

液相色谱耦联 Q Exactive 台式轨道阱质谱仪 分析利妥昔单克隆抗体

Martin Samonig^{1,2}, Christian Huber^{1,2} and Kai Scheffler^{2,3}

赛默飞世尔科技（中国）有限公司

关键词

单克隆抗体，完整蛋白质的质量测定，序列确定，蛋白质去卷积，自上而下的测序

目的

优化液相色谱质谱工作流程，采用整体柱在线耦联 Thermo Scientific Q Exactive 台式轨道阱质谱仪对单克隆抗体进行分析和表征。

前言

单克隆抗体 (mAb) 是增长需求最快的医药产品之一。他们在多种病症如癌症、感染性疾病、过敏症、炎症和自身免疫疾病的治疗中发挥着主要作用。由于单克隆抗体可以表现出显著的异质性，作为治疗产品，新的 mAb 要获得批准需要进行大量的分析表征。质谱法已经成为表征单克隆抗体的重要工具，能测定完整蛋白及分离的轻链和重链的分子量、解析糖基化和糖链的结构，确证正确的氨基酸序列，鉴定杂质，如生产过程中固有的宿主细胞蛋白质 (HCP)。

利妥昔单抗是一个拮抗蛋白质 CD20 的重组单克隆抗体，其在美国的商品名为 Rituxan® (Biogen Idec 公司 / 基因泰克公司)，在欧洲的商品名为 MabThera® (罗氏)。它是第一代肿瘤免疫治疗药物之一。利妥昔单抗分别在 1997 年被美国食品和药物监督管理局和 1998 年由欧盟委员会批准用于恶性淋巴瘤的治疗。这个抗体的可变区域作用靶点为细胞表面分子 CD20，这些分子存在于一些非霍奇金淋巴瘤中。

在这个应用报告中，结合使用自上而下和自下而上的方法，不管是分析完整和还原的利妥昔单抗，还是对其进行氨基酸序列分析，Q Exactive Orbitrap 质谱仪都具有优异的表现。



此外，评价了两套整体柱色谱系统在线耦合质谱仪的灵敏度。数据表明 Q Exactive 系统具有优异的分辨率和质量准确度，是一个具有高置信度的筛选工具，可快速精确地进行生物制药产品的开发和表征。

实验部分

样品制备

本研究所有实验均采用商品化的单克隆抗体利妥昔。利妥昔是一种无菌，透明，无色，不含防腐剂的静脉滴注用浓缩液。该产品浓度为 10 mg/mL，通常配制成 7.35 mg/mL 的柠檬酸钠缓冲注射液，其中含 0.7 mg/mL 聚山梨醇酯 80、9.0 mg/mL 氯化钠和无菌水，采用氢氧化钠或盐酸水溶液调 pH 值至 6.5。

由于利妥昔单抗中含聚山梨醇酯 80，在进行 LC / MS 分析之前需对样品进行透析。采用截留分子量为 3.5 kDa(MWCO) 的 Thermo Scientific™ Slide-A-Lyzer™ 透析盒透析。在 4°C 条件下，1 mL 利妥昔单抗在 2 L 20% 乙腈水溶液中透析 48 h。

对利妥昔单抗的轻链和重链进行分析时，样品与 5 mM 三羧甲基磷酸 (TCEP) 在 60°C 条件下孵育 30 min，还原样品中的二硫键。

为了对酶切后的 mAb 进行自下而上的分析，样品还原后，继续在室温暗处与 20 mM 碘乙酰胺 (IAA) 进行烷基化反应 30 min。样品采用 Thermo Scientific™ Pierce™ C18 枪头进行纯化，采用 Thermo Scientific™ SpeedVacc™ 旋转浓缩仪进行干燥，然后溶解在 0.5 M 三乙基碳酸氢铵缓冲液 (TEAB) 中。在 0 h 和 1.5 h 先后两次，按照 1:15 (w/w) 的比例加入测序级修饰胰蛋白酶 (Promega)，在 37 °C 条件下酶切 2.5 h。最后，加入适量三氟乙酸 (TFA)，调 pH 值至约为 3。

所有样品放在含玻璃内插管的自动进样器小瓶中 (0.1 ml 微量内插管，透明玻璃，VWR)。

液相色谱

实验使用的色谱柱为按照文献制备的 160 × 0.20 mm 内径聚 (苯乙烯 - 二乙烯基苯) (PS-DVB) 共聚物毛细管柱【1】和 Thermo Scientific™ PepSwift™ 250 × 0.20 mm 内径的 PS-DVB 毛细管柱。蛋白质的分离使用 Thermo Scientific™ Dionex™ UltiMate™ 3000 RSLC nano 系统，该系统配置 3 nL z-型微量检测池。分离过程中，柱温：55°C，流速：1 μL/min，色谱梯度：乙腈 (ACN) 溶液 (含 0.05% 三氟乙酸 TFA) 的浓度从 20% 升高至 60%，运行时间：10 min。对于胰蛋白酶的酶切混合物，色谱梯度：B 相浓度从 0% 升高至 50%，运行时间：30 min。对于还原型抗体，色谱梯度：B 相浓度从 35% 升高至 45%，运行时间：15 min。

采用 Thermo Scientific™ ProSwift™ RP-10R 50 mm × 1.0 mm 内径的整体柱在更高流速分离蛋白质，使用的仪器为 Thermo Scientific™ Dionex™ UltiMate™ 3000 RSLC nano 系统，该系统配置 45 nL 检测池。分离过程中，柱温设置为 55°C，流速为 60 μL/min。色谱梯度：B 相浓度从 26% 升高至 80%，运行时间：20 min。对于还原型抗体，分离轻链和重链的色谱

梯度：B 相浓度从 26% 升高至 56%，运行时间：20 min。

在上述梯度条件下，PepSwift 250 mm × 0.2 mm 内径的整体柱压力在 190 bar 到 260 bar 之间，ProSwift RP-10R 50 mm × 1 mm 内径的整体柱压力在 120 bar 到 180 bar 之间。

在本研究的所有实验中，流动相 A 是 0.05% 三氟乙酸水溶液，流动相 B 是 0.05% 三氟乙酸的乙腈溶液。色谱梯度见表 1 和表 2。

表 1. 本实验所用的液相色谱梯度，

色谱柱：PepSwift 250 mm × 0.2 mm 内径，流速：1 μL/min

时间 (min)	完整单克隆抗体 [%B]	时间 (min)	还原型单克隆抗体 [%B]	时间 (min)	单克隆抗体酶切混合物 [%B]
0.0	20	0.0	35	0.0	0
10.0	60	15.0	45	30.0	50
10.1	85	15.1	85	30.1	85
16.0	85	21.0	85	40.0	85
16.1	20	21.1	35	40.1	0
30.0	20	30.0	35	50.0	0

表 2. 本实验所用的液相色谱梯度，

色谱柱：ProSwift RP-10R 50 mm × 1 mm 内径，流速：60 μL/min

时间 (min)	完整单克隆抗体 [%B]	时间 (min)	还原型单克隆抗体 [%B]
0	26	0	26
15	80	15	56
20	80	15.1	80
20.1	26	20	80
30	26	20.1	26
		30	26

质谱

本研究所有实验均采用 Q Exactive 台式静电场轨道阱质谱仪。采用 ProSwift RP-10R 50 mm × 1 mm 内径色谱柱进行实验时配合使用 Thermo Scientific™ IonMax™ 加热电喷雾离子源，喷雾电压为 4 KV，鞘气和辅助气流速分别为 15 和 5 个单位。

其他实验均采用 Thermo Scientific™ NanoFlex™ 离子源进行，离子源上配置 15 cm 的 PicoTip® 喷雾针 (New Objective, Woburn, USA; 20 μm 内径，360 μm 外径，10 μm 针尖)，流速为 1 μL/min，电压为 1.5 KV。

详细方法设置见表 3。

表 3. 本研究所有实验用到的质谱参数

	完整单克隆抗体	还原型单克隆抗体	自上而下 AIF	5-plex 目标离子二级质谱	抗体酶切物
方法类型	全扫描	全扫描 (分 2 段)	全扫描 -AIF	目标离子二级质谱	全扫描 - 前 10 个高能碎裂解
总运行时间	30 min	0–15.8/15.8–30 min	25 min	25 min	40 min
扫描范围 m/z	1800–5000	800–3500/700–2500	300–2500	固定首个质量数 300	350–2000
分辨率 (一级 / 二级质谱)	17,500/x	140,000/17,500	70,000	n.a./70,000	70,000/17,500
全扫描 AGC	3.00×10^6	3.00×10^6	3.00×10^6	5.00×10^5	3.00×10^6 (MS)/ 1.00×10^5 (MS ²)
最大注入时间	150 ms	150 ms/200 ms	150 ms	150 ms	100 ms/100 ms
隔离窗口	n.a.	n.a.	n.a.	10 Th	2 Th
微扫描	10	5	5	5	1
毛细管温度	275 °C				
S-lens RF 值	80	80	50	50	50
SID [eV]	80	0/60	n.a.	LC O/HC 20	n.a.
NCE [%]	n.a.	n.a.	10–30	10–30	25

源内 CID

源内 CID (SID) 是一个 DC 补偿参数 (0–100 eV)，通常加在源内 DC 补偿电压上。源内 DC 补偿电压由三个电压组成：毛细管 DC, S-lens DC 和 S-lens 的出口透镜。通过设置源内 CID 参数来应用 DC 补偿电压，应用该电压会导致多极杆进样池内的分析物与仪器离子源区域残留的气体分子发生碰撞。

全离子裂解

全离子裂解 (AIF) 是裂解模式的一种，源内产生的所有离子由质谱仪的离子光学组件引导，在 C-trap 内累积，然后一起送入高能裂解池进行裂解。在这种情况下，四极杆的设置并不是选择某一特定母离子，而是采用所有离子均通过的模式 (RF-only)。在分析完整蛋白时，这是一种非常有用的方法，因为不同电荷态往往有不同的裂解模式，因此不容易预测哪一种模式最好。

数据分析

采用 Thermo Scientific™ Protein Deconvolution™ 软件 2.0 版对一级质谱图进行去卷积。分辨率为 17500 时得到的完整抗体和完整重链质谱图，用 ReSpect™ 算法去卷积。分辨率为 140000 时获得的完整轻链质谱图和分辨率为 70000 时得到的自上而下的谱图用 Xtract 算法去卷积。为了鉴定完整抗体和还原后的完整重链的糖型，我们将已知各种常见糖型组合的分子量与测定分子量进行比较。

采用 Thermo Scientific™ Qual Browser™ 仪器中的 Xtract 算法对自上而下的 HCD 和 AIF 质谱图去卷积。使用 Thermo Scientific™ ProSightPC™ 软件 3.0 版本的单个蛋白质模式对碎片离子进行归属，碎片离子质量误差设为 5ppm。

采用 Thermo Scientific™ Proteome Discoverer™ 软件 1.4 版本

SEQUEST® 算法对胰蛋白酶酶切物的数据进行检索分析。

数据库由三种蛋白组成：胰蛋白酶、轻链和重链的两个异构体（在 219 位为 Ala 或者 Val）。母离子质量误差设为 10 ppm，碎片离子的质量误差设为 20 mmu。考虑四个可变修饰：烷基化（半胱氨酸）、氧化（甲硫氨酸）、脱酰胺（天冬酰胺，谷氨酰胺）、谷氨酰胺转换为焦谷氨酸，和 N, N- 二甲基化（赖氨酸）（仅与胰蛋白酶自酶切产物鉴定相关）。

结果与讨论

利妥昔单抗是拮抗 CD20 蛋白的 IgG1 类单克隆抗体，由两条轻链 (213 个氨基酸) 和两条重链 (451 个氨基酸) 组成。轻链和重链通过 12 个链内和 4 个链间的二硫键连接（图 1）。在该抗体两个重链的 Asn³⁰¹ 处有糖链结构。连接的糖链组成和长度是变化多端的，从而导致该分子具有微观异质性。作为一种生物药，保持不同糖结构的种类和相对丰度是发挥抗体疗效必不可少的。附着在抗体上的常见糖的命名如图 2 所示。

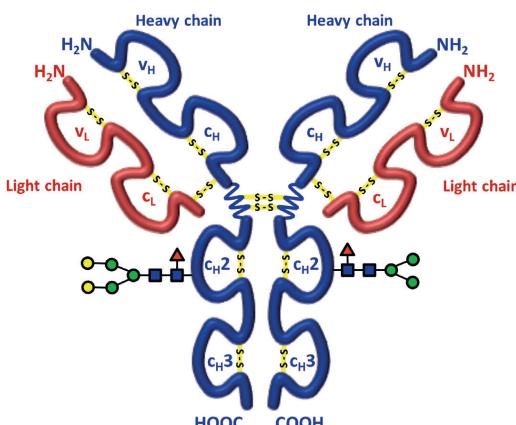


图 1. 人源 IgG1 属的单克隆抗体利妥昔的分子结构示意图

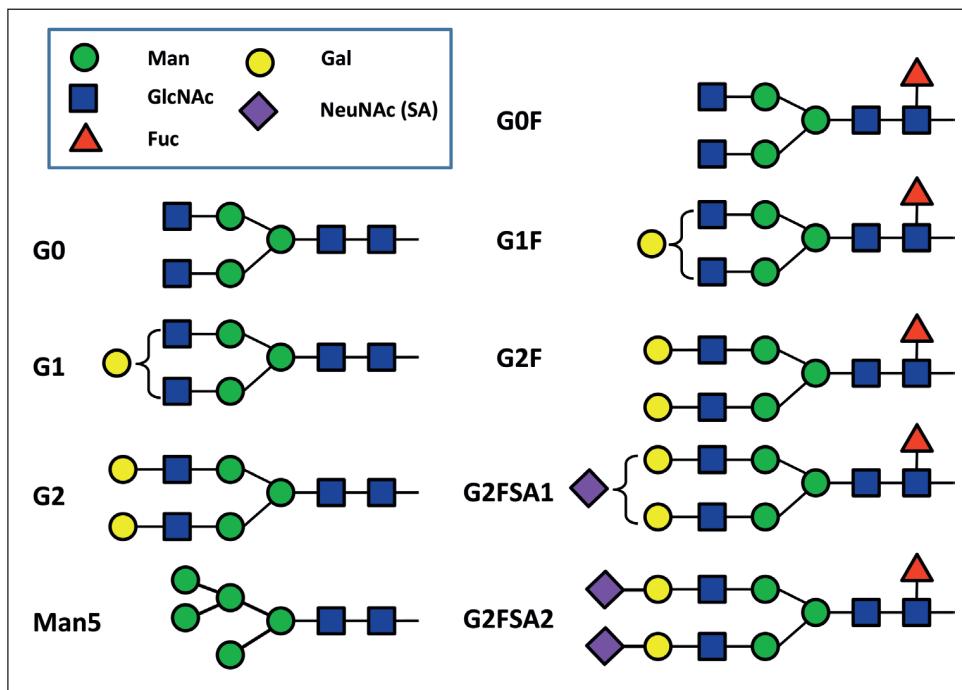
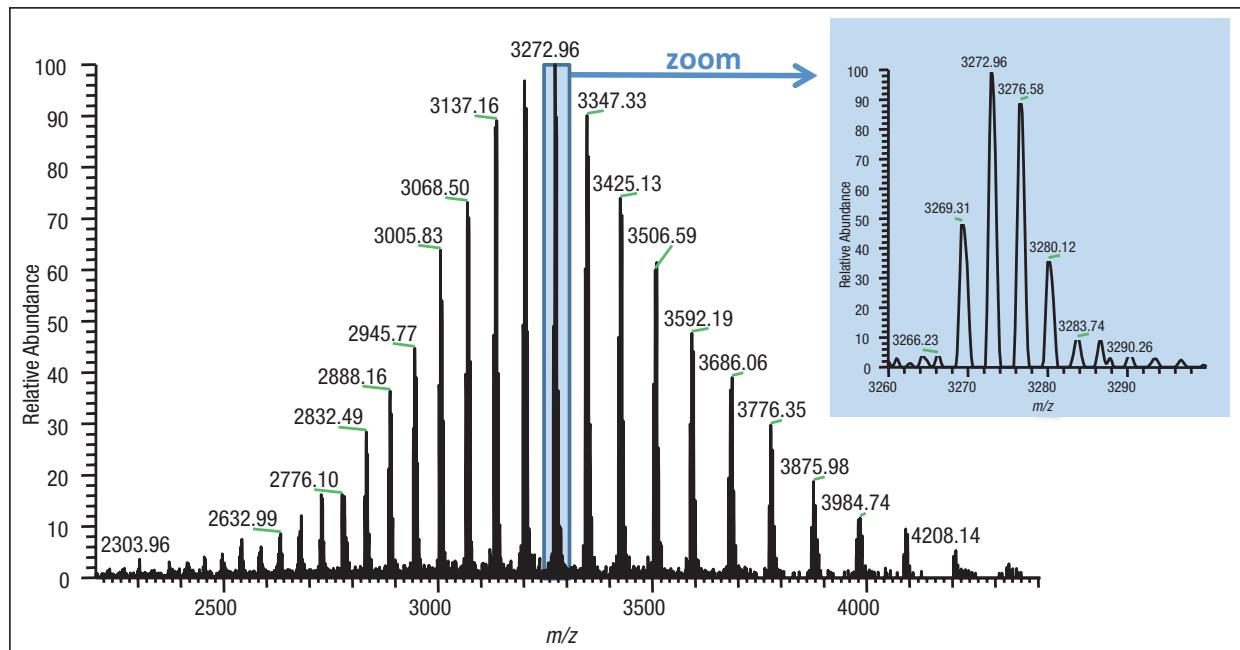


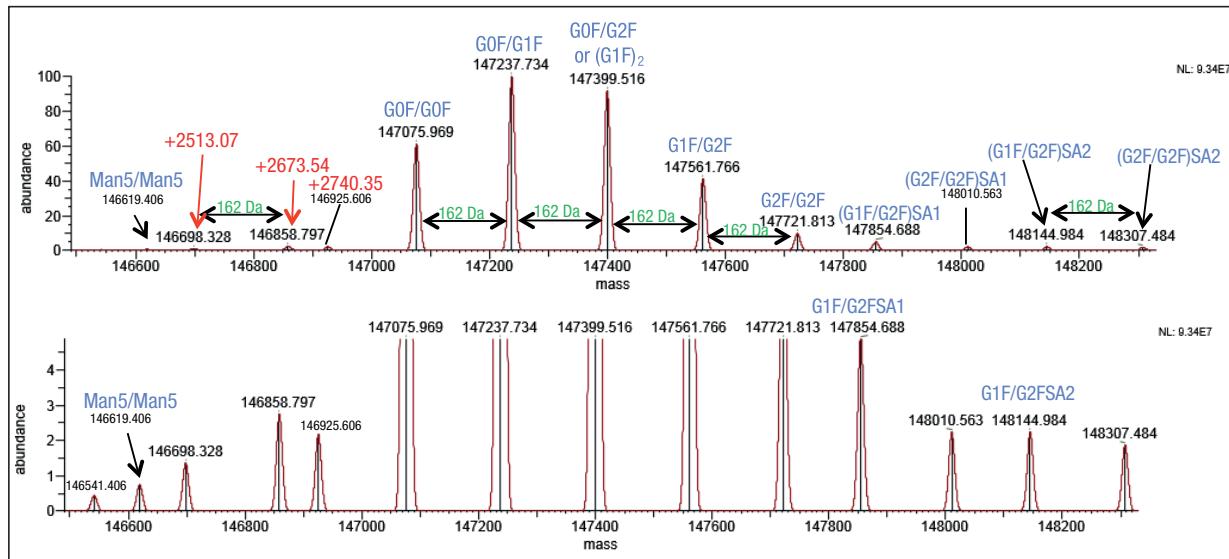
图 2. 抗体上常见的糖链结构的命名

图 3 显示的是用 $250 \times 0.2 \text{ mm}$ 内径整体柱分析 20 ng 利妥昔抗体得到的一级质谱图。在 m/z $1800\text{--}5000$ 范围内采集的质谱图显示了大分子蛋白质的典型电荷分布。插图显示的是 m/z 3269 时，放大的最大丰度电荷态 ($z=+45$)，其很好地描绘了完整抗体 4 个丰度最高的糖型。

通过将完整一级质谱图去卷积，获得了四个最强丰度糖型的完整质量数和一系列较低丰度糖型的质量数（如图 4 所示）。根据蛋白序列计算得到的理论分子量并结合图 2 中各种预测糖链的结构进行峰归属。

图 3. 利妥昔抗体的单张全扫描质谱图 (10 个 microscan)，用 $250 \times 0.2 \text{ mm}$ 内径色谱柱分析 10 ng 利妥昔抗体。

插图显示的是放大的最强丰度电荷态 ($z=+45$)，观察到的峰型代表该分子的不同糖型。



	GOF/GOF	GOF/G1F	GOF/G2F 或 (G1F) ₂	G1F/G2F	G2F/G2F
测量值	147,075.969	147,237.734	147,399.516	147,561.766	147,721.813
理论值	147,074.985	147,237.126	147,399.267	147,561.408	147,723.549
Δ ppm	-6.7	-4.1	-1.7	-2.0	11.8

图 4. 利妥昔抗体去卷积的质谱图和不同糖型归属(上图)，5个最强丰度糖型的理论质量数和测定质量数的对比(表)

为了获得完整抗体的一级谱图，我们对源内 CID 的设置参数进行了优化。该参数对获得高质量的谱图至关重要。

对于大多数蛋白质来说，源内 CID 一般在 25%-90% 较为适合，对于这个样品来说，最优设置为 80% SID。

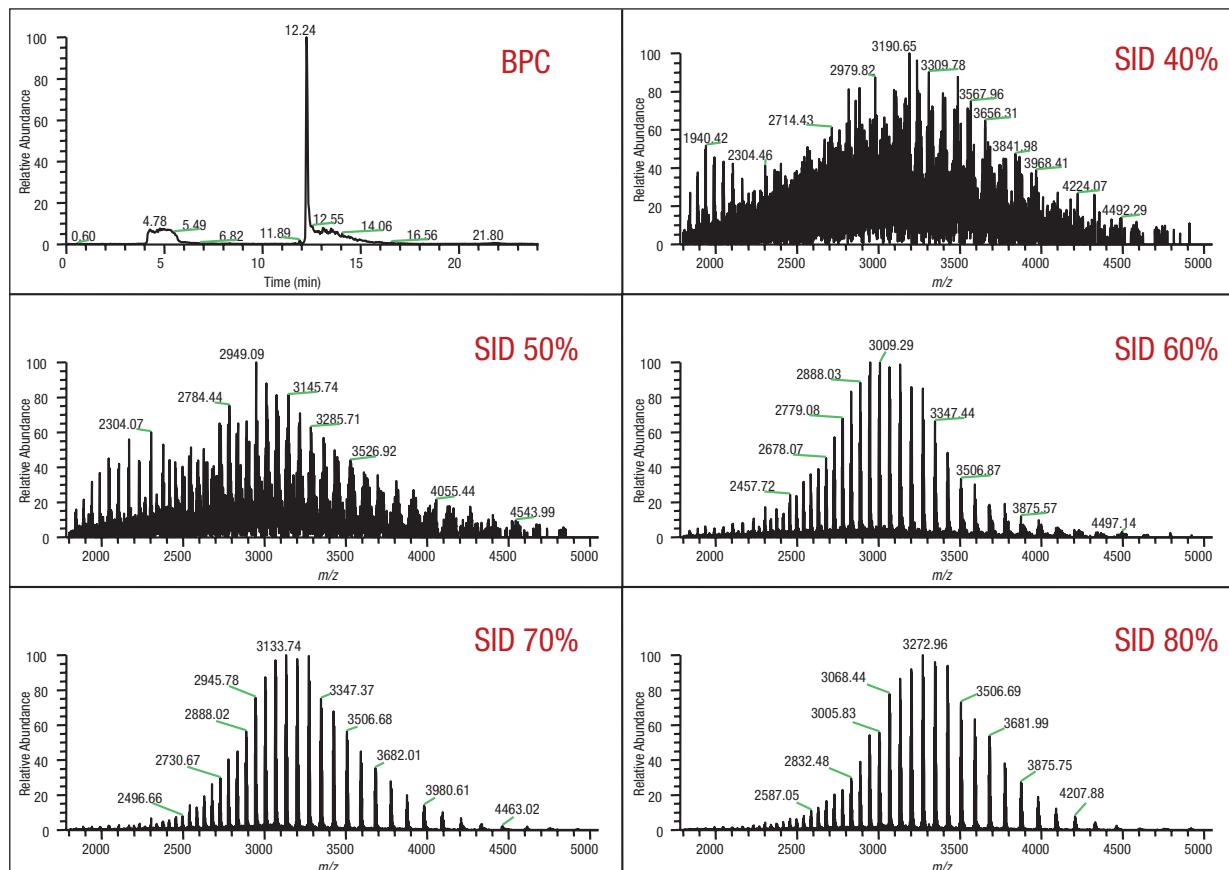


图 5. 在不同 SID 设置条件下，1 ng 完整利妥昔抗体的一级质谱图

表 4 显示的是轻链、非糖基化重链和完整抗体的计算质量数，分别从轻链的 213 和重链的 451 位氨基酸开始逐步计算。两个蛋白序列都包含的 N- 末端谷氨酰胺应会被修饰为焦谷氨酸，这将导致质量数丢失 17.0265 Da。此外，重链中 C- 末端赖氨酸有可能被切掉，导致质量数丢失 128.09497 Da。在构成完整抗体时，形成 16 个二硫键会丢失

32 个质子。两条重链上携带的糖链导致其质量数将会增加 1217.1-2352.1 Da。两条重链上可以携带不同的糖链，从而形成混合组成物，例如 G01/G2F。各个糖链的化学组成和质量数列在表 5 中。用于计算表 4 和表 5 中分子量的元素的单同位素和平均原子质量列于表 6 中。

表 4. 分别对带有不同修饰及不同糖型的利妥昔单克隆抗体的轻链和重链进行化学组成分析
以及单同位素分子量和平均分子量的逐步计算。图 4、图 6 和图 7 的蓝色格子显示了检测到的分子量。

元素组成	碳	氢	氮	氧	硫	分子量 (单同位素分子量)	分子量 (平均分子量)
轻链, 全序列 氨基酸 1-213	1016	1577	273	328	6	23,042.34369	23,056.5
N-terminal pyro Glutamic acid	1016	1574	272	328	6	23,025.31714	23,039.4
N-terminal pyro Glutamic acid, 2 intrachain S-S bonds	1016	1570	272	328	6	23,021.28584	23,035.4
2 × LC (N-term. pyroGlu)	2032	3148	544	656	12	46,050.63428	46,078.9
2 × LC (N-term. pyroGlu, 2 intrachain S-S bonds each)	2032	3140	544	656	12	46,042.57168	46,070.8
重链, 全序列 氨基酸 1-451	2197	3389	577	676	16	49,183.40813	49,214.0
N-terminal pyro Glutamic acid	2197	3386	576	676	16	49,166.38158	49,197.0
minus C-term. K (aa 1-450)	2191	3374	574	675	16	49,038.28661	49,068.8
minus 4 intrachain S-S bonds	2191	3366	574	675	16	49,030.22401	49,060.8
HC-G0F (pyro-Glu, - K, fully reduced)	2247	3466	578	714	16	50,482.82048	50,514.2
HC-G1F (pyro-Glu, - K, fully reduced)	2253	3476	578	719	16	50,644.87330	50,676.3
HC-G2F (pyro-Glu, - K, fully reduced)	2259	3486	578	724	16	50,806.92613	50,838.5
HC minus 4 intrachain S-S bonds + G0F	2247	3458	578	714	16	50,474.75788	50,506.1
2 × HC (pyroGlu, - K)	4382	6748	1148	1350	32	98,076.57323	98,137.7
2 × HC (pyroGlu, - K, 4 intrachain S-S bonds each)	4382	6732	1148	1350	32	98,060.44803	98,121.6
Man5 (HexNAc)2 (Hex)5	46	76	2	35	0	1216.42286	1217.1
G0 (HexNAc)4 (Hex)3	50	82	4	35	0	1298.47596	1299.2
G0F (HexNAc)4 (Hex)3 Fuc	56	92	4	39	0	1444.53387	1445.3
G1 (HexNAc)4 (Hex)4	56	92	4	40	0	1460.52878	1461.3
G1F (HexNAc)4 (Hex)4 Fuc	62	102	4	44	0	1606.58669	1607.5
G2 (HexNAc)4 (Hex)5	62	102	4	45	0	1622.58161	1623.5
G2F (HexNAc)4 (Hex)5 Fuc	68	112	4	49	0	1768.63951	1769.6
G1FSA (HexNAc)4 (Hex)4 Fuc SA	73	119	5	52	0	1897.68211	1898.7
G1FSA2 (HexNAc)4 (Hex)4 Fuc (SA)2	84	136	6	60	0	2188.77752	2190.0
G2FSA (HexNAc)4 (Hex)5 Fuc SA	79	129	5	57	0	2059.73493	2060.9
G2FSA2 (HexNAc)4 (Hex)5 Fuc (SA)2	90	146	6	65	0	2350.83035	2352.1
Man5/Man5 (HexNAc)4 (Hex)10	92	152	4	70	0	2432.84572	2434.2
G0F/G0F (HexNAc)8 (Hex)6 (Fuc)2	112	184	8	78	0	2889.06774	2890.7
G0F/G1F (HexNAc)8 (Hex)7 (Fuc)2	118	194	8	83	0	3051.12056	3052.8
G1F/G1F (HexNAc)8 (Hex)8 (Fuc)2	124	204	8	88	0	3213.17338	3215.0
G1F/G2F (HexNAc)8 (Hex)9 (Fuc)2	130	214	8	93	0	3375.22621	3377.1
G2F/G2F (HexNAc)8 (Hex)10 (Fuc)2	136	224	8	98	0	3537.27903	3539.2
G1F/G2FSA (HexNAc)8 (Hex)9 (Fuc)2 SA	141	231	9	101	0	3666.32162	3668.3
G1F/G2FSA2 (HexNAc)8 (Hex)9 (Fuc)2 (SA)2	152	248	10	109	0	3957.41704	3959.6

元素组成	碳	氢	氮	氧	硫	分子量 (单同位素分子量)	分子量 (平均分子量)
G2F/G2FSA (HexNAc)8 (Hex)10 (Fuc)2 SA	147	241	9	106	0	3828.37445	3830.5
G2F/G2FSA2 (HexNAc)8 (Hex)10 (Fuc)2 (SA)2	158	258	10	114	0	4119.46986	4121.7
Sum 2 × HC + 2 × LC (4 × pyroGlu, -2K)	6414	9896	1692	2006	44	144,127.20750	144,216.6
minus 32 S-S bond protons	6414	9864	1692	2006	44	144,094.95710	144,184.3
2HC + 2LC - 16 S-S bonds + (Man5)2	6506	10016	1696	2076	44	146,527.80282	146,618.5
2HC + 2LC - 16 S-S bonds + (GOF)2	6526	10048	1700	2084	44	146,984.02484	147,075.0
2HC + 2LC - 16 S-S bonds + GOF/G1F	6532	10058	1700	2089	44	147,146.07766	147,237.1
2HC + 2LC - 16 S-S bonds + GOF/G2F or (G1F)2	6538	10068	1700	2094	44	147,308.13049	147,399.3
2HC + 2LC - 16 S-S bonds + G1F/G2F	6544	10078	1700	2099	44	147,470.18331	147,561.4
2HC + 2LC - 16 S-S bonds + G2F/G2F	6550	10088	1700	2104	44	147,632.23613	147,723.5
2HC + 2LC - 16 S-S bonds + G1F/G2F SA	6555	10095	1701	2107	44	147,761.27872	147,852.7
2HC + 2LC - 16 S-S bonds + G1F/G2F (SA)2	6566	10112	1702	2115	44	148,052.37414	148,143.9
2HC + 2LC - 16 S-S bonds + G2F/G2F SA	6561	10105	1701	2112	44	148,923.33155	148,014.8
2HC + 2LC - 16 S-S bonds + G2F/G2F (SA)2	6572	10122	1702	2120	44	148,214.42696	148,306.1

表 5. 单糖的化学组成和分子量

	分子式	单同位素分子量	平均分子量	碳	氧	氮	氢
唾液酸	C ₁₁ O ₉ NH ₁₇	291.09542	291.3	11	8	1	17
半乳糖	C ₆ O ₅ H ₁₀	162.05282	162.1	6	5	0	10
N-乙酰葡萄糖胺	C ₈ O ₅ NH ₁₃	203.07937	203.2	8	5	1	13
甘露糖	C ₆ O ₅ H ₁₀	162.05282	162.1	6	5	0	10
岩藻糖	C ₆ O ₄ H ₁₀	146.05791	146.1	6	4	0	10

表 6. 用于计算表 4 和表 5 分子量的元素的单同位素原子量和平均原子量

	单同位素质量	平均质量
碳	12.000000	12.01074
氢	1.00782503	1.00794
氮	14.0030740	14.00674
氧	15.9949146	15.99940
硫	31.9720707	32.06608

最初的计算是根据发表在 DrugBank 数据库【2】的序列进行的，然而计算结果与我们实验测得的质量数并不匹配。通过将 DrugBank 中的序列与之前文献发表过的序列【3】进行比对，发现位于重链保守区域 219 位上的一个氨基酸不同，是丙氨酸而不是缬氨酸。使用 219 位为丙氨酸的序列的计算结果可以与完整蛋白质量数测定的结果以及先前报道的结果完全吻合【4】。为了进一步确认这点，我们进行了一系列的补充实验。

将抗体还原后（无烷基化），采用设置不同分辨率的方法分析已分离的轻链和重链，来观察在此分子量下能否实现同位素解析。由于轻链分子量较小，有可能得到同位素解析谱图，而重链无法得到，因为重链质量是轻链

质量的两倍。若要采用不同的分辨率设置，方法设置分为两个阶段（从 0-15.8 min 分辨率设为 140 K，从 15.8-30 min 分辨率设为 17.5 K）。通过优化梯度，实现轻链和重链的色谱峰完全分离。

在这项研究中，我们评价了两种整体柱（PepSwift 250 mm × 0.2 mm 内径和 ProSwift RP-10R 50 mm × 0.1 mm 内径），轻链和重链两峰分离的保留时间之差均可大于 2 min（图 6）。分别对轻链和重链的质谱图（图 6b 和 6c）去卷积。轻链的同位素解析质谱图用 Xtract 算法去卷积，得到了单一同位素分子量 23,025.3758 Da，与计算分子量只差 2.5 ppm。重链用 ReSpect 算法去卷积，产生三个峰，每一个峰代表一种主要的糖型，GOF，G1F 和 G2F（图 7）。

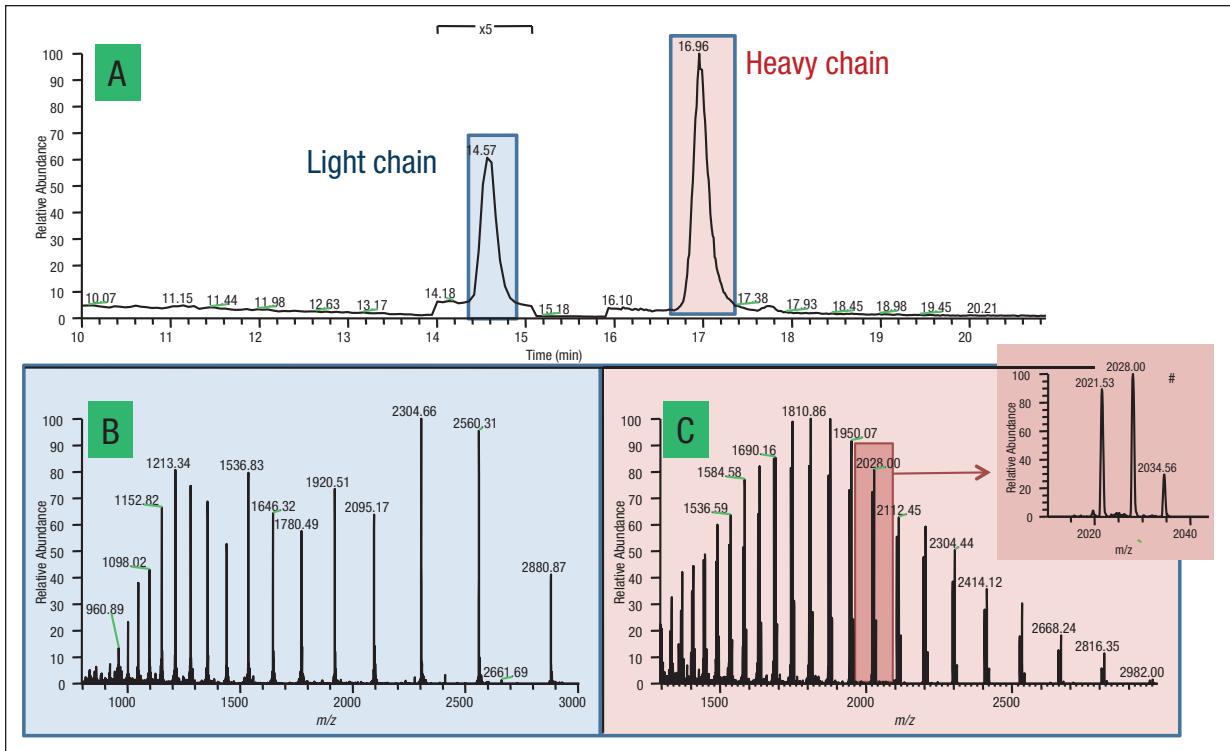


图 6. 还原型利妥昔抗体的色谱图 (A) , 轻链的一级全扫描质谱图 (B) 和重链的一级全扫描质谱图 (C) 。
C 中的插图显示的是电荷态 $z=+25$ 的放大图, 三个峰代表三种不同的糖型。

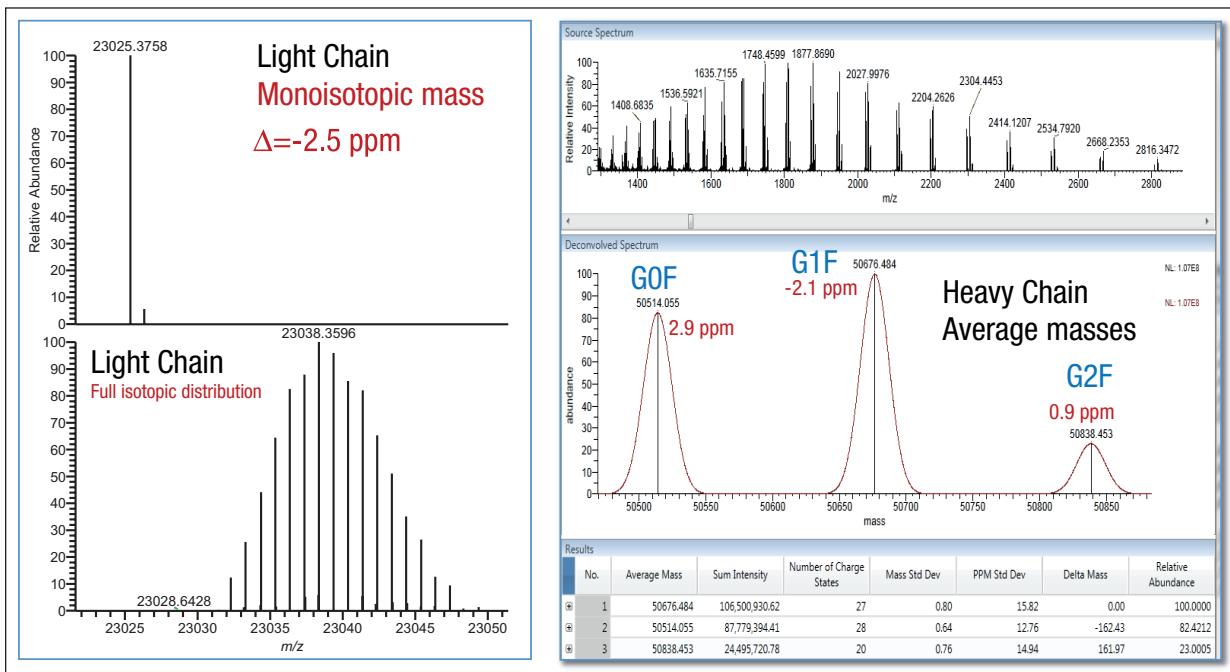


图 7. 轻链和重链的去卷积结果。在分辨率 140,000 全扫描模式下获得轻链的质谱图, 通过 Xtract 算法去卷积, 得到准确的单同位素分子量和完整的同位素分布图 (左图)。在 17,500 的分辨率下测得重链的质谱图, 通过 ReSpect 算法去卷积, 得到了平均分子量 (右图)。

为了评价使用 $250 \times 0.2\text{ mm}$ PepSwift 整体柱时仪器的检测限, 采集了一系列进样量的 LC/MS 样品。将 50pg-20ng 的完整抗体从最低浓度开始注入色谱柱 (图 8)。在序列开始之前和各样品之间运行两针空白样品, 以排除残留的影响。通过这种方法, 我们发现完整抗体的最低检测限

是 500 pg, 在进样量为 500 pg 的时候, 仍然可以得到较好的谱图, 从谱图上可以看到完整抗体中丰度最高的糖型。需要指出的是, 最低浓度的样品必须现配, 分析前不能在自动进样器中存放数小时。

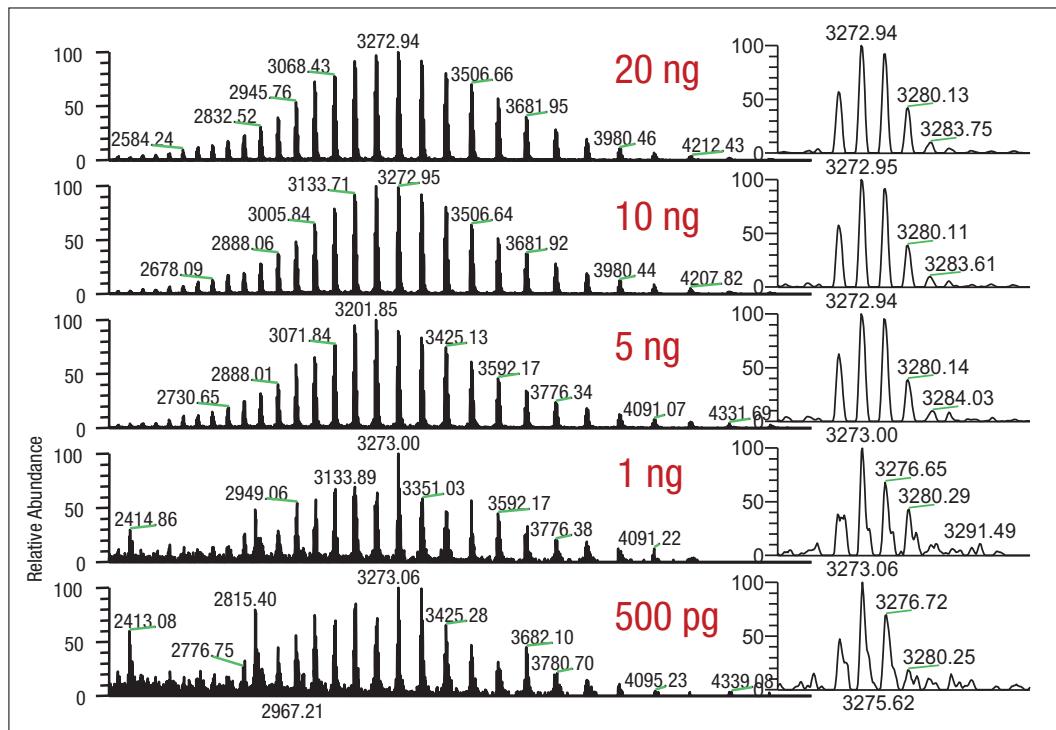


图 8. 用 $250\text{ mm} \times 0.2\text{ mm}$ PepSwift 整体柱分析完整的利妥昔抗体，上样量为 20 ng - 500 pg 一系列稀释浓度得到的一级全扫描质谱图。

在 $50\text{ mm} \times 1\text{ mm}$ 内径 ProSwift 整体柱上，进样量为 30 ng 和 150 ng 的时候，完整抗体都可以得到高质量的谱图(图 9)，

根据 30 ng 的上样量，我们预测最低进样量在 5 ng - 10 ng 之间，在此范围内仍然可以得到优质的谱图。

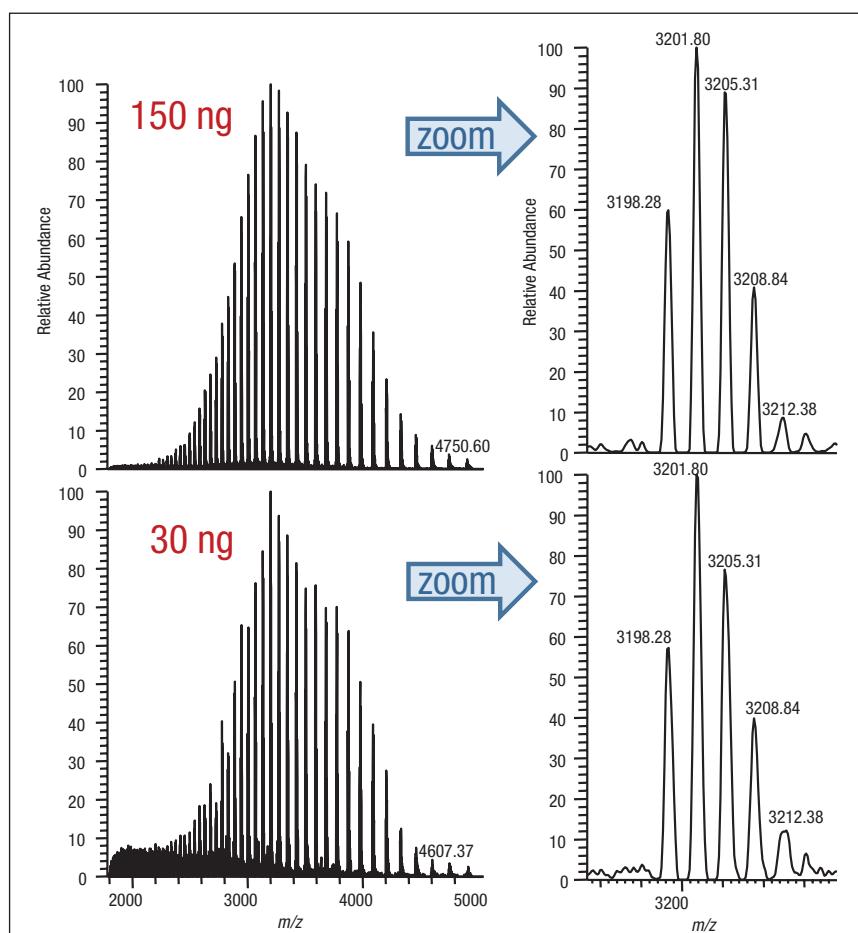


图 9. 用 $50\text{ mm} \times 1\text{ mm}$ ProSwift 整体柱分析完整的利妥昔抗体，上样量分别为 150 ng 和 30 ng 时的一级全扫描质谱图 (左侧) 和最高丰度电荷态放大图 (右侧)。

接下来，我们试图进一步确认轻链和重链的序列，进行了两类自上而下的实验：在 HCD 池中进行全离子裂解 (AIF)，或选择轻链和重链的 5 个电荷态进行多重 (5-plex) 目标离子二级质谱实验。所有谱图都是在分辨率为 70000 下采集。目标离子谱图的采集使用基于保留时间的质量数列表，包括先流出的轻链质荷比（保留时间 13.16 min: m/z 1536.96、1646.6、1773.3、1920.7、2095.4）和后流出的重链（保留时间 16-20min: m/z 1584.6、1635.7、1684.7、1748.5、1810.9）。在这类实验中，内含物列表中的第一个电荷态被选择并送到 HCD 池中进行裂解，子离子储存在 HCD 池中，同时分离第二个电荷态送到 HCD 池，裂解并贮存在该池中，直到第 5 个电荷态被裂解。5 个单独分离和裂解步骤产生的所有离子被一起送到 Orbitrap 分析器中，产生一个单一的碎片离子谱图。

所有归为轻链的碎片离子（如图 10 所示）在 N- 和 C- 末端都有好的覆盖率，在序列的中间也有一些碎片离子匹配，最终的覆盖率为 28%，占理论碎片的 15%。对于重链来说，两种方法的裂解效率都不是特别好，约产生 20 个碎片离子，大多数都位于序列的末端。

为了进一步确认序列，采取了自下而上的方法，将蛋白进行还原、烷基化、并用胰蛋白酶进行酶切。图 11 显示的是酶切后肽段混合物的色谱图。对数据进行数据库检索，数据库由轻链、两个重链变体以及胰蛋白酶四个条目组成，结果显示轻链的序列覆盖率达 96%，重链的序列覆盖率达 78.8%（图 12）。两个来自轻链的、较短的、没有检测到的肽段 (LEIK 和 EAK) 只能在一级质谱图中以完整质量数的形式检测到，而肽段 EAK 可以根据含漏切位点的肽段 EAKVQWK 所对应的准确质量数进行鉴定。综合二级质谱图鉴定的肽段和完整短肽的准确质量数，轻链的序列覆盖率为 100%。

b1	- Q - I - V { L } S } Q - S } P - A } I - L - S - A - S - P - G - E - K - V - T - M - T - C } R - A -	y189
b26	- S - S - S - V - S - Y - I - H - W - F - Q - Q - K - P - G - S - S } P - K - P - W - I - Y - A - T -	y164
b51	- S } N } L } A } S - G - V } P - V - R - F - S - G - S - G - T - S - Y - S - L - T - I - S -	y139
b76	- R - V - E } A - E - D - A - A - T } Y - Y - C - Q - Q } W } T - S - N } P - P - T - F - G - G - G -	y114
b101	- T - K - L - E - I - K - R - T - V - A - A - P - S - V - F - I t F t P - P - S - D - E - Q - L - K -	y89
b126	- S - G - T - A - S - V - V { C - L - L - N - N - F - Y - P - R - E - A - K - V - Q - W - K - V - D -	y64
b151	- N - A - L t Q t S - G t N t S t Q t E t S t V t T t E t Q - D t S - K - D t S - T t Y t S t L t S -	y39
b176	t S - T t L t T t L t S t K t A - D t Y - E t K t H t K - V t Y - A t C - E t V - T t H t Q t G t L -	y14
b201	t S t S t P t V t T t K t S t F t N - R - G - E - C -	y1

图 10. 根据通过 AIF 实验得到的碎片离子匹配利妥昔抗体轻链的序列覆盖率。鉴定到 17 个 b 离子和 50 个 y 离子，对应 15.7% 的理论碎片离子数。（426 个理论碎片离子中鉴定到 67 个离子）

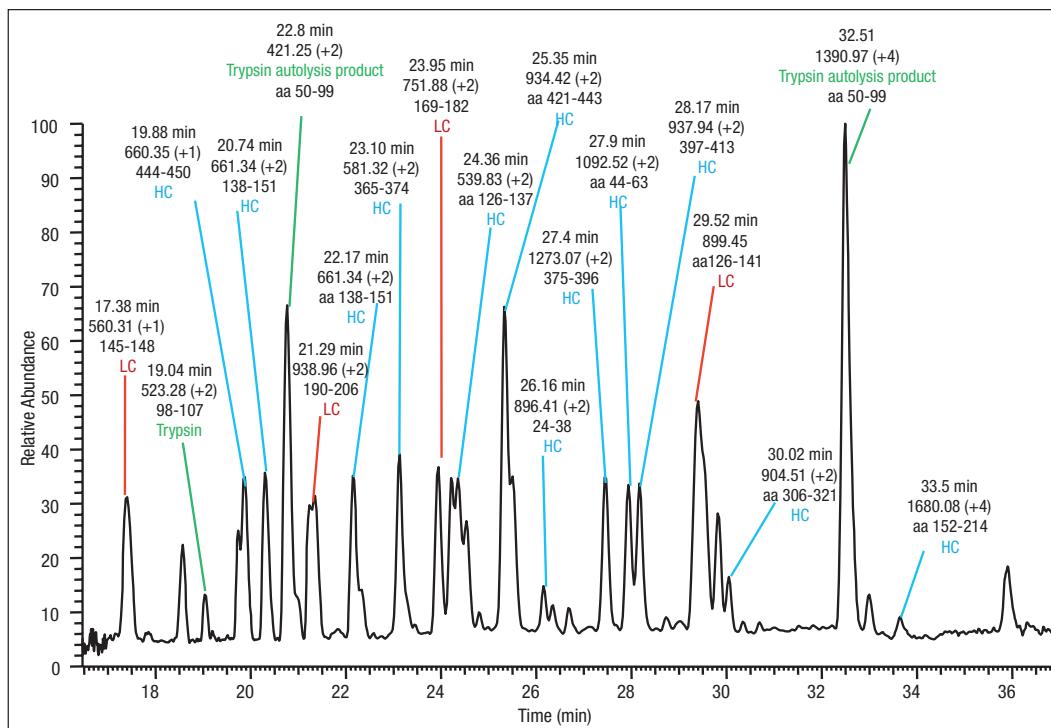


图 11. 将利妥昔抗体还原、烷基化、胰蛋白酶切后，酶切后肽段的基峰色谱图。

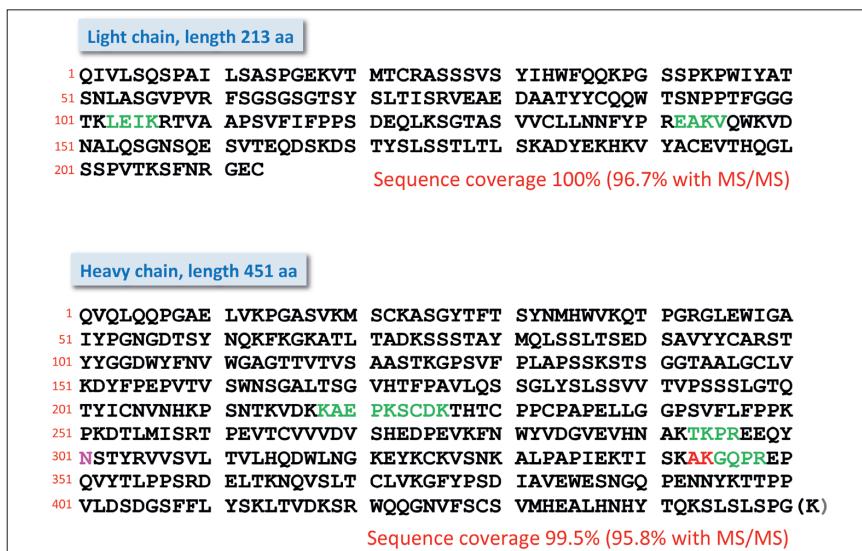


图 12. 利妥昔单抗轻链和重链的氨基酸序列。黑色字母显示的氨基酸代表通过二级质谱图鉴定到的部分。

绿色显示的序列是仅通过完整肽段一级全扫描数据鉴定到的。红色显示的是重链部分的两个氨基酸（AK），他们既不能用一级也不能用二级质谱鉴定到。对于轻链来说，最终的序列覆盖率为 100%（二级质谱鉴定覆盖率为 96%），对于重链来说是 99.5%（二级质谱鉴定的覆盖率为 98.8%）。重链中天冬酰胺³⁰¹表示糖基化位点。

对于重链，也可以根据完整肽段的准确质量数鉴定肽段 QOPR。最后，根据二级谱图，含糖基化位点 Asn³⁰¹ 的肽段没有鉴定到其未糖基化的形式。而对其进行基于理论糖基化修饰的数据库检索则很成功。此外，很容易在全扫描谱图中检测到糖肽，可以看到不同的糖基化形式和不同的电荷态。由于存在特征峰，二级谱图上也很容易看到。图 13 显示的是一个含 GOF 肽的谱图举例，显示了完整母

离子和用 HCD 裂解糖肽得到的典型裂解类型：质量数为 204 的 hexonium 离子（HexNAc）和质量数为 366 的离子（Hex-HexNAc）以及阶梯式碎片离子：己糖 (*m/z* 162)、N-乙酰基氨基糖 (203) 和岩藻糖 (146)。综合考虑基于所有肽段的一级全扫描和基于数据库搜索的二级质谱图，重链的序列覆盖率达 99.5%，只有两个氨基酸（氨基酸 343-344）没有检测到。

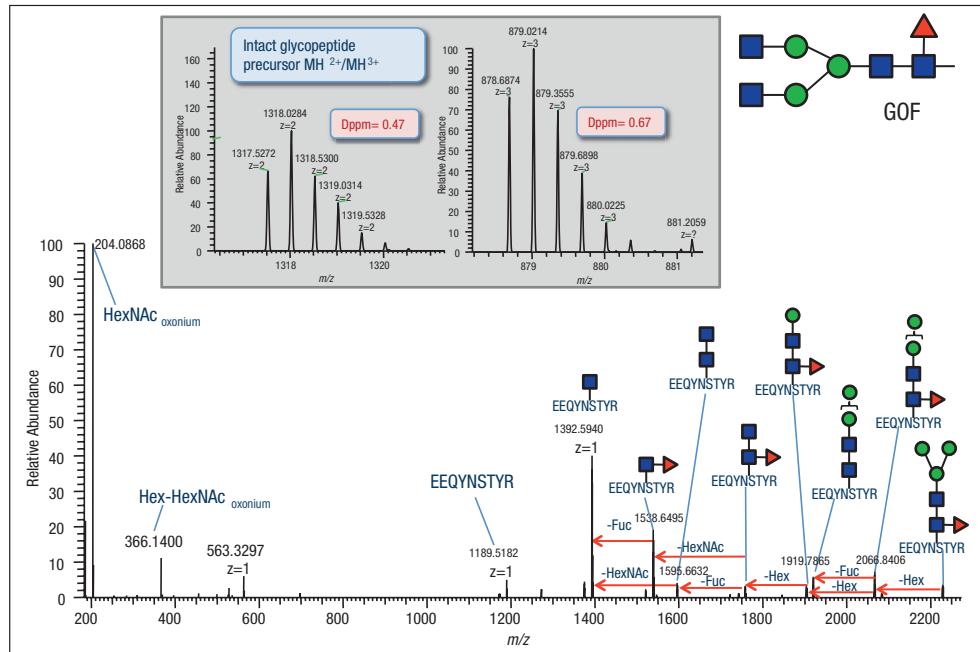


图 13. 糖肽氨基酸序列 297-305 (EEQYN^{*}STYR, ^{*}=GOF) 的二级质谱图，其母离子是三价电荷。

插图显示的是在全扫描质谱图中二价和三价母离子的同位素分布。

结论

在本研究中，我们提供一种新流程，结合快速色谱，利用两种规格的整体柱和高分辨 Orbitrap 质谱仪，对完整及还原利妥昔进行了分析，采用 AIF, Multiplexed HCD 自上而下的裂解，以及自下而上的分析方法进行序列验证。

即使抗体的上样量低至 500 pg，我们仍可以得到高质量的质谱图，该数据证明了 LC-MS 仪器的高灵敏度。此外，为了分析还原型单克隆抗体，我们采用色谱的方法将抗体的重链和轻链分开，设置不同的分辨率对其进行检测。

通过这个流程，我们可以测定完整抗体的分子量，对轻链和重链的氨基酸序列进行确认和验证，对该抗体不同糖基化形式的相对丰度进行鉴定和评价。

致谢

感谢 Daniel Pürstinger 在样品制备过程中给予的帮助，感谢 Remco Swart 提供本研究当中用到的 PepSwift 和 ProSwift 整体柱。感谢奥地利联邦经济、家庭、青年部和全国研发技术部提供的财政支持。

参考文献

- Premstaller, A.; Oberacher, H. and Huber, C.G. High-Performance Liquid Chromatography-Electrospray Ionization Mass Spectrometry of Single- and Double Stranded Nucleic Acids Using Monolithic Capillary Columns. *Anal. Chem.* **2000**, *72*, 4386-4393.
- <http://www.drugbank.ca/drugs/DB00073>
- Nebija, D.; Kopelet-Frank, H.; Urban, E.; Noe, C. R. and Lachmann, B. Comparison of two-dimensional gel electrophoresis patterns and MALDI-TOF MS analysis of therapeutic recombinant monoclonal antibodies trastuzumab and rituximab. *Journal of Pharmaceutical and Biomedical Analysis*, **2011**, *56*, 684–91.
- Kurabayashi, R.; Hashii, N.; Harazono, A. and Kawasaki, N. Rapid evaluation for heterogeneities in monoclonal antibodies by liquid chromatography/mass spectrometry with a column-switching system. *Journal of Pharmaceutical and Biomedical Analysis*, **2012**, *67-68*, 1–9.