

应用 IP-2D nano-HPLC-MALDI-TOF-TOF 鉴定蛋白质泛素化修饰*

毛建华^{1, 2, 3)} 张群业²⁾ 黄秋花²⁾ 陈竺²⁾ 陈隼¹⁾ 陈赛娟^{1, 2)**}

(¹⁾中国科学院动物研究所, 北京 100101; ²⁾上海交通大学医学院附属瑞金医院上海血液学研究所,
医学基因组学国家重点实验室, 上海 200025; ³⁾中国科学院研究生院, 北京 100049)

摘要 近年来, 在蛋白质研究中, 特别是在蛋白质翻译后修饰(PTM)的研究中, 生物质谱技术的应用越来越广泛, 与纳升级 HPLC 的联合应用, 使这一技术手段更加有效. 针对泛素化在细胞功能调控中发挥关键作用的 PTM 的特点, 将免疫沉淀、2D nano-HPLC 和基质辅助激光解吸/电离串联飞行时间质谱(IP-2D nano-HPLC-MALDI-TOF-TOF)有机整合, 建立了天然状态下蛋白质泛素化位点的鉴定方法, 并应用这一方法确定出 K562 细胞内具有酪氨酸激酶活性的蛋白 c-ABL 的泛素化位点. 为定性鉴定生理和病理状态下内源性蛋白的泛素化修饰提供了借鉴.

关键词 泛素化, 蛋白质翻译后修饰, IP-2D nano-HPLC-MALDI-TOF-TOF

学科分类号 Q7, R3

高效液相色谱 (high performance liquid chromatography, HPLC 或 LC) 问世于 20 世纪 70 年代, 经过三十多年的发展, 已成为分离分析领域的核心技术手段^[1], 近年来在生物医学领域得到了广泛应用, 在分离包括蛋白质和多肽在内的生物大分子方面发挥着独特的优势^[2]. 生物质谱技术在 20 世纪 90 年代早期出现后, 极为迅速地用于生命科学领域, 不仅在蛋白质鉴定和分子质量精确测定等方面发挥了重要作用^[3, 4], 而且, 为蛋白质翻译后修饰的精细研究提供了强大的技术支持^[5, 6]. HPLC 与质谱 (MS) 尤其是串联质谱 (MS/MS) 联用实现了分离与鉴定的有效结合, 极大地推动了蛋白质功能的研究、生物标志的筛选和药物的研发等领域的发展^[7~9].

泛素 (ubiquitin, Ub) 是一类在真核细胞中普遍存在的由 76 个氨基酸组成的高度保守的小分子, 编码蛋白分子质量为 8.5 ku. 泛素-蛋白酶体系统是细胞内绝大多数蛋白质降解的主要方式之一^[10~12], 由于在这一领域的重要贡献, Ciechanover A, Hershko A 和 Rose I 等三位科学家获得了 2004 年的诺贝尔化学奖. 泛素单体和 / 或寡聚泛素链能够在 E1 (ubiquitin activation enzyme), E2 (ubiquitin-

conjugation enzyme) 和底物特异的 E3 (ubiquitin ligation enzyme) 作用下通过其 C 端甘氨酸的氨基酸以共价键连接到底物蛋白 (包括泛素分子本身) 赖氨酸的 ϵ 氨基上. 泛素分子第 48 位赖氨酸介导的多聚链所结合的底物蛋白通常会被蛋白酶体识别而发生降解^[10, 13]. 细胞内的蛋白质时刻处在产生、发挥功能和降解的动态平衡中, 蛋白质的泛素化修饰是与蛋白质降解和功能调控密切相关的翻译后修饰形式, 对它的深入研究能够帮助我们更好地认识细胞周期调控、增殖与分化、凋亡以及蛋白质周转代谢的机制^[14, 15]. 因此, 人们对研究蛋白质泛素化的方法和技术进行了很多探索, 如 Wang 和 Kirkpatrick 等^[16, 17]建立了针对体外合成多肽或过表达富集的蛋白质泛素化位点的鉴定方法, 这些方法的建立极大地推动了泛素化位点鉴定的研究, 但尚未见报道用于鉴定体内天然状态下蛋白质的这种特异修饰. 由

* 国家自然科学基金资助项目 (30600262, 30772744, 30572143, 30623010) 和上海市教育委员会重点学科资助项目 (Y0201).

** 通讯联系人. Tel: 021-64370045-610601, Fax: 021-64743026

E-mail: sjchen@stn.sh.cn

收稿日期: 2008-03-02, 接受日期: 2008-04-08

于细胞内蛋白质的泛素化修饰处于动态变化之中, 给其信息的捕捉带来了一定的难度, 因此, 需要建立新的方法来攻克这一难关.

发生泛素化的底物蛋白在胰蛋白酶的作用下, 连接泛素链的赖氨酸残基上会产生特异的分子质量相差 114 u 的 GG-K 信号肽^[18, 19], 可被质谱鉴定. 根据这一特点, 我们利用免疫沉淀 (immunoprecipitation, IP) 对低丰度蛋白的纯化和富集的优势, 结合 2D nano-HPLC 多肽分离和 MALDI-TOF-TOF 对蛋白质及修饰鉴定的优势, 建立了较低丰度内源性蛋白泛素化修饰的鉴定方法, 并利用这一方法在细胞内具有酪氨酸激酶活性的 c-ABL 蛋白上检测到了一个泛素化位点, 为认识 c-ABL 蛋白的转归提供了帮助, 同时为其他形式的蛋白质翻译后修饰 (SUMO 化、甲基化及乙酰化等) 以及定性鉴定生理及病理状态下内源性蛋白的泛素化修饰提供了借鉴.

1 材料与方法

1.1 材料

1.1.1 主要试剂. Poly-Ub chain K48 (Bostonbiochem 公司); 甲苯磺酰苯丙氨酰氯甲酮 - 胰蛋白酶 (TPCK-Trypsin) 购自 Promega 公司; c-ABL 单克隆抗体 (AbCam 公司); Protein G sepharose 4 Fast Flow (GE Healthcare 公司); YM-50 超滤柱 (Millipore 公司); 二硫苏糖醇 (DTT)、碘乙酰胺 (IAA) 购自 Promega 公司; 甲酸胺、甲酸购自 Fluka 公司; 乙腈 (Fisher 公司); HPLC 用水 (Merck 公司); 基质 α -氰基 -4- 羟基肉桂酸 (CHCA) 购自 Sigma 公司; 蛋白酶抑制剂 cocktail tablets (Roche 公司); 去泛素化酶抑制剂 (N-ethylmaleimide, NEM) 及蛋白酶体抑制剂 (MG132) 均购自 Calbiochem 公司; 胎牛血清购自 Hyclone 公司.

1.1.2 主要仪器及软件. SHIMADZU 公司 2D nano-HPLC 20 型色谱仪, AB 公司 4700 型 MALDI-TOF-TOF 质谱仪, LCsolution software, GPS Explorer™ V3.5 software, 4000 Series Explorer™ software.

1.1.3 细胞株. K562 (人慢性粒细胞白血病细胞株) 购自 ATCC 公司.

1.2 方法

1.2.1 MS/MS 鉴定 Poly Ub K48 的泛素化修饰位点. 将 Poly Ub K48 用 5 mmol/L NH_4HCO_3 配成 8.4×10^{-2} mmol/L 的储液, 取 336 pmol/L 按胰蛋白

酶与蛋白质的质量比为 1 : 50 的比例加入胰蛋白酶, 37℃ 酶解过夜. 取酶解后的 Poly Ub K48 稀释至 840 fmol/L 与 CHCA 按 1 : 1 的比例混匀点于 AB 4700 靶板上, 真空抽干样品后用 4700 质谱仪进行鉴定. 校正仪器质量误差在 0.1 u, 先进行肽质量指纹图 (peptide mass fingerprint, PMF) 分析, 将质谱实际鉴定到的肽段与 <http://prospector.ucsf.edu> 网站上对 Ub 进行电子酶解的肽段进行比对, 发现包含 K48 的肽段 1 346.742 7 变成了 1 460.769 4, 二者刚好相差 114 u, 因此选择 1 460.769 4 作为母离子进行串联质谱鉴定. 同时在 GPS 数据库的搜索引擎 Mascot 的 modification 文件中添加了 GG-K (Residues: K 242.09497) 修饰, 根据 GPS 搜索提供的 Peptide view 的信息对串联质谱图进行数据库自动搜索辅助人工解谱.

1.2.2 IP 分离纯化 c-ABL 蛋白. 悬浮培养 K562 细胞, 收细胞前 4 h 加入 5 $\mu\text{mol/L}$ 蛋白酶体抑制剂 MG132 处理, 收集约 5×10^7 细胞, PBS 洗 3 次后, 加入 RIPA 裂解液 (50 mmol/L Tris-HCl pH7.4, 150 mmol/L NaCl, 1% NP-40, 0.25% 脱氧胆酸钠, 0.1% SDS) 添加蛋白酶抑制剂 cocktail 以及 10 mmol/L NEM, 细胞充分裂解后, 经 4℃, 12 000 g 离心 15 min, 收集上清, 加入 Protein G sepharose 预处理 4 h, 12 000 g 离心 2 min, 收集上清, 加入 c-ABL 单克隆抗体, 4℃ 旋转混匀过夜. 次日, 用 0.5% NP-40 裂解液 (50 mmol/L Tris-HCl pH7.4, 150 mmol/L NaCl, 0.5% NP-40) 洗 Sepharose 3 次, 5 000 g 离心 2 min, 弃上清, 最后一次洗完, 加入 2 $\times$ 上样缓冲液 (1 mol/L Tris-HCl pH7.4, 4% SDS, 100 mmol/L DTT) 100℃ 煮沸 15 min 将结合的蛋白质洗脱下来, 12 000 g 离心 2 min 收集上清, 将上清转移至 YM-50 超滤离心管中, 10 000 g 离心 10 min, 去除免疫重链、轻链及小分子杂蛋白.

1.2.3 c-ABL 蛋白的酶解. IP 得到的蛋白质组分经 100 mmol/L DTT 还原 30 min, 400 mmol/L IAA 烷基化避光处理 30 min, 蛋白质样品经甲醇 - 氯仿法沉淀除盐 (150 μl 蛋白质 : 600 μl 甲醇 : 150 μl 氯仿 : 450 μl 三蒸水), 12 000 g 离心 5 min, 吸取中间的白色沉淀, 用 600 μl 甲醇洗 2 次后, 重悬于 30 μl 5 mmol/L NH_4HCO_3 , 按 1.2.1 提到的比例加入胰蛋白酶, 37℃ 消化过夜.

1.2.4 2D nano-HPLC 分离、收集多肽. 酶解后多肽经 0.22 μm 滤膜 (Millipore) 过滤后进行 2D nano-HPLC 分离. HPLC 流动相的配制: 第一

相强阳离子(SCX)交换, SCX 缓冲液 A(5 mmol/L 甲酸胺, 0.02%甲酸), SCX 缓冲液 B(300 mmol/L 甲酸胺, 1.2%甲酸); 第二相反相分离(RP), RP 缓冲液 A(0.1%甲酸, 95%水, 5%乙腈), RP 缓冲液 B(0.1%甲酸, 5%水, 95%乙腈). 样品先经过 SCX 分离, SCX 梯度洗脱盐浓度为: 0%、15%、40% 及 100%. SCX 洗脱下来的组分在 trap 柱上收集后再经第二相 RP 柱分离, 2%缓冲液 B 冲洗 12 min, 在 12~50 min 内进行 2%~55%缓冲液 B 梯度洗脱, 50~53 min 进行 55%~70%缓冲液 B 梯度洗脱, 53~58 min 保留在 70%缓冲液 B, 58~65 min 内快速由 70%降至 2%的缓冲液 B, 然后用 2%缓冲液 B 冲洗平衡系统 10 min. SCX 0%的组分经 RP 分离后, 再分离 SCX 15%的组分, 依次分离 SCX 的 4 个组分. 根据 UV 检测器检测到的峰, 自动点靶仪(Accuspot)将出峰的组分自动地点到 AB 4700 靶板上.

1.2.5 MS/MS 鉴定 c-ABL 及泛素化修饰位点. 将 CHCA 基质点在样品上, 真空抽干样品后用 4700 质谱仪进行鉴定, 用 GPS 远程控制的方法设置 LC-MALDI 模式的 MS 及 MS/MS 的方法文件, 使用正离子反射模式测定, 每张谱图 MS 设置打击 1 000 次, 检测器采集时间间隔设为 0.5 ns, 离子源加速电压为 1 kV, 采集质量范围为 800~4 000, 聚焦质量为 2 000, 采用 4700 专用的质量标准作外标校正; MS/MS 设置打击 4 000 次, TIS 离子选择门设为 60, 根据 PMF 的信息, 自动选取 S/N(信噪比)大于 70 的肽段作为母离子进行 MS/MS 串联质

谱, 然后自动搜索 Swissprot(2007 年 4 月下载)数据库, 可变修饰除了选择半胱氨酸为脲甲基半胱氨酸(carbamidomethyl-Cys)、甲硫氨酸氧化和 N 端的谷氨酸焦性没食子酸(Pyro-glu)化外, 还选择自己添加进去的 GG-K 进行搜索, 根据 Mascot 搜索引擎得到的 Peptide view 的信息, 对串联谱图进行数据库自动搜索辅助人工解谱.

2 结 果

2.1 MS/MS 鉴定 Poly Ub K48 的泛素化修饰位点

2.1.1 质谱鉴定蛋白质发生泛素化修饰的特征性 GG-K 信号. Ub 分子的氨基酸序列如图 1a, 其中共有 7 个赖氨酸(K)已用黑体加下划线标记, C 端的两个氨基酸为甘氨酸(GG), 用黑体斜体标记. 图 1b 为多聚泛素链或连有多聚泛素链的靶蛋白经胰蛋白酶酶解后的模式图, 胰蛋白酶可以特异地识别赖氨酸和精氨酸(R)并且在羧基端将肽键切断. 如图 1b 所示, Ub 分子 C 端第 76 位的 G 与 Ub 分子的赖氨酸(K)或者 Ub 分子与靶蛋白的赖氨酸(K)通过共价键形成稳固的结合, 在胰蛋白酶的作用下, 氨基酸链在 R 所在的位置被切断, 这样发生泛素化修饰的氨基酸链上就会多出 GG-K(图 1b 中 GG 用黑体加下划线标记)的信号肽^[19, 20], 作为判定靶蛋白发生泛素化修饰特异的识别信号, 经 MS/MS 分析可以确定发生泛素化修饰的 K 的位置. 本文就是以此作为依据, 以 Poly Ub^[2~7] K48 作为研究对象, 建立了鉴定蛋白质泛素化修饰的技术方法.

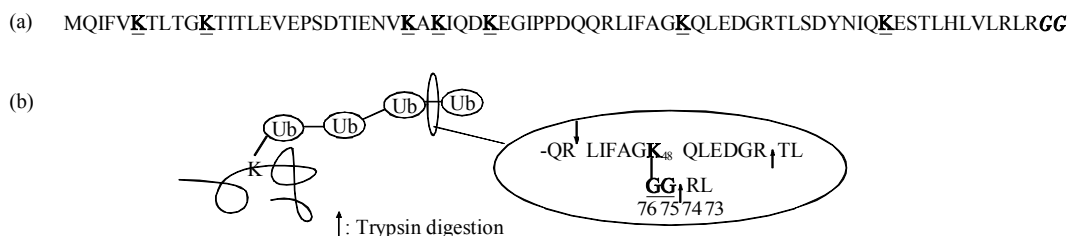


Fig. 1 Amino acid sequence of Ub molecular and model of trypsin digestion of poly Ub chain linked on targeted protein

(a) Amino acid sequence of Ub molecule, the lysine of Ub molecule was marked by black and underline, and the C terminal glycine was marked by black and italic. (b) Model of trypsin digestion of poly Ub chain linked on targeted protein, the signal peptide of ubiquitination was marked by black and underline.

2.1.2 PMF 上找到 1 460.769 4 母离子. 表 1 为 Ub 分子经电子酶解得到的肽段中包含 K48 的肽段, 这些肽段上的 K 都有连接 GG(114 u, $M(G) =$

57.05 u)的机会. 图 2a 为 4700 型 MALDI-TOF 质谱仪对 Poly Ub K48 进行鉴定得到的 PMF 谱图, 胰蛋白酶酶解的片段主要集中分布在 $m/z=800 \sim$

2 000 之间, 这是由于 Ub 的分子质量只有 8.5 ku 的缘故, 图 2 中峰的数量较少, 说明样品较纯, 峰信号较强, 信噪比高, 可以清晰地看到 1 460.769 4 这个肽段, 恰好为 1 346.742 7 加上两个 GG 的分子质量(114)得到的肽段, 说明发生泛素化修饰的

K 就在 1 460.769 4 这个母离子上. 图 2b 为 PMF 经 Mascot 数据库搜索的结果, 蛋白质的分数大于 30 分是可信的, 而搜索到 Ub 分子的得分为 57 分, 说明鉴定的结果可靠, 即 1 460.769 4 母离子是来自 Ub 分子.

Table 1 Trypsin digestion of Ub molecular based on prediction

<i>m/z</i>	Start	End	Missed cleavages	Sequence
1 346.742 7	43	54	1	(R)LIFAGKQLEDGR(T)
1 668.906 8	34	48	1	(K)EGIPPDQQRLLIFAGK(Q)
2 153.171 3	30	48	2	(K)IQDKEGIPPDQQRLLIFAGK(Q)
2 367.241 5	34	54	2	(K)EGIPPDQQRLLIFAGKQLEDGR(T)
2 409.277 3	43	63	2	(R)LIFAGKQLEDGRTLSDYNIQK(E)

Only the peptides containing K48 were shown in this table.

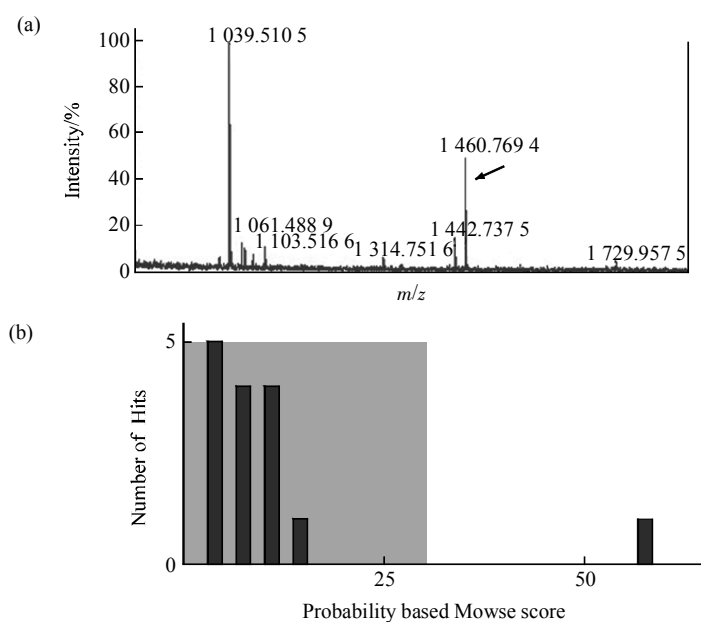


Fig. 2 PMF and database query result of Poly Ub K48

(a) PMF of poly Ub K48. The peak of 1 460.769 4 was observed and marked. (b) Database query result. Ion score is $-10 \lg P$, where P is the probability showing the observed match is a random event. Protein scores greater than 30 indicate identity or extensive homology ($P < 0.05$), here the ion score of poly Ub K48 is 57.

2.1.3 MS/MS 谱图上检测到 Poly Ub K48 发生泛素化修饰. 在离子化或碰撞过程中所产生的能量会使肽键发生断裂, 断裂后靠近 N 端的碎片称为 b 离子, 靠近 C 端的碎片称为 y 离子, 根据 b 离子系列或 y 离子系列提供的信息, 可以获得从 N 端到 C 端或者从 C 端到 N 端的氨基酸序列信息, 这就是 MS/MS 能够对氨基酸进行测序的原理^[20]. 基于这一原理, 选择包含 K48-GG 的肽段 1 460.769 4

作为母离子进行 MS/MS 及 Mascot 数据库搜索, 得到了 y 离子和 b 离子两个系列的离子, 由表 2 可以看出从 y1~y12 的 11 个离子谱图中可以看到 8 个 (y1, y2, y3, y4, y5, y6, y9, y10)(黑体标记), 而 y7 和 y8 脱掉一个水分子的离子(942.42 和 999.44)也可以在谱图上看到, 说明在能量的作用下离子碎裂很均匀, 产生了各相差一个氨基酸的离子系列, 因此 MS/MS 谱图质量较高, 可以用作数据库自动搜索

辅助人工解谱. 图 3a 为 MS/MS 谱图进行 GPS 数据库搜索的信息, 由于 Mascot 的 modification 文件中已经添加 GG-K 的修饰信息, 因此自动检索到了可变修饰 GG-K. 图 3b 为根据 MS/MS 谱图和 Peptide view 提供的信息进行数据库自动搜索辅助人工解谱, 图中可见, $y_{11}-y_{10}(1\,347.66-1\,234.58=113)$, 根据氨基酸的分子质量, 可以判定这个位置为异亮氨酸(I), 按照同样的方法可以得到这个肽段的氨基酸序列信息. 而 $y_7-y_6(942.42-717.35=$

225), 恰好为赖氨酸($M(K)=128.17$)的分子质量加上 GG(114)减去一个水分子 ($M(H_2O)=18$)再加上一个质子的分子质量, 由此可以断定 y_6 对应的 K 发生了泛素化修饰, 根据解谱得到的氨基酸序列与 Ub 分子的氨基酸序列比对, 就可以明确是 Ub 分子的 K48 发生了泛素化修饰, 而我们使用的 Poly Ub 就是通过 K48 形成的多聚泛素链. 至此, 我们利用已知的多聚泛素链建立了鉴定蛋白质泛素化修饰的技术方法.

Table 2 GPS search and peptide view of Poly Ub K48

	Immon	a	a*	a ⁰	b	b*	b ⁰	Seq	y	y*	y ⁰	#
1	86.1	86.1			114.09			L				12
2	86.1	199.18			227.18			I	1 347.66	1 330.63	1 329.65	11
3	120.08	346.25			374.24			F	1 234.58	1 217.55	1 216.56	10
4	44.05	417.29			445.28			A	1 087.51	1 070.48	1 069.5	9
5	30.03	474.31			502.3			G	1 016.47	999.44	998.46	8
6	215.11	716.4	699.38		744.4	727.37		K	959.45	942.42	941.44	7
7	101.07	844.46	827.44		872.46	855.43		Q	717.35	700.33	699.34	6
8	86.1	957.55	940.52		985.54	968.51		L	589.29	572.27	571.28	5
9	102.06	1 086.59	1 069.56	1 068.58	1 114.58	1 097.56	1 096.57	E	476.21	459.18	458.2	4
10	88.04	1 201.62	1 184.59	1 183.6	1 229.61	1 212.58	1 211.6	D	347.17	330.14	329.16	3
11	30.03	1 258.64	1 241.61	1 240.63	1 286.63	1 269.6	1 268.62	G	232.14	215.11		2
12	129.11							R	175.12	158.09		1

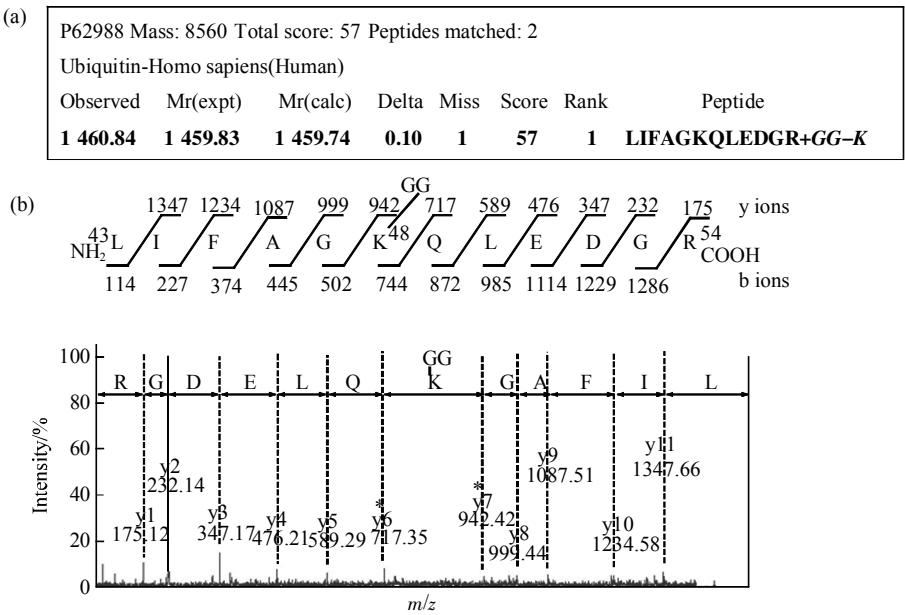


Fig. 3 Mascot search results and MS/MS analysis of Poly Ub K48 linked GG

(a) Mascot search results. Variable modification of GG-K was observed from the parent ion 1 460.84. (b) MS/MS analysis results. The signal peptide of ubiquitination was marked.

2.2 IP-2D-nano-HPLC 技术分离 c-ABL 蛋白

应用 SHIMADZU 公司的 2D nano-HPLC 液相色谱仪建立了多肽分离技术平台, 通过优化条件, 得到了分布均匀, 分离效率较好的色谱峰. 在此基础上, 利用 c-ABL 单克隆抗体通过 IP 技术分离纯化包含 c-ABL 蛋白在内的复合物, 经胰酶酶解后用优化条件的 2D nano-HPLC 分离, 得到的色谱峰如图 4, 组分 1 为 0% 盐离子梯度洗脱后经第二维

反相(C18)柱分离的组分, 组分 1 分离的峰并不是很多, 而且强度也不是很强, 因为盐离子强度较弱的原因, 组分 2、3、4 分别为 15%、40% 和 100% 的盐离子梯度洗脱后经反相柱分离的组分, 可见这 3 个盐梯度洗脱峰数量较多, 峰分离效率较好, 强度较强, 说明包含的多肽数量较多, 分离到的每一个组分自动点到 ABI 4700 的靶板上, 用于接下来的质谱鉴定.

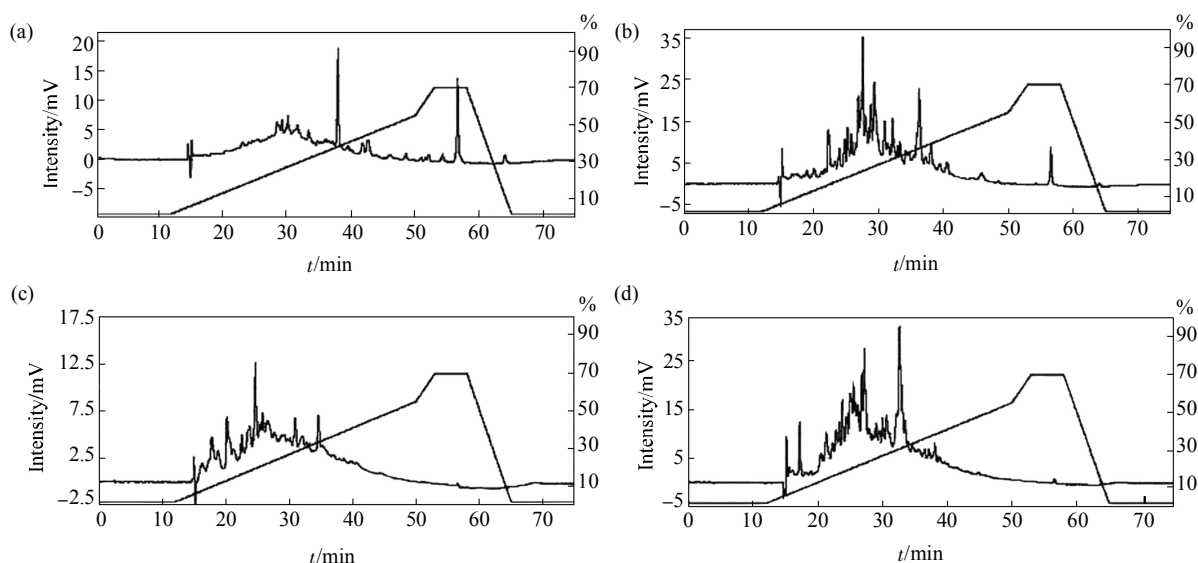


Fig. 4 2D-nano-HPLC separation of c-ABL IP complex

(a) Fraction 1: c-abl IP and 2D HPLC separation (0% gradient for SCX and gradient elution for RP). (b) Fraction 2: c-abl IP and 2D HPLC separation (15% gradient for SCX and gradient elution for RP). (c) Fraction 3: c-abl IP and 2D HPLC separation (40% gradient for SCX and gradient elution for RP). (d) Fraction 4: c-abl IP and 2D HPLC separation (100% gradient for SCX and gradient elution for RP).

2.3 MS/MS 鉴定 c-ABL 上的泛素化修饰位点

液相色谱分离到的组分经 4700 型 MALDI-TOF-TOF 进行质谱鉴定, 采用 GPS 远程控制的方法, 先进行 PMF 检测, 然后按照设定的参数符合条件的肽段自动选作母离子进行第二级质谱鉴定, 最后自动搜索 GPS 数据库, 可变修饰选择 GG-K. 如图 5a, 在经 IP 纯化, HPLC 分离的组分中找到了 c-ABL (Proto-oncogene tyrosine-protein kinase ABL1), 虽然序列覆盖率不是很高, 但是实际上得到了 N 端从 1 到 12 这段氨基酸序列(图中加黑体加下划线的序列), 因此结果是可信的, 并且在赖氨酸上发现了 GG-K(图 5a 中黑体加下划线)修饰. 为

了进一步明确发生泛素化修饰的 K 的位置, 根据 GPS 搜索信息从 4700 质谱仪调取 c-ABL 的 PMF 和 MS/MS 谱图, 如图 5b, 在 PMF 谱图上我们可以清晰地观察到 1 537.671 1 这个峰, 由于质谱仪的质量误差设为 0.1 u, 因此这个峰就是图 5a 中黑体加下划线标记出来的 1 537.77 包含了 GG-K 肽段所对应的峰, 然后调取以 1 537.671 1 作为母离子的 MS/MS 谱图, 如图 5c, y 离子系列中的 y4, y5, y8 及 b 离子系列中的 b7 在谱图可以找到对应的峰, 结合 GPS 提供的氨基酸序列信息进行数据库自动搜索辅助人工解谱, 可以判定发生泛素化修饰的为 K12.

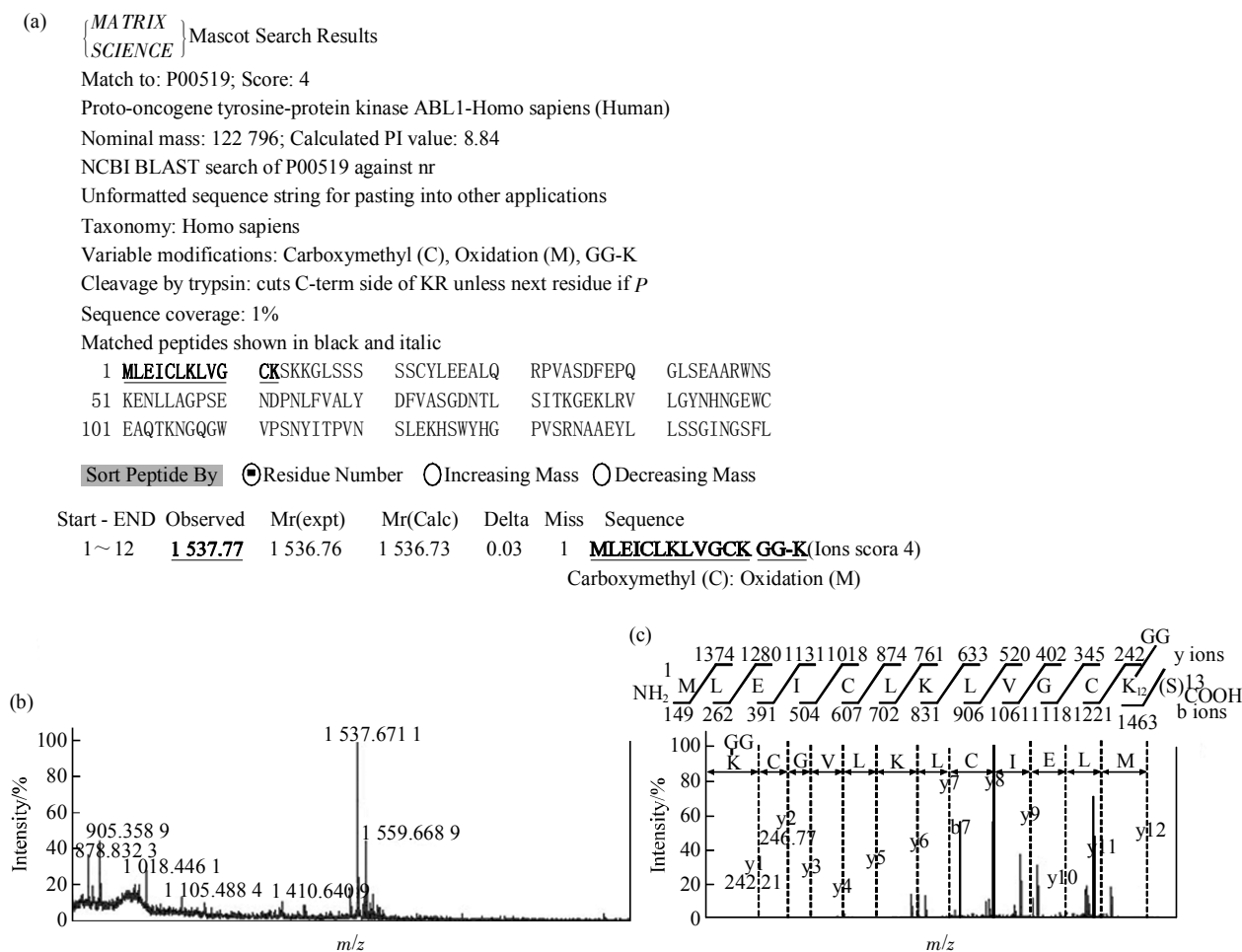


Fig. 5 Identification of ubiquitination modification site on c-ABL

(a) Mascot search results. GG-K on 1 537.671 1 was observed. (b) PMF of c-ABL. The peak 1 537.671 1 was selected as the parent ion for MS/MS. (c) MS/MS analysis results. The signal peptide from K12 was observed and marked.

3 讨 论

蛋白质组学的发展经历了从单纯的表达谱鉴定^[21], 到先进行预分离的亚蛋白质组学^[22], 再到功能蛋白质组学的演变^[23], 尤其是目前对蛋白质翻译后修饰的研究越来越受到人们的重视^[24]. 单纯数据驱动的研究目前在技术上还有难以克服的困难, 但是将其所具有的对完全未知研究对象的强大探索能力和经典的细致功能学研究技术结合, 必将有助于克服很多常规功能学研究中的困难, 加深人们对于许多科学问题的认识. 蛋白质翻译后修饰在细胞各种功能的行使中发挥着重要作用. 常规功能学研究方法通常只能确定少数蛋白质发生的某些修饰, 然而, 作为细胞功能调控的最重要手段之一, 一种蛋白质常常同时存在着多种修饰(如磷酸化、泛素化、SUMO 化和乙酰化等修饰常可发生于同一蛋白

质)^[25], 这就给修饰类型和位点的确定带来了很大困难. 由于蛋白质发生修饰后其分子质量会发生相应的改变, 而质谱能够精确测定蛋白质或多肽分子质量, 根据这一特性, 人们开始将质谱用于蛋白质翻译后修饰的鉴定^[3-6]. 目前对蛋白质磷酸化和糖基化修饰鉴定的研究较多^[26,27], 而对蛋白质泛素化^[28]等修饰的研究还相对较少. Wang 和 Bauer 等^[16,29]通过对合成多肽进行体外磺酸化反应, 使其易于产生信 / 噪比(S/N)较高的完整的 y 离子系列谱图, 提高了质谱鉴定的效率, 为蛋白质泛素化位点的鉴定提供了技术上的借鉴. 但是在这类有机化学修饰反应中, 需要的蛋白质量往往较大, 并存在一些副反应, 对鉴定可能造成一定干扰, 目前并未见到有应用此方法对内源性蛋白进行鉴定的报道, 因此本文在前人研究的基础上建立了内源性蛋白泛素化修饰的鉴定方法, 并借助该方法对细胞内具有酪氨酸激

酶活性的蛋白 c-ABL 的泛素化位点进行了鉴定。最近基于 Paragon 新算法的数据库搜索软件 ProteinPilot™ 可以自动检测 150 多种蛋白质的翻译后修饰, 给蛋白质的翻译后修饰研究带来了极大的便利^[30]。

IP 是一种传统的蛋白质分离、纯化和富集的技术, 但是纯化得到的蛋白质总量较少, 成分仍然比较复杂, 需要进一步的分离和富集。传统的双向电泳可以用于 IP 产物的分离, 但是对分子质量极大、极小和一些酸碱度极酸、极碱的蛋白质分离效果不佳, 对丰度较低的蛋白质的检测灵敏度不够。与之相比, 2D nano-HPLC 检测灵敏度大为提高, 而且可以根据所选柱子的特点, 如凝胶过滤柱、离子交换柱及反相柱等对蛋白质进行多维分离, 与传统的双向电泳技术具有一定的互补性, 并且可以将所分离的蛋白质或多肽自动收集到 MALDI 板上或者直接经 ESI 质谱进行鉴定, 减少了蛋白质的损失, 在一定程度上实现了对某些低丰度蛋白的富集。将 HPLC 用于 IP 纯化得到的蛋白质或酶解多肽的再分离, 可以大大地提高质谱鉴定的成功率。

综上所述, 我们利用 IP 纯化蛋白的特性、HPLC 富集和分离多肽的优势及质谱鉴定蛋白质翻译后修饰的优势, 建立了对内源性蛋白, 尤其是某些低丰度蛋白或多肽的高效分离、纯化和鉴定方法, 并且应用相应的数据库自动搜索辅以人工解读, 对蛋白质泛素化修饰作出精确解读。利用此方法, 我们对 c-ABL 蛋白的泛素化位点进行了鉴定, 同时为其他形式的蛋白质翻译后修饰以及生理、病理状态下内源性蛋白的泛素化修饰的定性鉴定提供了借鉴。

致谢 感谢翁林教授在论文撰写和修改过程中给予的指导 and 帮助。

参 考 文 献

- 1 陈立仁, 蒋生祥, 刘霞, 等. 高效液相色谱基础与实践. 北京: 科学出版社, 2001. 2~5
- 2 Chen L R, Jiang S X, Liu X, *et al.* Base and Practice of HPLC. Beijing: Science Press, 2001. 2~5
- 3 Aguilar M I. HPLC of peptides and proteins: methods and protocols. Humana Press, 2004, **251**: 3~7
- 4 Mann M, Højrup P, Roepstorff P. Use of mass spectrometric molecular weight information to identify proteins in sequence databases. Biol Mass Spectrom, 1993, **22** (6): 338~345
- 5 Yates J R 3rd, Speicher S, Griffin P R, *et al.* Peptide mass maps: a highly informative approach to protein identification. Anal Biochem, 1993, **214** (2): 397~408
- 6 Mann M, Jensen O N. Proteomic analysis of post-translational modifications. Nat Biotechnol, 2003, **21**(3): 255~261
- 7 Jensen O N. Modification-specific proteomics: characterization of post-translational modifications by mass spectrometry. Curr Opin Chem Biol, 2004, **8**(1): 33~41
- 8 Beck H C, Nielsen E C, Matthiesen R, *et al.* Quantitative proteomic analysis of post-translational modifications of human histones. Mol Cell Proteomics, 2006, **5**(7): 1314~1325
- 9 Haskins W E, Kobeissy F H, Wolper R A, *et al.* Rapid discovery of putative protein biomarkers of traumatic brain injury by SDS-PAGE-capillary liquid chromatography-tandem mass spectrometry. J Neurotrauma, 2005, **22**(6): 629~644
- 10 Chen E I, Hewel J, Felding-Habermann B, *et al.* Large scale protein profiling by combination of protein fractionation and multidimensional protein identification technology (MudPIT). Mol Cell Proteomics, 2006, **5**(1): 53~56
- 11 Herskko A, Ciechanover A. The ubiquitin system. Annu Rev Biochem, 1998, **67**: 425~479
- 12 Bochtler M, Ditzel L, Groll M, *et al.* The proteasome. Annu Rev Biophys Biomol Struct, 1999, **28**: 295~317
- 13 Ciechanover A, Elias S, Heller H, *et al.* Characterization of the heat-stable polypeptide of the ATP-dependent proteolytic system from reticulocytes. J Biol Chem, 1980, **255**(16): 7525~7528
- 14 Pickart C M. Mechanisms underlying ubiquitination. Annu Rev Biochem, 2001, **70**: 503~533
- 15 Ciechanover A, Iwai K. The ubiquitin system: from basic mechanisms to the patient bed. IUBMB Life, 2004, **56**(4): 193~201
- 16 Varshavsky A. Regulated protein degradation. Trends Biochem Sci, 2005, **30**(6): 283~286
- 17 Wang D, Kalume D, Pickart C, *et al.* Identification of protein ubiquitylation by electrospray ionization tandem mass spectrometric analysis of sulfonated tryptic peptides. Anal Chem, 2006, **78**(11): 3681~3687
- 18 Kirkpatrick D S, Weldon S F, Tsapralis G, *et al.* Proteomic identification of ubiquitinated proteins from human cells expressing His-tagged ubiquitin. Proteomics, 2005, **5** (8): 2104~2111
- 19 Marotti L A Jr, Newitt R, Wang Y, *et al.* Direct identification of a G protein ubiquitination site by mass spectrometry. Biochemistry, 2002, **41**(16): 5067~5074
- 20 Cooper H J, Heath J K, Jaffray E, *et al.* Identification of sites of ubiquitination in proteins: a fourier transform ion cyclotron resonance mass spectrometry approach. Anal Chem, 2004, **76**(23): 6982~6988
- 21 Carpenter K H, Wiley V. Application of tandem mass spectrometry to biochemical genetics and newborn screening. Clin Chim Acta, 2002, **322** (1-2): 1~10
- 22 Florens L, Washburn M P, Raine J D, *et al.* A proteomic view of the Plasmodium falciparum life cycle. Nature, 2002, **419**(6906): 520~526
- 23 Casey T M, Meade J L, Hewitt E W. Organelle proteomics: identification of the exocytic machinery associated with the natural

- killer cell secretory lysosome. *Mol Cell Proteomics*, 2007, **6** (5): 767~780
- 23 Kanemaki M, Sanchez-Diaz A, Gambus A, *et al.* Functional proteomic identification of DNA replication proteins by induced proteolysis *in vivo*. *Nature*, 2003, **423**(6941): 720~724
- 24 Unwin R D, Smith D L, Blinco D, *et al.* Quantitative proteomics reveals posttranslational control as a regulatory factor in primary hematopoietic stem cells. *Blood*, 2006, **107**(12): 4687~4694
- 25 Levkowitz G, Waterman H, Ettenberg S A, *et al.* Ubiquitin ligase activity and tyrosine phosphorylation underlie suppression of growth factor signaling by c-Cbl/Sli-1. *Molecular Cell*, 1999, **4**(6): 1029~1040
- 26 Benschop J J, Mohammed S, O'Flaherty M, *et al.* Quantitative phosphoproteomics of early elicitor signaling in *Arabidopsis*. *Mol Cell Proteomics*, 2007, **6**(7): 1198~1214
- 27 Kawasaki N, Itoh S, Ohta M, *et al.* Microanalysis of N-linked oligosaccharides in a glycoprotein by capillary liquid chromatography/mass spectrometry and liquid chromatography/tandem mass spectrometry. *Anal Biochem*, 2003, **316**(1): 15~22
- 28 Warren M R, Parker C E, Mocanu V, *et al.* Electrospray ionization tandem mass spectrometry of model peptides reveals diagnostic fragment ions for protein ubiquitination. *Rapid Commun Mass Spectrom*, 2005, **19**(4): 429~437
- 29 Bauer M D, Sun Y, Keough T, *et al.* Sequencing of sulfonic acid derivatized peptides by electrospray mass spectrometry. *Rapid Commun Mass Spectrom*, 2000, **14**(10): 924~929
- 30 Shilov I V, Seymour S L, Patel A A, *et al.* The paragon algorithm: a next generation search engine that uses sequence temperature values and feature probabilities to identify peptides from tandem mass spectra. *Mol Cell Proteomics*, 2007, **6** (9): 1638~1655

Identification of Protein Ubiquitination by IP-2D nano-HPLC-MALDI-TOF-TOF*

MAO Jian-Hua^{1,2,3)}, ZHANG Qun-Ye²⁾, HUANG Qiu-Hua²⁾, CHEN Zhu²⁾, CHEN Quan¹⁾, CHEN Sai-Juan^{1,2)**}

¹⁾*Institute of Zoology, The Chinese Academy of Sciences, Beijing 100101, China;*

²⁾*State Key Laboratory of Medical Genomics and Shanghai Institute of Hematology, Ruijin Hospital,*

Shanghai Jiao Tong University, School of Medicine, Shanghai 200025, China;

³⁾*Graduate School of The Chinese Academy of Sciences, Beijing 100049, China)*

Abstract Mass spectrometry has been extensively used in the identification of protein and its posttranslational modification (PTM). Mass spectrometry coupled with nano high performance liquid chromatography (HPLC) could improve its resolution and efficiency. Ubiquitination, one kind of the important protein posttranslational modifications (PTM), plays a pivotal role in dynamic balance and functional regulation of protein. A new strategy taking the advantage of immunoprecipitation, 2D nano HPLC and matrix-assisted laser desorption/ionization-time of flight mass spectrometry was established to identify the ubiquitination of endogenous protein in mammalian cells. By using this strategy, the ubiquitinated sites of c-ABL in K562 leukemia cells was successfully identified. Thus, it may provide a valuable tool to characterizing the ubiquitination of endogenous proteins under physiologic and pathologic conditions.

Key words ubiquitination, PTM, IP-2D nano-HPLC-MALDI-TOF-TOF

*This work was supported by a grant from The National Natural Science Foundation of China (30600262, 30772744, 30572143, 30623010) and The Key Discipline Program of Shanghai Municipal Education Commission (Y0201).

**Corresponding author.

Tel: 86-21-64370045-610601, Fax: 86-21-64743206, E-mail: sjchen@stn.sh.cn

Received: March 2, 2008 Accepted: April 8, 2008