Biopharmaceutical follicle-stimulating hormone (FSH) characterisation: monitoring Glycosylation Critical Quality Attributes (GCQAs) using a procainamide labelling system for structural glycan analysis and LC-ESI-QTOF for glycopeptide mapping

Rebecca I. Thomson[1], Richard A. Gardner[2], Radoslaw P. Kozak[2], Julia Smith[3], Katja Strohfeldt[1], Helen M. I. Osborn[1], and Daniel I. R. Spencer[2]
1 University of Reading, Whiteknights, Reading RG6 6AH, United Kingdom
2 Ludger Ltd., Culham Science Centre, Oxfordshire, OX14 3EB, United Kingdom
3 Bruker Ltd. Coventry, Coventry, United Kingdom

rad.kozak@ludger.com

Introduction

FSH is used clinically to stimulate follicular maturation for in vitro fertilisation (IVF) and treatment of anovulatory women. FSH glycosylation can significantly influence implantation and pregnancy rates in patients undergoing IVF. Given the potential impact on patients, biopharmaceutical developers producing FSH products must carefully optimise, accurately measure and tightly control glycosylation throughout the production lifetime of their drug. There are many Glycosylation Critical Quality Attributes (GCQAs) to consider including the presence of sialylation (which affects half-life), *O*-acetylation of sialic acids and oligomannoses (with possible unwanted effects).

In general, there are three main ways to analyse glycosylation:

- characterisation of intact glycoproteins
- characterisation of protease-digested glycopeptides
- structural analysis of glycans released from proteins

Our focus for this has been on:

- **Producing a practical workflow suitable for QbD-based drug realisation**
- **Compliance with emerging regulations from the FDA, EMA, KFDA and cFDA [1-3]**

Ludger's strategy for detection and quantitation of N-glycans in biologics

Figure 1 outlines the workflows we use for detection and quantitation of N-glycans in biopharmaceuticals. N-glycans are released from the glycoprotein using PNGase F endoglycosidase (QABio Cat # E-PNG01 or Ludger Cat # LZ-PNGASEF-96) then derivatised with a fluorescent tag – procainamide (Cat # LT-KPROC-24). Labelled glycans are run on two orthogonal platforms – HILIC (hydrophilic Interaction Liquid Chromatography) UHPLC and ESI-MS/MS – in hyphenated configuration.

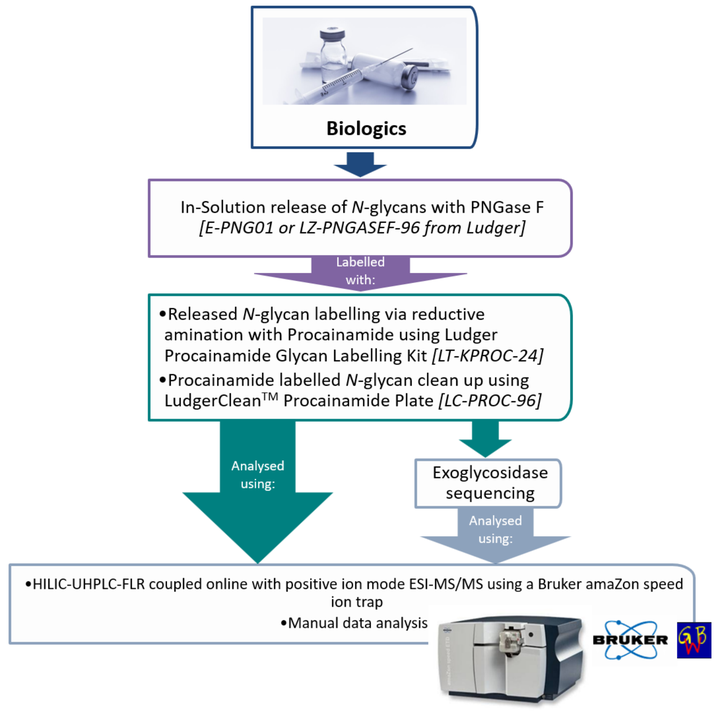

Figure 1: General glycoprofiling scheme for quantitative analysis of N-glycans in biologics.

Procainamide system for characterisation and quantitation of N-glycans in drugs

Analysis of procainamide labelled N-glycans has been shown to give significantly better performance for detailed, quantitative glycan structure determination than methods based on analysis of 2-aminobenzamide (2-AB) labelled glycans commonly used in the biopharma industry [4].

One of the GCQAs to consider is sialylation which greatly influences the clinical performance of biological drugs, the main effect being on therapeutic efficacy. Sialic acid *O*-acetylation is seen in FSH but, so far, has been relatively unexplored due to lack of suitable analytical tools.

The procainamide system is flexible and can be tuned to provide accurate measurements of N-glycans and their sialic acid *O*-acetylation using HILIC-UHPLC-ESI-MS combined with sialate-*O*-acetylesterase digest (see Fig 4). Together, the two workflows allow production of high resolution glycan maps that can be used to monitor GCQAs and assess any potential risks or issues with a particular product batch.

Figure 2 shows HILIC-UHPLC-FLR-ESI-MS profiles of procainamide labelled N-glycans from FSH. The glycan structure assignments and relative molar abundances determined from the HILIC-UHPLC -ESI-MS/MS analyses are given in Table 1.

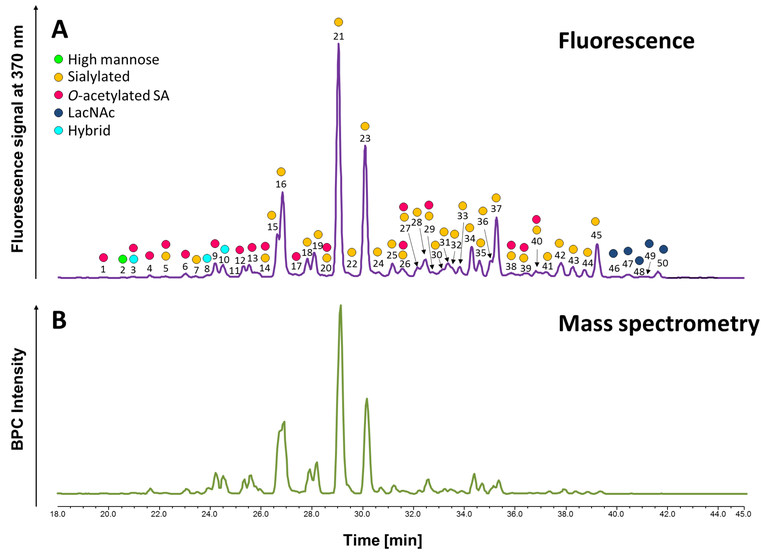

Figure 2: HILIC-UHPLC-ESI-MS profiles of procainamide labelled N-glycans released from FSH.
A: HILIC-FLR of procainamide labelled FSH N-glycans
B: HILIC-ESI-MS Base Peak Chromatogram (BPC) of procainamide labelled FSH N-glycans

Table 1 shows example glycan structure topologies that are consistent with the MS and MS/MS fragmentation data (see Fig 4). However, these must be considered as tentative - further analyses (e.g. exoglycosidase sequencing would be needed to increase confidence in the detailed structures proposed.

UHPLC peak ID	GU value	% Area	Possible structure	Composition: Hex (H)	HexNAc (N)	Fuc (F)	0 OAc	1 OAc	2 OAc	H+ adducts: [M+2]2+ calculated	[M+2]2+ observed	[M+3]3+ calculated	[M+3]3+ observed	[M+4]4+ calculated	[M+4]4+ observed
1	5.87	0.07	A2G2S2(Ac)4	5	4	0	0	0	2	1306.00	nd	871.01	870.93	653.51	653.43
2	6.03	0.03	Man5	5	1	0	0	0	0	727.81	727.78	585.54	585.87	-	-
3	6.12	0.19	A2G1	4	4	0	0	0	0	849.86	849.80	566.91	566.63	425.43	425.31
			A2G2S2(Ac)3	5	4	0	0	1	1	1285.00	1285.37	857.00	856.98	643.00	643.34
			FA2G2S2(Ac)4	5	4	1	0	0	2	1379.03	nd	919.69	919.60	690.02	690.03
4	6.29	0.31	A2G2S1(Ac)2	5	4	0	0	0	1	1118.45	1118.35	745.97	745.93	559.73	nd
5	6.44	0.26	A1G1S1	4	3	0	1	0	0	893.87	893.79	596.25	596.27	447.44	447.09
			A2G2S2(Ac)2	5	4	0	0	2	0	1263.99	nd	843.00	842.94	632.50	632.57
6	6.64	0.49	A2G2S1(Ac)1	5	4	0	0	1	0	1097.44	1097.33	731.96	731.93	549.22	nd
			FA2G2S1(Ac)2	5	4	1	0	0	1	1191.48	1191.50	794.65	794.59	596.24	596.90
7	6.74	0.19	A2G1S1	4	4	0	1	0	0	995.41	995.29	663.94	663.92	498.21	498.67
8	6.87	0.30	Polyhexose	8	0	0	0	0	0	767.81	767.77	512.21	512.59	-	-
9	6.94	1.52	A2G2S2(Ac)2	5	4	0	0	2	0	1263.99	1264.86	843.00	842.93	632.50	632.77
10	7.02	1.49	A2G2	5	4	0	0	0	0	930.89	930.81	620.93	620.93	465.95	465.88
11	7.15	0.15	Polyhexose	8	0	0	0	0	0	767.81	767.75	512.21	512.41	-	-
12	7.24	1.13	FA2G2S2(Ac)2	5	4	1	0	2	0	1337.02	1336.92	891.68	891.69	669.02	669.03
13	7.30	1.26	A2G2S2(Ac)1	5	4	0	1	1	0	1242.99	1242.87	829.00	828.93	622.00	622.06
14	7.37	0.87	FA2G2	5	4	1	0	0	0	1003.92	1003.79	669.61	669.58	502.46	nd
			A2G2S2(Ac)1	5	4	0	1	1	0	1242.99	1242.86	829.00	828.97	622.00	622.01
15	7.60	3.45	A2G2S1	5	4	0	1	0	0	1076.44	1076.33	717.96	717.95	538.72	538.59
16	7.66	8.68	A2G2S1	5	4	0	1	0	0	1076.44	1076.34	717.96	717.95	538.72	nd
17	7.80	0.43	A2G2S2(Ac)1	5	4	0	1	1	0	1242.99	1242.91	829.00	828.94	622.00	nd
18	7.93	2.17	FA2G2S1	5	4	1	1	0	0	1149.47	1149.33	766.65	766.66	575.24	575.28
19	8.01	2.58	FA2G2S1	5	4	1	1	0	0	1149.47	1149.88	766.65	766.60	575.24	nd
20	8.10	0.30	FA2G2S2(Ac)1	5	4	1	1	1	0	1316.02	nd	877.68	877.65	658.51	658.56
			FA2G2S1	5	4	1	1	0	0	1149.47	1149.20	766.65	766.62	575.24	575.87
21	8.29	22.08	A2G2S2	5	4	0	2	0	0	1221.98	1221.84	814.99	814.92	611.50	611.25
22	8.41	0.60	A2G2S2	5	4	0	2	0	0	1221.98	1221.84	814.99	814.92	611.50	611.25
23	8.61	13.00	FA2G2S2	5	4	1	2	0	0	1295.01	1295.25	863.68	863.63	648.01	648.22
24	8.77	0.71	A3G3S1	6	5	0	1	0	0	1259.00	1258.85	839.67	839.64	630.01	629.88
25	8.94	1.76	A3G3S1	6	5	0	1	0	0	1259.00	1259.38	839.67	839.61	630.01	nd
			FA2G2S2	5	4	1	2	0	0	1295.01	1295.36	863.68	863.62	648.01	647.87
			A3G3S3(Ac)1	6	5	0	2	1	0	1571.10	nd	1047.04	1047.64	786.06	786.04
26	9.06	1.52	A3G3S1	6	5	0	1	0	0	1259.00	1259.40	839.67	839.61	630.01	630.00
			FA3G3S1	6	5	1	1	0	0	1332.03	nd	888.36	888.29	666.52	666.50

UHPLC peak ID	GU value	% Area	Possible structure	Composition: Hex (H)	HexNAc (N)	Fuc (F)	0 OAc	1 OAc	2 OAc	H+ adducts: [M+2]2+ calculated	[M+2]2+ observed	[M+3]3+ calculated	[M+3]3+ observed	[M+4]4+ calculated	[M+4]4+ observed
27	9.26	1.04	FA3G3S1	6	5	1	1	0	0	1332.03	nd	888.36	888.32	666.52	666.69
			FFA2G2S2	5	4	2	2	0	0	1368.04	nd	912.36	912.31	684.53	684.06
			FA3G3S3(Ac)1	6	5	1	2	1	0	1644.13	nd	1096.42	1096.32	822.57	822.29
28	9.36	2.60	A3G3S2	6	5	0	1	0	0	1404.55	1404.28	936.70	936.66	702.78	702.77
29	9.45	0.64	FA3G3S3(Ac)1	6	5	1	1	0	0	1644.13	nd	1096.42	1096.32	822.57	822.53
			A3G3S2	6	5	0	1	0	0	1404.55	nd	936.70	936.62	702.78	702.79
30	9.59	2.56	A3G3S2	6	5	0	1	0	0	1404.55	nd	936.70	936.65	702.78	702.78
			FA3G3S2	6	5	1	1	0	0	1477.58	nd	985.39	985.29	739.29	739.28
31	9.66	0.82	FA3G3S2	6	5	1	1	0	0	1477.58	nd	985.39	985.31	739.29	739.29
32	9.71		FA3G3S2	6	5	1	1	0	0	1477.58	nd	985.39	985.31	739.29	739.29
33	9.81	1.34	FA3G3S2	6	5	1	1	0	0	1477.58	1477.94	985.39	985.30	739.29	739.29
34	9.98	3.01	A3G3S3	6	5	0	1	0	0	1550.10	nd	1033.73	1033.64	775.55	775.54
35	10.08	1.75	A3G3S3	6	5	0	1	0	0	1550.10	nd	1033.73	1033.64	775.55	775.55
36	10.24	1.43	FA3G3S3	6	5	1	1	0	0	1623.13	nd	1082.42	1082.33	812.07	812.07
37	10.32	5.96	FA2G2S2	5	4	1	2	0	0	1295.01	nd	863.68	863.69	648.01	nd
			FA3G3S3	6	5	1	1	0	0	1623.13	nd	1082.42	1082.28	812.07	812.03
38	10.53	1.17	FA3G3S3	6	5	1	1	0	0	1623.13	nd	1082.42	1082.32	812.07	812.02
			A3G3S3(Ac)1	6	5	0	2	1	0	1587.10	nd	1058.40	1058.30	794.05	794.02
39	10.76	0.82	FA2G4S2	7	6	1	2	0	0	1660.15	nd	1107.10	1107.04	830.58	830.54
			A3G3S3(Ac)1	6	5	0	2	1	0	1587.10	nd	1058.40	1058.33	794.05	794.04
40	10.90	1.10	FA4G4S2	7	6	1	2	0	0	1660.15	1660.44	1107.10	1107.65	830.58	830.53
			FA4G4S4(Ac)1	7	6	1	3	1	0	-	-	1315.17	nd	986.63	986.51
41	11.06	0.83	A4G4S3	7	6	0	3	0	0	-	-	1155.45	1155.73	866.84	866.80
42	11.28	2.06	A4G4S3	7	6	0	3	0	0	-	-	1155.45	1155.25	866.84	866.79
			FA4G4S3	7	6	1	3	0	0	-	-	1204.13	1204.00	903.35	903.31
43	11.47	1.28	FA4G4S3	7	6	1	3	0	0	-	-	1204.13	1204.04	903.35	903.30
44	11.65	0.87	A4G4S4	7	6	0	4	0	0	-	-	1252.48	1252.65	939.61	939.58
45	11.86	3.36	FA3G3S3	6	5	1	3	0	0	1623.13	-	1082.42	1082.36	812.07	nd
			FA4G4S4	7	6	1	0	0	0	-	-	1301.16	1301.00	976.12	976.09
46	12.15	0.27	A4G4S3(LacNAc)1	8	7	0	3	0	0	-	-	1277.16	nd	958.12	957.99
47	12.37	0.52	FA4G4S3(LacNAc)1	8	7	1	3	0	0	-	-	1325.84	nd	994.63	994.53
48	12.59	0.15	FA4G4S3(LacNAc)1	8	7	1	3	0	0	-	-	1325.84	1326.34	994.63	994.38
49	12.69	0.15	A4G4S4(LacNAc)1	8	7	0	4	0	0	-	-	1374.19	nd	1030.89	1030.81
50	12.89	0.70	FA4G4S4(LacNAc)1	8	7	1	4	0	0	-	-	1422.87	nd	1067.41	1067.31

Table 1: Summary glycan assignment for FSH.

Figures 3 and 4 and Table 1 summarise the results of this glycoprofiling scheme applied to quantitative characterisation of the N-glycans of FSH. This drug contains a complex mixture of glycan structures, several of which co-elute on HILIC-LC and/or have the same mass composition.

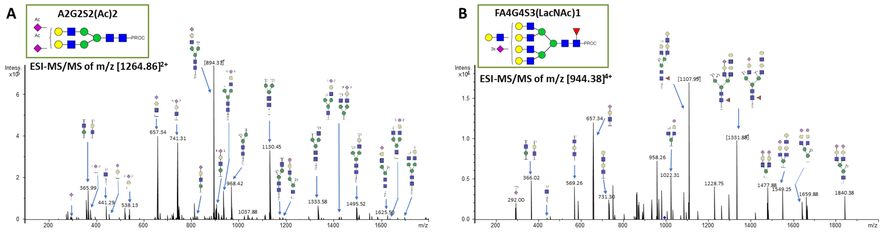

Figure 3: Example of MS/MS fragmentation patterns of procainamide labelled N-glycans released from FSH

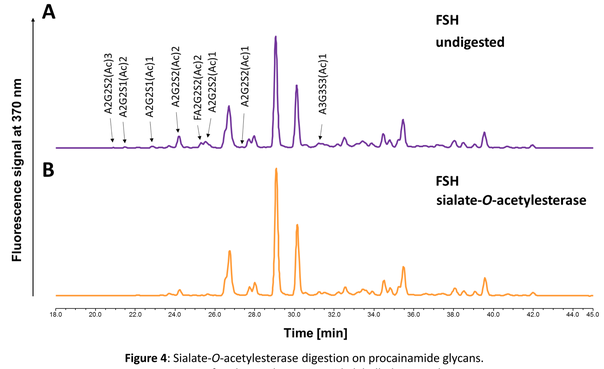

Figure 4: Sialate-*O*-acetylesterase digestion on procainamide glycans.
A: HILIC-FLR of undigested procainamide labelled FSH N-glycans
B: Identities of N-glycans containing *O*-acetylated sialic acids were confirmed after treatment with sialate-*O*-acetyl esterase (Ludger Cat # LZ-ACASE-KIT)

Ludger's strategy for glycopeptide analysis

Figure 5 outlines the workflow we use for glycopeptide analysis. The FSH sample was subjected to reduction and alkylation followed by overnight trypsin digestion. Glycopeptides were separated by an acetonitrile gradient on a Ionopticks C18 nano UHPLC, 1.6 µm, 75µm x 250mm Column. Spectra were acquired using a modified standard Instant Expertise method on a Bruker impact II ESI QTOF, with a fixed MS duty cycle of 3.5 s at 2 Hz and variable MSMS at 1.5 – 4 Hz depending upon precursor intensity. Peak lists were generated in DataAnalysis 4.3 and glycopeptide spectra identified in ProteinScape 4.0.

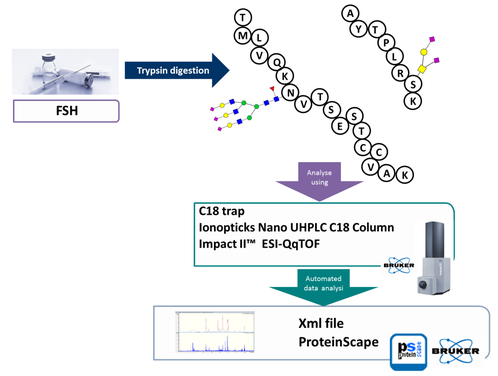

Figure 5: Scheme for glycopeptide analysis.

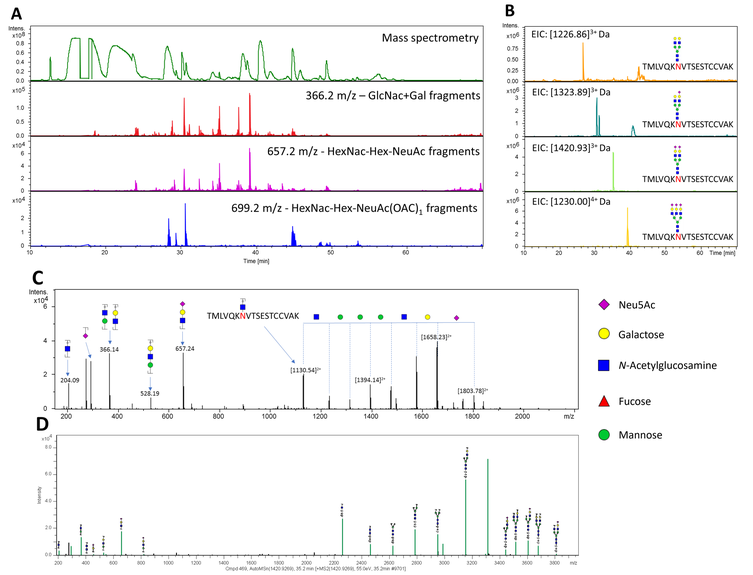

Figure 6: The FSH sample - glycopeptide chromatography on the Ionopticks C18 column and analysis on impact II MS.
A: BPC of sample, TIC and EIC for glycan fragments: 366.2 m/z – GlcNac+Gal fragments; 657.2 m/z - HexNac-Hex-NeuAc fragments; 699.2 m/z - HexNac-Hex-NeuAc(OAC)$_1$ fragments
B: EIC showing the FSH α-subunit region glycopeptides with different glycans attached
C: Manual identification and labelling of TMLVQKNVTSESTCCVAK glycopeptide + A2G2S2 glycan
D: Proteinscape 4.0 automatic identification and labelling of TMLVQKNVTSESTCCVAK glycopeptide + A2G2S2 glycan.

Conclusion

The glycan workflow (Figure 1) is optimised for GMP grade analyses. The fluorescence data provides relative quantitation; and in combination with exoglycosidase sequencing and MS/MS workflows can provide high confidence N-glycan structural information.

The workflow for glycopeptide analysis provides information of which glycans are at specific amino acid sites. As more glycopeptide fragmentation information becomes available, and is entered into the ProteinScape 4.0 database, this analysis will become more efficient.

Acknowledgements

This work was funded by the BBSRC Case Award BB/K012673/1 and Ludger Ltd.

References

1. Fernandes, DL (2012) 'A Systematic Approach to Optimisation and Comparability of Biopharmaceutical Glycosylation Throughout the Drug Life Cycle' in Subramanian, G (Ed) Biopharmaceutical Production Technology' published by Wiley-VCH Verlag GmbH & Co KGaA.
2. Fernandes, DL (2012) 'Assessing Comparability of Glycosylation between a Biosimilar Antibody and the Innovator Product', Antibodies Asia 2012, Shanghai, China
3. Tharmalingam T, Adamczyk B, Doherty MA, Royle L, Rudd PM. (2013) 'Strategies for the profiling, characterisation and detailed structural analysis of N-linked oligosaccharides' Glycoconj J. 30:137-146
4. Kozak RP, Tortosa CB, Fernandes DL, Spencer DI. (2015) 'Comparison of procainamide and 2-aminobenzamide labeling for profiling and identification of glycans by liquid chromatography with fluorescence detection coupled to electrospray ionization-mass spectrometry' Anal Biochem 486:38-40

Contact for more information

Thank you for viewing my poster. If you'd like a copy or want to know more about our glycomics workflow then please email me (**rad.kozak@ludger.com**) or connect with me on LinkedIn. Thanks, Rad.

Ludger
www.ludger.com