维普资讯 http://www.cqvip.com

论 著 文章编号:0253-3626(2008)07-0769-04

采用 MS-Fit 进行肽质量指纹谱鉴定蛋白质时主要参数的优化

彭咏波[1],母昭德[1],马永平[2],夏永鹏[1,3],邱宗荫[1]

(1. 重庆医科大学药学院药物分析教研室,重庆 400016;2. 重庆医科大学生物化学与分子生物学教研室,重庆 400016;
3. 重庆市食品药品监督管理局,重庆 400014)

【摘 要】目的:为了优化肽质量指纹谱(Peptide mass fingerprinting,PMF)鉴定蛋白质时采用 MS-Fit 引擎搜索的分析参数。方法:将 2-DE 分离后的 Carbonic anhydrase-2 和 BSA 进行胶内充分酶解,肽段经过 MALDI-TOF-MS 分析得到 PMF 数据。选择 Swissprot 数据库,以 Carbonic anhydrase2 为模型优化搜索参数。结果:主要搜索参数的最佳设置为:半胱氨酸修饰为 Carbamidomethylation,肽质量容错模型为百分数,在研究中 0.1%容错数为最好。结论:本文通过标准蛋白对搜索主要参数的优化,建立了方便、可靠的 MS-Fit 搜索参数模式。

【关键词】MALDI-TOF-MS;PMF 搜索;蛋白质;MS-Fit
【中国图书分类法分类号】R917 【文献标识码】A 【收稿日期】2007-10-09

Optimization of the major parameters of MS-Fit in peptide mass fingerprint identify proteins

PENG Yong-bo, et al

(*Department of Pharmaceutical Analysis, College of Pharmacy, Chongqing Medical University*)

【Abstract】*Objective*: To optimize the major parameters of bioinformatics analysis conditions in process of PMF(Peptide Mass Fingerprinting) identify target proteins. *Methods*: In this process of experiment design and analysis, we made the Bovine Carbonic anhydrase-2 and BSA from 2-DE gel, after proteolysis with porcine trypsin and gathered the peptides, subsequently identified with MALDI-TOF MS and got the peptides mass data. Then, we selected database of Swissprot and MS-Fit search engine, the standard Carbonic anhydrase-2 protein used as model to set up protein search standard conditions. *Results*: Our study showed that the cysteine was modified by carbamidomethylation more convenient in protein search, at the same time, the better mode of mass tolerance was percentage and 0.1% was suitable tolerance in our search. Then we used BSA mass data to check the accuracy of MS search standard conditions. *Conclusions*: Results show that the optimized parameters are reliable.

【Key words】 MALADI-TOF-MS; PMF search; Protein; MS-Fit

通过 PMF 数据进行生物信息学分析是蛋白质组分析中关键的技术环节之一。由于质谱仪本身的分析精度存在差别,加上研究人员对生物信息学搜索软件的认识和掌握程度不一,导致搜索结果的差别很大,以致不同实验室所得到的结果难以相互认同,甚至还可能对同一个事件(如某一种病理过程的比较蛋白质组学研究结果)得出完全不同的结论,这一事实已引起蛋白质组学研究人员越来越多的的关注[1-3]。本文以牛 Carbonic anhydrase-2 蛋白为模型优化并建立搜索标准参数,最后用蛋白 BSA 确证优化参数,对采用 MS-Fit 引擎进行 PMF 鉴定蛋白途径的主要参数进行了较为全面的考究。

作者介绍:彭咏波(1983-),男,硕士,
研究方向:药物分析与新药开发。
通讯作者:邱宗荫,男,教授,E-mail: zongyin@sohu.com。
基金项目:国家自然科学基金资助项目(30476021),教育部博士点基金资助项目(教特发中心函[2004]166 号)。

1 材料与方法

1.1 实验仪器、药品和试剂

MALADI-TOF-MS 质谱仪(ABI Applied Biosystems, USA);双向电泳系统(PROTEAN IEF Cell, PRPTEAN Ⅱ Xi Cell)(Bio-Rad, USA)。

牛碳酸酐酶 2(Bovine Carbonic anhydrase 2)(pI 6.4; MW

维普资讯 http://www.cqvip.com

29ku);牛血清白蛋白(BSA)(PI 5.8;MW 69ku)(D5323A,TaKaRa,大连);质谱用胰酶 V5280(20606101,Promega,USA);NH_4HCO_3(CAS1006-33-7);$Na_2S_2O_3$(CAS1012-17-7);测序级 $K_3Fe(CN)_6$(CAS1047-19-5);测序级 α-腈-4-羟基肉桂酸(α-cyano-4-hydroxycinnamic acid,CHCA)(CAS2201-09-1);色谱纯三氟乙酸(TFA)(CAS1157-25-7)(SIGMA,USA);色谱纯无水 CH_3CN(200604126,江苏汉邦科技有限公司);超纯水(Millipore,USA);碘乙酰胺(0915853,SANGON,上海);二硫苏糖醇(L23619,SANGON,上海);低熔点琼脂糖(YY0118B1007Y,SANGON,上海)。

1.2 实验方法

1.2.1 对标准蛋白 Carbonic anhydrase 2 和 BSA 的前处理 从双向电泳分离后的银染凝胶上用硅化取样管精确取得。进行质谱分析前,样品按下列程序准备:①脱色:将胶块放入 1ml 规格的硅烷化 EP 管中。胶块先用超纯水涡漩洗涤 1min×3 次;接着用 50μl 30mmol/L $K_3Fe(CN)_6$/100mmol/L $Na_2S_2O_3$(1∶1)溶液脱色至银置换完全;水洗 3 遍至透明;用 5 倍胶块体积的 CH_3CN 脱水 3 次,弃去 CH_3CN,以 50μl/gel 100mmol/L NH_4HCO_3 使胶块溶胀,时间为 20min/gel;吸出液体后再用无水 CH_3CN 脱水完全;最后将 EP 管中胶块冰冻干燥。②酶解和萃取肽段:将冰冻干燥的胶块在 4℃冰浴中吸收 20μl 10ng/μl 的胰酶溶液/gel,时间为 1h;接着向每一胶块上层的液体中加入 30μl 40mmol/L NH_4HCO_3/CH_3CN 溶液(9∶1,v/v),在 37℃水浴中酶解 24h;吸出上清液,在水浴中用 30μl 50% CH_3CN/5% TFA 水溶液萃取 1h×2 次;分别将萃取液合并后冰冻干燥[4,5]。

1.2.2 CHCA 饱和溶液及样品/基质溶液制备 质谱分析前,以 50%CH_3CN/5% TFA 水溶液(1∶1,v/v)配制成 3mg/ml。用 1μl 50%CH_3CN/5% TFA 水溶液(1∶1,v/v)将肽混合物溶解并与 CHCA 基质饱和溶液 1μl 混合。

1.2.3 点样 取混合均匀的样品/基质溶液 1μl 点在质谱仪样品板上,在室温、洁净条件室下自然干燥 8h 后再对样品基质结晶进行质谱分析。

1.2.4 质谱分析条件 337nm 氮激光,强度(Intensity):2 000~2 200 坎德拉;反射模式(Reflectron Model);加速电压(Accelerating voltage):20 000V;延迟提取时间(Delay extraction time):120ns;栅格电压(Grid voltage):75%;导向网(Guide wire):0.3%;质谱图肽段质量范围 800~6 000u。

2 结果

2.1 质谱数据的获得

2.1.1 质谱数据校正 由于随着肽段质量数的增加,质谱分辨率明显下降,峰变宽,精确度也随之下降[6],为了保证 1 200~2 500 坎德拉范围内肽段质荷比的准确性,采用 Calibration Mix2 肽混合物[血管紧张素Ⅰ(1 297.3),促肾上腺皮质激素片段 1~17(2 093.2),促肾上腺皮质激素片段 18~39(2 465.2),促皮质激素片段 7~38(3 657.5),牛胰岛素(5 730.9;2 867.0)]外标校正;同时用胰酶自切峰[842.5]做内标校正以保证数据的可靠性。

2.1.2 标准蛋白 Carbonic anhydrase2 和 BSA 的质谱分析 把制备好的肽混合物/CHCA 基质溶液点上样品板,经过 MALADI-TOF-MS 分析获得如下 MS 结果(图 1)。将处理后得到的 MS 原始数据经过质量校正、扣除角蛋白峰污染峰和胰酶自切峰后,获得用于数据库搜索的 MS 数据。

2.2 PMF 搜索条件的优化

2.2.1 稳定参数的设置 在使用 MS-Fit 搜索时,选择具有蛋白专一性的 Swissprot 作为筛选数据库;Peptide masses 模式根据数据处理时采用同位素校正故设置为同位素(Mononisotopic);种属按蛋白来源选择牛(Bos taurus);PI 设置 6.2~6.6;MW 设置为 26~32ku;我们参考猪胰蛋白酶的酶解充分时间条件,故漏切位点数选择 1 个;Homology Mode 设置为 homology;可能的化学修饰选择默认模式下的 Peptide N-terminal Gln to pyroGlu | Oxidation of M | Protein N-terminus Acetylated |。

2.2.2 主要参数的设置 在质量容错(Mass Tolerance)模式下,引擎提供了 4 种选择,按大小依次为 u、%、ppm 和 mmu,我们分别设置了 0.1u(%)、0.5u

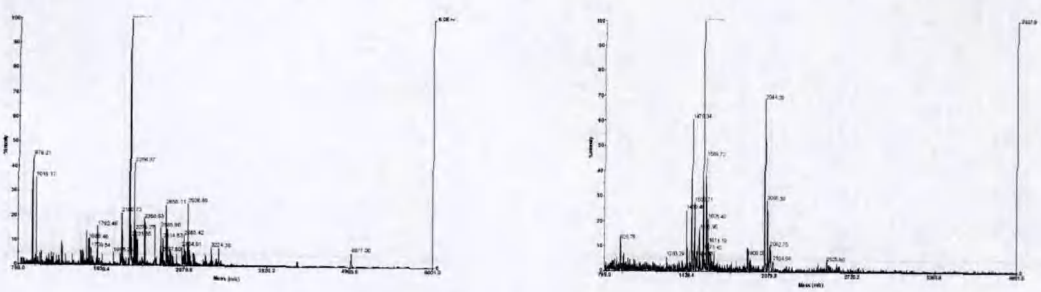


图 1 Carbonic anhydrase2 和 BSA 的 MS 图

维普资讯 http://www.cqvip.com

表 1 MS-Fit 搜索参数设置

方法名	肽质量容错类型	半胱氨酸修饰类型	质量容错数
D1,D4,D7	u	acrylamide	0.1
D2,D5,D8		/carboxymethylation	0.5
D3,D6,D9		/ Carbamidomethylation	1.0
B1,B4,B7	%		0.1
B2,B5,B8			0.5
B3,B6,B9			1.0
P1,P4,P7	ppm		100
P2,P5,P8			500
P3,P6,P9			1 000
M1,M4,M7	mmu		100
M2,M5,M8			500
M3,M6,M9			1 000

说明:1、2、3 编码的是对应 acrylamide 修饰;4、5、6 编码的是对应 carboxymethylation 修饰;7、8、9 编码的是对应 Carbamidomethylation 修饰。

表 3 BSA MS-FIT 检测结果

方法	分子量 / 等电点		蛋白质	登记号	MOWSE 得分	序号	序列覆盖率(%)	肽覆盖率(%)	检出数
	标准值	期望值							
B7	69ku/5.8	69294/5.8	BSA	P02769	1.59e+6	1	53.6	85.0(17/20)	1
D8		69294/5.8	BSA	P02769	25.8	1	8.2	20.0(4/20)	1
P8		69294/5.8	BSA	P02769	45.2	1	10.9	25.0(5/20)	1
M8		69294/5.8	BSA	P02769	25.8	1	8.2	20.0(4/20)	1

(%)、1.0u(%)、100ppm(mmu)、500ppm(mmu)、1 000 ppm(mmu)值来考察对搜索的影响;同时,考察了半胱氨酸 acrylamide、carboxymethylation、Carbamidomethylation 修饰对搜索结果产生的影响。具体参数设置见表 1。

2.3 PMF 搜索结果

2.3.1 MS-Fit 的搜索结果 对 Bovine carbonic anhydrase 2(CA-Ⅱ)的 PMF 数据,按照表 1 所列的各种不同的筛选参数组合在 Swissprot 数据库中进行搜索的结果列于表 2。36 种组合中,当半胱氨酸修饰设为 acrylamide 修饰时,在质量容错值很小的条件下,有 9 种(D1、D4、D7、P1、P4、P7、M1、M4、M7)没搜索到蛋白数据;同时 Carbamidomethylation 比 carboxymethylation 更利于检索蛋白数据;根据序列覆盖率、肽段覆盖率以及检出候选蛋白数等结果表明,容错值 0.1%为最佳,其次是 1u 和 500ppm(mmu)。此外,我们还得出选择百分数作为质量容错模式时,肽段匹配率和序列覆盖率均获得了很好的结果。

2.3.2 MS-Fit 搜索的优化参数 根据以上研究结果可知:(1)当对胶块充分酶解,漏切位点设为 1,选择 Swissprot 数据库做蛋白质 PMF 分析时,Mass tolerance 采用百分数时比较方便筛选候选蛋白。(2)在选择半胱氨酸的修饰时,Carbamidomethylation 比 carboxymethylation 和 acrylamide 相对要容易筛选出候选蛋白。

2.4 蛋白 BSA 检验优化参数结果

将 MS-Fit 搜索条件下的 MW 设为 68~72ku,PI 设为 5.6~6.0。然后使用优化好的参数设置筛选 BSA,均得到了很好的结果,证明了优化的参数是可靠、便捷的。结果见表 3。

3 讨 论

维普资讯 http://www.cqvip.com

表 2 MS-Fit 的搜索结果

方法	MW/PI 分子量 / 等电点		蛋白	登记号	MOWSE 得分	序号	序列覆盖率 (%)	肽段覆盖率 (%)	检出数
	标准值	期望值							
D1	29ku/6.4	29 114/6.4	–	–	–	–	–	–	0
D2		29 114/6.4	CA-Ⅱ	P00921	93926	1	31.9	61(8/13)	1
D3		29 114/6.4	CA-Ⅱ	P00921	1.74e+8	1	56.9	92(12/13)	1
D4		29 114/6.4	–	–	–	–	–	–	0
D5		29 114/6.4	CA-Ⅱ	P00921	117514	1	31.9	61(8/13)	1
D6		29 114/6.4	CA-Ⅱ	P00921	2.54e+8	1	56.9	92(12/13)	1
D7		29 114/6.4	–	–	–	–	–	–	0
D8		29 114/6.4	CA-Ⅱ	P00921	117514	1	31.9	61(8/13)	1
D9		29 114/6.4	CA-Ⅱ	P00921	2.54e+8	1	56.9	92(12/13)	1
B1		29 114/6.4	CA-Ⅱ	P00921	2.27e+9	1	66.9	100(13/13)	1
B2		29 114/6.4	CA-Ⅱ	P00921	5.78e+8	1	70.8	100(13/13)	5
B3		29 114/6.4	CA-Ⅱ	P00921	4.97e+7	1	77.7	100(13/13)	15
B4		29 114/6.4	CA-Ⅱ	P00921	2.30e+9	1	66.9	100(13/13)	2
B5		29 114/6.4	CA-Ⅱ	P00921	5.70e+8	1	70.8	100(13/13)	4
B6		29 114/6.4	CA-Ⅱ	P00921	4.12e+7	1	77.7	100(13/13)	13
B7		29 114/6.4	CA-Ⅱ	P00921	2.72e+9	1	66.9	100(13/13)	1
B8		29 114/6.4	CA-Ⅱ	P00921	6.73e+8	1	70.8	100(13/13)	6
B9		29 114/6.4	CA-Ⅱ	P00921	4.87e+7	1	77.7	100(13/13)	13
P1		29 114/6.4	–	–	–	–	–	–	0
P2		29 114/6.4	CA-Ⅱ	P00921	5.69e+9	1	66.5	100(13/13)	1
P3		29 114/6.4	CA-Ⅱ	P00921	2.27e+9	1	66.9	100(13/13)	1
P4		29 114/6.4	–	–	–	–	–	–	0
P5		29 114/6.4	CA-Ⅱ	P00921	5.75e+9	1	66.5	100(13/13)	1
P6		29 114/6.4	CA-Ⅱ	P00921	2.30e+9	1	66.9	100(13/13)	2
P7		29 114/6.4	–	–	–	–	–	–	0
P8		29 114/6.4	CA-Ⅱ	P00921	6.80e+9	1	66.5	100(13/13)	1
P9		29 114/6.4	CA-Ⅱ	P00921	2.72e+9	1	66.9	100(13/13)	1
M1		29 114/6.4	–	–	–	–	–	–	0
M2		29 114/6.4	CA-Ⅱ	P00921	93926	1	31.9	61(8/13)	1
M3		29 114/6.4	CA-Ⅱ	P00921	1.74e+8	1	56.9	92(12/13)	1
M4		29 114/6.4	–	–	–	–	–	–	0
M5		29 114/6.4	CA-Ⅱ	P00921	117514	1	31.9	61(8/13)	1
M6		29 114/6.4	CA-Ⅱ	P00921	2.54e+8	1	56.9	92(12/13)	1
M7		29 114/6.4	–	–	–	–	–	–	0
M8		29 114/6.4	CA-Ⅱ	P00921	117514	1	31.9	61(8/13)	1
M9		29 114/6.4	CA-Ⅱ	P00921	2.54e+8	1	56.9	92(12/13)	1

蛋白质组学研究中的标准化问题已引起研究人员的高度重视。如果不同实验室间的数据不能互认,其研究结果是没有意义的。为此,人类蛋白质组组织(HUPO)于 2002 年专门成立了一个委员会(HUPO Proteomics standards initiative,PSI)来研究和解决这些问题[1,2]。与经典的蛋白质研究不同,在蛋白质组学的研究工作中,是通过质谱数据与蛋白

(下转第 787 页)

维普资讯 http://www.cqvip.com

参 考 文 献

[1] 江 涌,孙晓川.载脂蛋白 E 多态性与颅脑损伤[J]. 中华创伤杂),2005;21(2):154-157.

[2] Lomnitski L,Oron L,Sklan D,et al. Distinct alterations in phospholipid metabolism in brains of apolipoprotein E-deficient mice [J]. J Neurosci Res,1999;58(4):586-592.

[3] Horsburgh K,McCulloch J,Nilsen M,et al. Intraventricular infusion of apolipoprotein E ameliorates acute neuronal damage after global cerebral ischemia in mice[J]. J Cereb Blood Flow Metab, 2000;20(3):458-462.

[4] Teasdale GM,Murray GD,Nicoll JA,et al. The association between APOE ε4,age and outcome after head injury: a prospective cohort study[J]. Brain,2005;128(11):2556-2561.

[5] Ariza M,Pueyo R,Matarin MM,et al. Influence of APOE polymorphism on cognitive and behavioural outcome in moderate and severe traumatic brain injury[J]. J Neurol Neurosurg Psychiatry, 2006;77(10):1191-1193.

[6] Isoniemi H,Tenovuo O,Portin R,et al. Outcome of traumatic brain injury after three decades--relationship to ApoE genotype[J]. J Neurotrauma,2006;23(11):1600-1608.

[7] Jiang Y,Sun X,Xia Y,et al. Effect of APOE polymorphisms on early responses to traumatic brain injury[J]. Neuroscience letters,2006; 408(2):155-158.

[8] Jiang Y,Sun X,Xia Y,et al. Correlation be tween APOE -491AA promoter in ε4 carriers and severity in early stage of traumatic brain injury(abstract)[J]. J Neurotrauma,2007;24(7):1249.

[9] 江 涌,孙晓川,夏玉先,等. 载脂蛋白 E 基因多态性与颅脑损伤的相关性[J]. 中华创伤杂志,2005;21(7):520-523.

[10] 江 涌,孙晓川,夏玉先,等. 载脂蛋白 E 基因型及启动子区多态性与颅脑损伤病情发展的相关性研究[J]. 中国神经精神疾病杂志,2006;32(5):431-435.

[11] Hatters DM,Peters-Libeu CA,Weisgraber KH. Apolipoprotein E structure:insights into function[J]. Trends Biochem Sci,2006;31 (8):445-454.

(责任编辑:冉明会)

(上接第 772 页)

质数据库的比对来获得鉴定结果的,如果在蛋白质数据库比对时的参数设置不合理,将对搜索结果产生很大的影响。因此,PSI 的第 1 项工作就是建立质谱数据的交换规则和搜索引擎的规范化。

本文正是根据这种思想对采用 MS-Fit 引擎对肽质量指纹谱鉴定蛋白质时搜索条件的优化进行了探索。通过用牛 Carbonic anhydrase 2 蛋白进行搜索方法优化,得出:①在蛋白酶解充分,漏切位点设为 1,选择 Swissprot 数据库做蛋白质 PMF 分析时,质量容错值模式采用百分数(%)时比较 u、ppm、mmu 更加方便、科学,其原因在于 MALDI-TOF-MS 鉴定质量数越大的蛋白或者肽段时,其校正误差的能力更加薄弱,肽段准确性随质量大小呈现类线性远离标准值趋势。因此采用质量百分数较为理想。②容错值大小的选择应该根据质谱仪准确度校正的最大误差和标准蛋白筛选的最佳数值为准,不应随机选择,如本文筛选出了 0.1 个百分数值时为最佳。③在选择半胱氨酸的修饰时,Carbamidomethylation 比 carboxymethylation 和 acrylamide 相对要容易筛选出候选蛋白,这个跟蛋白在尿素存在情况下极易脲基甲酰化相符合的,同时也说明了丙烯酰胺化和羧甲基化对半胱氨酸的修饰作用不如脲基甲酰化显著。蛋白质搜索引擎主要参数的优化,为后续鉴定蛋白质的筛选提供了可靠、快捷、高效的保证。

参 考 文 献

[1] Orchard S,Zhu W,Julian RKJ,et al. Further advances in the development of a data interchange standard for proteomics data[J]. Proteomics,2003;3(10):2065-2066.

[2] RaiA J,Gelfand CA,Haywood BC,et al. HUPO Plasma Proteome Project specimen collection and handling:Towards the standardization of parameters for plasma proteome samples[J]. Proteomics, 2005;5(13):3262-3277.

[3] 母昭德,彭咏波,易发平,等.血清蛋白质组双向凝胶电泳技术的建立[J]. 重庆医科大学学报,2006;31(6):821-823.

[4] Rosenfeld J,Capdevielle J,Guillemot JC,et al. In-gel digestion of proteins for internal sequence analysis after one- or two-dimensional gel electrophoresis[J]. Anal. Biochem,1992;203(1):173-179.

[5] Hellman U,Wernstedt C,Gonez J,et al. Improvement of an "In-Gel" digestion procedure for the micropreparation of internal protein fragments for amino acid sequencing[J]. Anal. Biochem,1995; 224(1):451-455.

[6] 夏其昌,曾 嵘.蛋白质化学与蛋白质组学[M]. 北京:科学出版社,2004.318-320.

(责任编辑:谭乐勤)