修饰蛋白质组学分离鉴定新技术新方法

宋春侠^{①②}, 孙珍^①, 秦洪强^①, 叶明亮^①, 张丽华^①, 张玉奎^①, 邹汉法^{①*}

① 中国科学院大连化学物理研究所, 中国科学院分离分析化学重点实验室, 国家色谱研究中心, 大连 116023;

② 中国石油化工股份有限公司石油化工科学研究院, 北京 100083

* 联系人, E-mail: hanfazou@dicp.ac.cn

2013-05-13 收稿, 2013-07-09 接受, 2013-09-06 网络版发表

国家自然科学基金委员会创新研究群体科学基金(21021004)和国家重点基础研究发展计划(2013CB911202)资助

摘要 蛋白质翻译后修饰(PTMs)是对基因产物蛋白质进行共价加工的过程, 它是生命体在有限基因数目的条件下进行复杂生理学过程调控的重要原因. 生物体内多达 300 种的翻译后修饰在细胞周期调控、信号传导及免疫应答等众多生命活动中发挥着关键作用. 由于大多数翻译后修饰的蛋白质通常具有丰度低、动态范围宽以及结构复杂等特点, 因此对蛋白质翻译后修饰进行规模化分析面临着巨大的挑战. 基于生物质谱技术的修饰蛋白质组学分析方法的出现为翻译后修饰蛋白质的高通量分析提供了前所未有的发展契机, 翻译后修饰蛋白质的规模化分析鉴定, 已经成为蛋白质组学研究的核心内容之一. 本文主要针对修饰蛋白质组学常用分析方法, 包括翻译后修饰蛋白质的富集策略、分级方法、质谱鉴定技术和定量分析技术的最新研究进展进行介绍, 并对其发展前景进行展望.

关键词

蛋白质翻译后
修饰
磷酸化
糖基化
内源性多肽
蛋白质组学

蛋白质翻译后修饰(PTMs)是通过对蛋白质侧链和末端进行化学修饰或者是被蛋白水解酶酶切来改变蛋白质性质的共价加工过程^[1]. 生物体内多达 300 种的翻译后修饰在细胞周期调控、信号传导及免疫应答等众多生命活动中发挥着关键作用, 它也是生命体在有限基因数目的条件下进行复杂生理学过程调控的重要原因. 正是由于基因产物蛋白质的不同剪切类型以及众多翻译后修饰类型的存在, 才使得蛋白质的功能更加完善, 调控更为精细^[1~3]. 大多数发生翻译后修饰的蛋白质通常具有丰度低、动态范围宽以及结构复杂等特点, 因此对蛋白质翻译后修饰进行规模化分析面临着巨大的挑战.

由于蛋白质翻译后修饰发生在转录之后, 因此基因组学方法无法实现翻译后修饰类型以及结构的检测. 近年来, 随着生物质谱技术的进步和高效分离方法的发展, 蛋白质组学技术为蛋白质翻译后修饰的系统分析提供了前所未有的发展契机, 由此也派

生出一门新的学科: 修饰蛋白质组学^[1,4]. 它主要利用蛋白质组学分析技术对修饰蛋白质及其位点进行鉴定, 对修饰结构进行确认以及对不同生理状态下翻译后修饰的变化情况进行定量分析. 该方法的基本分析流程如图 1 所示: 生物样品中的蛋白质经样品提取后, 在蛋白质层次上进行富集或者直接酶解, 产生的肽段混合物经过液相色谱预分级, 针对不同修饰肽段进行选择富集后进入质谱分析, 通过数据库检索的方法得到最终鉴定的肽段和蛋白序列, 并根据质谱峰强度或者色谱峰面积进行定量. 目前, 对翻译后修饰蛋白质进行系统分析已经成为蛋白质组学研究的核心内容之一. 本文主要针对修饰蛋白质组学常用分析方法及最新研究进展进行介绍, 并对其发展前景进行展望.

1 修饰蛋白质组学分析中的富集策略

生物体内大多数发生翻译后修饰的蛋白质通常

引用格式: 宋春侠, 孙珍, 秦洪强, 等. 修饰蛋白质组学分离鉴定新技术新方法. 科学通报, 2013, 58: 3007–3016

Song C X, Sun Z, Qin H Q, et al. The new method and technology for the characterization and identification of post-translational modifications of proteins (in Chinese). Chin Sci Bull (Chin Ver), 2013, 58: 3007–3016, doi: 10.1360/972013-515

图 1 基于生物质谱的修饰蛋白质组学基本分析流程

具有丰度低、动态范围宽和结构复杂等特点, 因此, 对复杂生物样品中修饰蛋白质和肽段进行选择性的富集, 是实现修饰蛋白质组学规模化分析鉴定最重要的步骤. 作为一种最重要的蛋白质翻译后修饰, 蛋白质磷酸化在细胞信号传导、细胞生长、增殖与分化等方面发挥着至关重要的作用^[5]. 目前, 在磷酸化蛋白质组学分析中应用最为广泛的富集方法主要包括固定化金属离子亲和色谱法和金属氧化物色谱法. 两种富集策略分别利用固定在色谱基质上的金属离子或者过渡金属氧化物与磷酸肽磷酸基团之间的相互作用而实现磷酸肽的富集. 为了降低酸性肽段的非特异性吸附, 通常可以通过优化上样及洗脱条件来改善磷酸肽的富集效果. 例如, 可以在上样过程中加入各种竞争试剂(2,5-二羟基苯甲酸(DHB)和乳酸等含有羧基的试剂)来提高 TiO_2 材料富集的特异性^[6,7]. 另一方面, 还可以通过发展具有更高特异性的新型螯合配基来实现磷酸肽的选择性富集. Zou 课题组^[8-10]根据磷酸酯基团与过渡金属离子所形成的独特配位结构, 合成了一系列高特异性的磷酸肽富集材料(图 2), 并在磷酸肽鉴定和规模化分析中取得了优异的效果.

蛋白质糖基化修饰也是一种非常重要的蛋白质

图 2 基于磷酸酯配基的固定化金属离子亲和色谱材料用于磷酸肽富集的原理图

翻译后修饰, 它参与了新生肽段的折叠和缔合以及糖蛋白的转运和分泌等众多重要的生理学过程^[11,12]. 根据糖链连接的氨基酸残基不同, 蛋白质糖基化可分为 N-糖基化、O-糖基化、C-糖基化和糖基化磷脂酰乙醇胺. 由于 N-糖基化发生在特定的氨基酸序列中(NXS/T/C, X 可以是除了脯氨酸以外的任意氨基酸), 因此成为目前研究最为广泛的糖基化之一. N-糖基化蛋白常用的富集方法有凝集素法、酰肼化学法、亲水作用色谱法、金属氧化物/氢氧化物亲和色谱法(MOAC)和固相金属离子亲和色谱法(IMAC)等. 凝集素法是最早被用于糖基化蛋白质分析的富集方法, 该方法的缺点是凝集素与糖链间的作用相对较弱, 非特异性吸附较多, 导致富集效率不高. 最近, Zielinska 等人^[13]发展了基于超滤膜过滤的糖基化肽段富集方法, 通过超滤去除凝集素表面的非特异性吸附, 使得糖基化肽段富集特异性达到 40%, 极大地提高了糖基化位点鉴定的效率以及规模. 酰肼化学法是通过高碘酸钠将糖链上的顺式邻二醇氧化成醛基, 再与含有酰肼官能团的材料键合, 从而实现糖基化蛋白质和肽段的富集. 2003 年, Zhang 等人^[14]首次将该方法用于糖基化蛋白质的富集, 随后被广泛用于血浆/血清、细胞、组织等样品中糖基化蛋白质和肽段的定性定量分析中^[15,16]. 该方法富集特异性较高, 能够广泛富集各种糖链结构的蛋白质和肽段, 不足之处是糖链被氧化破坏, 因此无法得到有关糖链结构方面的信息. 亲水作用色谱法是利用糖链的亲水性质, 通过亲水相互作用使糖基化肽段与普通肽段实现分离的富集方法. 该方法可以保持糖链结构的完整性, 因此可以获得糖基化肽段的完整信息^[17-19]. Zhu 等人^[20]发展了一种离心辅助的亲水作用色谱糖蛋白富集微反应器, 通过将蛋白质酶解和亲水作用色谱富集以及糖基化肽段的去糖基化等技术集成到微反应器中, 实现微量样品中糖基化蛋白质

的快速、高灵敏度检测。近几年来,基于过渡金属离子与糖肽之间较强相互作用的 MOAC 和 IMAC 法也被广泛用于糖肽富集中。Palmisano 等人^[21]利用 TiO_2 在特定缓冲溶液条件下对糖链末端唾液酸的高度亲和性,并结合亲水作用色谱的高效分离能力,实现了 817 个唾液酸化糖蛋白和 1632 个唾液酸化糖肽的鉴定。Zhu 等人^[22]发展了基于 Ti^{4+} -IMAC 材料的糖肽微反应器用于含唾液酸的糖肽富集,在 1 μL 人血清中成功鉴定到 217 个含唾液酸的糖肽以及 194 个糖基化位点。

O-糖基化的富集方法与 N-糖基化的富集方法类似。例如,利用麦胚凝集素(WGA)对 N-乙酰氨基葡萄糖(GlcNAc)的亲合作用,可以有效地富集带有 O-GlcNAc 的糖蛋白^[23];将传统的酰肼化学法进行适当的改动,通过加长高碘酸钠氧化时间和升高氧化温度,可以实现酰肼微球对带有 O-GlcNAc 肽段的富集^[24]。另外,还可以利用酮基-生物素标签在半乳糖转移酶的作用下,选择性标记带有 O-GlcNAc 的蛋白,再利用生物素-亲和素之间的作用进一步分离富集^[25];也可以利用 O-GlcNAc 叠氮类似物进行代谢标记,然后再通过点击化学反应以及生物素-亲和素之间的相互作用,使得带有 O-GlcNAc 的蛋白得到分离富集^[26]。

在蛋白质的周期循环过程中,蛋白质水解酶参与了蛋白质前体的剪切和蛋白质降解过程,因此,蛋白质水解酶的剪切也被认为是一种极其重要的翻译后修饰^[1]。蛋白质在水解酶作用下产生的多肽碎片蕴含着丰富的生物信息,为了解复杂的生命过程提供不可或缺的重要资源^[27]。目前,内源性多肽的富集策略主要有离心超滤法^[28,29]、有机溶剂沉淀法^[30,31]

和纳米材料富集法^[32,33]。近年来,基于多肽分子与纳米材料之间的体积排阻效应以及疏水相互作用的纳米材料富集法,由于其操作简便、富集效率高等优点而引起人们的广泛关注。

Tian 等人^[34]采用具有高度有序孔道结构的硅材料(MCM-41, 2 nm)选择性富集血浆中的内源性多肽,可以将多肽检测的信噪比提高 100 倍左右,并且操作更加简便,非常适合复杂生物样品中内源性多肽的富集。为了增加多肽与硅基材料间的相互作用,提高内源性多肽的富集效率, Liu 等人^[35]在磁性介孔硅材料的表面引入疏水性的 C8 官能团,进而增加多肽在材料上的保留,最终能够从小鼠的脑提取物中鉴定到 267 条内源性多肽,比之前发展的 $\text{Fe}_3\text{O}_4@\text{nSiO}_2\text{-mSiO}_2$ 微球的效果提高了 3.45 倍。Qin 等人^[36]发展了基于高度有序碳介孔材料(OMC)的内源性多肽富集方法(图 3),可以从极微量血清(20 μL)样品中鉴定到 3402 条非冗余内源性多肽,实现了内源性多肽的高灵敏度分析检测。除了直接利用材料与多肽分子的相互作用外,还可以通过对材料表面改性,引入离子交换基团^[37]或者金属离子^[38],实现内源性多肽的选择性富集。Hu 等人^[38]制备了固定化 Ti^{4+} 修饰的 MCM-41 材料用于血清中内源性磷酸肽的富集,该方法同时结合了介孔材料的体积排阻效应和 Ti^{4+} 离子对磷酸肽的亲合作用,具有较高的富集效率。

除了上述常用富集方法外,基于目标修饰蛋白质与特异性抗体相互作用的抗体富集法也被广泛用于修饰蛋白质组学分析中。特别是对于缺乏特异性化学富集材料的其他翻译后修饰类型,如乙酰化、甲基化和泛素化等,特异性的抗体富集方法一直是实现其有效分离鉴定的通用技术手段^[39,40]。此外,针对

图 3 有序碳介孔材料(OMC)高选择性富集血浆中内源性多肽的原理图^[36]

酪氨酸磷酸化蛋白质和肽段的抗体富集法也在酪氨酸磷酸化蛋白质分析检测中发挥着重要作用^[41]。

2 修饰蛋白质组学分析中的分离方法

由于生物样品的高度复杂性, 经过富集的修饰蛋白/肽段样品仍然具有较高的动态范围, 因此, 在蛋白质或肽段层次上对样品进行预分级, 成为修饰蛋白质组学规模化分析鉴定不可或缺的重要内容。传统的一维或二维凝胶电泳(1DE 或 2DE)常常被用于实现蛋白质层次上的预分级。Han 等人^[42]采用十二烷基硫酸钠-聚丙烯酰胺凝胶电泳(SDS-PAGE)对人肝样品中的蛋白质进行预分级, 结合肽段层次上的 Ti-IMAC 富集策略, 最终从正常人肝组织中鉴定到 1023 个磷酸化蛋白质和 2225 个磷酸化位点, 极大地提高了人肝磷酸化蛋白质组学规模化分析水平。近年来, 肽段固定 pH 梯度等电聚焦法(IPG-IEF)由于其良好的重复性和较高的分辨率(~ 0.1 pH), 在蛋白质组学分析领域得到了广泛的关注。Cao 等人^[43]利用糖链末端唾液酸的电负性, 采用 IPG-IEF 法对带有唾液酸的糖肽进行预富集, 使其在 IEF 结束时集中在 IPG 较低的 pH 条带处, 然后再利用 TiO_2 对含有唾液酸的肽段进行亲和富集, 在鼠肝中鉴定到 582 个唾液酸化的糖肽和 614 个糖基化位点(对应于 322 个唾液酸化的糖蛋白), 这是迄今为止得到的最大规模鼠肝唾液酸糖基化蛋白质组学数据集。

近年来, 随着鸟枪法蛋白质组学技术(shotgun proteomics)的发展, 基于肽段层次上的多维液相色谱分离方法在翻译后修饰蛋白质组学规模化分析中得到了广泛应用。为了提高修饰肽段鉴定的覆盖度, 在第一维分级时选择与第二维反相色谱分离具有良好正交性的分离模式尤为重要。强阳离子交换色谱(SCX)^[44]、强阴离子交换色谱(SAX)^[45]、亲水作用色谱(HILIC)^[46]、高 pH 反相色谱^[47]和体积排阻色谱(SEC)^[13]等多种液相色谱分离模式均被成功用于修饰肽段的规模化分析鉴定。其中, 以离线的 SCX 分离模式应用最为广泛。Huttlin 等人^[44]将 SCX 方法与蛋白质层次上的 SDS-PAGE 分级策略相结合, 可以从小鼠的 9 种组织中鉴定到 36000 个磷酸化位点, 显示了其在磷酸肽鉴定方面强大的分析性能。Liu 等人^[48]采用酰肼化学法对血浆中的糖基化蛋白质进行富集, 通过对 PNGaseF 酶酶切后产生的去糖基化肽段进行 SCX 预分级, 实现糖基化蛋白质的规模化鉴

定。Tan 等人^[49]也通过 SCX-RP(反相色谱)二维分析技术有效提高了泛素化蛋白质的鉴定覆盖度。

由于与第二维反相色谱具有良好的正交性, 近年来, 亲水作用色谱在翻译后修饰蛋白质组学的规模化分析中受到了广泛关注。它不仅可以作为糖基化肽段的富集手段, 还可以在第二维液相色谱分离时实现肽段的有效预分级, 提高糖基化蛋白质组学的鉴定覆盖度。Zhang 等人^[50]利用静电排斥-亲水作用色谱(ERLIC)分离模式实现糖基化肽段的分级和富集, 通过收集 7 个混合馏分, 可以从 1 mg 小鼠脑组织细胞膜蛋白质酶解液中鉴定到 922 个糖基化位点。McNulty 等人^[46]考察了 HILIC 用于磷酸肽预分级方法的可行性, 证实了该方法具有与 SCX 分级策略相当的磷酸肽鉴定能力。HILIC 分级策略也被用于提高甲基化蛋白质组学分析通量的研究中, 并取得了优于 SCX 和 IEF 分级方法的鉴定效果^[51]。

最近, Song 等人^[47,52]将高 pH 流动相下的反相色谱作为第一维预分级方法, 通过与间隔组分混合的馏分制备方法相结合, 对正常人肝组织中磷酸化位点进行规模化鉴定, 共鉴定到 2998 个磷酸化蛋白质和 9719 个磷酸化位点, 建立了迄今为止最大的人类肝脏磷酸化位点数据集。Zhu 等人^[53]利用该方法进行内源性磷酸肽的预分级, 通过与固定化 Ti^{4+} 的 MCM-41 富集材料相结合, 并根据目标数据库进行数据库检索来实现内源性磷酸肽的规模化鉴定。Choudhary 等人^[39]利用 IEF 方法对经过乙酰化抗体富集的肽段样品进行分级, 结合高分辨质谱分析方法, 从 3 个不同的细胞系中鉴定到 3600 个乙酰化位点, 对应 1700 个乙酰化蛋白质, 将蛋白质乙酰化分析通量提高了一个数量级。

3 修饰蛋白质组学分析中的质谱鉴定技术

作为蛋白质组学研究的核心技术, 质谱分析对于准确获得修饰蛋白质及其位点修饰信息至关重要。过去 20 年中, 高性能质谱仪的出现和新型质谱碎裂方式的发明, 极大地推动了蛋白质组学技术的发展。对于修饰蛋白质组学而言, 由于发生翻译后修饰的蛋白质和肽段具有与普通肽段不同的碎裂特性, 因此, 发展针对修饰肽段的新型质谱碎裂技术显得尤为重要。目前, 修饰蛋白质组学研究领域最常用的质谱碎裂方式仍然是碰撞诱导解离(CID/CAD), 该种碎裂模式下肽段主链上的酰胺键断裂, 产生含有丰富

鉴定信息的 *b*, *y* 离子. 而对于一些蛋白质翻译后修饰, 由于发生修饰的化学键比酰胺键更容易断裂, 因此容易在碰撞过程中发生中性丢失(例如, 磷酸肽丢失一分子 H_3PO_4 , 含有核心岩藻糖的糖基化肽段发生岩藻糖的中性丢失). 另一方面, 由于中性丢失峰的出现也是修饰肽段存在的重要信号, 因此, 可以通过对发生中性丢失的碎片离子进行三级质谱碎裂(MS3)来获得更多的碎裂信息, 改进修饰肽段序列的鉴定效果^[54,55]. Jia 等人^[55]利用 EndoF3 酶酶切来提高对核心岩藻糖糖基化位点的选择性, 并通过发展的中性丢失诱导三级质谱(MS3)方法实现含有核心岩藻糖的糖基化肽段的规模化鉴定.

近年来, 较为温和的电子捕获解离(ECD)和电子转移解离(ETD)技术在翻译后修饰肽段的鉴定方面发挥了重要作用. 该种碎裂方式通过向带正电荷的肽段引入一个活泼的电子而诱发肽段主链上 N-C α 键的断裂, 产生互补的 *c*, *z* 离子用于肽段序列的鉴定^[56]. 而侧链上的修饰基团(磷酸基团和糖链)由于不带正电, 因此可以在碎裂过程中保持完整的修饰信息, 有利于翻译后修饰肽段结构的鉴定. 与倾向于解离 +2/+3 价离子的 CID 技术所不同, ETD 技术对带有较高正电荷的肽段离子($z>3$)具有更好的碎裂效果. 因此, 采用多种酶酶切并结合 2 种碎裂模式的方法被认为是提高修饰肽段鉴定覆盖度的有效手段^[57,58]. 近几年, 新型的高能碰撞诱导解离技术(HCD)在修饰蛋白质组学规模化分析中显示了优异性能. 在该种碎裂模式下, 产生的碎片离子不仅具有较高的质量精度, 还可以消除常规离子阱质谱中的 1/3 质量排除效应, 有效实现低分子量端离子离子的检测(例如酪氨酸离子^[59]和单糖/寡糖离子^[60]), 极大地提高了翻译后修饰位点鉴定结果的可信度. 另外, 采用 HCD 碎裂技术的糖肽还能够产生更为“简化”的二级质谱图(MS2)(图 4), 可以很清晰地表明单糖从糖链末端到核心的丢失顺序, 从而有利于糖链结构的自动化解析.

由于修饰肽段在质谱分析时具有与普通肽段不同的碎裂特性, 因此非常需要发展针对修饰肽段的自动化数据处理方法. 在磷酸化蛋白质组学数据处理方面, Jiang 等人^[61]针对 MS2/MS3 结合的质谱碎裂方式分别提出了基于 MS2/MS3 匹配的自动化磷酸肽鉴定方法和基于分类筛选策略的磷酸肽鉴定方法^[62], 有效提高了磷酸肽鉴定结果的覆盖度和规模化数据

图 4 完整糖基化肽段的 HCD 二级质谱图

分析的处理效率. 而对于糖基化蛋白质组学来说, 由于糖链结构的高度复杂性, 使得大规模糖基化蛋白质分析目前还主要集中于糖基化位点的鉴定上. 尽管目前已经发展了一些糖基化蛋白质结构鉴定的自动化数据处理方法, 但分析对象还仅仅局限于单个或少数几个高丰度蛋白质^[63,64]. 因此, 发展针对蛋白质糖基化修饰的自动化数据处理方法, 成为提高糖基化蛋白质组学分析通量的重要途径.

4 修饰蛋白质组学分析中的定量方法

在生物体内, 质的改变通常需要量的积累, 因此, 对不同生理状态下修饰蛋白质的表达情况进行定量分析, 成为深入了解不同蛋白质翻译后修饰生物学功能的重要途径. 目前, 在修饰蛋白质组学定量分析领域, 主要包括无标记定量和稳定同位素标记定量 2 种策略. 由于采用的定量方法以及样品处理步骤不同, 导致不同的定量方法具有不同的特点(表 1)^[65], 因此需要根据实验目的以及样品的不同, 选择最合适的定量方法.

无标记定量方法主要通过比较同一蛋白质在不同质谱分析时肽段的鉴定次数(spectral counts)或者肽段的峰面积来实现蛋白质的相对定量. Chen 等人^[66]发展了基于非糖基化肽段的无标记糖基化蛋白质定量分析方法, 通过酰肼化学法对糖基化蛋白质进行富集, 然后利用胰蛋白酶酶切释放的非糖基化肽段进行质谱定量, PNGaseF 酶酶切释放的糖基化肽段进行位点鉴定. 将该方法用于正常人与肝癌患者血清中糖基化蛋白质的定量分析中, 发现了 38 个发生显著变化的糖基化蛋白质, 并选择了其中 3 个进行大样本验证, 充分证实了该方法用于糖基化蛋白质

表 1 修饰蛋白质组学常用定量方法性能比较 a)[65]

	无标记定量	同位素标记定量			
		SILAC 标记	iTRAQ/TMT 标记	二甲基同位素标记	O ¹⁸ 标记
降低操作误差的能力	差	优	良	良	良
成本效率	优	差	差	优	优
分析通量	优	优	良	优	差
不同类型质谱仪的兼容性	优	良	差	良	良
多重样品分析能力	优	良	优	良	差
定量准确性	良	优	优	优	良

a) SILAC, 细胞系氨基酸稳定同位素标记法; iTRAQ, 同重同位素标签标记定量法; TMT, 串联质谱标签

定量结果的可靠性. 尽管无标记定量方法操作简单, 但是由于不同样品的质谱分析需要单独进行, 因此很容易受到液质联用系统稳定性的影响. 而对于修饰肽段样品来说, 还需要进行单独的富集操作, 有可能引入新的定量误差^[67]. 因此, 可以在一次质谱分析中实现多个样品同时定量的稳定同位素标记定量方法成为修饰蛋白质组学定量分析的最佳选择.

目前应用较为广泛的同位素标记定量方法主要包括 SILAC, iTRAQ, O¹⁸ 标记法和二甲基稳定同位素标记定量方法. 其中, SILAC 技术由于标记过程发生在早期的样品处理阶段(细胞培养过程中), 因此通常被认为具有最好的定量准确性^[68]. 目前, 该技术不仅可以标记细胞样品, 还可以标记哺乳动物组织样品, 极大地扩展了其应用范围^[69]. 在糖基化蛋白质组学分析中, 可以通过在 H₂O¹⁸ 中切除糖链来提高糖基化位点鉴定结果的可信度^[70]. Liu 等人^[71]巧妙地利用了该策略, 通过将不同样品在 H₂O¹⁶ 和 H₂O¹⁸ 中分别进行胰蛋白酶酶切和 PNGaseF 酶酶切过程, 使得不同样品的同一糖基化位点具有 6 Da 的质量偏差, 从而实现了糖基化位点的准确鉴定和定量(图 5). iTRAQ 技术是目前应用最为普遍的一种化学标记技术, 由于肽段定量结果来自于相对简单的 MS2 图谱, 因此可以显著提高定量结果的灵敏度和可靠性^[72]. Hu 等人^[38]利用 iTRAQ 试剂对正常与肝癌血清样品进行标记, 通过与 Ti⁴⁺-MCM-41 富集材料相结合, 有效实现了血清样品中内源性磷酸肽的定量分析.

与需要消耗昂贵商品化试剂的 iTRAQ 方法相比, 二甲基稳定同位素标记技术由于其反应效率高、试剂便宜和操作简单等优点被广泛用于各种蛋白质翻译后修饰的定量分析中. 基于该种标记定量技术, Song 等人^[73]发展了一种拟三重标二甲基同位素标记定量

图 5 串联 ¹⁸O 稳定同位素标记定量方法流程图^[71]

方法, 用于改进磷酸化蛋白质组学定量结果的准确性和分析通量. 通过在相同样品中标记上不同重的同位素标签, 可以在仅消耗一半样品量及分析时间的情况下, 将准确定量结果的数目提高 50%. 在溶液中的二甲基化学标记反应通常需要溶液置换、除盐等处理步骤, 容易造成样品的损失, 从而影响定量结果的准确性, 而采用固相化学标记反应可以很好地解决这一问题. 因此, Sun 等人^[74]发展了一种在酰肼微球上糖肽固相富集和二甲基标记一体化的定量分析方法, 通过将富集特异性高达 90% 的酰肼化学法与二甲基标记方法相结合, 可以使糖肽的定量结果在两个数量级内保持良好线性, 并具有很好的富集回收率和检测灵敏度. 针对更低丰度的内源性磷酸肽, Qin 等人^[75]发展了基于磷酸酯钛杂化介孔材料的原位富集-标记方法, 通过在介孔孔道内实现富集、除盐、标记等步骤的集成, 有效减少了标记过程中样品的损失, 提高了分析结果的检测灵敏度和分析通量.

除了对不同样品中相同蛋白质进行相对定量外, 利用多反应监测方法(MRM)对特定修饰肽段的绝对浓度进行测定, 也是修饰蛋白质组学定量分析的重要内容. 该方法主要通过加入已知浓度的重标合成肽段作为内标物, 来实现样品中具有相同序列肽段的浓度测定. Gerber 等人^[76]利用绝对定量方法

(AQUA)实现蛋白样品中非磷酸肽和磷酸肽的同时定量分析,并对人源 separase 蛋白 S1126 位点的磷酸化占有率进行检测.最近,Zhao 等人^[77]在发展的针对核心岩藻糖糖基化肽段规模化分析方法的基础上,将 O¹⁸ 内标校正与 MRM 定量方法相结合,有效实现了血清样品中含有核心岩藻糖的糖基化蛋白质的绝对定量分析.

5 前景展望

在生物体内,各种翻译后修饰不是孤立存在的,它们之间的相互协调与相互影响才使得复杂生命活动得以顺利进行^[3].例如,在细胞信号转导过程中,往往需要糖基化蛋白质和磷酸化蛋白质的协同作用才能实现外界刺激信号到细胞核内的传递.而对于发生在同一蛋白质上的不同翻译后修饰,彼此之间的位点竞争与相对丰度变化对于蛋白质不同生理学功能的发挥也具有重要意义.基于生物质谱技术的蛋白质组学分析方法为在分子水平上揭示蛋白质翻

译后修饰过程以及理解细胞信号网络调控等复杂机制提供了有力工具.

与以全部蛋白质酶解产生的肽段混合物为分析对象的自下而上研究策略(bottom-up)相比,以完整蛋白质分子为分析对象的自上而下策略(top-down)可以有效避免搜库过程中产生的假阳性鉴定结果,并且可以实现多种翻译后修饰位点的同时检测,为修饰蛋白质组学的系统分析和功能研究提供强有力的技术手段.目前,由于缺乏蛋白质层次上的高效分离方法以及质谱鉴定技术,使得 top-down 蛋白质组学研究还停留在中等复杂程度样品的小规模分析中.随着各种新型富集材料、高效分离方法、新型质谱仪和谱图解析相关的生物信息学等相关技术的不断发展, top-down 蛋白质组学技术将在翻译后修饰蛋白质的规模化分析、低灵敏度检测和位点的准确定位等方面取得巨大进步,它将与 bottom-up 策略相结合,共同在人类疾病机理研究以及疾病标志物的高通量筛选方面发挥越来越重要的作用.

参考文献

- 1 Mann M, Jensen O N. Proteomic analysis of post-translational modifications. *Nat Biotechnol*, 2003, 21: 255–261
- 2 Seo J, Lee K J. Post-translational modifications and their biological functions: Proteomic analysis and systematic approaches. *J Biochem Mol Biol*, 2004, 37: 35–44
- 3 胡箭, 郭燕婷, 李艳梅. 蛋白质翻译后修饰研究进展. *科学通报*, 2005, 50: 1061–1072
- 4 Witte E S, Old W M, Resing K A, et al. Mapping protein post-translational modifications with mass spectrometry. *Nat Methods*, 2007, 4: 798–806
- 5 Paradela A, Albar J P. Advances in the analysis of protein phosphorylation. *J Proteome Res*, 2008, 7: 1809–1818
- 6 Larsen M R, Thingholm T E, Jensen O N, et al. Highly selective enrichment of phosphorylated peptides from peptide mixtures using titanium dioxide microcolumns. *Mol Cell Proteomics*, 2005, 4: 873–886
- 7 Sugiyama N, Masuda T, Shinoda K, et al. Phosphopeptide enrichment by aliphatic hydroxy acid-modified metal oxide chromatography for nano-LC-MS/MS in proteomics applications. *Mol Cell Proteomics*, 2007, 6: 1103–1109
- 8 Yu Z Y, Han G H, Sun S T, et al. Preparation of monodisperse immobilized Ti⁴⁺ affinity chromatography microspheres for specific enrichment of phosphopeptides. *Anal Chim Acta*, 2009, 636: 34–41
- 9 Zhou H J, Ye M L, Dong J, et al. Robust phosphoproteome enrichment using monodisperse microsphere-based immobilized titanium (IV) ion affinity chromatography. *Nat Protoc*, 2013, 8: 461–480
- 10 Zhou H J, Ye M L, Dong J, et al. Specific phosphopeptide enrichment with immobilized titanium ion affinity chromatography adsorbent for phosphoproteome analysis. *J Proteome Res*, 2008, 7: 3957–3967
- 11 Lowe J B. Glycosylation, immunity, and autoimmunity. *Cell*, 2001, 104: 809–812
- 12 Glinsky G V. Antigen presentation, aberrant glycosylation and tumor progression. *Crit Rev Oncol Hematol*, 1994, 17: 27–51
- 13 Zielinska D F, Gnad F, Wisniewski J R, et al. Precision mapping of an *in vivo* N-glycoproteome reveals rigid topological and sequence constraints. *Cell*, 2010, 141: 897–907
- 14 Zhang H, Li X J, Martin D B, et al. Identification and quantification of N-linked glycoproteins using hydrazide chemistry, stable isotope labeling and mass spectrometry. *Nat Biotechnol*, 2003, 21: 660–666
- 15 Zhang H, Liu A Y, Loriaux P, et al. Mass spectrometric detection of tissue proteins in plasma. *Mol Cell Proteomics*, 2007, 6: 64–71
- 16 Zhou Y, Aebersold R, Zhang H. Isolation of N-linked glycopeptides from plasma. *Anal Chem*, 2007, 79: 5826–5837

- 17 Alpert A J. Electrostatic repulsion hydrophilic interaction chromatography for isocratic separation of charged solutes and selective isolation of phosphopeptides. *Anal Chem*, 2008, 80: 62–76
- 18 Fu Q, Guo Z, Liang T, et al. Chemically bonded maltose via click chemistry as stationary phase for HILIC. *Anal Methods*, 2010, 2: 217–224
- 19 Boersema P J, Dvecha N, Heck A J R, et al. Evaluation and optimization of ZIC-HILIC-RP as an alternative MudPIT strategy. *J Proteome Res*, 2007, 6: 937–946
- 20 Zhu J, Wang F, Chen R, et al. Centrifugation assisted microreactor enables facile integration of trypsin digestion, hydrophilic interaction chromatography enrichment, and on-column deglycosylation for rapid and sensitive *N*-glycoproteome analysis. *Anal Chem*, 2012, 84: 5146–5153
- 21 Palmisano G, Lendal S E, Engholm-Keller K, et al. Selective enrichment of sialic acid-containing glycopeptides using titanium dioxide chromatography with analysis by HILIC and mass spectrometry. *Nat Protoc*, 2010, 5: 1974–1982
- 22 Zhu J, Wang F, Cheng K, et al. A simple integrated system for rapid analysis of sialic-acid-containing *N*-glycopeptides from human serum. *Proteomics*, 2013, 13: 1306–1313
- 23 Vosseller K, Trinidad J C, Chalkley R J, et al. O-Linked N-acetylglucosamine proteomics of postsynaptic density preparations using lectin weak affinity chromatography and mass spectrometry. *Mol Cell Proteomics*, 2006, 5: 923–934
- 24 Klement E, Lipinski Z N, Kupihár Z N, et al. Enrichment of O-GlcNAc modified proteins by the periodate oxidation-hydrazide resin capture approach. *J Proteome Res*, 2010, 9: 2200–2206
- 25 Khidekel N, Ficarro S B, Peters E C, et al. Exploring the O-GlcNAc proteome: Direct identification of O-GlcNAc-modified proteins from the brain. *Proc Natl Acad Sci USA*, 2004, 101: 13132–13137
- 26 Nandi A, Sprung R, Barma D K, et al. Global identification of O-GlcNAc-modified proteins. *Anal Chem*, 2005, 78: 452–458
- 27 Anderson N L, Anderson N G. The human plasma proteome: History, character, and diagnostic prospects. *Mol Cell Proteomics*, 2002, 1: 845–867
- 28 Zheng X Y, Baker H, Hancock W S. Analysis of the low molecular weight serum peptidome using ultrafiltration and a hybrid ion trap-Fourier transform mass spectrometer. *J Chromatogr A*, 2006, 1120: 173–184
- 29 Johnson K L, Mason C J, Muddiman D C, et al. Analysis of the low molecular weight fraction of serum by LC-dual ESI-FT-ICR mass spectrometry: Precision of retention time, mass, and ion abundance. *Anal Chem*, 2004, 76: 5097–5103
- 30 Chen J Z, Anderson M, Misek D E, et al. Characterization of apolipoprotein and apolipoprotein precursors in pancreatic cancer serum samples via two-dimensional liquid chromatography and mass spectrometry. *J Chromatogr A*, 2007, 1162: 117–125
- 31 Chen T, Gagliardo R, Walker B, et al. Partial structure of the phylloxera gene from the giant monkey frog, *Phyllomedusa bicolor*: Parallel cloning of precursor cDNA and genomic DNA from lyophilized skin secretion. *Peptides*, 2005, 26: 2624–2628
- 32 Villanueva J, Philip J, Entenberg D, et al. Serum peptide profiling by magnetic particle-assisted, automated sample processing and MALDI-TOF mass spectrometry. *Anal Chem*, 2004, 76: 1560–1570
- 33 Villanueva J, Lawlor K, Toledo-Crow R, et al. Automated serum peptide profiling. *Nat Protoc*, 2006, 1: 880–891
- 34 Tian R, Zhang H, Ye M, et al. Selective extraction of peptides from human plasma by highly ordered mesoporous silica particles for peptidome analysis. *Angew Chem Int Ed*, 2007, 46: 962–965
- 35 Liu S, Li Y, Deng C, et al. Preparation of magnetic core-mesoporous shell microspheres with C8-modified interior pore-walls and their application in selective enrichment and analysis of mouse brain peptidome. *Proteomics*, 2011, 11: 4503–4513
- 36 Qin H, Gao P, Wang F, et al. Highly efficient extraction of serum peptides by ordered mesoporous carbon. *Angew Chem Int Ed*, 2011, 50: 12218–12221
- 37 Tian R J, Ren L B, Ma H J, et al. Selective enrichment of endogenous peptides by chemically modified porous nanoparticles for peptidome analysis. *J Chromatogr A*, 2009, 1216: 1270–1278
- 38 Hu L H, Zhou H J, Li Y H, et al. Profiling of endogenous serum phosphorylated peptides by titanium (IV) immobilized mesoporous silica particles enrichment and MALDI-TOFMS detection. *Anal Chem*, 2009, 81: 94–104
- 39 Choudhary C, Kumar C, Gnad F, et al. Lysine acetylation targets protein complexes and co-regulates major cellular functions. *Science*, 2009, 325: 834–840
- 40 Ong S E, Mittler G, Mann M. Identifying and quantifying *in vivo* methylation sites by heavy methyl SILAC. *Nat Methods*, 2004, 1: 119–126
- 41 Boersema P J, Foong L Y, Ding V M Y, et al. In-depth qualitative and quantitative profiling of tyrosine phosphorylation using a combination of phosphopeptide immunoaffinity purification and stable isotope dimethyl labeling. *Mol Cell Proteomics*, 2010, 9: 84–99
- 42 Han G H, Ye M L, Liu H W, et al. Phosphoproteome analysis of human liver tissue by long-gradient nanoflow LC coupled with multiple stage MS analysis. *Electrophoresis*, 2010, 31: 1080–1089

- 43 Cao W, Cao J, Huang J, et al. Enhanced N-glycosylation site exploitation of sialoglycopeptides by peptide IPG-IEF assisted TiO₂ chromatography. *Glycoconj J*, 2012, 29: 433–443
- 44 Huttlin E L, Jedrychowski M P, Elias J E, et al. A tissue-specific atlas of mouse protein phosphorylation and expression. *Cell*, 2010, 143: 1174–1189
- 45 Han G H, Ye M L, Zhou H J, et al. Large-scale phosphoproteome analysis of human liver tissue by enrichment and fractionation of phosphopeptides with strong anion exchange chromatography. *Proteomics*, 2008, 8: 1346–1361
- 46 McNulty D E, Annan R S. Hydrophilic interaction chromatography reduces the complexity of the phosphoproteome and improves global phosphopeptide isolation and detection. *Mol Cell Proteomics*, 2008, 7: 971–980
- 47 Song C X, Ye M L, Han G H, et al. Reversed-phase-reversed-phase liquid chromatography approach with high orthogonality for multidimensional separation of phosphopeptides. *Anal Chem*, 2010, 82: 53–56
- 48 Liu T, Qian W J, Gritsenko M A, et al. Human plasma N-glycoproteome analysis by immunoaffinity subtraction, hydrazide chemistry, and mass spectrometry. *J Proteome Res*, 2005, 4: 2070–2080
- 49 Tan F W, Lu L F, Cai Y, et al. Proteomic analysis of ubiquitinated proteins in normal hepatocyte cell line Chang liver cells. *Proteomics*, 2008, 8: 2885–2896
- 50 Zhang H M, Guo T N, Li X, et al. Simultaneous characterization of glyco- and phosphoproteomes of mouse brain membrane proteome with electrostatic repulsion hydrophilic interaction chromatography. *Mol Cell Proteomics*, 2010, 9: 635–647
- 51 Uhlmann T, Geoghegan V L, Thomas B, et al. A lation. *Mol Cell Proteomics*, 2012, 11: 1489–1499
- 52 Song C X, Ye M L, Liu Z X, et al. Systematic analysis of protein phosphorylation networks from phosphoproteomic data. *Mol Cell Proteomics*, 2012, 11: 1070–1083
- 53 Zhu J, Wang F, Cheng K, et al. Analysis of human serum phosphopeptidome by a focused database searching strategy. *J Proteomics*, 2013, 78: 389–397
- 54 Lee J, Xu Y D, Chen Y, et al. Mitochondrial phosphoproteome revealed by an improved IMAC method and MS/MS/MS. *Mol Cell Proteomics*, 2007, 6: 669–676
- 55 Jia W, Lu Z, Fu Y, et al. A strategy for precise and large scale identification of core fucosylated glycoproteins. *Mol Cell Proteomics*, 2009, 8: 913–923
- 56 Coon J J. Collisions or electrons? Protein sequence analysis in the 21st century. *Anal Chem*, 2009, 81: 3208–3215
- 57 Molina H, Horn D M, Tang N, et al. Global proteomic profiling of phosphopeptides using electron transfer dissociation tandem mass spectrometry. *Proc Natl Acad Sci USA*, 2007, 104: 2199–2204
- 58 Alley W R, Mechref Y, Novotny M V. Characterization of glycopeptides by combining collision-induced dissociation and electron-transfer dissociation mass spectrometry data. *Rapid Commun Mass Spectrom*, 2009, 23: 161–170
- 59 Olsen J V, Macek B, Lange O, et al. Higher-energy C-trap dissociation for peptide modification analysis. *Nat Methods*, 2007, 4: 709–712
- 60 Segu Z M, Mechref Y. Characterizing protein glycosylation sites through higher-energy C-trap dissociation. *Rapid Commun Mass Spectrom*, 2010, 24: 1217–1225
- 61 Jiang X N, Han G H, Feng S, et al. Automatic validation of phosphopeptide identifications by the MS2/MS3 target-decoy search strategy. *J Proteome Res*, 2008, 7: 1640–1649
- 62 Jiang X N, Ye M L, Han G H, et al. Classification filtering strategy to improve the coverage and sensitivity of phosphoproteome analysis. *Anal Chem*, 2010, 82: 6168–6175
- 63 Cooper C A, Gasteiger E, Packer N H. GlycoMod—A software tool for determining glycosylation compositions from mass spectrometric data. *Proteomics*, 2001, 1: 340–349
- 64 An H J, Tillinghast J S, Woodruff D L, et al. A new computer program (GlycoX) to determine simultaneously the glycosylation sites and oligosaccharide heterogeneity of glycoproteins. *J Proteome Res*, 2006, 5: 2800–2808
- 65 Wang F J, Song C X, Cheng K, et al. Perspectives of comprehensive phosphoproteome analysis using shotgun strategy. *Anal Chem*, 2011, 83: 8078–8085
- 66 Chen R, Tan Y, Wang M, et al. Development of glycoprotein capture-based label-free method for the high-throughput screening of differential glycoproteins in hepatocellular carcinoma. *Mol Cell Proteomics*, 2011, 10: M110.006445
- 67 Soderblom E J, Philipp M, Thompson J W, et al. Quantitative label-free phosphoproteomics strategy for multifaceted experimental designs. *Anal Chem*, 2011, 83: 3758–3764
- 68 Ong S E, Blagoev B, Kratchmarova I, et al. Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. *Mol Cell Proteomics*, 2002, 1: 376–386
- 69 Kruger M, Moser M, Ussar S, et al. SILAC mouse for quantitative proteomics uncovers kindlin-3 as an essential factor for red blood cell function. *Cell*, 2008, 134: 353–364

- 70 Kuster B, Mann M. O-18-labeling of N-glycosylation sites to improve the identification of gel-separated glycoproteins using peptide mass mapping and database searching. *Anal Chem*, 1999, 71: 1431–1440
- 71 Liu Z, Cao L, He Y F, et al. Tandem O-18 stable isotope labeling for quantification of N-glycoproteome. *J Proteome Res*, 2010, 9: 227–236
- 72 Ross P L, Huang Y L N, Marchese J N, et al. Multiplexed protein quantitation in *Saccharomyces cerevisiae* using amine-reactive isobaric tagging reagents. *Mol Cell Proteomics*, 2004, 3: 1154–1169
- 73 Song C X, Wang F J, Ye M L, et al. Improvement of the quantification accuracy and throughput for phosphoproteome analysis by a pseudo triplex stable isotope dimethyl labeling approach. *Anal Chem*, 2011, 83: 7755–7762
- 74 Sun Z, Chen R, Cheng K, et al. A new method for quantitative analysis of cell surface glycoproteome. *Proteomics*, 2012, 12: 3328–3337
- 75 Qin H, Wang F, Wang P, et al. Phosphoric acid functionalized mesoporous organo-silica (EPO) as the adsorbent for *in situ* enrichment and isotope labeling of endogenous phosphopeptides. *Chem Commun*, 2012, 48: 961–963
- 76 Gerber S A, Rush J, Stemman O, et al. Absolute quantification of proteins and phosphoproteins from cell lysates by tandem MS. *Proc Natl Acad Sci USA*, 2003, 100: 6940–6945
- 77 Zhao Y, Jia W, Wang J F, et al. Fragmentation and site-specific quantification of core fucosylated glycoprotein by multiple reaction monitoring-mass spectrometry. *Anal Chem*, 2011, 83: 8802–8809

The new method and technology for the characterization and identification of post-translational modifications of proteins

SONG ChunXia^{1,2}, SUN Zhen¹, QIN HongQiang¹, YE MingLiang¹, ZHANG LiHua¹, ZHANG YuKui¹ & ZOU HanFa¹

¹National Chromatographic R&A Center, Chinese Academy of Sciences Key Laboratory of Separation Science for Analytical Chemistry, Dalian Institute of Chemical Physics, Chinese Academy of Sciences, Dalian 116023, China;

²SINOPEC Research Institute of Petroleum Processing, Beijing 100083, China

Post-translational modifications (PTMs) are covalent processing events of proteins. It is the reason that the complex biological process can be regulated under the limited gene number. As many as 300 PTMs of proteins play important roles in many living activities such as regulation of cell cycle, signal transduction and immune response. Most of the protein PTMs have the characteristics such as low abundance, high dynamic range and complex structures. Therefore, there are enormous challenges for the large scale analysis of protein PTMs. The PTMs proteomics approach which is based on the mass spectrometry emerges as an ideal strategy for the global analysis of protein PTMs. Now, the large scale analysis of protein PTMs has been the key research area of proteomics. In this review, we focus on the development of new method and technology of PTMs proteomics, including the enrichment strategies, separation method, mass spectrometry techniques and quantitative methods of the protein PTMs.

post-translational modifications, phosphorylation, glycosylation, endogenous peptides, proteomics

doi: 10.1360/972013-515