

基于生物质谱进行新药靶标筛选的研究策略及方法开发

刘宁¹, 王倩倩²

(南开大学 药物化学生物学国家重点实验室, 天津 300350)

摘要: 药物靶标的鉴定是一直困扰着药物发现和化学生物学研究的技术瓶颈, 发展准确和高通量的鉴定方法为发现新的治疗药物提供巨大的机会。但是目前普遍存在靶点质谱鉴定中杂蛋白干扰多、假阳性高、数据重现性低等问题。本文为了更好地探测药物靶标, 最大限度地减少临床试验的时间和成本并提高新药的成功率, 开发基于生物质谱技术高通量、高灵敏度、具有普遍适用性的高通量筛选方法, 通过蛋白质组学样品前处理、生物质谱一级全扫描和二级碰撞模式参数等设定、利用生物信息学计算进行药物靶标鉴定, 可为药物筛选提供突破口, 为新药研发中先导化合物的发现和设计提供重要的理论依据。

关键词: 生物质谱; 靶标筛选; 靶向蛋白质组学; 液相色谱-串联质谱

Strategy and Method Development of New Drug Target Screening Based on Biological Mass Spectrometry

Liu Ning, Wang Qianqian

(Nankai University, Tianjin 300350, China)

Abstract: The identification of drug targets has always been a technical bottleneck in drug discovery and chemical biology research. The development of accurate and high-throughput identification methods provides great opportunities for the discovery of new therapeutic drugs. However, there are many problems in the identification of target mass spectrometry, such as multiple interference of miscellaneous proteins, high false positive rate and low data reproducibility. In order to better detect drug targets, minimize the time and cost of clinical trials and improve the success rate of new drugs, this paper develops a high-throughput screening method based on bio-mass spectrometry with high throughput, high sensitivity and universal

¹ 作者简介: 刘宁 (1986-), 女, 河北衡水, 硕士, 实验师, 主要研究方向为生物质谱及仪器设备维护管理, liuning@nankai.edu.cn

² 通讯作者: 王倩倩 (1986-), 女, 河北唐山, 博士, 实验师, 主要研究方向为生物质谱及仪器设备维护管理, lx_wangqianqian@163.com

applicability. Through the pretreatment of proteomics samples, the setting of the parameters of the first and second collision modes of bio-mass spectrometry, and the identification of drug targets by bioinformatics calculation, it can provide a breakthrough for drug screening and provide an important theoretical basis for the discovery and design of lead compounds.

Keywords: biological mass spectrometry; target screening; targeted proteomics; liquid chromatography-tandem mass spectrometry

准确检测与疾病发生发展相关的蛋白靶点将有助于理解疾病的发病机制。蛋白质是基因表达产物，是生物体生命功能的执行者，是机体生理和病理状态的重要标志物，其表达变化可以直接反映机体的不同状态，有助于对疾病进行及时诊断和干预。Cohen 等^[1-2]研究表明，较之传统的基因检测，将蛋白质含量检测与基因检测相结合可以大大提高癌症(卵巢、肝、胃、胰腺、食道、结直肠、肺和乳腺)诊断的准确性。传统蛋白质定量方法主要基于放射性或荧光发光的免疫化学技术/免疫印迹技术^[3-4]，其中酶联免疫吸附试验(ELISA)技术^[5]可以定量分析血清中 ng/L 水平的蛋白质，是目前蛋白质研究的常用方法之一。但是，ELISA 技术存在诸多限制，比如蛋白质抗体难以获得、不同抗体批次之间存在差异、基于抗体的检测存在交叉反应等问题^[6-7]。因此，发展不依赖于抗体的蛋白质/多肽定量技术，将有助于提高生物标志物验证工作的效率。近年来，随着基于质谱法的蛋白质组学的发展，已发现大量与疾病相关的潜在蛋白靶点^[8-10]，这些潜在新药靶标需要经过多中心的临床病例验证有可能应用于创新药物的转化中^[11]。

1 样品前处理及参数设定

为了在复杂生物体系中对目标蛋白质/多肽的准确定量，需要在质谱分析前配置合适的液相色谱仪，并根据被检测肽段的性质选择合适的色谱分离方法，实现对目标物的准确检测。在蛋白质和多肽领域，与质谱联用的色谱仪按照流速主要分为常规流速液相色谱、微升液相色谱和纳升液相色谱。其中，纳升液相色谱可以配备内径小于 0.1 mm 的色谱柱和 nL/min 的低流速，低流速可以改善肽段在电喷雾中的离子化效率，排除色谱峰的干扰，提高检测灵敏度。因此，纳升液相色谱-质谱是有限量蛋白质样品研究的首选，常用于新药靶标早期的发现和确证^[12-14]。本次方法开发所使用的的科研仪器 Easy-nlc1000 纳升液相-Orbitrap Fusion 四极杆离子阱杂合轨道阱三合一质谱仪。

利用生物质谱进行靶标筛选试验一般流程包括:收集样本后进行蛋白质和含溴分子探针的孵育→pull down 实验后分段胶内酶切→优化液相梯度并进行 HPLC 分离→分级进入 MS 电场进一步离子化→通过优化质谱参数获得各离子质荷比和峰型信息→数据库检索比对获得蛋白质的定性和序列信息进行靶点确证。

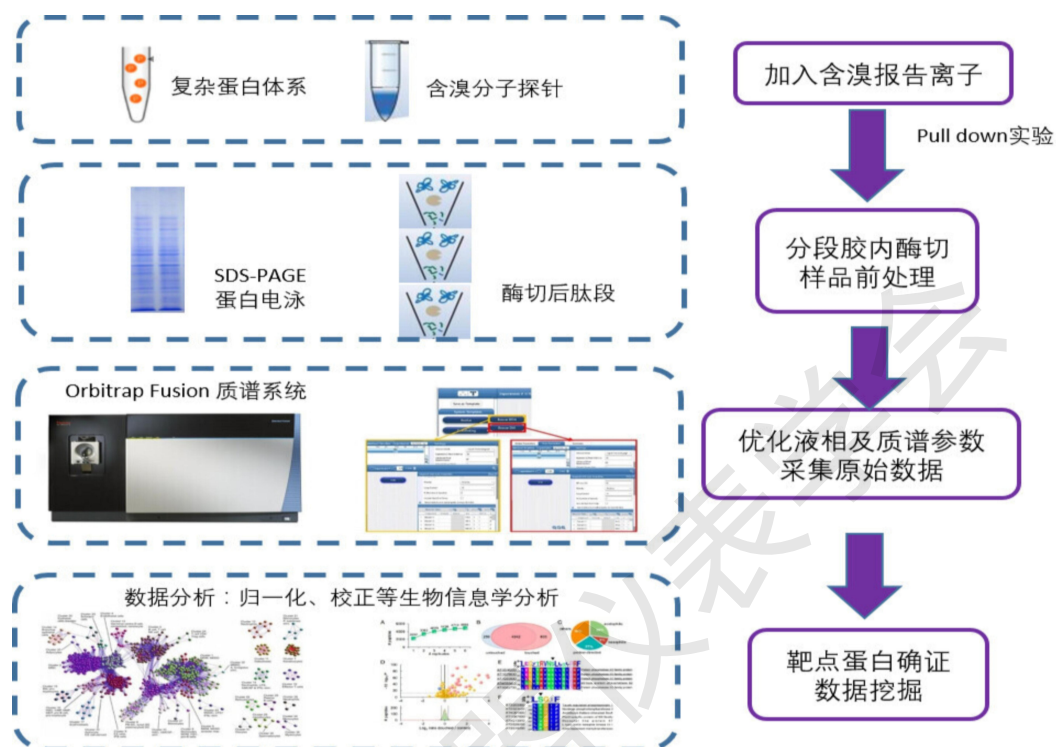


图 1 基于生物质谱技术提高药物靶标鉴定准确率的流程图

1.1 分段胶内酶切样品处理具体内容

1) 切下目标条带或将条带分成 3~6 组分 (条带宽度约为 2mm 左右), 置于 500ul EP 管中;

2) 每管加入脱色液 500ul 震荡, 视情况更换脱色液至脱色完全 (如颜色较深可加热 37 度), 用台式真空泵吸去上清液;

3) 干胶: 加 100%乙腈(ACN)振荡脱水至胶粒变白, 吸去液体, 真空干燥 5min。

4) 加 10mM DTT/25mM NH_4HCO_3 (1M 母液以 25mM NH_4HCO_3 稀释) 淹没胶块, 振荡混匀至胶块泡胀透明。56°C 1h。

5) 冷却至室温, 吸干, 快速加 55mM 碘代乙酰胺 (IAA) /25mM NH_4HCO_3 淹没胶块, 振荡混匀。(快速, 避光) 暗处放置 45min。

6) 25mM NH_4HCO_3 洗两次, 移去液体, 100%乙腈 ACN 脱水至胶粒变白。真空干燥 5min。

7) 每管加入 50mM 碳酸氢铵溶液 20ul, 再加入 1-2ul 胰酶 (或者将两者混合后再加入样品中, 淹没胶条), 将凝胶挤碎, 37 度温育 6 小时以上;

8) 每管加入 200ul (含 0.1%甲酸) 乙腈震荡 5 分钟, 吸取上清液至干净的 EP 管中;

9)凝胶中再加入 30ul 0.1%甲酸质谱水,震荡 5 分钟,再加入 200ul 100% (含 0.1%甲酸)乙腈震荡 5 分钟,吸取上清将上清合并;

10) 45 度真空旋干。

1.2 液相梯度及参数设置具体内容

纳升液相色谱在低流速进样过程中需要配备能够长时间保持较好重复性的色谱柱,以保持系统的稳定性,同时实际操作过程中容易出现色谱过载、质谱饱和等情况^[15-16],这些因素影响了肽段鉴定和定量的重复性,在分析一些浓度范围跨度较大的蛋白质样品(如临床组织和体液样本)时尤为明显。为解决以上问题,本次方法开发所使用的液相梯度及参数设置见表-1。

表-1 液相梯度及参数设置内容

参数	设置内容
Sample pickup	Sample pickup volume:1ul,flow: 20ul/min
Sample loading	Sample loading: 10ul,600bar
Pre-column	Pre-column:2ul,600bar
Analytical column	Analytical column:5ul,600bar
Auto-sampler	Standard:100ul
Flow	300nl/min
Duration 00:00	3%B
Duration 05:00	8%B
Duration 70:00	18%B
Duration 28:00	28%B
Duration 12:00	90%B
Duration 5:00	90%B

1.3 质谱参数设置具体内容

在本项目中,由于需要实时监测进入质谱的总离子流,保证在一定时间内进入质谱二级碎片的数量既不能过多又不能过少,同时选择合适的离子隔离窗口和碰撞能量,以保证得到的二级碎片能够充分被质谱检测器捕获到,为实现数据快速灵敏扫描,需要优化六部分参数,在实际实验过程中,需要观察得到的质谱原始谱图的二级碎片信息和实时扫描信号,调整质谱一级和二级的关键参数,如:离子隔离模式、HCD 碰撞能、最大注入时间、离子选择窗口等,在生物质谱应用于新药靶标筛选所使用的关键技术参数内容见表-2。

表-2 高分辨质谱进行药物靶标筛选使用的关键技术参数

参数名称	设置阈值
Detector Type	Orbitrap
Orbitrap Resolution	120000-200000
Mass Range	Normal
Use Quadrupole Isolation	TRUE
Scan Rang(m/z)	300-1500
RF lens(%)	60
AGC Target	Custom
Maximum Injection time(ms)	50
Maximum Injection mode	custom
Data Type	profile
Polarity	Postive
Filter Type	Intensity Threshold
Intensity Threshold	5.00E+03
Indude charge state(s)	2-7
Exclude after n times	1
Exclusion duration(s)	18
mass tolerance	10ppm
Isolation window(m/z)	1.6
Activation Type	HCD
Collision Energy Mode	Fixed
HCD Collision Energy(%)	30
Detector Type	HCD
Ion Trap Scan Rate	Rapid

2 质量控制

在蛋白质组学自下而上的研究策略中，样品前处理包括蛋白质提取、溶解、还原、烷基化和酶解等多个步骤，这些步骤的系统误差将在一定程度上影响定量准确性。Shuford 等对血清中甲状腺球蛋白定量结果的误差来源进行详细评估，结果表明，不完全酶解和样品间的酶解差异是定量误差的重要来源，对人源性和重组的甲状腺球蛋白进行详细比较，在各种消化条件下，这两种蛋白的定量结果都存在一定差异，所以应尽量选择与分析物类似的标准品作为校准品，以提高定量精确度。有报道指出，对研究的生物标志物的生物学性质有足够的了解，蛋白质生物标记物分析方法中非常重要，包括对正常或疾病状态下目标蛋白生理学的了解，以选择具有代表性的校准品^[17,18]为了获得蛋白质的准确定量结果，有研究者将与目的蛋白理化性质最为接近的全长重标重组蛋白插入实际样品中作为内部校准样品^[19,20]。这种样品前处理过程中的质控方法可以同其他的定量策略结合，得到更精确的蛋白/肽段定量结果。

本项目中利用 HeLa 细胞系的酶切肽段作为质控，减少测试过程中的系统误差。

3 成果应用及效益

3.1 成果应用情况

纳升液相色谱-四极杆离子阱杂合轨道阱三合一质谱仪于 2015 年安装使用，目前新测试方法服务科研项 45 项，机时达 3800 小时/年、服务科研项目数合计 96 项、培训师生人数超过 119 人、服务京津冀地区课题组超过 40 个，测试样品数 1790 个，成功筛选得到 KRAS、EGFR、AnnexinA2 等关键靶点，发表文章 40 篇，累计影响因子大于 180，并利用该技术方法对到中药复杂组分多靶点进行验证，基于生物质谱提高药物靶标筛选技术获得教育部、天津市科技局及南开大学设备处立项支持，为 Science, JACS, JMC 等权威期刊论文发表提供了高水平数据支撑。

3.2 经济效益和社会效益

利用生物质谱技术进行 AccRate-全球首创抗癌靶向药敏感性检测技术定义者，获得第七届中国国际“互联网+”大学生创新创业大赛高教主赛道金奖；该生物质谱测试方法在编号为 ACT001 的抗脑胶质瘤药物靶标筛选中发挥了重要作用，ACT001 获得美国药监局 FDA 批准的儿童罕见病资格（简称 RPDD）。在我国原创候选新药中，ACT001 是首个获得美国 RPDD 资格的品种，这也是 ACT001 继 2017 年获得针对胶质母细胞瘤（GBM）的美国孤儿药资格，2018 年获得 GBM 的欧盟孤儿药资格之后的又一重大突破 RPDD 也是申请和获得优先评审券（PRV）的资格证，而 PRV 是美国 FDA 为了鼓励公司针对儿童罕见疾病开发新药，授予新药上市快速评审的优先券。此次 ACT001 获批的 RPDD 主要针对一种儿童罕见疾病：带 H3K27 突变的弥漫性中线胶质瘤（DMG），以往往往称为更小范围的弥漫性内生型脑桥胶质瘤（简称 DIPG）。和传统蛋白质鉴定方法相比，基于生物质谱进行新药靶标筛选的研究策略可以应用于中药、天然产物来源药物等复杂成分靶标筛选，最终鉴定的有效蛋白数量提高 20%，数据重现性预期提高 20%以上，显著提高未知靶点鉴定的准确性。

4 结语

本文建立了基于纳升液相色谱-四极杆离子阱杂合轨道阱三合一质谱联用技术进行药物靶标筛选的方法。靶标蛋白的高级结构在与小分子配体结合后，其配体结合域及诱导变构区域的氨基酸残基的溶剂可及性会发生显著变化，该变化能够被还原烷基化等化学共价标记反应所捕捉，再利用蛋白质组学手段对细胞内的全蛋白组进行分析，能够获得真实生物体系内小分子孵育后全体蛋白组的可及性变化信息，方法操作便捷，质谱数据信息的后期分析简单

方便，能够普适性的运用于复杂生物体系中小分子药物配体的高通量靶标发现,为阐述药物及内源性代谢物等活性分子对细胞的调控机制奠定了基础,为创新药物研发先导化合物的发现提供了理论支持。

参考文献:

- [1] Cohen JD, Javed A, Thoburn C, *et al.* Combined circulating tumor DNA and protein biomarker-based liquid biopsy for the earlier detection of pancreatic cancers [J]. Proceedings of the National Academy of Sciences of the United States of America, 2017, 114(38): 10202-10207.
- [2] Cohen J D, Li L, WANG Y, *et al.* Detection and localization of surgically resectable cancers with a multi-analyte blood test [J]. Science, 2018, 359 (6378): 926-930.
- [3] Yalow R S, Berson S A. Immunoassay of endogenous plasma insulin in man [J]. Journal of Clinical Investigation, 1960, 39 (7): 1157-1175.
- [4] Macbeath G. Protein microarrays and proteomics [J]. Nature Genetics, 2002, 32: 526-532.
- [5] Engvall E, Perlmann. Enzyme-linked immunosorbent assay (ELISA). Quantitative assay of immunoglobulin G [J]. Immunochemistry, 1971, 8(9):871-874.
- [6] Prassas I, Brinc D, Farkona S, *et al.* False biomarker discovery due to reactivity of a commercial ELISA for CUZD1 with cancer antigen CA125[J]. Clinical Chemistry, 2014, 60(2): 381-388.
- [7] Rifai N, Watson I D, Miller W G. Commercial immunoassays in biomarkers studies : researchers beware [J]. Clinical Chemistry, 2012, 58(10): 1387-1388.
- [8] Lehallier B, Gate D, Schaum N, *et al.* Undulating changes in human plasma proteome profiles across the lifespan [J]. Nature Medicine, 2019, 25(12)1843-1850.
- [9] Egerstedt A, Berntsson J, Smith M L, *et al.* Profiling of the plasma proteome across different stages of human heart failure [J]. Nature Communications 2019 10 (1): 5830-5842.
- [10] Captur G, Heywood W E, Coats C, *et al.* Identification of a multiplex biomarker panel for hypertrophic cardiomyopathy using quantitative proteomics and machine learning[J]. Molecular and Cellular Proteomics, 2020, 19(1): 114-127.
- [11] Geyer P E, Holdt L M, Teupser, *et al.* Revisiting biomarker discovery by plasma proteomics[J]. Molecular Systems Biology, 2017,13(9): 942-956.

- [12] Cao W Q, Jiang B Y, Huang J M, *et al.* Straightforward and highly efficient strategy for hepatocellular carcinoma glycoprotein biomarker discovery using a nonglycopeptide-based mass Spectrometry pipeline [J]. *Analytical Chemistry*, 2019, 91(19): 12435-12443.
- [13] Liu X, Zheng W, Wang W, *et al.* A new panel of pancreatic cancer biomarkers discovered using a mass spectrometry-based pipeline [J]. *British Journal of Cancer*, 2017, 117(12): 1846-1854.
- [14] Huttenhin R, Choi M, Martin, *et al.* Mass spectrometry-based proteomics : existing capabilities and future directions [J]. *Chem SocRev*, 2012, 41(10): 3912-3928.
- [15] Falkenby LG, Such S G, Larsen M R, *et al.* Integrated solid-phase extraction-capillary liquid chromatography (speLC) interfaced to ESI-MS/MS for fast characterization and quantification of protein and proteomes [J]. *Journal of Proteome Research*, 2014, 13(12): 6169-6175.
- [16] Wilsons R, Vehus T, Berg H S, *et al.* Nano-LC in proteomics : recent advances and approaches [J]. *Bioanalysis*, 2015, 7(14): 1799-1815.
- [17] Cowan K J, Amaravadi L, Cameron M J, *et al.* Recommendations for selection and characterization of protein biomarker assay calibrator material [J]. *Aaps Journal*, 2017, 19(6): 1550-1563.
- [18] Minkel E V, Kuhn E, Cocco A R, *et al.* Domain-specific quantification of prion protein in cerebrospinal fluid by targeted mass spectrometry [J]. *Molecular and Cellular Proteomics*, 2019, 18 (12): 2388-2400.
- [19] Weber D M, Tran D, Goldman S M, *et al.* High-throughput mass spectrometry assay for Quantifying α -amyloid 40 and 42 in cerebrospinal fluid [J]. *Clinical Chemistry*, 2019, 65(12): 1572-1580.
- [20] Shuford C M, Walters J J, Holland P M, *et al.* Absolute protein quantification by mass spectrometry: not as simple as advertised [J]. *Analytical Chemistry*, 2017, 89(14): 7406-7415.