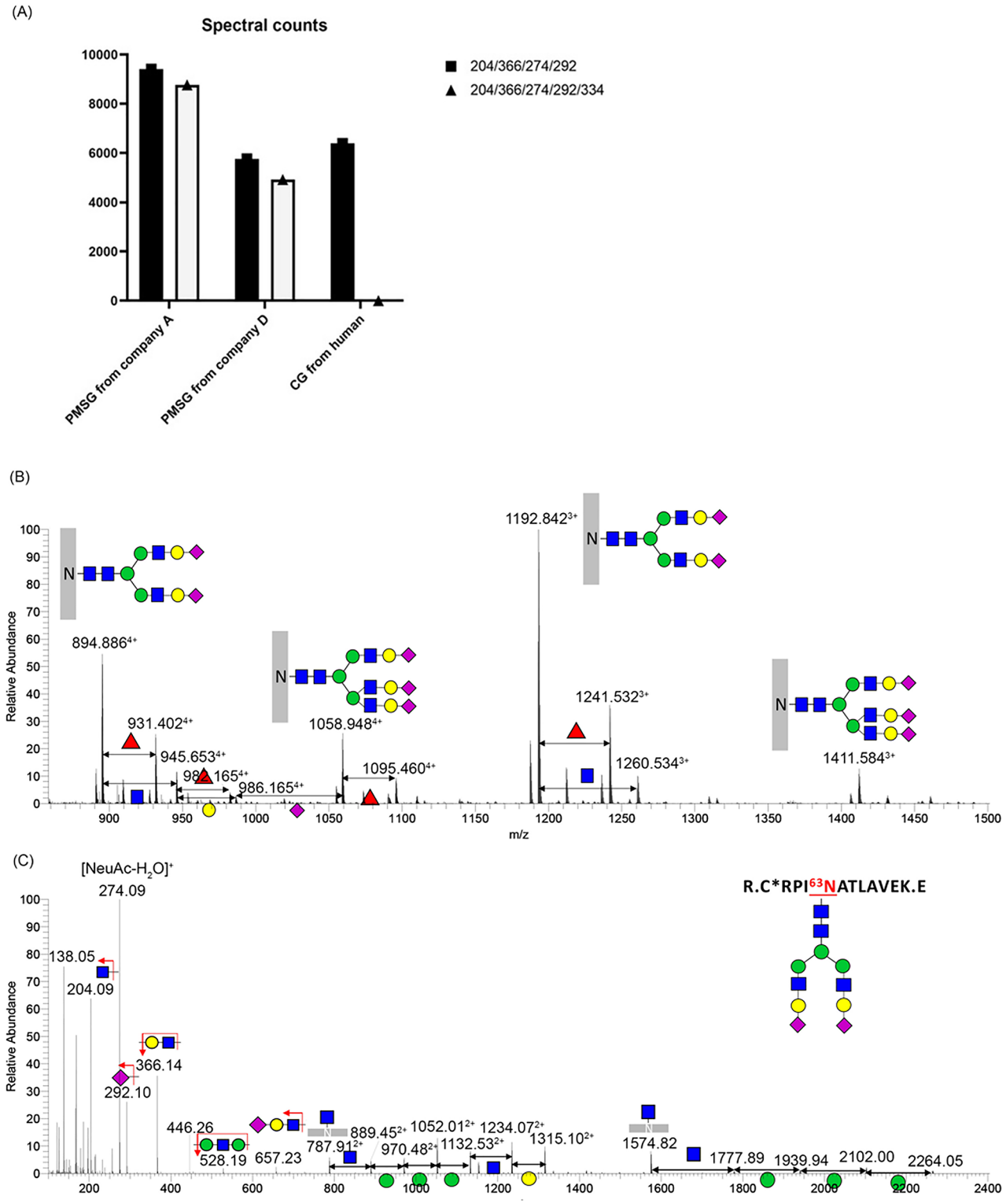


Fig. 4. Terminal epitopes among different species. (A) The spectral counts of MS/MS scans containing N-glycan with sialic acid were shown in black and N-glycan with acetylated sialic acid were marked in white. The oxonium ions that were utilized to filter their represent structures were $[HexNac]^+$= 204.087; $[NeuAc-H_2O]^+$= 274.092; $[NeuAc]^+$= 292.103; $[acetylNeuAc]^+$= 334.113 $[HexNac + Hex]^+$= 366.139. The required ions are shown in the figure legend and the mass tolerance was 10 ppm. CG stands for choriogonadotropin from Sigma. The data was first processed by Byonic with MS/MS filtering function and the spectra counts was counted by inhouse script. (B) MS profile demonstrating the glycoforms on site N31 from human choriogonadotropin subunit beta variant 2 (Swissprot accession#: Q6NT52). (C) The MS/MS spectrum of m/z 894.886 (+4) showing that this tryptic peptide was carried with biantenna structures with two sialic acids. The products A and B are shown exemplary. Products C and D do not differ and are shown in supplementary Figure S5.

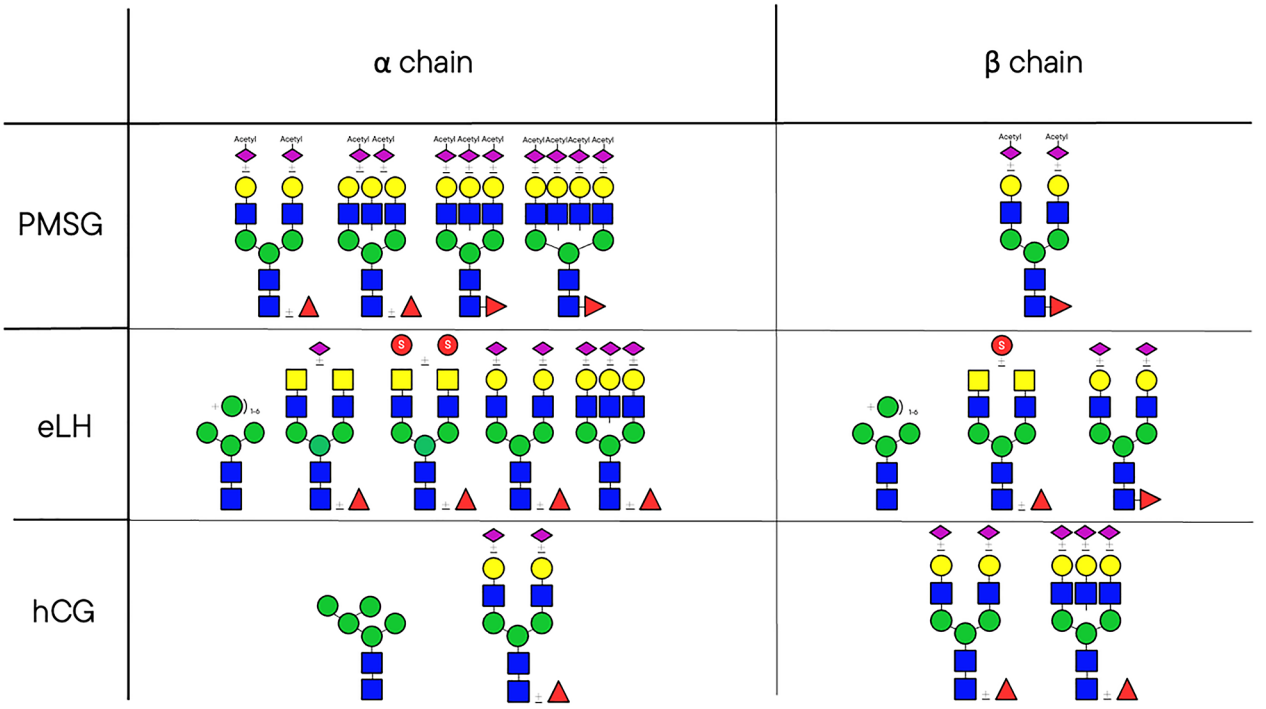


Fig. 5. Glycosylation profile of α- and β-chains from PMSG, eLH and human CG. In PMSG, O-aceylation on sialic acids was observed on both α- and β-chain. `S`-Sulfated N-glycans with polyLacNac were exclusively observed in horse luteinizing hormone. For human CG, bi- and tri-antenna N-glycans were observed in both α- and β-chain. All assignments were based on the biosynthesis pathway and previous studies.

PMSG displayed variations in their glycosylation patterns. Notably, differences in the sialic acid content were observed between two commercial PMSG preparations, potentially affecting their biological potency. These findings emphasise the significance of implementing a standardized production process[23]. Both works highlight the importance of the physicochemical profile and the glycosylation of the products. As the exact compositions of the natural PMSG extracts were unknown, we first performed a proteomic analysis to understand the protein composition of various commercial PMSG products. We supplemented this analysis with a thorough assessment of the N-glycosylation of these products.

In this comparison of four different PMSG products, we show that even though we identified a multitude of proteins in the extracts only 14 proteins were found across all samples, among these the α- and β-chains of eCG/eLH. Most of the others are known serum proteins, such as e.g. alpha-1-microglobulin, maltase-glucoamylase, apolipoprotein, Serpin family (alpha-antitrypsin), transcobalamin, and transthyretin[24]. Thus, even though all PMSG products supposedly contain the same bioactive hormone by description, the proteomic results proved evidence for significant differences among these products. Furthermore, and most astonishingly, the product from company B is the same as the product from company D, just sold under a different brand name (personal communication). Yet we identified significant differences among them, possibly as a result of different production lots.

It appears entirely likely that the background proteome co-determines the efficacy of each PMSG product – possibly by specific influences on pharmacokinetics. At the same time, changes in the pharmacodynamics of the hormone due to binding to other proteins in the preparations cannot be completely ruled out either. The background proteome of PMSG preparations is controlled by different factors of three basic types, namely those intrinsic to the animal, those extrinsic to the animal, and those in the *ex vivo* production phase. Intrinsic factors are horse species, age, exercise, state of gravidity at collection, and parity. Extrinsic factors are food, stress (including collection method), and pollutants. The ex vivo factors include sample preparation and storage, purification, formulation, and transport[12,25]. In previous studies, proteomic data also showed that there are differences in the composition of the equine oviduct fluid obtained from the oviducts ipsilateral to a pre-ovulatory follicle and that obtained after ovulation, coinciding with previous results in bovine[26,27]. It shows that equine fertilization is a complicated process, and the protein expression is one of important factors that promote fertilization and early embryo development in horses. In conclusion, in terms of protein constitution, the four assessed PMSG preparations differed significantly, with unknown consequences for pharmacological properties such as efficacy and side effects in the various species in which the drug is used. Further investigations into the differences between PMSG products and their pharmacological consequences appear necessary.

In addition to the proteomic analysis, we also assessed in detail the N-glycosylation profile of PMSG/eCG, which has until today not been compared between products, even though it is known for being essential for the

pharmacodynamic and pharmacokinetic properties of eCG[28,29]. We furthermore compared in a standardized manner the glycosylation profile of eCG with that of eLH and human CG.

Previous studies examining the LH-like activity of glycoprotein hormones have revealed that the Asn56 and Asn82 glycosylation sites of the eCG α-chain, counting without signal peptides (In this study, the labeling was with signal peptides on both α and β-chains.), play an essential role in the induction of a signal by its receptor. Similarly, the glycosylation sites Asn56 of the eFSH α-chain[30] and the hFSH α-chain[31] are critical for signal transduction. These findings indicated that the glycosylation sites Asn56 and Asn82 of the α-chain mediate the LH-like and FSH-like activities of eCG, eFSH, hCG[29], emphasizing the relevance of these glycosylation sites in the α-chain. eCG binds to FSH receptors of virtually all mammalian species, in which it has been tested and will produce biological effects specific to FSH. It has a similarly potent interaction with LH receptors. The structural basis of this duality is not known but is discussed to be related to the region 90–110 of the β-chain[32]. In our analysis, we found complex bi- and tri-antenna N-glycan structures on peptides of both α- and β-chains and in the position described in the literature. Taken together, we can confirm that all tested products contain eCG with the same, inconsistent with a previous study, complex N-glycans with core-fucosylation and sialylation were mainly observed in eCG, and the sulfated LacdiNAc N-glycans were observed in eLH[8].

Furthermore, we observed O-acetylation of sialic acids on eCG. Since we could not find a similar O-acetylation on hCG, our results indicate that this O-acetylation is, to some extent, species-specific. O-acetylation of sialic acid is regulated by two enzymes, sialic acid-specific O-acetyl transferases (SOATs), and O-acetyl esterases (SIAEs), which add and remove O-acetyl groups, respectively. This is important and is underlined in a recently published study, showing sialic acid types and their *O*-acetylation patterns have large differences among serum glycoproteins cross different species. The study of Liu 2024 shows that based on intact *N*-glycopeptide analyses, all sialoglycopeptides in human sera were modified by Neu5Ac without any *O*-acetylation; 90% of sialoglycopeptides in rat sera were also modified by Neu5Ac, with more than 60% that were further *O*-acetylated. In contrast, 99% of sialoglycopeptides in mouse sera contained Neu5Gc including 12% in *O*-acetylated forms. This study enhances the understanding of *O*-acetylated sialoglycan diversities and underscores the necessity of considering glycosylation profiles when choosing animal models[33]. Sialic acid-specific O-acetyl transferases have been identified and characterized in mammalian cells, invertebrates, bacteria, and viruses[34]. Hence the mammalian SOAT CAS1 domain containing 1 (CASD1) was also found in the horse genome and presents with high protein sequence identity to human CASD1, albeit containing an N-terminal extension resulting from a further 5´-located additional start codon in exon 1 (Supplementary Fig. S5). Similarly, an SIAE orthologue is located in the horse genome, albeit featuring a slightly lower protein sequence identity to the human protein (Supplementary Fig.S5). While it appears possible that this O-acetylation plays a role in facilitating the peculiar features of eCG, we cannot, at this point, provide experimental evidence thereof. Yet, the importance of sialic acid moieties for serum half-life of gonadotropic hormones has been known for a long time[35] Indicative is, however, the discrepancy in the efficacy of eCG produced in vitro compared to ex vivo PMSG[7]. It remains to be seen whether recently reported, functionally active in vitro produced eCG[36] sports this sialic acid attachment. Putting focus onto this O-acetylation may mean that it can now be tested in glycoengineering of glycoprotein hormones.

In a first approach we wanted to determine the biological function of the O-acetylation by treatment of PMSG with SIAE. Due to limited sources of SIAE and a lack of reliable SIAE during the study, we were unable to demonstrate the importance of O-acetylation of sialic acid. Instead, we tried two different sialidases (Neuraminidases from *Clostridium perfringens* and *Arthrobacter ureafaciens*), but cleavage of sialic acids was resisted as confirmed by MS. The reason might be acetylation at position −4, which cannot be enzymatically digested by sialidases[19]. Between positions C-7, C-8 and C-9 the acetyl-group can migrate and may allow cleavage when at particular positions, the C-4 position is static and resistant to most esterases[37] While we cannot prove this with current methods, the acetylation at this position might directly influence the half-life of the hormone in other mammals apart from horses. The missing esterases (SOATs) for hydrolisation[38], the hindrance of detection via immune-receptors[39] or the additional inhibition of esterases via methyl-phosphodiester at position C-4 and C-9[34] might influence this outcome.

In this study, we performed a standardized comparison of four PMSG products, of eLH and human CG via MS-based methods. In conclusion, our characterization of the commercially available products shows, that PMSG is an only crudely purified product. The four products shared < 6% of commonalities. Nevertheless, these products show consistently similar N-glycan structures including an acetylation of sialic acid in the α-chain and β chain as described before by Damm[18], that further distinguishes PMSG from eLH, which instead features a sulfation, and hCG. It is remarkable that identical glycosylation patterns are observed despite the highly variable protein content. The functional relevance of these difference in the N-glycan structures remain to be shown in the future. In an additional attempt we analysed commercial horse serum, which shows a great overlap in protein content, but no PMSG was detected (Supplemental table SVI) Our findings underscore the importance of standardizing products. The wide variability in co-injected proteins may contribute to allergies or adverse reactions and could influence biological activity or efficacy, complicating comparisons between different products or even batches. Notably, PMSG concentrations are not listed on product labels due to the difficulty in accurately determining the actual PMSG content. Whether these additional proteins impact receptor binding or half-life remains to be studied. These unresolved questions emphasize the need for detailed characterization of marketed products to better understand the mechanisms of the PMSG hormone.

Materials and methods

Materials

Mare serum gonadotropin was purchased from different sources, namely Apollo Scientific (Tamworth, UK Product.No 9002-70-4; Lot AS478456), Aviva Systems Biology Corporation (San Diego, USA Product.No

OPPA01037; Lot 917PGSMP07), MSD Animal Health (Luzern, Switzerland Product.No n.a; Lot A239A02) and Genway Biotech INC. (San Diego, USA Product.No GWB-2AE30A; Legacy DB ID 45962) companies. Horse Serum HI was acquired from GIBCO (Life Technologies, NZ). Luteinizing hormone from the equine pituitary, Human chorionic gonadotropin, dithiothreitol (DTT), and iodoacetamide (IAA) were purchased from Sigma-Aldrich (St. Louis, USA).

The samples were treated in the same way throughout the whole analysis, therefore, by all applicable scientific reproducibility measures we can assume that any incomplete digestion, loss of analytes during extraction, and degradation during preparation are the identical between samples. This is a standard approach in proteomics[40]. We cannot exclude the matrix effects among products since different vendors produced the products.

Sample preparation for site-specific glycan profiling analysis by mass spectrometry

All PMSGs used in this study are standardized and certified veterinary products. Before analysing these products via mass spectrometry, the concentration of proteins was measured via Nanodrop and normalized to the same amount. 50 µg of each PMSG sample was processed using the filter-assisted sample preparation protocol (FASP) with some modifications[41]. In brief, the samples were loaded into an Amicon device (Millipore Inc, USA) with 3 K MWCO membrane and washed three times with 50 mM ammonium bicarbonate buffer (ABC) at pH 8.5. Proteins were reduced with 25 mM (DTT) in ABC buffer for 1 h at 37 ° C and subsequently alkylated with 55 mM (IAA) for 1 h at 37 ° C in the dark. The sample was digested with 1 µg of sequencing-grade modified trypsin (Promega AG, USA) or GluC endopeptidase (Roche AG, Switzerland) and incubated for 16 h at 37 °C. Both samples were dried and dissolved in 2.5% acetonitrile with 0.1% formic acid. The samples were desalted with C18 Zip-tip (Millipore Inc, USA).

Proteomic and glycopeptide analysis by nanoUPLC-MS/MS mass spectrometry

All tryptic or GluC-digested (glyco)peptide mixtures were analyzed on a calibrated Q Exactive™ mass spectrometer (Thermo Fischer Scientific, Bremen, Germany) coupled to a Waters nanoAcquity UPLC system (Waters, Milford, USA) with a Picoview™ nanospray source 500 model (New Objective). The (glyco)peptide mixtures were dissolved in 2.5% acetonitrile (ACN)/0.1% FA (formic acid), loaded onto an Acquity UPLC M-Class Symmetry C18 trap column (180 µm × 20 mm, 100 Å, 5 µm particle size) and separated on an ACQUITY UPLC M-class HSS T3 C18 column (75 µm × 250 mm, 100 Å, 1.8 µm particle size), at a constant flow rate of 300 nL/min, with a column temperature of 50 °C and a linear gradient of 1 – 35% acetonitrile (ACN)/0.1% formic acid (FA) in 42 min, followed by a sharp increase to 98% acetonitrile in 2 min and then held isocratically for another 10 min. All samples were analyzed in Data-Dependent Acquisition (DDA) mode; one scan cycle comprised of a full scan MS survey spectrum, followed by up to 12 sequential HCD scans based on the intensity. The parameters above were applied to both proteomics and glycopeptide analysis. For the proteomics study, full scan MS spectra (400–2000 m/z) were acquired in the FT-Orbitrap at a resolution of 70,000 at 400 m/z, while HCD MS/MS spectra were recorded in the FT-Orbitrap at a resolution of 35,000 at 400 m/z. For glycopeptide analysis, full scan MS spectra (600–2000 m/z) were acquired in the FT-Orbitrap at a resolution of 70,000 at 400 m/z, while HCD MS/MS spectra were recorded in the FT-Orbitrap at a resolution of 35,000 at 400 m/z. HCD MS/MS spectra were performed with a autogain control value (AGC) of 500,000 by the collision energy setup at normalized collision energy (NCE) 25.

Data analysis for protein identification and glycopeptide profiling

For protein identification, all raw data were loaded into Byonic™ 3.11 (Protein Metrics, USA), and searched against Uniprot horse database (202101) with the following parameters: peptide tolerance of 10 ppm and an MS/MS tolerance of 0.03 Da for fragment ions with the consideration of carbamidomethylation at cysteine and oxidation at methionine (fixed modification). The proteomics data were available in the PRIDE server (PXD042535)[42] and filtered based on the following criteria: peptide FDR less than 1% and the minimum number of peptides per protein was 2. After general proteomic search, Byonic™ generated a focus database containing only the identified proteins. For site-specific N-glycosylation analysis, all raw data were then searched against the focused database with default 132 common biantenna N-glycan database by Byonic. The identified MS/MS spectra were inspected manually as described previously[41]. In addition, the raw data were then searched again with the consideration of N-glycans with acetyl groups. For the LH sample, the database was again searched with biantennary structures with LacdiNAc with sulfate groups. For the human sample, the database was searched against human database with biantennary structures and 132 common human glycan structures from Byonic.

Data availability

All data generated or analysed during this study are included in this published article (and its Supplementary Information files), and/or are available in the [Swissprot accession] repository, and indicated as such in the text.

Received: 8 October 2024; Accepted: 17 February 2025
Published online: 25 February 2025

References

1. McClamrock, H. D. Recombinant gonadotropins. *Clin. Obstet. Gynecol.* **46**, 298–316. https://doi.org/10.1097/00003081-200306000-00009 (2003).
2. Saint-Dizier, M., Foulon-Gauze, F., Lecompte, F., Combarnous, Y. & Chopineau, M. Cloning and functional expression of the equine luteinizing hormone/chorionic gonadotrophin receptor. *J. Endocrinol.* **183**, 551–559. https://doi.org/10.1677/joe.1.05888 (2004).

3. Allen, W. R. & Moor, R. M. The origin of the equine endometrial cups. I. Production of PMSG by fetal trophoblast cells. *J. Reprod. Fertil.* **29**, 313–316. https://doi.org/10.1530/jrf.0.0290313 (1972).
4. Squires, E. L., Garcia, M. C. & Ginther, O. J. Effects of pregnancy and hysterectomy on the ovaries of pony mares. *J. Anim. Sci.* **38**, 823–830. https://doi.org/10.2527/jas1974.384823x (1974).
5. Allen, W. R. Luteal deficiency and embryo mortality in the mare. *Reprod. Domest. Anim.* **36**, 121–131 (2001).
6. Murphy, B. D. in *Anim Reprod* Vol. 9 223–230 (2012).
7. Legardinier, S., Cahoreau, C., Klett, D. & Combarnous, Y. Involvement of equine chorionic gonadotropin (eCG) carbohydrate side chains in its bioactivity; lessons from recombinant hormone expressed in insect cells. *Reprod. Nutr. Dev.* **45**, 255–259. https://doi.org/10.1051/rnd:2005018 (2005).
8. Legardinier, S., Duonor-Cerutti, M., Devauchelle, G., Combarnous, Y. & Cahoreau, C. Biological activities of recombinant equine luteinizing hormone/chorionic gonadotropin (eLH/CG) expressed in Sf9 and mimic insect cell lines. *J. Mol. Endocrinol.* **34**, 47–60. https://doi.org/10.1677/jme.1.01624 (2005).
9. Smith, P. L., Bousfield, G. R., Kumar, S., Fiete, D. & Baenziger, J. U. Equine lutropin and chorionic gonadotropin bear oligosaccharides terminating with SO4-4-GalNAc and Sia alpha 2,3Gal, respectively. *J. Biol. Chem.* **268**, 795–802 (1993).
10. Aggarwal, B. B. & Papkoff, H. Plasma-clearance and tissue uptake of native and desialylated equine gonadotropins. *Domest. Anim. Endocrin.* **2**, 173–181 (1985). doi:Doi 10.1016/0739–7240(85)90013-X.
11. Martinuk, S. D., Manning, A. W., Black, W. D. & Murphy, B. D. Effects of carbohydrates on the pharmacokinetics and biological activity of equine chorionic gonadotropin in vivo. *Biol. Reprod.* **45**, 598–604. https://doi.org/10.1095/biolreprod45.4.598 (1991).
12. Manteca Vilanova, X., De Briyne, N., Beaver, B. & Turner, P. V. Horse Welfare during equine Chorionic Gonadotropin (eCG) production. *Anim. (Basel).* **9** https://doi.org/10.3390/ani9121053 (2019).
13. Hoppen, H. O. The equine placenta and equine chorionic gonadotrophin–an overview. *Exp. Clin. Endocrinol.* **102**, 235–243. https://doi.org/10.1055/s-0029-1211287 (1994).
14. Europe, C. o. in *European pharmacopoeia* Vol. 01/2008:0719 Ch. 01/2008:0719, 2121–2123 (Council Of Europe: European Directorate for the Quality of Medicines and Healthcare, Strasbourg, ©, (2010).
15. Cole, H. H. E., 48-HOUR, J. & Assay test for equine gonadotropin with results expressed in international units. *Endocrinology* **29**, 514–519 (1941).
16. Ciller, U. A., C., I. M. & McFarlane, J. R. in *SRB* : 39th Annual Conference of the Society for Reproductive Biology, Melbourne, Australia, 25th – 28th August, 2008 (CSIRO Publishing, Australia, 2008). (2008).
17. Bousfield, G. R., Butnev, V. Y., Gotschall, R. R., Baker, V. L. & Moore, W. T. Structural features of mammalian gonadotropins. *Mol. Cell. Endocrinol.* **125**, 3–19. https://doi.org/10.1016/s0303-7207(96)03945-7 (1996).
18. Damm, J. B., Voshol, H., Hard, K., Kamerling, J. P. & Vliegenthart, J. F. Analysis of N-acetyl-4-o-acetylneuraminic-acid-containing N-linked carbohydrate chains released by peptide-N4-(N-acetyl-beta-glucosaminyl)asparagine amidase F. Application to the structure determination of the carbohydrate chains of equine fibrinogen. *Eur. J. Biochem.* **180**, 101–110. https://doi.org/10.1111/j.1432-1033.1989.tb14620.x (1989).
19. Damm, J. B., Hard, K., Kamerling, J. P., van Dedem, G. W. & Vliegenthart, J. F. Structure determination of the major N- and O-linked carbohydrate chains of the beta subunit from equine chorionic gonadotropin. *Eur. J. Biochem.* **189**, 175–183. https://doi.org/10.1111/j.1432-1033.1990.tb15474.x (1990).
20. Schams, D. & Papkoff, H. Chemical and immunochemical studies on pregnant mare serum gonadotropin. *Biochim. Biophys. Acta.* **263**, 139–148. https://doi.org/10.1016/0005-2795(72)90168-7 (1972).
21. Sharif, M. K., Shahneh, H., Zolfagharian, A. Z., Kiasari, H. & Hedayati, B. A. Production and purification of equine Chorionic Gonadotropin hormone using polyclonal antibody. *Iran. J. Biotechnol.* **12**, e15137. https://doi.org/10.5812/ijb.15137 (2014).
22. Alvarez, R. H. et al. Physical-chemical and biological characterization of different preparations of equine chorionic gonadotropin. *J. Vet. Sci.* **17**, 459–465. https://doi.org/10.4142/jvs.2016.17.4.459 (2016).
23. Rodriguez, M. C. et al. Physicochemical characterization of a recombinant eCG and comparative studies with PMSG commercial preparations. *Protein J.* **42**, 24–36. https://doi.org/10.1007/s10930-023-10092-x (2023).
24. Adkins, J. N. et al. Toward a human blood serum proteome: analysis by multidimensional separation coupled with mass spectrometry. *Mol. Cell. Proteom.* **1**, 947–955. https://doi.org/10.1074/mcp.m200066-mcp200 (2002).
25. Wilsher, S. & Allen, W. R. Factors influencing equine chorionic gonadotrophin production in the mare. *Equine Vet. J.* **43**, 430–438. https://doi.org/10.1111/j.2042-3306.2010.00309.x (2011).
26. Lamy, J. et al. Regulation of the bovine oviductal fluid proteome. *Reproduction* **152**, 629–644. https://doi.org/10.1530/REP-16-0397 (2016).
27. Fernandez-Hernandez, P. et al. The proteome of Equine Oviductal Fluid Varies before and after ovulation: a comparative study. *Front. Vet. Sci.* **8**, 694247. https://doi.org/10.3389/fvets.2021.694247 (2021).
28. Min, K. S., Park, J. J., Byambaragchaa, M. & Kang, M. H. Characterization of tethered equine chorionic gonadotropin and its deglycosylated mutants by ovulation stimulation in mice. *BMC Biotechnol.* **19**, 60. https://doi.org/10.1186/s12896-019-0550-6 (2019).
29. Byambaragchaa, M. et al. Luteinizing hormone-like and follicle-stimulating hormone-like activities of equine chorionic gonadotropin beta-subunit mutants in cells expressing rat luteinizing hormone/chorionic gonadotropin receptor and rat follicle-stimulating hormone receptor. *Anim. Cells Syst. (Seoul).* **25**, 171–181. https://doi.org/10.1080/19768354.2021.1943708 (2021).
30. Saneyoshi, T. et al. Equine follicle-stimulating hormone: molecular cloning of beta subunit and biological role of the asparagine-linked oligosaccharide at asparagine(56) of alpha subunit. *Biol. Reprod.* **65**, 1686–1690. https://doi.org/10.1095/biolreprod65.6.1686 (2001).
31. Valove, F. M., Finch, C., Anasti, J. N., Froehlich, J. & Flack, M. R. Receptor binding and signal transduction are dissociable functions requiring different sites on follicle-stimulating hormone. *Endocrinology* **135**, 2657–2661. https://doi.org/10.1210/endo.135.6.7988456 (1994).
32. Murphy, B. D. & Martinuk, S. D. Equine chorionic gonadotropin. *Endocr. Rev.* **12**, 27–44. https://doi.org/10.1210/edrv-12-1-27 (1991).
33. Liu, D. et al. Distinct O-Acetylation patterns of serum glycoproteins among humans, mice, and rats. *J. Proteome Res.* **23**, 5511–5519. https://doi.org/10.1021/acs.jproteome.4c00653 (2024).
34. Visser, E. A. et al. Sialic acid O-acetylation: from biosynthesis to roles in health and disease. *J. Biol. Chem.* **297**, 100906. https://doi.org/10.1016/j.jbc.2021.100906 (2021).
35. Morell, A. G., Gregoriadis, G., Scheinberg, I. H., Hickman, J. & Ashwell, G. The role of sialic acid in determining the survival of glycoproteins in the circulation. *J. Biol. Chem.* **246**, 1461–1467 (1971).
36. Villarraza, C. J. et al. Development of a suitable manufacturing process for production of a bioactive recombinant equine chorionic gonadotropin (reCG) in CHO-K1 cells. *Theriogenology* **172**, 8–19. https://doi.org/10.1016/j.theriogenology.2021.05.013 (2021).
37. Schauer, R., Reuter, G., Stoll, S. & Sialate O-acetylesterases: key enzymes in sialic acid catabolism. *Biochimie* **70**, 1511–1519. https://doi.org/10.1016/0300-9084(88)90288-x (1988).
38. Schauer, R. Sialic acids: fascinating sugars in higher animals and man. *Zool. (Jena).* **107**, 49–64. https://doi.org/10.1016/j.zool.2003.10.002 (2004).
39. Vos, G. M. et al. Sialic acid O-acetylation patterns and glycosidic linkage type determination by ion mobility-mass spectrometry. *Nat. Commun.* **14**, 6795. https://doi.org/10.1038/s41467-023-42575-x (2023).

40. McDonald, W. H. & Yates, J. R. 3rd. Shotgun proteomics and biomarker discovery. *Dis. Markers.* **18**, 99–105. https://doi.org/10.1155/2002/505397 (2002).
41. Hang, I. et al. Analysis of site-specific N-glycan remodeling in the endoplasmic reticulum and the Golgi. *Glycobiology* **25**, 1335–1349. https://doi.org/10.1093/glycob/cwv058 (2015).
42. Perez-Riverol, Y. et al. The PRIDE database resources in 2022: a hub for mass spectrometry-based proteomics evidences. *Nucleic Acids Res.* **50**, D543–D552. https://doi.org/10.1093/nar/gkab1038 (2022).
43. Roepstorff, P. & Fohlman, J. Proposal for a common nomenclature for sequence ions in mass spectra of peptides. *Biomed. Mass. Spectrom.* **11**, 601. https://doi.org/10.1002/bms.1200111109 (1984).
44. Dell, A. et al. Mass spectrometry of carbohydrate-containing biopolymers. *Methods Enzymol.* **230**, 108–132. https://doi.org/10.1016/0076-6879(94)30010-0 (1994).
45. Richter, W. J., Muller, D. R. & Domon, B. Tandem mass spectrometry in structural characterization of oligosaccharide residues in glycoconjugates. *Methods Enzymol.* **193**, 607–623. https://doi.org/10.1016/0076-6879(90)93441-m (1990).

Acknowledgements
We appreciate the technical help that we received from Hafida Atiqi. We are grateful for graphical support by Polina Mishchenko and Antonis Katsoulas.

Author contributions
Author Contributions C.W.L., J. B-W., M.L., M.A. and T.B. designed the research project. C.W.L. performed experiments, analyzed and interpreted results. C.W.L., M.L. and T.B. wrote the manuscript. All authors read and approved the final manuscript.

Funding
We acknowledge project funding by the Swiss Laboratory Animal Science Association (SGV), Vontobel Stiftung, Stiftung pro Pferd (PR2018-05), KTI-Projekt Nr. 27179.2 PFLS-LS, and the Swiss 3RCC (TC-2020-008).

Declarations

Competing interests
The authors declare no competing interests.

Additional information
Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-025-90833-3.

Correspondence and requests for materials should be addressed to T.B.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.