LCMS-9030 用于牛血清白蛋白的鉴定研究

LCMS-QTOF-040

摘要： 本文采用岛津 LCMS-9030 飞行时间高分辨液质联用仪对牛血清白蛋白 (BSA) 进行鉴定研究，并结合岛津 LabSolutions 和 pFind 软件对 BSA 酶解肽段进行匹配打分。结果显示，在胰蛋白酶酶切的情况下，使用两种计算软件对 BSA 二级结构确认的匹配度分别为 82.4% 和 85.7%。利用 Byonic 软件对肽段上的修饰基团进行解析，结果显示，除固定修饰乙酰胺化外，BSA 仅存在少量的 N- 脱酰胺（Deamidated）等可变修饰。该方法快速、准确，分析精确度高，为蛋白类样品的肽质谱图分析提供参考。

关键词： Q-TOF　蛋白组学　BSA

基于液质联用技术的蛋白鉴定研究主要有两种思路：自下而上 (Bottom-up) 和自上而下 (Top-down)。最常用的是“自下而上”，也就是我们常说的鸟枪法 (shotgun)，是一种依赖数据库的策略，它充分利用了蛋白质自身的特点，即蛋白质可以被特定的酶在特定的位点酶切，酶切所获得的特异性肽段在质谱二级碎片分析时，通常会形成三种特征离子 by 离子，ax 离子和 cz 离子。通过质谱测定这些特征的二级碎片信息，结合理论谱库相似度匹配，从而实现肽段的鉴定，再根据肽段序列推导出蛋白质序列。因此肽质谱图的匹配（Peptide Sprectrum Matching，PSM）是一个非常重要的指标。PSM 常常借助搜库软件实现，常见的如 SEQUEST、Mascot、X！ Tandem、Comet 以及 Max Quant、pFind 软件等。这些搜库软件所使用的统计学打分算法通常有两种：正反库策略 (target-decoy，TDA) 和基于经验贝叶斯 (EB) 的打分。对于打分结果大部分软件都要求进行假发现率（False Discovery Rate, FDR）和酶切特异性 (Digestion Specific) 控制。而 PSM 匹配的准确性与未知翻译后修饰，质谱数据的质量准确度，肽段丰度等因素有关。岛津 LCMS-9030 采用了 UF-FlightTube 飞行管精准温控技术，保证了亚 ppm 的质量精度，同时，传承了 LCMS-8060 旗舰级三重四极杆质谱的高效、快速离子传输系统，结合新型漏斗形 MCP 检测器，拥有 fg 级灵敏度，有利于低丰度肽段测定。

本研究采用胰蛋白酶对 BSA 进行酶切，MS Scan +DDA 进行数据采集，质谱数据经 MSConvert 软件转换为 mzML 格式，经 pFind 软件 (中国科学院计算所) 和 Byonic 软件（Protein Metrics Inc.）分别进行蛋白鉴定研究，并对翻译后修饰进行分析。

实验部分

1.1 仪器

本实验使用高效液相色谱仪 LC-30AD 与四极杆飞行时间质谱仪 LCMS-9030 联用系统。具体配置为：

系统控制器：CBM-20Alite
输液泵：LC-30AD×2
柱温箱：CTO-20AC
脱气机：DGU-20A$_5$
自动进样器：SIL-30AC
色谱工作站：LabSolutions Ver. 5.99

1.2 分析条件

液相条件

色谱柱：Shim-pack GISS C18（150 mm× 2.1 mm I.D., 1.9 μm），PN:227-30048-03，岛津（上海）实验器材有限公司
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
柱温：40°C
进样体积：1 μL
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　梯度洗脱时间程序

Time(min)	Module	Command	Value
2.00	Pump	B.Conc	2
30.00	Pump	B.Conc	25
33.00	Pump	B.Conc	60
36.00	Pump	B.Conc	60
37.00	Pump	B.Conc	2
40.00	Controller	Stop	

质谱条件

离子化模式：ESI^+
加热模块温度：400°C
雾化气流速：3.0 L/min
干燥气流速：10.0 L/min
加热气流速：10.0 L/min
扫描模式：MS 和 MS/MS (DDA)
接口温度：300°C
接口电压：4.5 kV
DL 温度：250°C
事件时间：0.1 s

样品前处理

用 50 mM 碳酸氢铵做溶剂，将 BSA 粉末制剂配制成 0.1 g/mL 的溶液，然后用碳酸氢铵稀释至 100 μg/mL，吸取 50 μL 该溶液样品，加入 47 μL 碳酸氢铵溶液，加 500 mM 二硫苏糖醇 (DTT)1 μL，在 37°C下震摇 60 min，结束后加入 1.5 μL 碘乙酰胺 (IAA) 溶液（1 M），在室温条件下避光震摇 30 min，然后加入 1 μL 胰蛋白酶溶液（含 1 μg 胰蛋白酶），在 37°C下反应 12 小时，酶解结束后加入 1 μL TFA 溶液终止酶解反应，然后转入低吸附 PP 材质的样品瓶，进行分析。

结果与讨论

3.1 BSA 肽段色谱图

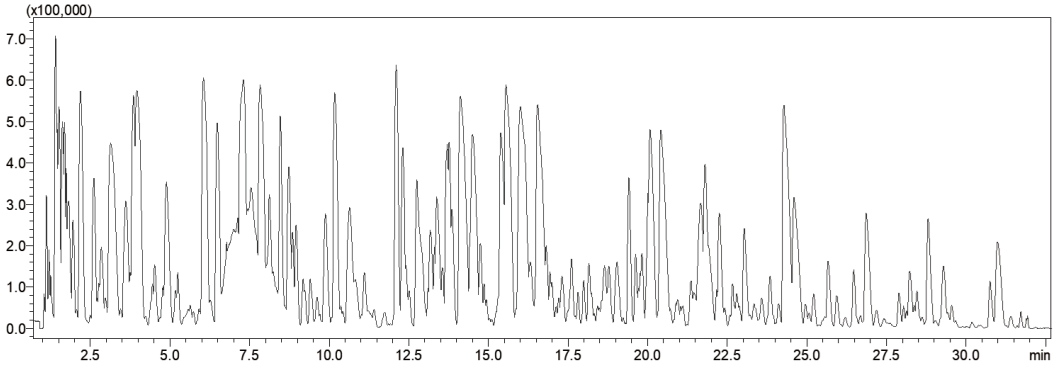

图 1　BSA 酶解后的 BPC 图

3.2 肽质谱图匹配

将 LabSolutions 采集的数据直接转换为 mzML 格式，导入 pFind 软件，并导入 BSA 的氨基酸序列文件 (P02769,UniProt)，进行多肽质谱匹配和翻译后修饰分析。pBuild 软件用于蛋白质搜库结构查看，在控制 FDR<1% 条件下，有 191 张库匹配二级质谱图，氨基酸序列覆盖率为 82.4%（图 2）。酶切特异性匹配良好，总体达到 98.3%。肽段 AA 长度分布 6~30 之间，其中 15 肽占比最高。实测一级质谱误差 1.58±3.58 ppm，质谱质量精度优异。识别到修饰类型 10 种，如半胱氨酸 C- 乙酰胺化，C- 乙酰胺化修饰 (Cam C，+57.021 Da) 比例高达 35.6%。除此之外，还存在赖氨酸 K 的乙酰胺化，N- 脱酰胺化 (deamidated)，可变修饰比例均低于 3%。

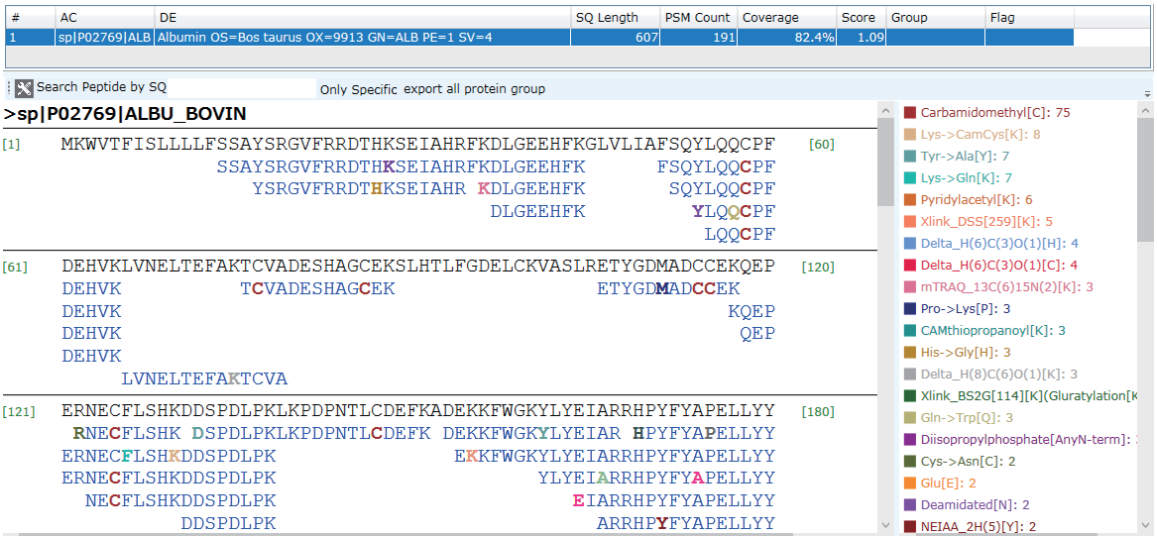

图 2　BSA 酶解后序列覆盖图

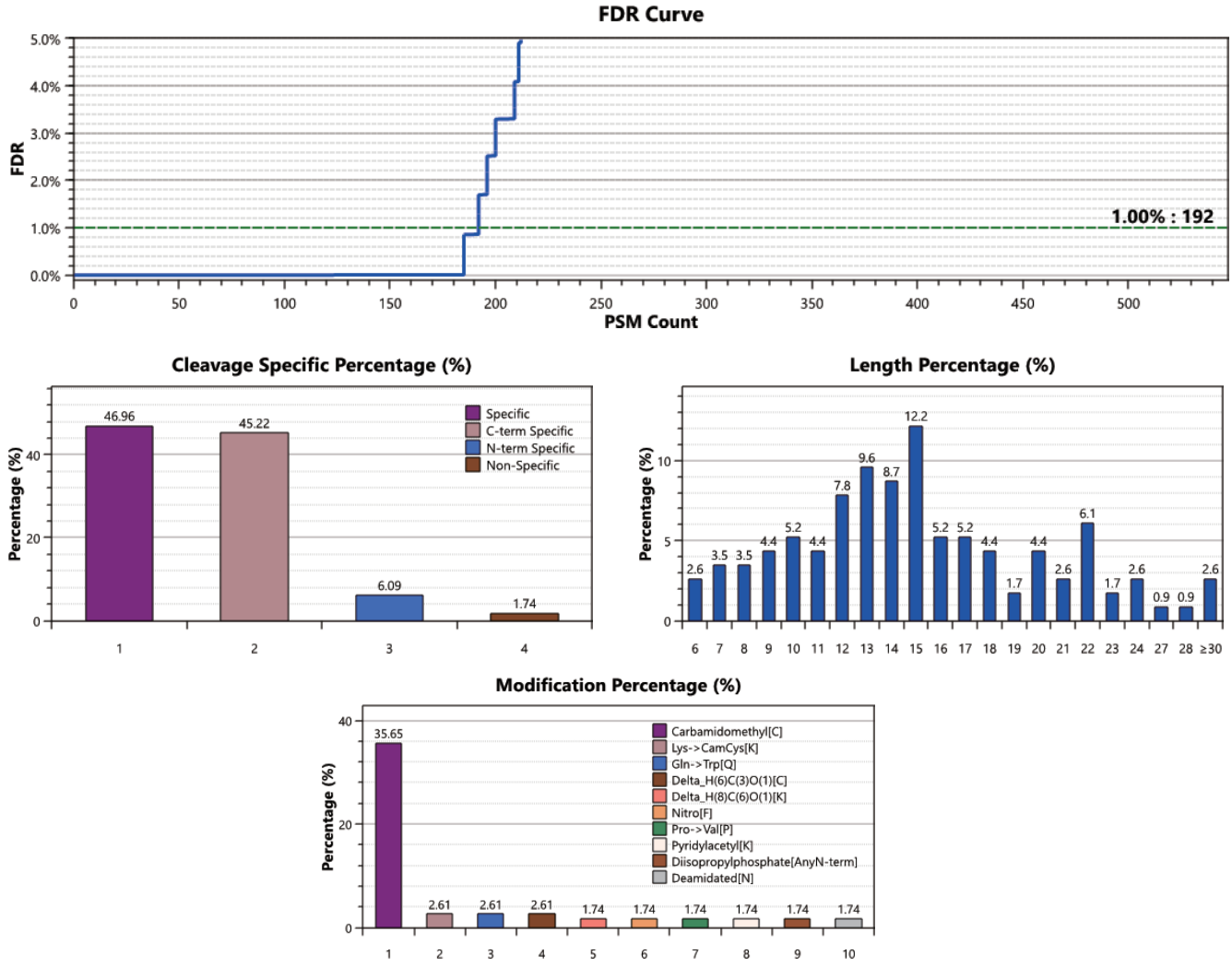

图 3　PSM 结果汇总图

为了交叉验证不同软件 PSM 打分结果，数据导入 Byonic 软件分析，氨基酸序列覆盖度可达 85.7%，与 pFind 软件计算的 PSM 结果 82.4% 相当。

```
>NP_851335.1 albumin precursor [Bos taurus]

        10        20        30        40        50        60        70        80        90       100       110
MKWVTFISLLLLFSSAYSRGVFRRDTHKSEIAHRFKDLGEEHFKGLVLIAFSQYLQQCPFDEHVKLVNELTEFAKTCVADESHAGCEKSLHTLFGDELCKVASLRETYGD
       120       130       140       150       160       170       180       190       200       210       220
MADCCEKQEPERNECFLSHKDDSPDLPKLKPDPNTLCDEFKADEKKFWGKYLYEIARRHPYFYAPELLYYANKYNGVFQECCQAEDKGACLLPKIETMREKVLTSSARQR
       230       240       250       260       270       280       290       300       310       320       330
LRCASIQKFGERALKAWSVARLSQKFPKAEFVEVTKLVTDLTKVHKECCHGDLLECADDRADLAKYICDNQDTISSKLKECCDKPLLEKSHCIAEVEKDAIPENLPPLTA
       340       350       360       370       380       390       400       410       420       430       440
DFAEDKDVCKNYQEAKDAFLGSFLYEYSRRHPEYAVSVLLRLAKEYEATLEECCAKDDPHACYSTVFDKLKHLVDEPQNLIKQNCDQFEKLGEYGFQNALIVRYTRKVPQ
       450       460       470       480       490       500       510       520       530       540       550
VSTPTLVEVSRSLGKVGTRCCTKPESERMPCTEDYLSLILNRLCVLHEKTPVSEKVTKCCTESLVNRRPCFSALTPDETYVPKAFDEKLFTFHADICTLPDTEKQIKKQT
       560       570       580       590       600
ALVELLKHKPKATEEQLKTVMENFVAFVDKCCAADDKEACFAVEGPKLVVSTQTALA
```

图 4　覆盖度图 (Byonic)

PMi-PTM 模块对氨基酸进行修饰研究，修饰主要有脱酰胺（deamidated）氧化（Oxidation/15.9949）、脱硫醇化（Dethiomethyl/-48.0034）、谷氨酰胺焦谷氨酸化（Gln->pyro-Glu/-17.0265）等。如下图 5 所示，在多肽 TVMENFVAFVDK 中，M 发生脱硫醇化修饰，比例达到 47.2%。

						MS Id ←	1
						MS Alias name ←	BSA DDA profile_005 (%) ↓
Sequence (unformatted)	Mod. Names	Protein name	Mod. AAs	Var. Pos. Protein	Labels		
TVMENFVAFVDK	Dethiomethyl/-48.0034	sp|P02769|ALBU_BOVIN Albumin OS=Bos taurus OX=9913 GN=ALB PE=1 SV=4	M	571			47.2
QEPER	Gln->pyro-Glu/-17.0265	sp|P02769|ALBU_BOVIN Albumin OS=Bos taurus OX=9913 GN=ALB PE=1 SV=4	Q	118			37.5
IETMR	Dethiomethyl/-48.0034	sp|P02769|ALBU_BOVIN Albumin OS=Bos taurus OX=9913 GN=ALB PE=1 SV=4	M	208			28.3
QTALVELLK	Gln->pyro-Glu/-17.0265	sp|P02769|ALBU_BOVIN Albumin OS=Bos taurus OX=9913 GN=ALB PE=1 SV=4	Q	549			19.1
QIK	Gln->pyro-Glu/-17.0265	sp|P02769|ALBU_BOVIN Albumin OS=Bos taurus OX=9913 GN=ALB PE=1 SV=4	Q	545			18.8
LGEYGFQNALIVR	Deamidated/0.9840	sp|P02769|ALBU_BOVIN Albumin OS=Bos taurus OX=9913 GN=ALB PE=1 SV=4	N	428			6.41
AWSVAR	Oxidation/15.9949	sp|P02769|ALBU_BOVIN Albumin OS=Bos taurus OX=9913 GN=ALB PE=1 SV=4	W	237			1.97
NYQEAK	Deamidated/0.9840	sp|P02769|ALBU_BOVIN Albumin OS=Bos taurus OX=9913 GN=ALB PE=1 SV=4	N	341			1.73
IETMR	Oxidation/15.9949	sp|P02769|ALBU_BOVIN Albumin OS=Bos taurus OX=9913 GN=ALB PE=1 SV=4	M	208			1.69

图 5　氨基酸修饰统计 (Byonic)

根据显示二级质谱图校对肽段的碎片，尤其是带修饰的碎片基团，以肽段 K.QNCDQFEK.L 为例，可以根据二级质谱图清晰判断每个碎片的断裂位置以及修饰位点（图 6）。

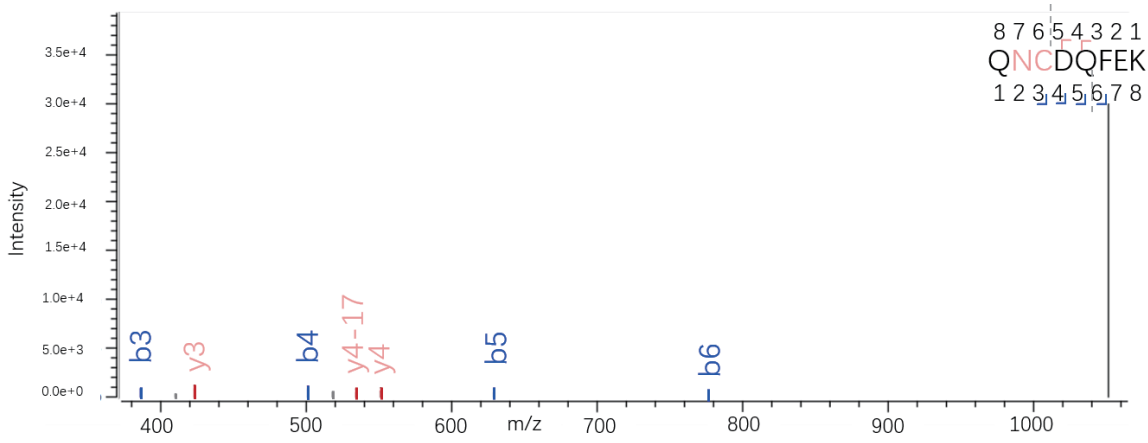

图 6　K.QNCDQFEK.L 二级质谱图

3.3 蛋白质鉴别

牛蛋白质组序列 (Uniprot proteome _UP000009136.fasta)，包含了牛的所有蛋白质组，将其用于二级质谱数据归属研究。经 pFind 软件计算，该样品中识别到三个蛋白，均指向牛血清白蛋白，多肽覆盖度均大于 68%，如下图 7。以上结果表明，在复杂的蛋白组成体系下，牛血清白蛋白鉴定结果仍然准确可靠。

#	AC	DE	SQ Length	PSM Count	Coverage
1	sp|P02769|ALBU_BOVIN	Albumin OS=Bos taurus OX=9913 GN=ALB PE=1 SV=4	607	1	69.7%
2	tr|A0A3Q1LNN7|A0A3Q1LNN7_BOVIN	Albumin OS=Bos taurus OX=9913 GN=ALB PE=4 SV=1	576	1	68.2%
3	tr|A0A140T897|A0A140T897_BOVIN	Albumin OS=Bos taurus OX=9913 GN=ALB PE=4 SV=1	607	1	71.2%

图 7　牛蛋白组匹配结果 (pFind)

同样地，采用 Byonic 软件进行计算，牛血清白蛋白（BOVIN Album）得分较高，获得了 74.8% 的匹配度，远远高于其他蛋白质，定性匹配结果良好。如下图 8。

Prot. Rank	Protein Name	|Log Prob|	Best |Log Prob|	Best Score	# Spectra	# Uniq. Peps.	# Mod Peps.	% Cov.	# AAs	Intensity
1 | 1	>tr|A0A140T897|A0A140T897_BOVIN Albumin OS...	288.72	11.37	841.3	855	68	6	74.8	607	1.760e+8
2 | 2	>sp|Q3SZR3|A1AG_BOVIN Alpha-1-acid glycoprot...	24.82	5.66	540.3	29	12	0	43.1	202	2.849e+5
3 | 3	>sp|P15497|APOA1_BOVIN Apolipoprotein A-I OS...	16.30	4.63	432.1	13	8	0	31.7	265	5.289e+4
4 | 4	>sp|Q2UVX4|CO3_BOVIN Complement C3 OS=Bos...	8.50	5.75	494.2	13	5	0	1.9	1661	1.692e+5
5 | 5	>sp|P41361|ANT3_BOVIN Antithrombin-III OS=Bo...	3.58	3.55	383.8	3	2	0	3.2	465	1.999e+4
6 | 6	>tr|F6RLA2|F6RLA2_BOVIN Testis specific serine ki...	2.91	2.91	412.2	1	1	0	2.5	326	3.322e+4
7 | 7	>tr|A5PJH7|A5PJH7_BOVIN LOC788112 protein OS...	2.90	2.88	253.2	2	1	0	5.2	172	7.158e+3
8 | 8	>sp|P02769|ALBU_BOVIN Albumin OS=Bos taurus ...	2.04	1.05	304.4	46	3	0	1.6	607	1.292e+6
9 | 9	>tr|E1BH06|E1BH06_BOVIN C4a anaphylatoxin OS...	1.92	1.84	274.2	6	4	1	1.6	1741	4.090e+4
10 | 10	>tr|A0A3Q1N2C2|A0A3Q1N2C2_BOVIN Uncharact...	1.85	1.83	259.3	2	1	0	2.2	316	8.088e+4
11 | 11	>sp|Q3T0E2|SYCP3_BOVIN Synaptonemal complex...	1.84	1.62	319.2	14	2	0	5.3	225	1.537e+6
12 | 12	>sp|Q58DT5|ELMD3_BOVIN ELMO domain-contain...	1.78	1.78	274.1	1	1	0	2.1	381	3.453e+4
13 | 13	>tr|F1MDS0|F1MDS0_BOVIN Proteoglycan 4 OS=B...	1.63	1.63	261.4	1	1	0	0.7	1195	7.079e+3

图 8　牛蛋白组匹配结果 (Byonic)

结论

本文采用自下而上 (Bottom-up) 的研究策略，对 BSA 进行了鉴定研究。BSA 样品经 LCMS-9030 分析，经两种软件计算，PSM 匹配得分均大于 80%，即使是在复杂蛋白组成体系中，PSM 也能获得 68% 或以上的匹配结果。同时，对 BSA 翻译后修饰进行了研究，除固定修饰乙酰胺化外，BSA 仅存在少量的 N- 脱酰胺 (Deamidated) 等可变修饰，可变修饰比例低于 3%。该方法简便，准确，适合蛋白组学或蛋白类药物的鉴别研究。