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

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

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

[2]Laboratory of PhysioPathology and PharmacoToxicology of the Human Placenta, UMR-S 1139, Inserm, Université de Paris, France.

[3]Sorbonne Université, France

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

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

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

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

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

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

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

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

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

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

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

Mass (Da)	Batch 1				Batch 2				Batch 3			
	<t_R> (min)	RSD(t_R) (%)	<relative area> (%)	RSD(relative area) (%)	<t_R> (min)	RSD(t_R) (%)	<relative area> (%)	RSD(relative area) (%)	<t_R> (min)	RSD(t_R (%)	<relative area> (%)	RSD(relative area) (%)
13848.9	18.67	0.43	13	14	18.31	0.14	13	14	18.63	1.11	5	9
13688.8	x	x	x	x	x	x	x	x	18.66	0.36	1	20
13724.8	18.67	0.34	17	9	18.33	0.19	18	35	18.83	1.21	1	26
14307.1	19.01	0.29	20	36	18.83	0.66	15	19	18.81	0.74	9	8
14032.0	18.81	0.44	2	12	x	x	x	x	18.60	0.86	7	14
14016.0	18.86	0.29	100	0	18.50	0.38	100	0	18.87	0.70	100	0
14381.0	18.80	0.29	16	10	18.46	0.34	19	29	18.88	0.59	2	23
14323.1	18.89	0.25	21	32	18.13	0.21	25	14	18.79	0.80	9	18
14349.0	19.29	0.43	23	28	18.84	0.77	26	31	18.82	0.24	6	37
14431.0	19.12	0.23	9	9	18.83	0.24	9	11	19.01	0.63	2	23
14397.1	19.28	0.77	8	21	18.85	0.68	14	15	18.80	0.24	2	12
14391.1	19.26	0.47	22	34	18.77	0.69	31	32	18.84	0.46	7	21
14672.2	19.09	0.13	35	1	18.65	0.37	39	11	18.84	0.75	35	31
14440.0	19.00	0.47	6	17	18.79	0.20	6	34	18.95	0.22	6	31

14963.3	19.28	0.29	86	5	19.02	0.13	87	17	19.08	0.40	37	34
14472.0	19.18	0.13	8	12	18.94	0.17	6	17	19.04	0.68	6	20
14979.3	18.91	1.13	24	16	18.74	1.90	20	16	19.03	0.65	8	19
13562.8	19.26	0.14	30	17	19.09	0.21	23	6	18.92	0.89	21	34
14640.1	19.37	0.21	8	6	19.02	0.30	10	20	19.27	0.33	5	39
13852.9	19.34	0.19	4	6	19.03	0.17	3	5	19.01	0.85	2	20
14512.0	19.26	0.65	13	26	18.87	0.32	14	35	19.11	0.82	5	21
15328.3	19.16	0.33	17	23	18.91	0.14	26	16	19.01	0.77	9	30
14614.2	19.28	0.21	48	4	19.05	0.33	42	13	19.24	0.35	13	26
14598.1	19.35	0.19	33	2	19.05	0.15	42	16	19.29	0.61	19	15
14682.2	19.43	0.24	60	4	19.09	0.00	60	8	19.37	0.75	18	11
15254.3	19.48	0.15	69	5	19.24	0.10	72	13	19.74	0.53	32	22
14680.2	19.41	0.34	8	8	19.09	0.16	9	27	19.39	0.60	4	37
15619.4	19.40	0.39	26	6	19.18	0.22	35	7	19.56	0.75	14	28
15338.4	19.53	0.21	22	2	19.28	0.33	24	11	19.87	0.38	12	37
15910.7	19.52	0.20	15	2	19.26	0.29	18	14	20.18	0.31	6	31

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