

基体辅助激光解吸质谱法在生物大分子质量研究中的应用

周国华 罗国安* 朱敏生**

(清华大学生命科学与工程学院, 北京 100084; **南京军区药品检验所, 南京 210002)

摘要 研究了基体辅助激光解析飞行时间质谱法(MALDI-TOF MS)测定蛋白和糖蛋白的实验条件, 实测了重组人白细胞介素 2、肿瘤坏死因子、粒细胞巨噬细胞集落刺激因子、白细胞干扰素 $\alpha 2b$ 、白细胞干扰素 α_1 、促红细胞生成素、钙调蛋白及其片段、大鼠脑型一氧化氮合成酶和天然牛精液蛋白提取物的分子量和纯度, 鉴定了蛋白质混合物中各单个成分, 结果表明 MALDI-TOF MS 可有效地研究基因工程蛋白质和天然蛋白提取物等生物大分子的质量。

关键词 基体辅助激光解析飞行时间质谱法; 蛋白质; 分子量

自 1988 年 Hillenkamp 等^[1]首先报道用基体辅助激光解析飞行时间质谱(matrix-assisted laser desorption/ionization time of flight mass spectrometry, 简称 MALDI-TOF MS)分析分子量超过 1.65×10^{-20} g (10 kDa) 的蛋白质以来, 因其有精度高(在 0.1% ~ 1% 之间)、速度快(几分钟测一个样)、对杂质的承受力高(允许一定量的 SDS、难挥发性盐存在)等优点, 近年来用于研究蛋白质、核苷酸和多糖等生物大分子的应用报道越来越多^[2~5]。基因工程药物大多是重组蛋白产品, 其常规质量控制主要是测定基因序列、N 端部分氨基酸序列和分子量等。但仅凭这些手段还无法了解产品的准确分子量、糖基化程度、二聚体形成及杂质信息等。本文系统研究了 MALDI-TOF MS 最佳实验条件的选择, 并用该法直接准确地测定了重组人白细胞介素 2(IL-2)、肿瘤坏死因子(TNF α)、粒细胞巨噬细胞集落刺激因子(GM-CSF)、白细胞干扰素 $\alpha 2b$ (INF α_{2b})、白细胞干扰素 α_1 (INF α_1)、促红细胞生成素(EPO)、钙调蛋白(CaM)及其片段、大鼠脑型一氧化氮合成酶

(nNOS)等 9 种基因工程蛋白质和一种天然牛精液蛋白提取物(BSE)的分子量和纯度, 同时还研究了蛋白质混合物中单组分的测定方法, 阐述了 MALDI-TOF MS 法在生物活性蛋白质质量研究中的应用潜力。

材料和方 法

材料 IL-2, TNF α , GM-CSF, INF α_{2b} , CaM 及其片段和 nNOS 均为大肠杆菌(*E. coli*)表达产物, INF α_1 为杆状病毒家蚕细胞表达产物, EPO 为中国仓鼠卵巢细胞(CHO)表达产物, BSE 为新鲜牛精液蛋白提取物, 以上产品均为实验室制备。肌红蛋白、 β -乳球蛋白、碳酸酐酶 II 和牛血清白蛋白购自 Sigma。

MALDI-TOF MS 靶的制备和测定法 所有实验均在 Finnigan Laser MAT 2000 飞行时间质谱分析仪(Finnigan MAT, Hemel Hempstead, Herts, UK)上完成, 采用 N_2 激光(337 nm, 2 ns 脉冲), 正离子方式检测, 不锈钢靶(Finnigan MAT)的靶芯直径为 2 mm, 面积约 3.14 mm^2 , 使用 Laser Mat 软件调节激光辐射的目标部位和次数, 激光能量和加速电压均可用该软件调节到最优。以肌红蛋白作外部校准。分别以芥子酸[简称 SA, $90 \text{ mmol} \cdot \text{L}^{-1}$], 乙

本文于 1997 年 6 月 18 日收到。

本文为国家自然科学基金重点资助项目(No. 296350220)

* 通讯联系人

腈-乙醇-水(60:4:36)溶解], 2,5-二羟基苯甲酸[简称 DHB, $100 \text{ mmol} \cdot \text{L}^{-1}$, 甲酸-水(9:1)溶解], α -氰基-4-羟基肉桂酸[简称 A-CHC, $20 \text{ mg} \cdot \text{ml}^{-1}$, 乙腈-乙醇-水(40:1:9)溶解]为基体。测定时取样品 $1 \mu\text{l}$ 、基体溶液 $5 \mu\text{l}$ 与 0.5% 三氟醋酸(TFA)水溶液 $1 \mu\text{l}$ 混匀, 取混合液 $1 \mu\text{l}$ 滴于靶片上, 在空气中室温吹干放入质谱仪中测试。

结果和讨论

1 实验条件的确定

1.1 基体的选择 MALDI-TOF/MS 实验中, 样品的制备非常重要。由于基体辅助解吸的机理尚不十分清楚, 故仅能凭经验选择合适的基体。本文选择 3 种常用基体进行研究, 分别是 DHB, SA 及 A-CHC。在测定糖蛋白 EPO 时, 发现 DHB 基体给出的质谱峰宽, 信号弱, 二聚体单电荷离子比例高, 峰不尖锐, 卫星峰很多, 而 SA 和 A-CHC 基体给出的质谱峰明显优于 DHB。为验证这一结论, 用上述基体分别测定了未知糖蛋白-牛精液提取物, 得出了类似的结论, 进一步用非糖基化蛋白试验, 3 种基体均可给出较好的质谱峰, 但 SA 更好一些, 所以本文以下实验均以 SA 为基体。

1.2 样品的制备 为提高测定灵敏度, 本文在制备样品时, 加入了适量的 TFA。测定图谱的好坏, 主要取决于样品与基体形成的晶型, 如果样品中含有表面活性剂或其他难溶性盐, 则很难产生好的结晶。以 NOS 为例, 样品中含有

SDS, 不经处理测不出结果, 若在基体完成结晶之前, 滴加冷水于靶上洗涤, 则可测出 MS 图。所以制备未知样品时, 最好用显微镜监控晶型形成情况。在样品量少的情况下, 只要样品与基体能形成好的晶型, 通过仪器的信号累加功能, 同样可获得好的 MS 图。此外样品与基体的比例对测定也会产生影响, 以未知牛精液蛋白提取物为例, 将样品与基体的比例分别配成 1:1000 和 1:5000(摩尔比), 均得出较好的结果, 后者灵敏度较低, 但可通过仪器信号累加功能加以调整。实验中所需样品量极少, 通常样品的浓度在 $1 \text{ mg} \cdot \text{ml}^{-1}$ 左右, 若蛋白分子量为 $1.65 \times 10^{-20} \text{ g}$ (10 kDa) 和 $3.3 \times 10^{-20} \text{ g}$ (20 kDa), 则点样量分别为 6 和 12 ppm, 由此说明本法的灵敏度很高。

1.3 仪器参数的选择 样品制备是关键, 只有好的晶型才有可能通过仪器参数的优化得出好的 MS 图。实验中发现如在仪器自动设定条件下观察不到质谱峰, 则靠仪器参数的调节也难以奏效, 必须重新制样, 尤其是要考虑除去样品中的缓冲盐和表面活性剂。要获取最佳 MS 图, 经作者多次实验证明: 能量不宜设置过高, 调至最大能量的 30% 左右为宜; 叠加次数视情况而定, 通常设在 20 次左右。

2 单一蛋白质的分子量测定

2.1 非糖基化蛋白 采用 MALDI-TOF MS 测定了用大肠杆菌表达的非糖基化基因工程产品: IL-2, TNF α , GM-CSF, INF α_{2b} , Ca M 片段和 nNOS 的分子量。其氨基酸数目、理论分子量、测得分子量及测定误差见表 1。

Tab 1 Results of measured molecular weights (MW) by matrix assisted laser desorption ionization time of flight mass spectrometry (MALDI-TOF MS) for non glycoproteins

Protein	Number of amino acid	Theoretical MW ($\times 10^{-20} \text{ g}$)	Measured MW ($\times 10^{-20} \text{ g}$)	Relative error (%)
IL-2	134*	2.6	2.6	0.8
GM-CSF	127	2.4	2.5	5.56
TNF α	158*	2.9	2.9	0.18
nNOS	199*	3.7	3.7	0.26
INF α_{2b}	166*	3.2	3.2	0.79
Ca M fragment	95	1.8	1.8	0.05

* An additional methionine in NH₂-terminus. IL-2: interleukin 2; GM-CSF: granulocyte-macrophage colony stimulating factor; TNF α : tumor necrosis factor α ; nNOS: neural nitric oxide synthase; INF α_{2b} : interferon α_{2b} ; Ca M: calmodulin.

从表 1 看出: GM-CSF 的测定误差为 5.56%, 不可能是由仪器误差引起的, 从其质谱图看, 峰很尖锐, 而且产品很纯。经 Edman 测序, 发现 N 端多 8 个氨基酸 MMKSDNSH, 将这多出的氨基酸加入到理论氨基酸序列中, 得到平均分子量为 2.5×10^{-20} g (15.42 kDa), 与测得分子量相比, 相对误差为 0.12%。IL-2 与 $IFN\alpha_{2b}$ 的测定误差接近 0.8%, 即比理论值大约多一个氨基酸, 这可能是蛋白质被修饰了, 如磷酸化、半胱氨酸化等, 本文未作进一步确证。在 $IFN\alpha_{2b}$ 质谱图中 (图 1) 半峰宽较大, 峰不对称, 可看出峰的左侧含有未分开的峰, 推测是 $IFN\alpha_{2b}$ 分子末端去除 MET 引起的。

2.2 糖蛋白 在真核细胞内表达的蛋白, 通常是糖蛋白, 由于糖的结构多样性, 使表达产物是一个非均相性的混合物, 有些糖对蛋白的生物学活性起重要作用, 如 CHO 细胞表达的基因工程人促红细胞生成素 (EPO) 中的唾液酸就决定着 EPO 的体内生物学活性, 所以测定蛋白质糖基化程度意义重大。本文用 MALDI-TOF MS 法测定了 EPO, $IFN\alpha_1$, BSE 3 种糖蛋白的分子量, 并根据氨基酸序列计算了糖基化程度, 结果见表 2。

图 2 为 BSE 的 MALDI-TOF MS 图谱。从图 2 可知, MS 峰很宽, 是一典型的非均一性糖蛋白, 否则峰很尖锐。由于本品是未知蛋白, 故

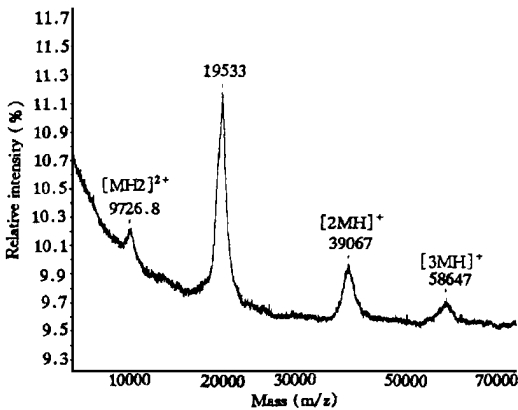


Fig 1 MALDI-TOF MS of interferon α_{2b} . [MH]⁺: Molecular ion; [mMnH]²⁺: Multiply charged monomer or oligomer, the same below.

Tab 2 Results of measured molecular weights (MW) by MALDI-TOF MS for glycoproteins

Protein	Number of amino acid	Measured MW ($\times 10^{-20}$ g)	Content of carbohydrate(%)
EPO	166	4.7	35.7
$IFN\alpha_1$	166	3.4	5.2
BSE	-	3.7	-

EPO: erythropoietin; BSE: bovine semen extract.

不能算出其糖基化程度。EPO 是一糖基化程度很高的糖蛋白, 用 SDS-PAGE 和 GPC-HPLC 法测得其分子量约为 5.6×10^{-20} g (34 kDa), 而本文的 MS 法测定结果为 4.7×10^{-20} g (28.60 kDa), 两者相差达 8.9×10^{-21} g (5.4 kDa), 由此提示: SDS-PAGE 和 GPC-HPLC 两种传统方法在测定糖基化程度较高的蛋白质分子量时, 仅能给出表观分子量, 结果误差较大。家蚕细胞表达的生物分子, 其糖基化程度较低, 从 $IFN\alpha_1$ 的测定结果看, 其糖的含量仅占 5.2%, 理论推算和实测结果相一致。由图 2 和 EPO 的 MS 图可知糖蛋白的 MS 半峰宽很大 [分别为 1.3×10^{-20} g (8 kDa) 和 4.9×10^{-21} g (3 kDa)]。通常 TOF-MS 半峰宽是由同位素引起的, 但并不会超过 8.3×10^{-22} g (0.5 kDa), 本文表 1、图 3 和图 4 中的非糖蛋白的测定结果均证明了这一点。故可用单电荷离子质谱峰的半峰宽大于 8.3×10^{-22} g (0.5 kDa) 作为糖蛋白的一个判据。

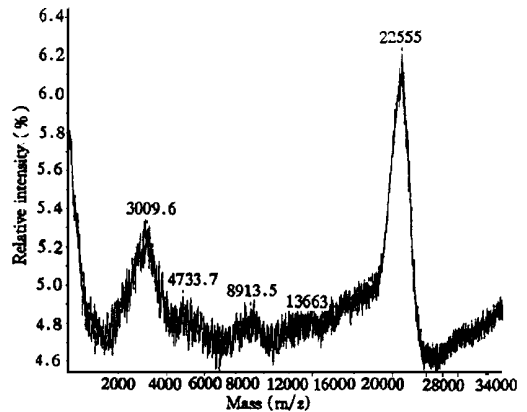


Fig 2 MALDI-TOF MS of bovine semen extract.

3 混合蛋白质的测定

3.1 已知标准混合物的测定 多种蛋白质的混合物不需分离就可同时进行分子量测定,对实际应用很有价值。MALDI-TOF MS 在混合物的分析中有突出的优点,可同时测出多组分的分子量。但在多种蛋白质存在下,会出现多电荷分子离子峰和多聚体分子离子峰,使质谱图变得复杂,由于它们分子量之间具有倍数关系,故可确定其峰的归属。

图3为肌红蛋白、 β 乳球蛋白、碳酸酐酶 II 和牛血清白蛋白 4 种蛋白质混合物的 MALDI-TOF MS 图谱,在图中不仅可观察到蛋白质的分子离子峰,还可观察到它们相应的多电荷和

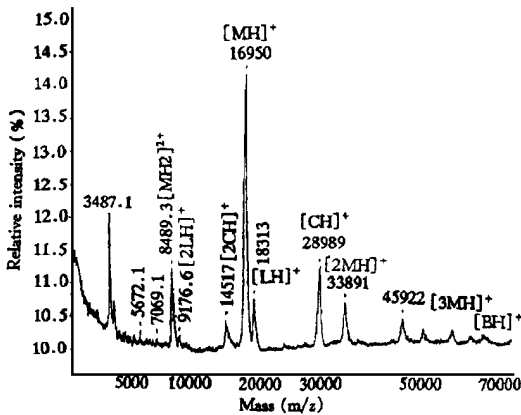


Fig 3 MALDI-TOF MS of protein mixtures. M: myoglobin; L: β -lactoglobulin; C: carbonic anhydrase II; B: bovine serum albumin.

3.3 表达产物纯度和杂质相对含量的测定

由于 MALDI-TOF MS 一次可测出混合物中所有物质的分子量,所以可用该法初步测定表达产物的纯度,同时可得到所含杂质的分子量信息。图5为在大肠杆菌中表达的钙调蛋白片段质谱图。从图中可观察到本品含有分子量为 1.6×10^{-21} g (9.64 kDa) 和 1.5×10^{-21} g (9.07 kDa) 的杂质,不过含量很少。从图2中可看到

多聚体的分子离子峰。各峰的归属直接标注于图上。4种蛋白的分子量测定准确度分别为 0.0%, 0.04%, 0.12% 和 0.18%。

3.2 混合物中目的蛋白质的确定 图4为含钙调蛋白的混合物质谱图,已知钙调蛋白的理论分子量约为 2.8×10^{-20} g (17 kDa),经测定混合物中含有分子量为 2.8×10^{-17} g (16934 kDa) 的蛋白,并且在质谱图中出现双电荷分子离子和单电荷二聚体离子峰。在该混合物中,钙调蛋白的含量是主要的,其它杂质的分子量均小于钙调蛋白。由于钙调蛋白很难形成二聚体,故二聚体离子峰极小。

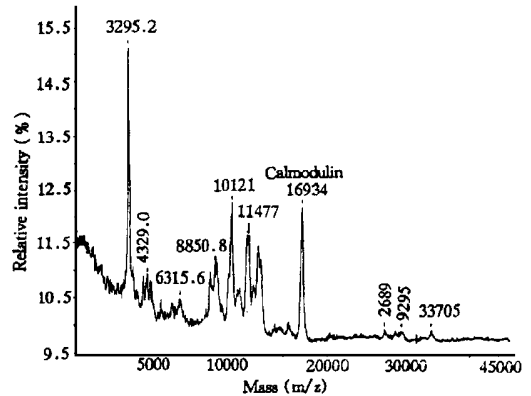


Fig 4 MALDI-TOF MS of mixtures containing calmodulin.

牛精液提取物中含有分子量为 4.9×10^{-21} g (3.01 kDa), 7.8×10^{-21} g (4.73 kDa), 1.5×10^{-21} g (8.91 kDa) 和 2.3×10^{-21} g (13.66 kDa) 的杂质,除 4.9×10^{-21} g (3.01 kDa) 的含量较高外,其余含量很低。此外在 IL-2, TNF α , GM-CSF, INF α_2 , nNOS, INF α_1 和 EPO 的质谱图中未见明显杂质。

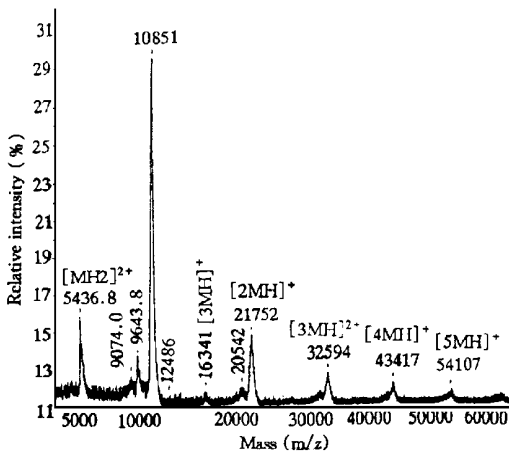


Fig 5 MALDI-TOF MS of calmodulin fragment.

4 结论

本文研究了 MALDI-TOF/MS 法在测定蛋白质时的基体选择、样品制备和仪器参数选择方法,并用于 6 种非糖蛋白、3 种糖蛋白的分子量测定、混合物中各组分的鉴别、表达产物中目的蛋白质的纯度和杂质含量的确定等工作中。MALDI-TOF MS 的实验结果证实了 SDS-PAGE 和 GPC-HPLC 法在测定糖蛋白分子量时存在误差,提出用质谱峰半峰宽大于 8.3×10^{-22} g(0.5 kDa) 作为糖蛋白的一个判据。本文通过大量的应用实例证明 MALDI-TOF MS 法在蛋白质分子量、纯度和杂质测定中具有结

果准确、速度快、样品消耗少等优点。适用于研究初期测定表达产物的分子量、纯度和杂质情况,是测定稀少样品或昂贵样品的经济方法。作者认为从事基因工程产品开发和蛋白提取、纯化研究时, MALDI-TOF MS 是首选的检测工具。但由于 MALDI-TOF MS 是新近出现的技术,到目前为止,我国新生物制品报批工作中还未对此作出要求,相信将来会成为我国分子生物学研究和蛋白质纯化监控的重要手段。

参 考 文 献

- 1 Karas M, Hillenkamp F. Laser desorption ionization of proteins with molecular mass exceeding 10000 daltons. *Anal Chem*, 1988, **60**: 2299
- 2 Jespersen S, Koedam JA, Hoogerbrugge CM, et al. Characterization of O-glycosylated precursors of insulin-like growth factor II by matrix-assisted laser desorption/ionization mass spectrometry. *J Mass Spectr*, 1996, **31**: 893
- 3 Karas M, Bahr U, Strupat K, et al. Matrix dependence of metastable fragmentation of glycoproteins in MALDI-TOF mass spectrometry. *Anal Chem*, 1995, **67**: 657
- 4 Ferstner GJ, Fall KF, Barofsky DF, et al. Mass spectrometric peptide and protein charting. *J Mass Spectr*, 1995, **30**: 519
- 5 Amado FML, Santana-Marques G, Ferrer-Correia AJ, et al. Analysis of peptide and protein samples containing surfactants by MALDI-MS. *Anal Chem*, 1997, **69**: 1102

APPLICATION OF MATRIX ASSISTED LASER DESORPTION IONIZATION TIME OF FLIGHT MASS SPECTROMETRY ON THE QUALITY STUDY OF BIOMACROMOLECULES

Zhou Guohua(Zhou GH), Luo Guoan(Luo GA) and Zhu Minsheng(Zhu MS)*

(Institute for Life Science and Engineering, Tsinghua University, Beijing 100084;

* Nanjing Military Area Institute for Drug Control, Nanjing 210002)

ABSTRACT The experimental conditions of matrix-assisted laser desorption/ionization time of flight mass spectrometry (MALDI-TOF MS) in the analysis of proteins and glycoproteins were studied and were used to analyze the molecular weights and purity of nine recombinant bioactive proteins, such

as interleukin 2 , tumor necrosis factor α , granulocyte-macrophage colony stimulating factor , interferon α_{2b} , interferon α_1 , erythropoietin , calmodulin and its fragments , neural nitric oxide synthase , and a natural protein extracted from bovine semen . It was also applied to determine the objective protein and calmodulin in an unknown mixture . In addition , the single ingredient in protein mixtures was characterized by the MALDI-TOF MS . The results showed that the technique of MALDI-TOF MS can be employed to determine the quality of recombinant and natural proteins effectively and accurately .

KEY WORDS Matrix-assisted laser desorption/ionization time of flight mass spectrometry ; Proteins ; Molecular weight