

利用 Q-TOF 结合 PEAKS Studio 软件确证司美格鲁肽氨基酸序列

LCMS-QTOF-111

摘要： 本文采用 LCMS-Q-TOF 高分辨液质联用系统分析司美格鲁肽样品，获得其二级质谱图，结合 PEAKS Studio 软件对司美格鲁肽的氨基酸序列进行了确证。实验结果显示，该方法可以用于快速确证司美格鲁肽的氨基酸序列。

关键词： 司美格鲁肽 氨基酸序列 Q-TOF PEAKS Studio 软件

技术特点：

- ❖ Glu-C 酶切相对于其他酶切（如胰蛋白酶、赖氨酸蛋白酶）在司美格鲁肽结构确认上具有明显优势，可实现 100% 的序列覆盖度。
- ❖ PEAKS Studio 软件可基于 sequence tag 的搜库算法，完成司美格鲁肽的氨基酸序列确认，保证鉴定结果的准确性。

司美格鲁肽是一种新型长效胰高血糖素样肽 -1 (GLP-1) 受体激动剂，广泛应用于 2 型糖尿病和肥胖症的治疗，其通过模拟天然 GLP-1 的作用机制，促进胰岛素分泌、抑制胰高血糖素释放，从而有效控制血糖水平并实现减重效果。对于肥胖人群，司美格鲁肽可通过调节中枢神经系统中的食欲控制通路，减少食物摄入量，帮助患者实现可持续的体重管理。目前，司美格鲁肽凭借其卓越的疗效和良好的安全性，已成为糖尿病和肥胖症治疗领域的一大热点，受到医药行业及患者群体的广泛关注。

但由于多肽结构的复杂性及修饰特性，在药物研发与生产过程中，序列确证尤为重要。准确的序列确证不仅确保了药物分子的一致性与纯度，为质量控制提供了科学依据，还进一步保障了药物在治疗糖尿病和减肥领域的安全性和有效性。

本文利用岛津 LCMS-Q-TOF 高分辨液质联用系统，结合 PEAKS Studio 软件对司美格鲁肽的质谱数据进行了深入分析。成功实现了对其氨基酸序列的快速确认，为司美格鲁肽在药物研发、质量控制及规模化生产中的序列一致性提供了可靠的科学依据。

■ 实验部分

1.1 仪器

岛津 LCMS-Q-TOF 高分辨液质联用系统，配置信息如下：

系统控制器	: CBM-40lite	自动进样器	: SIL-40C XS
脱气机	: DGU-403	柱温箱	: CTO-40S
输液泵	: LC-40D XS×2	检测器	: LCMS-9050
色谱工作站	: Labsolutions Ver. 5.120		
数据处理软件	: PEAKS Studio Ver.12.0		

1.2 分析条件

液相色谱条件

色谱柱：Shim-pack GIST C18 AQ (100 mm×2.1 mm I.D., 1.9 μm),
岛津（上海）实验器材有限公司，P/N: 227-30807-02

流 动 相 : A 相: 0.1% 甲酸水溶液; B 相: 0.1% 甲酸乙腈溶液
流 速 : 0.3 mL/min
进 样 体 积 : 1 μ L
柱 温 : 50 $^{\circ}$ C
洗 脱 方 式 : 梯度洗脱, B 相初始浓度为 2%, 时间程序见表 1。

表 1 梯度洗脱程序

时间	单元	处理命令	值
1.00	泵	B.conc	2
5.00	泵	B.conc	30
7.00	泵	B.conc	90
9.00	泵	B.conc	90
9.10	泵	B.conc	2
12.00	控制器	Stop	

质谱条件

离子化模式 : ESI(+)	加热模块温度 : 400 $^{\circ}$ C
雾化器流速 : 3.0 L/min	接 口 电 压 : 4.5 kV
加热气流速 : 10.0 L/min	碰 撞 电 压 : 30 V
干燥气流速 : 10.0 L/min	事 件 时 间 : 0.1 s
接 口 温 度 : 300 $^{\circ}$ C	扫 描 模 式 : MS+DDA
D L 温 度 : 250 $^{\circ}$ C	

1.3 样品前处理

稀释: 使用 50 mM 碳酸氢铵溶液将司美格鲁肽稀释至约 100 μ g/mL。



变性: 取 100 μ L 司美格鲁肽溶液于 98 $^{\circ}$ C 加热 15 min。



酶解: 在变性溶液中加入 1 μ g Glu-C 酶, 涡旋混合均匀, 37 $^{\circ}$ C 酶解 12 h。



终止: 在酶解溶液中加入 1 μ L 甲酸, 涡旋混合均匀, 酶解终止。

■ 结果与讨论

2.1 司美格鲁肽酶解

司美格鲁肽由 17 种共 31 个氨基酸组成，其中含一个非天然氨基酸 Aib（质量数相当于在一个丙氨酸结构中加了质量数约为 14 的修饰基团），且在赖氨酸处有一个 C18 脂肪二酸的修饰，司美格鲁肽具体结构见图 1。由于司美格鲁肽序列较长，修饰基团空间位阻大，为了提高序列的匹配度，需将其酶解为较短的肽段。常用的酶有胰蛋白酶、赖氨酸蛋白酶、Glu-C 等，其中胰蛋白酶酶切位点为精氨酸 R 和赖氨酸 K 的 C 端，从司美格鲁肽的结构可以看出，其中含有 1 个 K 和 2 个 R。K 上有一个大的修饰基团，会影响酶切效率，且两个 R 相隔很近，酶切后肽段仅含 1 个氨基酸，在质谱中响应较低，干扰较大。赖氨酸蛋白酶酶切位点为 K 的 C 端，同样会受到 K 上修饰基团的影响，酶切效率低。最终选用 Glu-C 酶，其酶切位点为氨基酸 E 和 D 的 C 端，理论上可将司美格鲁肽酶切为 5 个肽段，具体见图 2。

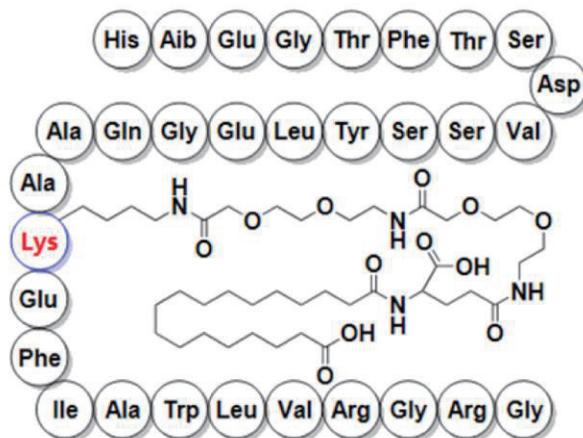


图 1 司美格鲁肽氨基酸序列



图 2 司美格鲁肽经 Glu-C 酶切后理论肽段

2.2 司美格鲁肽氨基酸序列确证

采用 LCMS-QTOF 的 DDA 模式采集 Glu-C 酶切后的司美格鲁肽样品，得到 TIC 图如图 3 所示。将此数据导入 PEAKS Studio 软件中，选择 DB Search 工作流程，并输入司美格鲁肽的理论氨基酸序列及其修饰基团，可自动进行氨基酸序列确证。

司美格鲁肽氨基酸序列覆盖率结果及具体匹配肽段信息如图 4 所示，从图可以看到氨基酸序列覆盖率为 100%，共匹配到 9 条肽段，除完全酶切得到的 5 条肽段外，还有未完全酶切的 4 条肽段。此外，PTM 为检测到的修饰，字母 A 和 S 分别代指非天然氨基酸 Aib 和 C18 脂肪二酸修饰。9 条肽段具体的质量、m/z、保留时间及偏差均给出，其中质量偏差均小于 5 ppm。根据软件提供的肽段 m/z 和保留时间信息，可将肽段与色谱峰对应起来，具体见图 3。

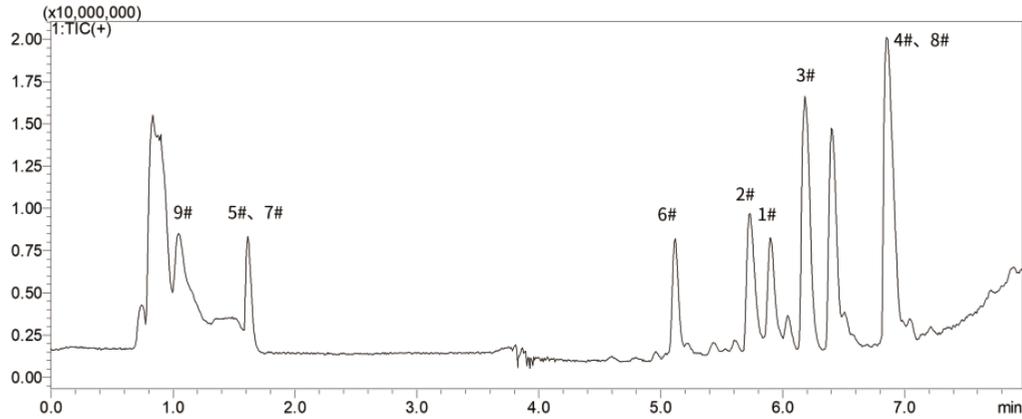
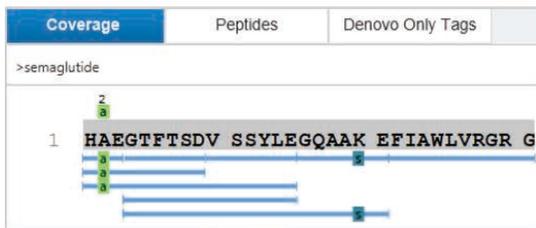


图3 LCMS-QTOF 采集酶切后司美格鲁肽 TIC 图及对应肽段

Accession	Gene	Cluster	Top	-10lgP	Coverage(%)	#Peptides	#Unique	PTM	Avg. Mass	Description
1	semaglutide	1	true	144.68	100.00%	9	9	A S	3384	>semaglutide

覆盖率

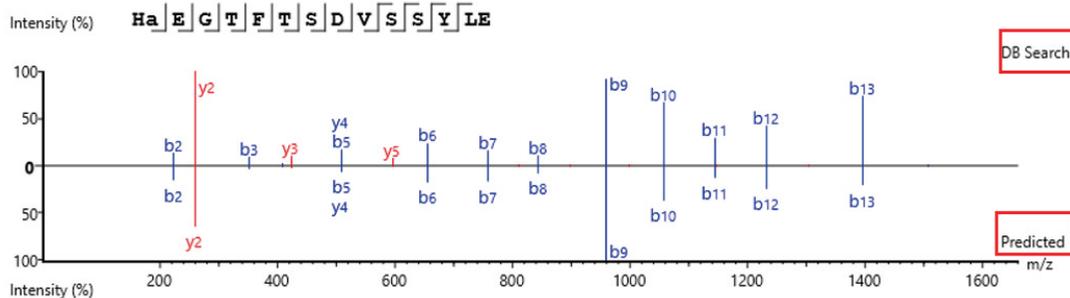


匹配肽段

Peptide	-10lgP	Mass	Length	ppm	m/z	RT	ΔRT	MS2 Correlation
1 HA(+14.02)EGTFTSDVSSYLE	87.39	1655.7314	15	-4.2	828.8696	5.91	-85.99	0.75
2 FIAWLVRGRG	36.58	1173.6770	10	-1.3	587.8450	5.73	-83.77	0.75
3 GTFTSDVSSYLE	34.53	1304.5771	12	-4.6	653.2929	6.23	-101.47	0.81
4 GTFTSDVSSYLEGQAAK(+715.43)E	31.86	2604.2947	18	-1.7	869.1039	6.88	-94.92	0.41
5 HA(+14.02)EGTFTSD	30.14	977.4090	9	-1.6	489.7110	1.62	2.02	0.19
6 VSSYLE	29.94	696.3330	6	-4.6	697.3371	5.10	-35.00	0.93
7 GTFTSD	27.38	626.2548	6	-1.8	627.2609	1.59	-11.81	0.45
8 GQAAK(+715.43)E	26.25	1317.7279	6	-3.4	1318.7307	6.90	30.00	0.34
9 HA(+14.02)E	11.89	369.1648	3	-1.3	370.1716	1.14	N/A	N/A

图4 肽段覆盖率及匹配肽段信息

点击具体肽段即可查看其匹配的氨基酸信息，图5为司美格鲁肽第一个匹配肽段的二级氨基酸碎片归属，其中，y1表示C端的第一个氨基酸碎裂产生的碎片离子，b1是N端的第一个氨基酸碎裂产生的碎片离子。以此类推，y2~y13对应C端第2~13个氨基酸处肽键断裂产生的碎片离子，b2~b13对应N端第2~13个氨基酸处肽键断裂产生的碎片离子。上图中DB Search部分为实际采集得到的二级质谱图，Predicted为软件根据肽段预测的氨基酸碎片及其强度，从图中可以看出，理论预测与实际采集到的氨基酸碎片均已匹配成功。下图为二级碎片对应的理论m/z及强度和实际采集得到的m/z。从图可知，强度大于0.89%的氨基酸碎片均检测到。



Ion Match				
#	Label	Predicted M/Z	Predicted Intensity (%) ▼	Query M/Z
14	b9	960.41	100.00	960.40
2	y2	261.14	64.28	261.14
16	b10	1059.47	36.66	1059.47
19	b12	1233.54	24.17	1233.54
21	b13	1396.60	20.56	1396.60
9	b6	657.30	17.65	657.30
10	b7	758.35	16.56	758.34
1	b2	223.12	15.26	223.12
17	b11	1146.51	12.76	1146.50
12	b8	845.38	7.92	845.38
6	b5	510.23	6.35	510.23
3	b3	352.16	3.26	352.16
5	y3	424.21	2.40	424.21
4	b4	409.18	1.55	409.20
8	y5	598.27	0.89	598.27

图 5 肽段的二级碎片归属

■ 结论

本文基于岛津 Q-TOF 高分辨液质联用仪 LCMS-9050 和 PEAKS Studio 数据分析软件，借助酶切手段，成功确证了司美格鲁肽氨基酸序列。LCMS-9050 具有分辨高、质量数准确度的特点，结合 PEAKS Studio 软件可快速完成多肽药物的氨基酸序列确证。

岛津应用云

