

基于液相色谱 - 超高分辨质谱的融合蛋白 O 糖基化位点分析策略

龙珍^a 胡馨月^b 祝翔^a 黄敏^a 连浩志^a 张劭阳^a

^a 赛默飞世尔科技（中国）有限公司

^b 中国食品药品检定研究院

关键字： O 糖肽， HCD 碎裂， Orbitrap， EThcD， nanoLC， O-glycan-specific Protein Digestion

1. 引言

N 糖基化和 O 糖基化修饰是蛋白糖基化修饰的两种主要形式。N 糖基化位点有特定的规律，通常包含在 NXS/T 这样的序列中 (X≠P)。此外，N 糖基化修饰的蛋白在经过 N 糖内切酶（如 PNGaseF）处理后，N 糖基化位点连接的糖链去除且该位点的脱酰胺比例显著增加。因此，用该策略可以对重组类蛋白的 N 糖基化位点做准确分析^[1]。对于 N 糖基化修饰比例较少的位点，或者复杂的 N 糖蛋白组学的样本，通常难以通过脱酰胺比例显著变化来判断是否是糖基化位点。对于这样的样本，还可以使用重水环境的 N 糖内切酶处理^[2]。重水环境下的 N 糖酶酶切后生成的肽段与脱酰胺肽段存在 2 Da 差异，在高分辨质谱中容易鉴定。

相较于 N 糖基化位点，O 糖基化位点的分析更为困难。第一，O 糖基化没有固定的序列特征，所有的 S/T 都可能带有 O 糖修饰。第二，没有 O 糖链内切酶；第三，多个 O 糖基化位点集中在同一肽段的情况，导致该肽段的质谱响应低、碎片少、鉴定结果的假阳性大。以依那西普为例，图 1 展示了依那西普理论 O 糖基化位点（红色显示）所在区域（T184、S199、T200、T205、

T208、S212、S213、S216、T217 和 S226）。如果用 trypsin 酶切处理该蛋白（黄颜色标记），这 10 个 O 糖基化位点将出现在 3 个肽段中，其中一个肽段包含多达 7 个 O 糖基化位点。T184 虽然可以酶切到单独的肽段中，但该肽段还含有 N 糖基化位点 N171（紫色文本显示），N 糖基化的加入，为该肽段的 O 糖基化鉴定带来复杂性。如果以 trypsin+chymotrypsin（绿色本底显示）+GluC（紫色本底显示）联合酶切，这些糖基化位点将分散在 5 个肽段中，其中糖基化位点最多的肽段包含 4 个 O 糖基化位点。

158PH QICNVVAIPG NASMDAVCTS TSPT^{RS}MAPG AVH^{LP}QP^{VST}
RSQHTQPTPE^{PT}STAPSTSFL^{LP}MGPSPPAE^{GS}TGDEPKS

图1 依那西普O糖基化位点序列图

本文以复杂的 O 糖基化修饰蛋白依那西普为样本，展示不同的分析策略对 O 糖基化位点的鉴定，供相关技术人员参考和选择。第一，PNGaseF+trypsin 酶切，UHPLC-HCD 鉴定；第二，PNGaseF+trypsin 酶切，nanoLC-EThcD 鉴定；第三，PNGaseF+trypsin+chymotrypsin+GluC 酶切，UHPLC-HCD 鉴定；第四，O-glycan-specific Protein Digestion 酶 / 与 trypsin 联合酶切，UHPLC-HCD 鉴定。

2. 实验部分

2.1 PNGaseF+trypsin 酶切

取适量样品用同等体积 RapiGest™ 于 60°C 变性 15 min 后，加入二硫苏糖醇 (Dithiothreitol, DTT) 溶液使得 DTT 的终浓度为 50 mM，混合均匀后于 60°C 反应 1 h，然后将样品冷却至室温；向上述样品中加入碘代乙酰胺 (Iodoacetamide, IAM) 溶液，使得 IAM 的终浓度为 100 mM，混合均匀后将样品放置在室温、避光环境中反应 30 min；取出样品，加入样品等体积的浓度为 50 mM 的碳酸氢铵溶液和 2 μL PNGaseF，混匀后于 37°C 反应过夜；然后加入 2 μg trypsin，混合均匀后于 37°C 反应 4 h；21,000 g，离心样品 10 min，取上清液分析。

2.2 PNGaseF+trypsin+chymotrypsin+GluC 酶切

取适量样品用同等体积 RapiGest™ 于 60°C 变性 15 min 后，加入 DTT 溶液使得其终浓度为 50 mM，混合均匀后于 60°C 反应 1 h，然后将样品冷却至室温；向上述样品中加入 IAM 使得其终浓度为 100 mM，混合均匀后将样品放置在室温、避光环境中反应 30 min；取出样品，加入样品等体积的浓度为 50 mM 的碳酸氢铵溶液和 2 μL PNGaseF 混匀后于 37°C 反应过夜；然后加入 2 μg GluC 混匀后于 37°C 反应 4 h；再加入 2 μg trypsin 混匀后于 37°C 反应 4 h；最后加入 2 μg 糜蛋白酶 (Chymotrypsin) 混匀后于 25°C 过夜反应；21,000 g，离心样品 10 min，取上清液分析。

2.3 O-glycan- specific Protein Digestion 酶切

取适量样品，加入 8 mol/L 尿素变性缓冲液使样品稀释至 1 mg/mL，57°C 加热 30 min 进行变性，冷却至室温。加入 DTT 溶液到终浓度 50 mM，57°C 加热 1 h；加入 1 mol/L IAM 溶液到终浓度 100 mM，室温避光反应 30 min；将反应完成后的样品转移到 10 kDa 超滤管中，用 21000 g 超滤离心 15 分钟；取出后再向超滤管中加入 100 μL 50mM Tris 缓冲液 (pH 6.8)，重复 2 次。用 100 μL 50mM Tris 缓冲液 (pH 6.8) 复溶样品后，添加 2.5 μL OPERATOR 以及 2.5 μL SialEXO，于 37°C 中孵育过夜 (16-18 h)，得到样品 1。将部分样品 1 用 21000 g，离心 5 min，取上清液分析 (样品 2)。取样品 1 约 50 μL，加入 2 μg trypsin 37°C 过夜酶切后，21000 g，离心 5 min，取上清液分析 (样品 3)。样品 2 和 3 用于液质分析。

3. 仪器及方法

质谱仪器：Orbitrap Exploris 240 和 Orbitrap Eclipse™ Tribrid™；

色谱仪器：Vaquish Flex 液相色谱系统和 UltiMate™ 3000 RSLCnano；

3.1 仪器方法

1) NanoLC-EThcD 分析

色谱仪：UltiMate™ 3000 RSLCnano；Trap 柱：Acclaim PepMap™ 100Å 75 μm x 20 mm, 2 μm，货号 164946；分析柱：Acclaim PepMap™ 100Å 75 μm x 250 mm, 2 μm，货号 164941；柱温：50°C；流动相：A: 0.1% FA-水；B: 0.1% FA: 乙腈(20:80, v/v)；上样流速：8 μL/min；分析流速：0.3 μL/min；上样量：1 μL；上样泵：100% A；

分析泵梯度：

时间 (min)	B/%
-5	2
0	2
3	8
98	28
113	35
116	99

质谱条件:

检测模式	+
喷雾电压	2.4 kV
毛细管温度	320°C
分辨率设置	一级 60,000@m/z 200; 二级 15,000@m/z 200
母离子扫描范围	m/z 350-1800
隔离窗口	1.6
MicroScans	1
HCD CE (%)	30
RF-Lens	60
ETD触发离子	CompoundName m/z 243.026 243.026 HexNAc 204.0867 HexNAcfragment 138.0545 HexNAcHex 366.1396 HexNAcfragment2 126.0551 HexNAcfragment3 144.0655 HexNAcfragment4 168.0655 HexNAcfragment5 186.0761

2) UHPLC-HCD 分析;

色谱仪: Vanquish Flex; 分析柱: BEHC18 130Å, 2.1 µm x 150 mm, 1.7 µm ; 柱温: 60°C

流动相: A: 0.1% FA- 水; B: 0.1% FA- 乙腈; 分析流速: 0.2 mL/min; 上样量: 10 µL

梯度:

时间 (min)	B/%
0	3
3	3
85	32
90	90
95	90
95.1	3
100	3

质谱条件:

检测模式	+
喷雾电压	3.8 kV
毛细管温度	320°C
雾化温度	350°C
分辨率设置	一级 120,000@m/z 200; 二级 15,000@m/z 200
母离子扫描范围	m/z 350-1800
AGC	一级 300%; 二级 100%
(n)CE	30%

4. 数据采集及分析

Thermo Xcalibur Qual Browser 定性浏览器

BioPharma Finder 5.2、pGlyco3.1 和 Byonic5.5 软件处理数据。pGlyco 软件为中科院计算机所贺思敏老师团队开发, 供研究人员免费使用。软件的下载、使用, 可参考以下链接, 或与贺思敏老师团队直接联系。

链接: https://pan.baidu.com/s/19M70Q_qm-c8AYe6SLDJRxQ?pwd=1234

提取码: 1234

5. 结果与讨论

5.1 PNGaseF+trypsin 酶解

在 N 糖脱除的情况下，trypsin 酶切后，UHPLC-HCD 检测可得到的较为确定的 O 糖基化位点是 T8、S186、S199、T245，如图 2 所示。该肽段虽然也包含 T200，但图 2 检出的肽段是 S199 糖基化，而 T200 未糖基化的情况。

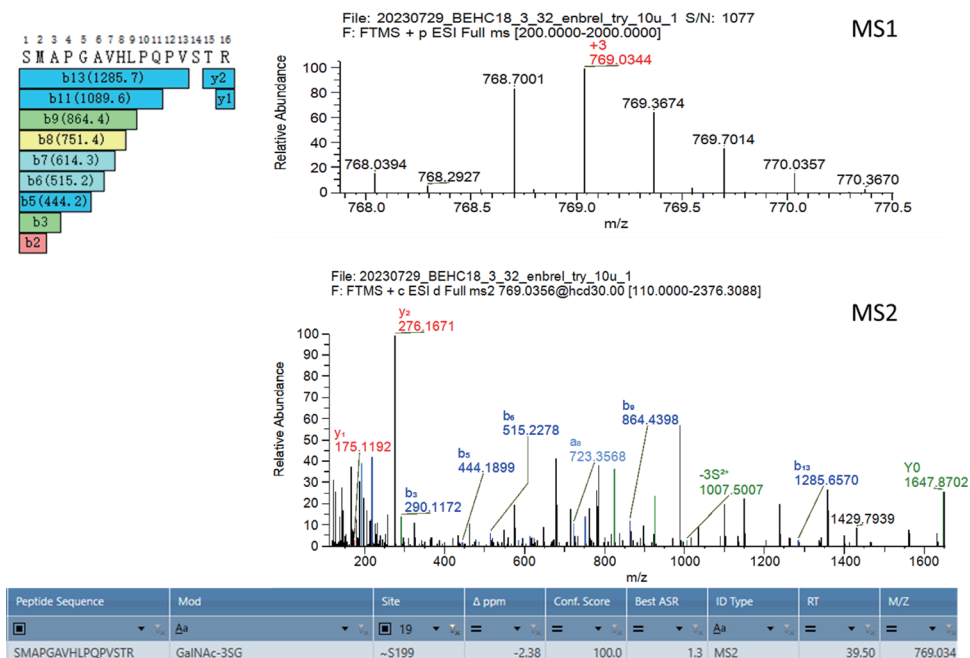


图2 trypsin所得酶解液中S199的鉴定

与 UHPLC 相比（色谱柱内径通常 2.1 mm），NanoLC 所用色谱柱内径更小（75 μ m 内径），可以使用更低的流速（300 nL/min），因此可以实现更好的雾化效率，进而提高检测的灵敏度，实现高糖基化比例的位点的多种糖链类型鉴定和低糖基化修饰比例位点的鉴定。例如，高比例糖基化修饰的位点 T8、S186、S245、S199 和 T200 可以鉴定到 GalNAc、GalNAc-3G、GalNAc-3SG 和 GalNAc-6S-3SG 等多种糖型。在常规源中难以检出的低 O 糖基化修饰位点，如 T243 也可以在 nano 分析中被鉴定。如图 3 所示，虽然该肽段响应很低，但 Orbitrap 的高灵敏度和高分辨率仍然可以实现该 O 糖基化位点的鉴定。首先，y23 离子可以排除 O 糖基化来自 T245。精确的一级质量数（-0.52 ppm）可以推测 O 糖基化发生在 T243。第二，尽管该肽段的一级和二级响应都较低，但仍然保持了高的质量准确度，所有二级碎片的质量偏差均小于 20 ppm。高的质量准确度可以有效区分噪音与信号，避免噪音带来的假阳性干扰。

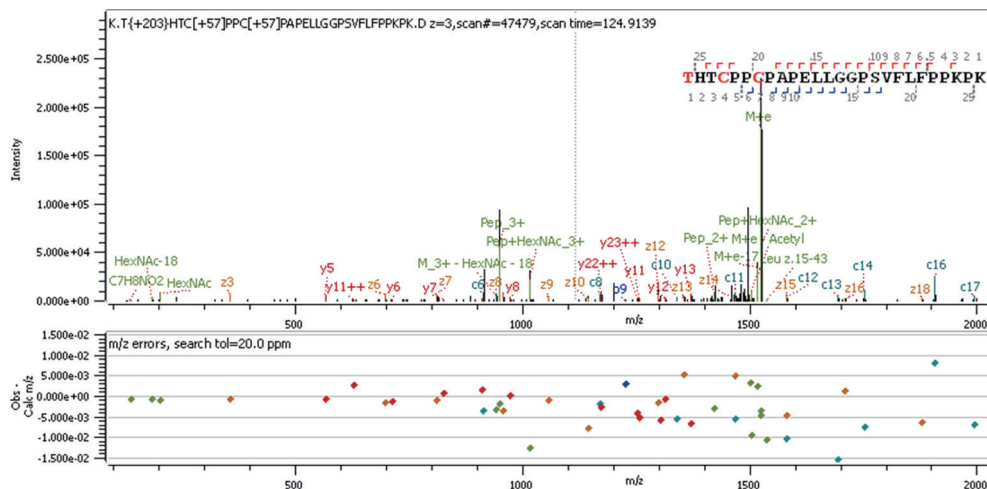


图3 nanoLC+ETHcD鉴定T243修饰

5.2 多酶联合酶切

虽然 nanoLC+ETHeD 具有强大的鉴定能力，但与常规生物药表征的质谱平台相比，需要额外增加设备。为了用尽可能少的成本并实现更好的鉴定，我们进一步优化前处理的方法，以此期待获得糖基化位点更单一的肽段，从而实现 O 糖基化位点的准确鉴定。首先，我们使用了脱除 N 糖后，多酶联合酶切的方法。如图 4 所示，在去掉 N 糖后，联合 trypsin、chymotrypsin 和 GluC 酶切后，成功实现 T184 位点的鉴定。值得注意的是 1) HCD 的高能碎裂模式，不仅可以得到糖链碎片，还可以得到含有糖链信息的二级碎片，只是这样的碎片响应较低，在普通的检测器（如 TOF）中无法被检测，但可以被 Orbitrap 的高灵敏度检测到。例如 y3 碎片是该糖基化位点鉴定的关键碎片，不仅具有较高的质量准确度（< 11 ppm 偏差），而且可以检测到同位素信号。2) 由于 BPF 标识二级碎片时卡值较高，低丰度的碎片可能会漏标记。可以尝试 pGlyco 检索 [3, 4]，然后 gLabel 标记。或者参考 BPF 给出的理论表，自行标记

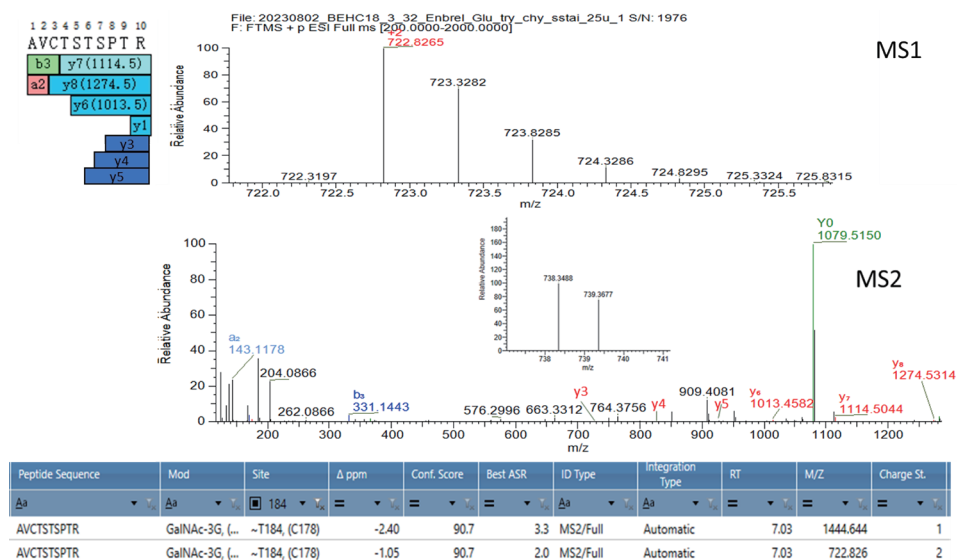


图4 联合酶切后UHPLC+HCD鉴定T184修饰

用同样的方式还可以鉴定到 T8、S186、S199、T205、T245、S199、T200 和 S226。以 S199 为例，同样可以鉴定到含有糖链的二级碎片（y6-y13）。BPF 软件不仅可以显示一级质量偏差、二级质量数、二级碎片覆盖度（ASR），还可以展示理论二级碎片检出情况和二级碎片的偏差（图 6）。

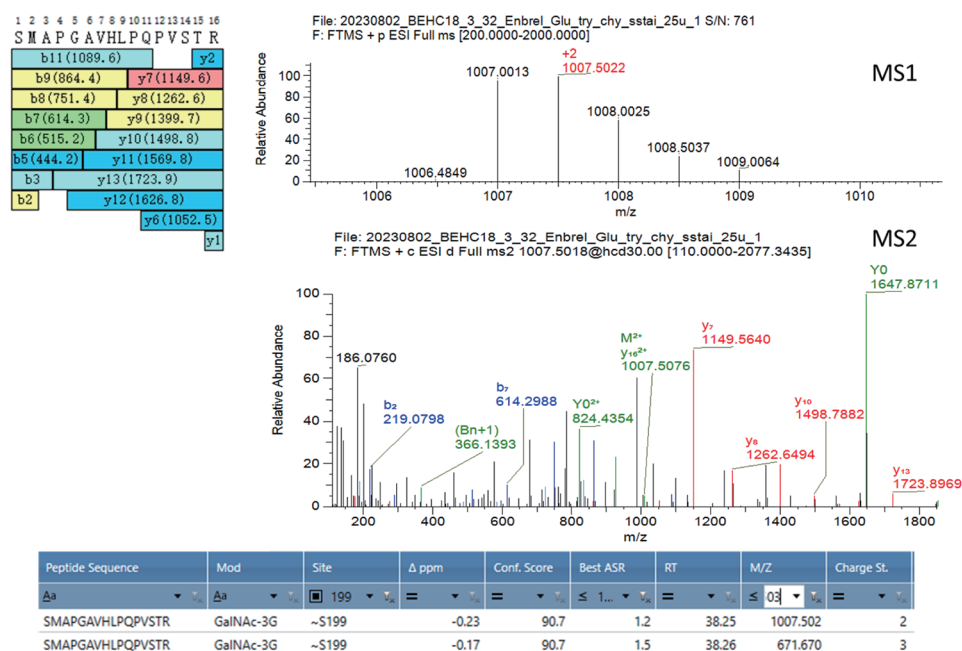


图5 联合酶切后UHPLC+HCD鉴定S199修饰

Ion Series									
Neutral Losses									
Internal Fragments									
Precursor Ions									
Glycan Ions									
#1	a+	a2+	b+	b2+	Sequence	y+	y2+	#2	
1					S			-3.19	16
2	-0.24		+0.22		M				15
3			+4.46		A				14
4					P	+6.24	+9.5		13
5			-6.94		G	+7.86			12
6	+2.96		-5.19		A	+15.39			11
7	+9.05		+3.36		V	+8.95			10
8	-1.58		+0.48		H	+1.41			9
9	+1.61		-0.24		L	+1.52			8
10					P	+0.49			7
11			+11		Q	+11.04			6
12					P				5
13					V				4
14					S-GalNAc-3G				3
15					T	+1.97			2
16					R	+2.89			1

图6 联合酶切后UHPLC+HCD鉴定S199修饰所得二级碎片偏差表

5.3 O-glycan- specific Protein Digestion 酶切

脱除 N 糖后的 trypsin+chymotrypsin+GluC 联合酶切，可以实现部分 O 糖基化位点的鉴定。但由于依那西普的 O 糖基化位点过于密集，这几种酶的联合酶切所得肽段包含多个 O 糖基化位点（T200-S217），为这些糖基化位点的鉴定带来极大的不确定性。本文尝试用一种以带有糖基化修饰的 S/T 为酶切位点的酶（O-glycan- specific Protein Digestion，商品名 OpeRATOR®）单独或者与 trypsin 联合处理依那西普，实现依那西普所有理论 O 糖基化位点的鉴定，鉴定结果如表 1 所示。典型谱图，如图 7-9 所示。如图 7，虽然该肽段仅包含两个氨基酸残基，保留极弱（死时间附近洗脱），受到其它强极性组分（如盐）的干扰，导致响应较弱。但 Orbitrap 高的一级灵敏度和二级灵敏度仍然能实现该肽段的鉴定，且一级偏差小于 3 ppm，二级偏差小于 5 ppm。除了 b、y 离子，在该肽段的二级碎片中还可以鉴定到糖链的大量碎片（图 7 绿色显示部分），为该糖基化位点的鉴定提供了可靠保障。

在这些 O 糖基化位点中 S212 和 T213 最难鉴定。以 S212 为例，如图 8 所示。pGlyco 可以标识出该肽段的所有碎片，结合该酶的酶切特点 1) N 端以糖链修饰的 S/T 开头；2) 该酶所得肽段在 HCD 碎裂中，距离糖基化位点较近的碎片没有含糖链的 b、y 碎片，距离该位点较远的肽键断裂时会保留碎片上的糖链碎片（图中 b3），推测糖基化位点为 S212。

仅仅使用 O-glycan- specific Protein Digestion 酶处理，S226 难以鉴定，因为在 S226 后没有其它 O 糖基化位点，会导致产生以 S226 为 N 端的蛋白碎片。将该碎片用 trypsin 再次酶切，可以得到质谱容易鉴定的多肽。如图 9 所示，采用联合酶切，可以实现 S226 的准确鉴定。

除理论 O 糖基化位点以外，该方法还可以鉴定到 T8、T89、T245 和 S186 等非理论 O 糖基化位点。三种方法均可检测到 O 糖基化位点 T8、T245 和 S186，为该糖蛋白中这几种 O 糖基化位点的鉴定提供了有力保障。

表1 O-glycan- specific Protein Digestion酶/与trypsin联合酶切实现依那西普理论O糖基化位点的鉴定

糖基化位点	肽段	糖型 ¹	m/z	ASR ²	价态	偏差(ppm)
T184	TR	GalNAc-3G	641.297	1.3	1	-2.32
S199	SMAPGAVHLPQPVS	GalNAc-3G	617.651	1.3	2	-2.15
T200	TRSQH	GalNAc	416.201	1.0	2	-1.58
T205	TQP	GalNAc-3G	710.308	1.5	1	-0.08
T208	TPEP	GalNAc-3G	808.344	1.0	1	-1.93
S212	STAP	GalNAc-3G	740.318	1.3	1	-1.9
T213	TAPSTSFLPMGP*	GalNAc-3G	842.405	1.7	2	-0.42
S216	STSFLPMGP	GalNAc-3G	707.836	1.4	2	-1.02
T217	TSFLPMGP	GalNAc-3G	664.320	1.3	2	-1.24
S226	SPPAEGSTGDEPK*	GalNAc-3G	546.239	1.0	3	-2.69

*O-glycan- specific Protein Digestion 与 trypsin 联合酶切，其余肽段为 O-glycan- specific Protein Digestion 酶单独酶切所得。

¹ 表格仅展示了典型糖核心，很多位点还有其它糖核心检出。之所以检测的是糖核心而不是糖型，是因为前处理中加了神经氨酸酶

² ASR= 实测肽键相关碎片数目 / 理论肽键碎片数目。

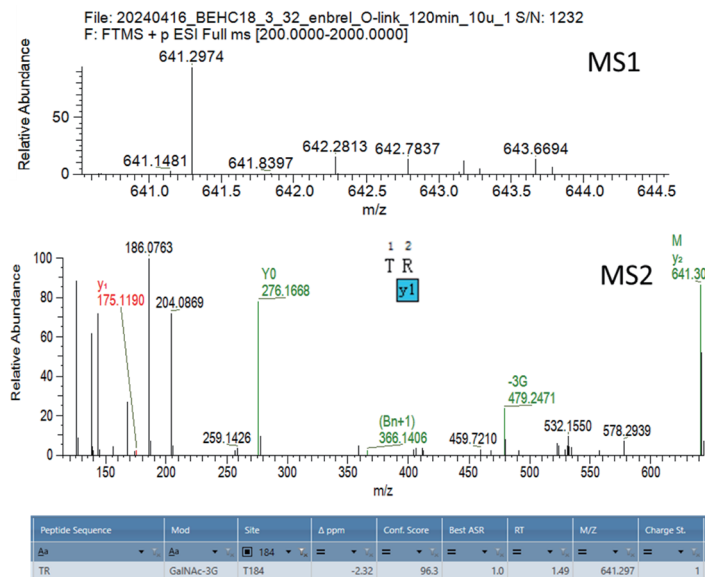


图7 O-glycan- specific Protein Digestion酶切鉴定T184

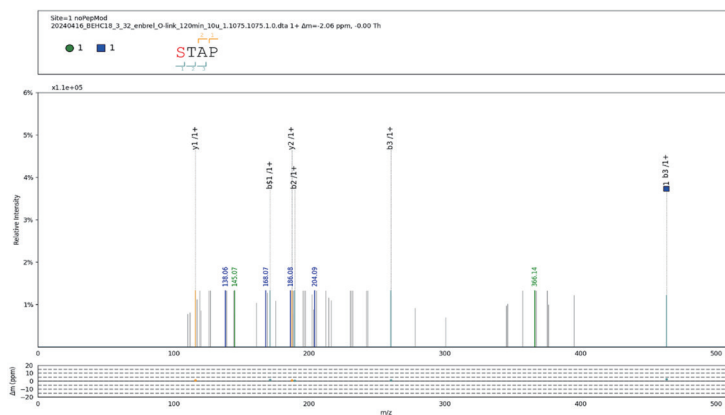


图8 O-glycan- specific Protein Digestion酶切鉴定S212

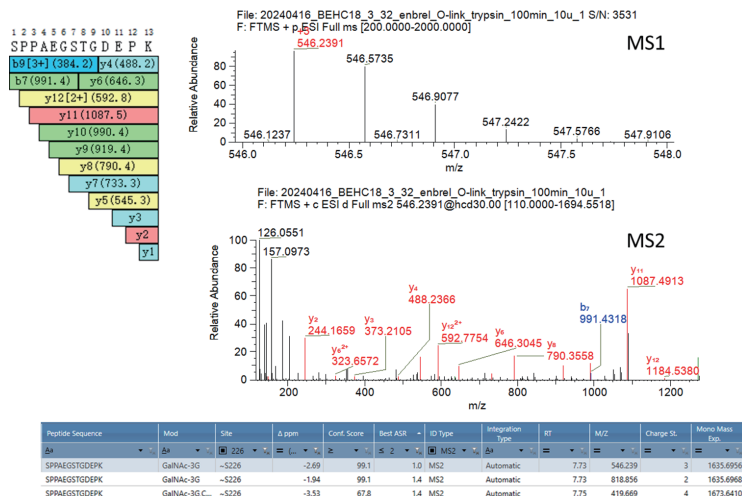


图9 O-glycan- specific Protein Digestion与trypsin联合酶切鉴定S226

结论

O糖基化修饰与重组蛋白药物的活性息息相关，但O糖基化位点的鉴定充满了挑战，尤其是O糖基化位点密集蛋白。本文以依那西普为例，介绍复杂O糖基化修饰蛋白的O糖基化位点鉴定策略。从常见的trypsin酶切，到联合酶切，再到特殊的O糖蛋白酶酶切。从结果来看，O糖蛋白酶如O-glycan-specific Protein Digestion的使用，可以极大的简化O糖基化位点鉴定的难题，且不需要特殊的碎裂方式。但如果较为全面的鉴定O糖基化位点的糖型和低修饰比例的O糖基化位点，建议使用nanoLC-EThcD鉴定，提高肽段的一级和二级响应，同时获取丰富的二级碎片，提高鉴定的可靠性。对于未报到的O糖基化位点，多方法的验证十分必要。在多种方法中均可以检出，可以提高此类糖基化位点鉴定的可靠性。在软件的使用方面，BPF软件可以实现绝大多数O糖肽的分析。但pGlyco软件和Byonic软件也是BPF软件较好的互补，提高原始数据的利用率。

参考文献

- [1] Y. Ohyama, K. Nakajima, M.B. Renfrow, J. Novak, K. Takahashi, Mass spectrometry for the identification and analysis of highly complex glycosylation of therapeutic or pathogenic proteins, *Expert Rev Proteomics*, 17 (2020) 275-296.
- [2] L. Cao, J.K. Diedrich, D.W. Kulp, M. Pauthner, L. He, S.-K.R. Park, D. Sok, C.Y. Su, C.M. Delahunty, S. Menis, R. Andrabi, J. Guenaga, E. Georgeson, M. Kubitz, Y. Adachi, D.R. Burton, W.R. Schief, J.R. Yates III, J.C. Paulson, Global site-specific N-glycosylation analysis of HIV envelope glycoprotein, *Nature Communications*, 8 (2017) 14954.
- [3] M.-Q. Liu, W.-F. Zeng, P. Fang, W.-Q. Cao, C. Liu, G.-Q. Yan, Y. Zhang, C. Peng, J.-Q. Wu, X.-J. Zhang, H.-J. Tu, H. Chi, R.-X. Sun, Y. Cao, M.-Q. Dong, B.-Y. Jiang, J.-M. Huang, H.-L. Shen, C.C.L. Wong, S.-M. He, P.-Y. Yang, pGlyco 2.0 enables precision N-glycoproteomics with comprehensive quality control and one-step mass spectrometry for intact glycopeptide identification, *Nature Communications*, 8 (2017) 438.
- [4] W.-F. Zeng, W.-Q. Cao, M.-Q. Liu, S.-M. He, P.-Y. Yang, Precise, fast and comprehensive analysis of intact glycopeptides and modified glycans with pGlyco3, *Nature Methods*, 18 (2021) 1515-1523.



赛默飞
官方微信

热线 800 810 5118
电话 400 650 5118
www.thermofisher.com

thermo scientific