

蛋白质组学研究中傅立叶变换离子回旋共振质谱数据采集方法的比较

曾家伟 贾伟 应万涛 钱小红*

(军事医学科学院放射与辐射医学研究所, 北京蛋白质组研究中心, 蛋白质组学国家重点实验室, 北京 102206)

摘 要 探讨了线性离子阱-傅立叶变换离子回旋共振质谱(LTQ-FT)的数据采集模式对蛋白质组鉴定结果的影响, 比较和分析了针对不同复杂程度样本的最佳采集模式。对于 α -乳清蛋白 4 种标准蛋白质酶切肽段混合物的简单体系, 在傅立叶变换离子回旋共振腔中选 3 个最强母离子进行选择离子监测扫描后, 再实施二级碎裂的方法(SIM3)得到的肽段覆盖率, 分别是直接选 10 个最强母离子进行二级碎裂(FT10)方法得到的肽段覆盖率的 1.5 ~ 1.9 倍。另外, 对酵母蛋白酶切肽段混合物的复杂体系鉴定时, 只采集双电荷和三电荷母离子进行二级扫描的方法(FT10_23)比单电荷、双电荷和三电荷都采集的方法(FT10_123)得到的图谱鉴定成功率提高 64.1%。实验还对不同数据采集模式下图谱的特点进行了考察。本研究表明, 针对不同复杂程度的样本, 应采取不同的质谱数据采集方法。

关键词 蛋白质组学; 质谱采集模式; 线性离子阱-傅立叶变换离子回旋共振质谱

1 引言

蛋白质组学分析的样本极其复杂, 准确度与扫描速度哪一个可以产生更为理想的数据结果是一个非常基础和非常重要的方法学问题^[1,2]。在以肽段为直接鉴定目标的 Shotgun 蛋白质组研究策略中, 离子阱型质谱以其高灵敏度、高扫描频率得到了广泛应用^[3,4]。新型的离子阱质谱层出不穷, 但由于其设计原理的局限, 准确度和分辨率还不理想。而傅立叶变换回旋共振质谱, 可以提供超高的准确度(1 ~ 2 ppm)和分辨率(500000 FWHM 以上)^[5]。但缓慢的扫描速度却使其长期以来无法应用于大规模的蛋白质组研究。混合型质谱 Linear quadrupole ion trap/FT-ICR(LTQ-FT)将离子阱质谱与傅立叶变换回旋共振质谱巧妙地结合, 使之兼具了这两种质谱仪的优点: 高扫描速度、高灵敏度、高准确度及高分辨率。这类混合型质谱成为大规模蛋白质组分析中最有力的工具之一^[6-9]。本研究针对 LTQ-FT 质谱不同采集方法的特点, 设计实验, 总结出适合不同复杂程度样本的采集方法。

2 实验部分

2.1 仪器与试剂

真空冰冻干燥机 SC100A Speedvac Plus(美国 Thermo Savant 公司); 线性离子阱-傅立叶变换离子回旋共振质谱仪(LTQ-FT 美国 Thermo Fisher 公司); 超纯水由 Millipore 纯水系统(美国 Millipore 公司)制备。

α -乳清蛋白、碳酸酐酶 2、3-磷酸甘油醛脱氢酶、卵白蛋白和啤酒酵母干粉蛋白(Sigma 公司); 乙腈(ACN, HPLC 级, 美国 Fisher 公司); 三氟乙酸(TFA)、甲酸(FA)、碘乙酰胺(IAA)和二硫苏糖醇(DTT)(比利时 ACROS 公司); 胰蛋白酶(Trypsin, Sequencing grade, Promega 公司); 其它试剂为国产分析纯。

2.2 实验方法

2.2.1 标准蛋白酶切 取 α -乳清蛋白、碳酸酐酶 2、3-磷酸甘油醛脱氢酶和卵白蛋白各 200 μg , 溶于 100 μL 含有 8 mol/L 尿素、10 mmol/L DTT 和 50 mmol/L NH_4HCO_3 溶液中, 置 37 $^\circ\text{C}$ 水浴中还原 4 h, 然后加入 5 μL 1 mol/L 碘乙酰胺放置暗处 1 h 进行烷基化, 加 900 μL 50 mmol/L NH_4HCO_3 溶液将尿素稀释到 0.8 mol/L, 最后加入 16 μg 胰蛋白酶, 在 37 $^\circ\text{C}$ 保温反应 16 h。

2009-12-21 收稿; 2010-02-08 接受

本文系国家重点基础研究计划项目(Nos. 2007CB914100, 2010CB912700)、国家重大创新药物创制专项(No. 2009ZX09301-002)、国家自然科学基金重点项目(Nos. 20735005, 20635010)资助项目

* E-mail: qianxh@nic.bmi.ac.cn

2.2.2 酵母蛋白酶切 取啤酒酵母干粉 200 mg 溶于 1 mL 50 mmol/L NH_4HCO_3 , 超声 15 min 40000 g 离心 1 h。取上清液, 冻干, 加入 Loading buffer (50 mmol/L Tris-HCl 2% SDS, 10% Glycerol, 0.1% 溴酚蓝) 煮沸 10 min。4% ~ 12% SDS-PAGE 分离, 考马斯亮蓝染色, 分子量 60 ~ 110 kDa 范围条带胶上酶切。

2.2.3 液相分离条件 将脱盐后的酶切肽段用 A 相复溶, 15000 g 离心 10 min 后上样。流动相 A: 0.1% FA-98% 水溶液; 流动相 B: 0.1% FA-80% ACN 溶液。采用 Agilent 1100 纳升级液相色谱系统。分析柱为 PicoFrit™ 反相柱 (BioBasic® C_{18} , 10 cm $\times$ 75 μm i. d., 5 μm , New Objective, Woburn, MA, USA), 流速为 300 nL/min。4 个标准蛋白酶切肽段液相分离梯度: 0 ~ 30 min, 流动相 B 线性升至 50%; 酵母胶上酶切蛋白液相分离梯度: 0 ~ 90 min, 流动相 B 线性升至 50%。

2.2.4 质谱采集条件 磁场强度为 7 T。喷雾电压 1.8 kV; 离子传输毛细管温度为 200 $^{\circ}\text{C}$; 归一化碰撞能量为 35%。

FT10 方法是一级质谱数据使用 FT 进行全扫描, 扫描范围 m/z 400 ~ 2000, 在 m/z 400 处的分辨率设置为 10^5 。离子自动增益控制 (AGC) 设置为: 一级质谱扫描为 5×10^5 , 二级串联质谱为 1×10^4 。依次选取一级质谱中最强的 10 个离子在 LTQ 中进行 CID 二级串联质谱分析。

SIM3 方法为一级质谱数据采用 FT 进行全扫描, 扫描质荷比范围是 m/z 400 ~ 2000, 在 m/z 400 处的分辨率为 2.5×10^4 。依次选取一级质谱中最强的 3 个离子在 FT 中再进行选择离子监控模式扫描, 在 m/z 400 处的分辨率为 5×10^4 。同时对这 3 个离子在 LTQ 中进行 CID 二级串联质谱分析。离子自动增益控制 (AGC) 设置为: FT 全扫描为 3×10^6 , 选择离子监控扫描 (SIM) 为 5×10^4 , 二级串联质谱为 1×10^4 。

2.2.5 数据检索 分别建立 4 种标准蛋白和酵母蛋白的正反库, 用 Bioworks 3.2 进行搜库。FT10 方法检索母离子质量误差范围为 10 ppm, SIM3 方法检索母离子质量误差范围为 3 ppm^[10]。二级离子检索质量误差范围为 1 Da。肽段鉴定结果采用 1% 假阳性率过滤。

3 结果与讨论

3.1 数据依赖型串联质谱采集方法产生数据特点的比较

在蛋白质组学研究中, LTQ-FT 型质谱仪常采用的串联质谱采集方法有: (1) 在傅立叶变换离子回旋共振腔中对某一时间点上所有母离子完成全扫描, 然后选择峰强度最高的前 10 个母离子在离子阱中进行串联质谱分析 (称为 FT10 方法); (2) 在傅立叶变换离子回旋共振腔中对某一时间点上所有母离子完成全扫描, 然后选择峰强度最高的前 3 个母离子进行选择离子监控扫描后再在离子阱中进行串联质谱分析 (称为 SIM3 方法)。两种方法的主要区别在于 SIM3 方法提供母离子质量准确度 (约 2 ppm) 优于 FT10 方法 (约 10 ppm)^[11]。但是在相同时间内, FT10 方法采集 5 ~ 10 张二级质谱数据 (约 3 s) 的时间, SIM3 方法只能采集 3 张二级质谱数据。因此理论上 FT10 方法可以获得更多的肽段信息。Haas 等^[12]的研究表明, 在对酵母蛋白全酶切得到的多肽复杂体系鉴定时, FT10 方法比 SIM3 方法得到的肽段多约 30%, 且鉴定到了更多的低丰度蛋白。这表明在复杂体系的研究中, 更快的二级串联扫描速度对提高多肽鉴定数量更重要。

近年来, 蛋白质组学的发展已经从最初以表达谱分析为目标的研究, 向专门化方向迈进, 如目标蛋白质组学。目标蛋白质组学更关注于相对有限体系中 (如蛋白质复合体等) 蛋白质的组成, 以及各组分在某种条件下相对和绝对量的变化。目标蛋白组的研究体系相对简单, 其样品往往是通过复杂的处理过程, 如亲和纯化或免疫共沉淀, 因此常常由于获得的目标样品量较少而较难鉴定^[13, 14]。本研究设计了 4 种标准蛋白质酶切肽段混合物的简单体系, 分别用 FT10 方法和 SIM3 方法进行鉴定。从表 1 可见, 采用 SIM3 的方法得到的肽段覆盖率是 FT10 方法的 1.5 ~ 1.9 倍。理论上, FT10 方法可以获得更多的肽段信息, 但在此实验中鉴定到的肽段覆盖率反而低于 SIM3 方法。这因为在 SIM3 方法中, 一级质谱全扫描的自动增益控制 (AGC) 值设置为 3×10^6 , 而在 FT10 方法中为 5×10^5 。较高的 AGC 值可以聚集更多的离子, 使一些低丰度或离子化效率低的肽段达到进行二级碎裂的要求, 并使最终鉴定肽段数增加。但离子数的增加直接导致空间电荷效应的提高, 使得质谱扫描的分辨率和准确度降低, 这也是 FT10 方法中 AGC 值设置较低的原因。但 SIM3 方法在全扫描后还要选择峰强度最高的前 3 个母离子

进行选择离子监控扫描,而这时的 AGC 设定值为 5×10^4 ,反而会提高母离子扫描的分辨率和准确度。本研究表明,在较简单的体系中,更高的离子积累值和准确度可以得到更多的鉴定结果。所以 SIM3 采集方法可能对有限体系中蛋白的鉴定更加适合。因此针对复杂程度不同的样品,应先分析 AGC 值对鉴定结果的影响,选择最佳的平衡点和采集模式。

表 1 4 种蛋白 LTQ-FT 鉴定的覆盖率统计
Table 1 Observed sequence coverage of four proteins

蛋白名称 Protein name	覆盖率 Coverage(%)		鉴定覆盖率的比值(SIM3 /FT10) Ratio of coverage (SIM3 vs FT10)
	FT10	SIM3	
α -乳清蛋白 Alpha-lactalbumin	8.13	15.45	1.90
碳酸酐酶 2 Carbonic anhydrase 2	11.58	19.69	1.70
3-磷酸甘油醛脱氢酶 Glyceraldehyde-3-phosphate dehydrogenase	17.17	25.90	1.51
卵白蛋白 Ovalbumin	21.04	39.35	1.87

每种蛋白的上样量为 20 fmol(Loading amount: 20 fmol)。

3.2 串联质谱采集时选择不同电荷离子作为母离子的比较

分辨率较低的离子阱型质谱如 LCQ 或 LTQ,不能分辨母离子带电状态,会将所有达到信号强度要求的离子都进行碰撞并记录串联质谱信号。通常,单电荷母离子的鉴定结果较少,但会过多地占用扫描时间。而 LTQ-FT 质谱仪的回旋共振系统具有极高准确度和分辨率,扫描得到的谱图可以清晰地辨别同位素峰簇,从而计算出离子的电荷数。为了确定有选择地只采集某种电荷的母离子的碎片是否有助于提高鉴定效率,本研究对采集单电荷、双电荷和三电荷的方法即 FT10_123 和只采集双电荷和三电荷的方法 FT10_23 进行比较。实验样本为酵母蛋白酶切肽段,每种采集方法重复 3 次。

由表 2 可见,FT10_123 方法所产生的谱图数量比 FT10_23 的方法要提高 58.5%。但 FT10_23 方法谱图鉴定成功的比例(59.4%) 明显高于 FT10_123 方法(36.2%)。因此只采集双电荷和三电荷母离子的方法比单电荷、双电荷和三电荷都采集的方法简单,且效率更高。两种方法有 78.6% 肽段交盖,说明 FT10_123 方法鉴定到的大部分肽段在 FT10_23 方法中也可鉴定到。

表 2 不同方法采集的谱图数和鉴定成功率
Table 2 Number of MS/MS spectra and confident peptide identifications from two different acquisition methods

采集方法 methods	第一次 First		第二次 Second		第三次 Third		平均采集的 谱图数 Average spectra	谱图鉴定成功率 MS/MS spectra identified rates (%)
	Acquired spectra	Peptide IDs	Acquired spectra	Peptide IDs	Acquired spectra	Peptide IDs		
FT10_123	771	52371	7111	2558	9505	3876	8110	36.2
FT10_23	3761	2344	4860	2763	6713	4008	5117	59.4

表 3 显示了鉴定肽段在不同电荷中的分布。在单电荷、双电荷和三电荷母离子都采集的方法中,单电荷所占比例极少,双电荷母离子占主要部分。可能是单电荷肽段较难碎裂产生二级离子,因此鉴定的成功率不高。另外在电喷雾源质谱中,许多单电荷离子属于背景离子,反而占用采集时间,降低质谱的采集效率。表 2 的结果也说明了这一点。

表 3 成功鉴定肽段在不同电荷中的分布情况
Table 3 Distribution of confident peptide identifications with different charges

采集方法 Methods	单电荷 Single charge	双电荷 Double charge	三电荷 Triple charge	平均鉴定的肽段数 Average Peptide IDs
FT10_123	231 (7.8%)	2477 (84.4%)	227 (7.8%)	2935
FT10_23		2803 (92.3%)	235 (7.7%)	3038

3.3 轮廓图和棒状图模式质谱数据的特点

本研究采用轮廓图(Profile) 和棒状图(Centroid) 两种采集模式并发现一些不同点。轮廓图所记录的信息较多,通常一个质谱峰在轮廓图中会保留多个信号采集点,但数据量较大;而棒状图经过中心化,只用一个信号值记录一个峰,数据量较小。在蛋白质组学研究中常采用棒状图记录离子阱中的串联质

谱信号。但文献 [15] 报道, 在其它条件相同的情况下, 采用轮廓图模式记录的数据比棒状图模式数据肽段鉴定的真阳性结果更多。本实验在人工分析一些质谱图时发现, 在棒状图中所记录的同位素质谱峰簇并不完整, 其中第偶数个同位素峰都缺失了, 见图 1A 图中 1 和 2 区域。为了判断棒状图和轮廓图哪种模式记录的结果更准确, 采用傅立叶变换回旋共振系统采集相同的离子, 同时获得了肽段同位素质谱峰簇的质荷比和所带电荷数目, 证明轮廓图模式数据正确。这提示手工分析离子阱数据时, 最好采集轮廓图模式数据, 否则很可能导致对质谱数据的误判。例如在图 1 中双电荷离子的同位素峰簇相邻同位素峰间 $m/z = 0.5$, 但是偶数个同位素峰的缺失使记录的相邻同位素峰间 $m/z = 1$ 。这样在读图时会误将双电荷离子误认为是单电荷离子, 以至无法确定离子真正质量数, 从而导致分析错误。由于轮廓图模式包含了更多的谱图信息, 因此在手工分析未知样品或研究某种新的碎裂模式时, 最好采用轮廓图模式, 使结果更准确可靠。

4 结 论

本研究运用高精度和高分辨率的 LTQ-FT 质谱, 对 α -乳清蛋白等 4 种标准蛋白和酵母蛋白的酶切肽段分别进行鉴定。发现对于 4 种标准蛋白的简单体系, 采用 SIM3 方法鉴定到的肽段覆盖率是 FT10 方法的 1.5 ~ 1.9 倍。在分析酵母蛋白样本时, 只采集双电荷和三电荷母离子串联碎裂的方法比单电荷、双电荷和三电荷都采集的方法鉴定效率更高。因此对复杂程度不同的样本, 应分析并选择最佳的数据采集模式。此外, 由于轮廓图模式包含了更多的谱图信息, 因此在手工分析未知样品或研究某种新的碎裂模式时, 最好采用轮廓图模式, 使结果更准确可靠。

References

- 1 Zubarev R, Mann M. *Mol. Cell. Proteomics.*, **2007**, 6(3): 377 ~ 381
- 2 Mann M, Kelleher N L. *PNAS*, **2008**, 105(47): 18132 ~ 18138
- 3 Schwartz J C, Senko M W, Syka J E. *J. Am. Soc. Mass. Spectrom.*, **2002**, 13(6): 659 ~ 669
- 4 LU Zhuang(卢庄), ZHAO Li-Yan(赵丽艳), ZHANG Yang-Jun(张养军), CAI Yun(蔡耘), DENG Yu-Lin(邓玉琳), ZHANG Yu-Kui(张玉奎), QIAN Xiao-Hong(钱小红). *Chinese J. Anal. Chem.* (分析化学), **2009**, 37(7): 950 ~ 954
- 5 Bernhard S. *J. Am. Soc. Mass. Spectrom.*, **2004**, 15(5): 703 ~ 714
- 6 Syka J E, Marto J A, Bai D L, Horning S, Senko M W, Schwartz J C, Ueberheide B, Garcia B, Busby S, Muratore T, Shabanowitz J, Hunt D F. *J. Proteome. Res.*, **2004**, 3(3): 621 ~ 626
- 7 Schaub T M, Hendrickson C L, Horning S, Quinn J P, Marshall A G. *Anal. Chem.*, **2008**, 80(11): 3985 ~ 3990
- 8 Marshall A G, Hendrickson C L. *Annu. Rev. Anal. Chem.*, **2008**, 1: 579 ~ 599
- 9 Cox J, Matic I, Hilger M, Nagaraj N, Selbach M, Olsen J V, Mann M. *Nat. Protoc.*, **2009**, 4(5): 698 ~ 705
- 10 Zhang J, Ma J, Dou L, Wu S, Qian X, Xie H, Zhu Y, He F. *J. Proteome. Res.*, **2009**, 8(2): 849 ~ 859
- 11 Olsen J V, Ong S E, Mann M. *Mol. Cell. Proteomics.*, **2004**, 3: 608 ~ 614
- 12 Haas W, Faherty B K, Gerber S A, Elias J E, Beausoleil S A, Bakalarski C E, Li X, Villén J, Gygi S P. *Mol. Cell. Proteomics.*, **2006**, 5(7): 1326 ~ 1337

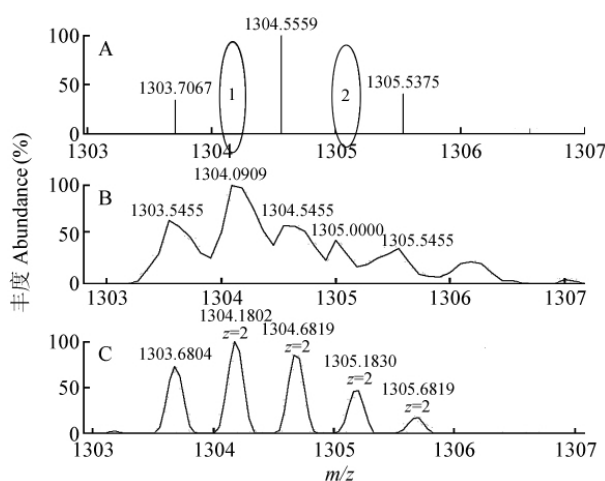


图 1 不同模式采集质荷比为 1303.6804 肽段的质谱图

Fig. 1 Direct comparison of peptide ($m/z = 1303.6804$) detected with different acquisition method

A 和 B 分别为离子阱采集的棒状图和轮廓图; C 是用傅立叶变换回旋共振系统采集的相同肽段 (Peptide was measured on an ion trap at Centroid format (A), Profile format (B) and on FT MS at Profile format (C)).

- 13 Yang L , Zhang H , Bruce J E. *Analyst* , **2009** , 4(4) : 755 ~ 762
14 Honda S , Selker E U. *Genetics* , **2009** , 182(1) : 11 ~ 23
15 Alves G , Wu W W , Wang G , Shen R F , Yu Y K. *J. Proteome. Res.* , **2008** , 7(8) : 3102 ~ 3113

Comparison of Acquisition Methods for Linear Quadrupole Ion Trap-Fourier Transform Cyclotron Resonance Mass Spectrometer in Shotgun Proteomics

ZENG Jia-Wei , JIA Wei , YING Wan-Tao , QIAN Xiao-Hong*

(State Key Lab of Proteomics , Beijing Proteome Research Center , Beijing Institute of Radiation Medicine , Beijing 102206)

Abstract Mass spectrometers with high accuracy and high resolution , such as FT-ICR have been widely used in proteomic studies in recent years. To compare the effectiveness of various data acquisition methods in shotgun proteomics , we used a linear quadrupole ion trap-Fourier transform ion cyclotron resonance (LTQ-FT) mass spectrometer to analyze two peptide mixtures with different complexity by different acquisition strategies. For the peptide mixture of four standard proteins , sequence coverage obtained by selected ion monitoring scans (SIM3) was 1. 51 to 1. 9 fold than that by survey MS scans (FT10) . For the peptide mixture of yeast proteins , the strategy of only acquiring double and triple charged ions (FT_23) got 64. 1% more MS/MS spectra identifications than that of total acquisition of single , double and triple charge ions (FT_123) . Finally , the features of the spectrum obtained with different acquisition mode were also compared. In conclusion , optimal acquisition methods are needed for the samples with different complexity.

Keywords Proteomics; Mass spectrometer acquisition methods; Linear quadrupole ion trap-Fourier transform ion cyclotron resonance mass spectrometer

(Received 21 December 2009; accepted 8 February 2010)

欢迎订阅 欢迎投稿 欢迎刊登广告
《分析试验室》

国内统一刊号: CN11 - 2017/TF

国际 CODEN 码: FENSE4

国外代号: M848

国际标准刊号: ISSN 1000 - 0720

邮发代号: 82 - 431

广告经营许可证: 京西工商广字第 0441 号

《分析试验室》是中文核心期刊 ,月刊 ,大 16 开 ,128 页 ,国内外公开发行。

《分析试验室》1982 年创刊 ,目前已成为我国著名的分析化学专业刊物。影响遍及冶金、地质、石油化工、环保、药物、食品、农业、商品检验和海关等社会各行业及各学科领域。《分析试验室》以突出创新性和实用性为办刊宗旨 ,作者来自全国各行业的生产、科研第一线; 已被列为全国中文核心期刊、中国科技论文统计用期刊、美国“CA 千种表”中我国化学化工类核心期刊、中国学术期刊(光盘版) 和中国期刊网全文数据库等国内外多家检索数据库、文摘收录 ,影响因子连续多年列化学类前列。本刊常设“研究报告”、“研究简报”、“仪器装置与设备”等栏目。“定期评述”栏目系统发布特邀知名专家学者撰写的国内外分析化学各领域的综合评述 ,连续跟踪学术发展前沿。“国际会议”栏目每期介绍影响广泛的分析化学领域国际学术交流会议。

2011 年《分析试验室》每期定价 18 元 ,全年 12 期 216 元。

全国各地邮局征订 邮发代号 82 - 431。漏订的读者可直接与编辑部联系。

编辑部地址: 北京新街口外大街 2 号

邮编: 100088。电话: 010 - 82013328; E-mail: analysislab@ 263. net; ana-info@ 263. net