SUPPORTING INFORMATION:

The discovery of the once weekly glucagon like peptide 1 (GLP-1) analog semaglutide

Jesper Lau*, Paw Bloch, Lauge Schäffer, Ingrid Pettersson, Jane Spetzler, Jacob Kofoed, Kjeld Madsen, Lotte Bjerre Knudsen, James McGuire, Dorte Bjerre Steensgaard, Holger Martin Strauss, Dorte X. Gram, Sanne Moller Knudsen, Flemming Seier Nielsen, Peter Thygesen, Steffen Reedtz-Runge and Thomas Kruse.

Contents of supporting information:

2S

Mass spectrometric methods and data.

MALDI-TOF:

Molecular weights of the peptides were determined using matrix-assisted laser desorption time of flight mass spectroscopy (MALDI-MS), recorded on a Microflex (Bruker). A matrix of alfa-cyano-4-hydroxy cinnamic acid was used.

Or altenatively LCMS with electrospray ionization using the system stated in the scheme and according to the list below:

Agilent TOF:

An Agilent Technologies LC/MSD TOF (G1969A) mass spectrometer was used to identify the mass of the sample after elution from an Agilent 1200 series HPLC system. The de-convolution of the protein spectra was calculated with Agilent's protein confirmation software. Eluents: A: 0.1% Trifluoro acetic acid in water; B: 0.1% Trifluoro acetic acid in acetonitrile . Column: Zorbax 5u, 300SB-C3, 4.8x50mm. Gradient: 25% - 95 % B over 15 min.

Sciex3000

A Perkin Elmer Sciex API 3000 mass spectrometer was used to identify the mass of the sample after elution from a Perkin Elmer Series 200 HPLC system. Eluents: A: 0.05% Trifluoro acetic acid in water; B: 0.05% Trifluoro acetic acid in acetonitrile. Column: Waters Xterra MS C-18 X 3 mm id 5 µm. Gradient: 5% - 90 % B over 7.5 min at 1.5ml/min

LCT premier

Performed on a setup consisting of Waters Acquity UPLC system and LCT Premier XE mass spectrometer from Micromass. Eluents: A: 0.1% Formic acid in water; B: 0.1% Formic acid in acetonitrile The analysis was performed at RT by injecting an appropriate volume of the sample (preferably 2-10µl) onto the column which was eluted with a gradient of A and B.The UPLC conditions, detector settings and mass spectrometer settings were: Column: Waters Acquity UPLC BEH, C-18, 1.7µm, 2.1mm x 50mm. Gradient: Linear 5% - 95% acetonitrile during 4.0 min (alternatively 8.0 min) at 0.4ml/min. Detection: 214 nm (analogue output from TUV (Tunable UV detector)) MS ionisation mode: API-ES; Scan: 100-2000 amu (alternatively 500-2000 amu), step 0.1 amu.

Sciex 100 or 150

LCMS was performed on a setup consisting of Sciex API 100 (or150) Single quadropole mass spectrometer, Perkin Elmer Series 200 Quard pump ,Perkin Elmer Series 200 autosampler, Applied Biosystems 785A UV detector, Sedex 75 evaporative light scattering detector.The instrument control and data acquisition were done by the Sciex Sample control software running on a Windows 2000 computer. The HPLC pump is connected to two eluent reservoirs containing: A: 0.05% Trifluoro acetic acid in water B: 0.05% Trifluoro acetic acid in acetonitrile. The analysis is performed at room temperature by injecting an appropriate volume of the sample (preferably

3S

20 μl) onto the column which is eluted with a gradient of acetonitrile.The HPLC conditions, detector settings and mass spectrometer settings used are giving in the following table.Column: Waters Xterra MS C-18 X 3 mm id 5 um; Gradient: : 5% - 90 % acetonitrile linear during 7.5 min at 1.5ml/min;Detection : 210 nm (analogue output from DAD);ELS (analogue output from ELS), 40°C.MS ionisation mode API-ES.

4S

Alalog #	MS method	Mol Weight Calc. (found)	Ions found m/z
GLP-1		3355,7	N/A
liraglutide	LCMS Low resolution (ESI) Sciex100API	3751,2 (3751)	1876.4 1251.8 938.8
semaglutide	LCMS Low resolution (ESI) Sciex100API	4113,6 (4113)	1372.3 1029.8
1	LCMS Low resolution (ESI) Sciex 100	3383,7 (3384)	1129 847
2	LCMS Low resolution (ESI) (LCT Premier);	3779,3 (3779.5)	1260.8 945.9
3	LCMS Low resolution (ESI) LCT premier	4097,6	1366.7 1025.2
4	LCMS Low resolution (ESI) Sciex100API	3652,1 (3652)	1826,4 1218 914
5	LCMS Low resolution (ESI) Sciex100API	3680,1 (3680)	1840 1228 922
6	LCMS Low resolution (ESI) Sciex100API	4099,6 (4099)	1367.3 1026.2

5S

7	LCMS Low resolution (ESI) Sciex100API	3397,7 (3397)	1133.5 850.3
8	LCMS Low resolution (ESI) LCT premier	3765,2 (3765)	1883.6 1255.9
9	LCMS Low resolution (ESI) Sciex150	3926,4 (3926))	1309,8.
10	LCMS Low resolution (ESI) Sciex100API	3910,4 (3910)	1304 979
11	LCMS Low resolution	4055,5 (4055)	1352,9
12	LCMS Low resolution	4200,7 (4200)	1051,7 1401,1
13	LCMS Low resolution (ESI) Sciex100API	4083,6 (4083)	1362 1021
14	LCMS Low resolution (ESI) LCT premier	4111,6	1371.4 1028.9
15	LCMS Low resolution (ESI)	4029,4 (4029)	1344 1009
16	LCMS Low resolution (ESI)	4057,5 (4056)	1353 1015
17	LCMS Low resolution (ESI) Sciex100API	3956,4 (3956)	1319,8

18	LCMS Low resolution (ESI)	4085,5 (4086)	1363 1022
19	LCMS Low resolution (ESI) Sciex100API	3694,2 (3693)	1232 924
20	LCMS Low resolution (ESI) Sciex100API	3823,3 (3824)	1275.8 957.3
21	LCMS Low resolution (ESI)	3968,4 (3968)	1984.6 1232.3 992.9
22	LCMS Low resolution (ESI)	4113,6 (4113)	1372 1029
23	LCMS Low resolution (ESI) Sciex100API	4112,6 (4112)	1371.7 1029.0 924.5
24	MALDI-TOF alpha-cyano-4-hydroxycinnamic acid	4258,7 (4256)	4256
25	LCMS Low resolution (ESI) Sciex100API	4246,8 (4248)	1416.4 1062.5
26	LCMS Low resolution (ESI) Sciex100API	4242,7 (4242)	1415.2 1062
27	LCMS Low resolution (ESI) Sciex100API	4371,8 (4372)	1458.4 1094.3

28	LCMS Low resolution (ESI) Sciex100API	4182,6 (4182)	1395.1 1046.6 837.5
29	LCMS Low resolution (ESI)	4069,6 (4069)	2035.5 1357.4 1018.2
30	LCMS Low resolution (ESI) Sciex100API	4053,5 (4053)	1014.3 1352.1
31	LCMS Low resolution (ESI) Sciex100API	4141,6 (4142)	1036,7 1381,8
32	LCMS Low resolution (ESI)	4169,7 (4169)	1390.7 1043.4
33	LCMS Low resolution (ESI) Sciex100API	3970,4 (3970)	1325 993
34	LCMS Low resolution (ESI) Sciex100API	3970,4 (3970)	993,8 1987,2
35	LCMS (ESI) Agilent TOF	4084,6 (4084.9)	2043.1 1362.4 1022.1

36	LCMS (ESI) Agilent TOF	4114,6 (4114.9	2058.1 1372.4 1029.5
37	LCMS Low resolution (ESI) Sciex100API	3937,5 (3938)	1969
38	MALDI-TOF alpha-cyano-4-hydroxycinnamic acid	4156,7 (4156.8)	4156.8
39	MALDI-TOF alpha-cyano-4-hydroxycinnamic acid	4142,7 (4141,9)	4141,9
40	LCMS Low resolution (ESI) Sciex100API	3694,2 (3694)	924,8 1232,4 1848,2
41	LCMS Low resolution (ESI) Sciex100API	3850,3 3851,1	1284,8 964,0
42	LCMS Low resolution (ESI) Sciex100API	4140,6 (4140,7)	829,5 1036, 1381,3
43	LCMS Low resolution (ESI) Sciex100API	4269,8 (4269)	1423.9 1068.8 855.3
44	LCMS Low resolution (ESI) Sciex100API	4126,6 (4126,7)	1376,6

UPLC and HPLC methods used for purity analysis:

General methods:

The RP-HPLC analysis may be performed using UV detection at 214 nm and e.g. a Vydac 218TP54 4.6mm x 250mm 5μ C-18 silica column (The Separations Group, Hesperia, USA) and eluted at e.g. 1 ml/min at 42 ^{o}C. Most often one of four different elution conditions are used:

A1: Equilibration of the column with a buffer consisting of 0.1M $(NH_4)_2SO_4$, which was adjusted to pH 2.5 with concentrated H_2SO_4 and elution by a gradient of 0% to 60% CH_3CN in the same buffer during 50 min.

B1: Equilibration of the column with 0.1% TFA / H_2O and elution by a gradient of 0% CH_3CN / 0.1% TFA / H_2O to 60% CH_3CN / 0.1% TFA / H_2O during 50 min.

B6: Equilibration of the column with 0.1% TFA / H_2O and elution by a gradient of 0% CH_3CN / 0.1% TFA / H_2O to 90% CH_3CN / 0.1% TFA / H_2O during 50 min.

Alternatively the RP-HPLC analysis was performed using UV detection at 214 nm and a Symmetry300, 3.6mm x 150mm, 3.5μ C-18 silica column (Waters) which was eluted at 1 ml/min at 42 ^{o}C.

B4: Equilibration of the column with 0.05% TFA / H_2O and elution by a gradient of 5% CH_3CN / 0.05% TFA / H_2O to 95% CH_3CN / 0.05% TFA / H_2O during 15 min.

Analytical methods, HPLC and UPLC

Method name	Eluent A	Eluent B	Linear gradient	column	Column temp ^{o}C.	Run time (min)	Flow rate ml/min
B4	acetonitrile	1% TFA in water	5-95%	Symmetry300 C18 , 5 um, 3.9 mm x 150 mm	42	15	1
RX	acetonitrile	1% TFA in water	5-95%	Symmetry300 C18 , 5 um, 3.9 mm x 250 mm	42	15	1
J	acetonitrile	1% TFA in water	20-65%	Jupiter proteo C12 , 4 um, 4.6 mm x 250 mm	42	50	1
B6	acetonitrile	0.1% TFA in water	0-90%	Vydac 218TP54 4.6mm x 250mm 5 C-18 silica column	42	50	1
02	99.95% Acetonitrile, 0.05% Trifluoroacetic acid	99.95% Water, 0.05% Trifluoroacetic acid	5-95%	ACQUITY UPLC BEH C18, 1.7um, 2.1 mm x 150 mm column	40	16	0.4
B2	99.95% Acetonitrile, 0.05% Trifluoroacetic acid	99.95% Water, 0.05% Trifluoroacetic acid	5-60	BEH130, C18, 1.7um, 2.1x150mm	40	16	0.40
07	20% Isopropanole, 20% Water and 60% Acetonitrile	0.09 M di-Ammonium Hydrogen Phosphate (aq) and 10% Acetonitrile, pH 3.6	Step gradient: 35% B over 2 minutes, then 35% to 65% B over 15 minutes, then 65% to 80% B over 3 minutes and finally 80% B over 1 minute	Phenomenex Kinetex C18, 1.7um, 2.1 mm x 150 mm column	60	23	0.5
A6	80%acetonitrile	10mM TRIS and 15mM ammonium sulphate pH 7.3 in 20%acetonitrile 80% water	5-90	HSS T3, 1.8um, 2.1 mm x 150 mm column	30	16	0.35
A3	70 % acetonitrile.	A: 90 % H_2O, 10 % CH_3CN, 0.25 M ammonium bicarbonate; B: 70 % CH_3CN, 30 % H_2O.	25-55	HSS T3, 1.8um, 2.1 mm x 150 mm column	30	16	0.35

List of analytical data:

Analog #	Retention time (min)	Purity (%)	Method
liraglutide	10.8	100	B4
semaglutide	9.6	100	B4
1	10.7	96	B2
2	46.6	97	J
3	11.4	91	B4
4	9.50	97	B4
5	9.89	96	B4
6	11.8	93	B4
7	7.8	100	B4
8	10.6	98	A6
9	35.2	92	B6
10	10.7	95	B4
11	9.9	95	02
12	34.5	100	B6
13	10.7	88	B4
14	11.4	91	B4
15	9.7	96	07
16	8.2	96	B4
17	30.4	99	B6
18	8.4	95	02
19	9.9	91	B4
20	9.6	97	B4
21	14.0	100	A3
22	8.8	91	02
23	8.7	100	02
24			
25	32.1	95	B6
26	10.1	100	B4
27	10.7	100	B4
28			
29	9.1	100	B4
30	8.87	89.5	02
31	9.6	100	B4
32	10.4	94	B4
33	12.1	100	B4
34	34.0	100	B6

12S

35	12.6	98	A3
36	8.5	99	B4
37	14.6	94	07
38	8.7 11.2	98 93	B4 A3
39	8.3 11.3	94 94	B4 A3
40	11.0	95	B4
41	9.6	99	B4
42	32.4	95	B6
43	9.3	100	B4
44	32.1	98	RX