

蛋白质翻译后修饰研究进展

阮班军¹ 代 鹏¹ 王 伟¹ 孙建斌¹ 张文涛¹ 颜 真^{1*} 杨静华^{2*}

(¹肿瘤生物学国家重点实验室, 第四军医大学药学院药物基因组学教研室, 西安 710032; ²美国波士顿大学
肿瘤外科实验室及VA临床蛋白组学实验室, 波士顿 02130, 美国)

摘要 蛋白质是执行细胞功能的基本功能单元, 其表达受基因组和表观遗传学的调控。通常, 蛋白质在表达以后还需要经过不同程度的修饰才能发挥所需要的功能。这种翻译后修饰过程受到一系列修饰酶和去修饰酶的严格调控, 使得在某一瞬间细胞中蛋白质表现出某种稳定或动态的特定功能。最新的研究表明, 真核细胞中存在着各种各样的蛋白质修饰过程, 其中大约70%目前还无法解释。有理由认为, 这种经过了特定修饰的蛋白质, 更客观地反映了细胞的各种生理以及病理过程。因此, 除了基因组所编码的“裸”蛋白质组的表达以外, 更需要对经过翻译后修饰的蛋白质及蛋白质组的调控过程进行深入的研究。该文对常见翻译后修饰以及研究方法进行了综述。

关键词 蛋白质组; 蛋白质翻译后修饰; 基因组; 分子鉴定

Progress on Post-translational Modification of Proteins

Ruan Banjun¹, Dai Peng¹, Wang Wei¹, Sun Jianbin¹, Zhang Wentao¹, Yan Zhen^{1*}, Yang Jinghua^{2*}

(*The State Key Laboratory of Cancer Biology, Department of Pharmacogenomics, School of Pharmacy, The Fourth Military Medical University, Xi'an 710032, China; ²Department of Surgery, VA Boston Healthcare System, Boston University School of Medicine, Boston MA02130, USA*)

Abstract Although protein expression is regulated through genetics and epi-genetics, its function is determined by many factors, of which post-translational modification (PTM) is particularly important. Protein modifications are precisely regulated through a variety of modifiers and de-modifiers, most of which are yet to be defined. Most physiological or pathological processes are usually regulated through typical but complex and dynamic PTM patterns and networks. Reasonably, proteome of PTM is more precise to describe the physiological or pathological processes of cells at the molecular level to clarify protein functions and easy to find out biomarkers and molecule targets for disease diagnosis and drug development. Here we reviewed some common types of modifications of proteins and methods in PTM research.

Key words proteomics; post-translational modification; genomics; molecular identification

蛋白质是执行细胞功能的基本功能单元, 其表达受基因组、表观遗传学和翻译后修饰多个层次的调控。基因组学的主要任务是研究生物基因组的组成以及组内各基因的精确结构、相互关系与表达调控; 表观基因组^[1]则指示了每种细胞类型及其基因

组中的独特基因表达程序。以上两种学科的研究都只是停留在基因的层面, 然而蛋白质作为生命活动的直接执行者, 还需要经过不同程度的加工和修饰才能发挥功能。最新的研究表明, 真核细胞中存在着各种各样的蛋白质修饰方式, 调控着细胞的各种

收稿日期: 2013-10-13 接受日期: 2014-03-26

国家自然科学基金(批准号: 81071369、30928031、81272276)和国家重点基础研究发展计划(973计划)(批准号: 2010CB933902)资助的课题

*通讯作者。Tel: 029-84772368, E-mail: yanzhen@fmmu.edu.cn; jyang@bu.edu.cn

Received: October 13, 2013 Accepted: March 26, 2014

This work was supported by the National Natural Science Foundation of China (Grant No. 81071369, 30928031, 81272276) and the National Key Basic Research Program of China (973 Program) (Grant No.2010CB933902)

*Corresponding authors. Tel: +86-29-84772368, E-mail: yanzhen@fmmu.edu.cn; jyang@bu.edu.cn

网络出版时间: 2014-07-02 10:25 URL: <http://www.cnki.net/kcms/doi/10.11844/cjcb.2014.07.0299.html>

生理以及病理过程, 因此对经过翻译后修饰的蛋白质和蛋白质组需要更为深入的研究。本文对常见的蛋白质翻译后修饰、多种修饰的功能协同作用对生理和病理的影响以及蛋白翻译后修饰的常用研究方法进行了综述。

1 蛋白质翻译后修饰概述

早在HGP实施期间的1995年, 蛋白质组(proteome)一词就由澳大利亚学者Williams和Wilkins^[2]提出, 该词源于蛋白质(protein)与基因组(genome)两个词的组合, 意指“proteins expressed by a genome”, 即“一个细胞或一个组织基因组所表达的全部蛋白质”。其对应于基因组的所有蛋白质构成的整体, 将研究从个别蛋白质扩展到从整体水平研究蛋白质。然而, 对于蛋白质组的功能实现却远比想象中的复杂。研究发现, 蛋白质组处于一个新陈代谢的动态变化过程中, 蛋白质的合成在不同时空下受到不同因素的调控, 同一种细胞的生长与活动, 因不同时期、不同条件其蛋白质组也在不断变化过程中, 病理或疾病治疗过程中的细胞蛋白质组及其变化也与正常生理过程不同^[3]。即虽然高等生物体内所有细胞仅拥有一个相同的基因组, 但不同分化的组织细胞却拥有不同的蛋白质组。

其中, 蛋白质翻译后修饰作为蛋白质功能调节

的一种重要方式, 对蛋白质的结构和功能至关重要, 也是蛋白质组复杂性的主要原因之一。据估计, 人体内50%~90%的蛋白质发生了翻译后修饰, 有的是肽链骨架的剪接, 有的是在特定氨基酸侧链上添加新的基团, 还有的是对已有基团进行化学修饰^[4]。生物体通过种类繁多的修饰方式直接调控蛋白质的活性, 也大大扩展了蛋白质的化学结构和功能, 显著增加了蛋白质的多样性和复杂性, 使可编码的蛋白质种类大大超过了20种天然氨基酸的组合限制。因此, 经典的一个基因一个蛋白的对应关系早已经被打破, 据估计人类基因组可编码蛋白质的2万个基因, 可表达超过20万个功能蛋白质。目前已经确定的翻译后修饰方式超过400种, 常见修饰过程有磷酸化、泛素化、甲基化、乙酰化、糖基化、SUMO化、亚硝基化、氧化等等, 表1列出了Swiss-Prot Knowledgebase更新至2014年6月所收集的以实验方法发现的蛋白质翻译后修饰的种类和频率。然而, 我们已掌握的蛋白修饰过程还非常有限, 至少70%还不为人知, 包括未知的修饰种类、未知的修饰蛋白质、未知的蛋白修饰位点等。蛋白质修饰分为可逆和不可逆两种方式, 不可逆的如O位的羧基端甲基化, 而磷酸化、N-位甲基化、N-乙酰化修饰均属于可逆的修饰, 会随着细胞的生理状态和外界环境变化而改变, 从而起到细胞内外信号传递、酶原激活的作用,

表1 蛋白质翻译后修饰频率(<http://selene.princeton.edu/PTMCuration/>)

Table 1 Frequency of post-translational modification of proteins (<http://selene.princeton.edu/PTMCuration/>)

频率	修饰类型	频率	修饰类型
Frequency	Modification	Frequency	Modification
39 818	Phosphorylation	156	ADP-ribosylation
9 119	Acetylation	140	Citrullination
5 976	N-linked glycosylation	82	Farnesylation
3 181	Amidation	81	S-nitrosylation
1 774	Methylation	63	Deamidation
1 698	Hydroxylation	62	Geranyl- geranylation
1 338	O-linked glycosylation	57	Formylation
1 136	Ubiquitylation	52	Nitration
921	Pyrrolidone Carboxylic Acid	39	GPI anchoring
595	Sulfation	33	Bromination
503	Sumoylation	29	S-diacylglycerol cysteine
413	Gamma-Carboxyglutamic Acid	21	FAD
348	Palmitoylation	9 571	Others
184	Myristoylation	67 975	Total characterized
156	C-linked glycosylation	77 546	Total processed

因此可以作为蛋白质构象和活性改变的调控开关,一旦发生异常的修饰常常导致疾病的发生。正因如此,对蛋白质序列及其多样化的修饰尚待更为深入的探究,以阐明蛋白质多样化的动态修饰对蛋白质功能和细胞功能的影响及其分子机制。

2 常见的蛋白质翻译后修饰

2.1 磷酸化

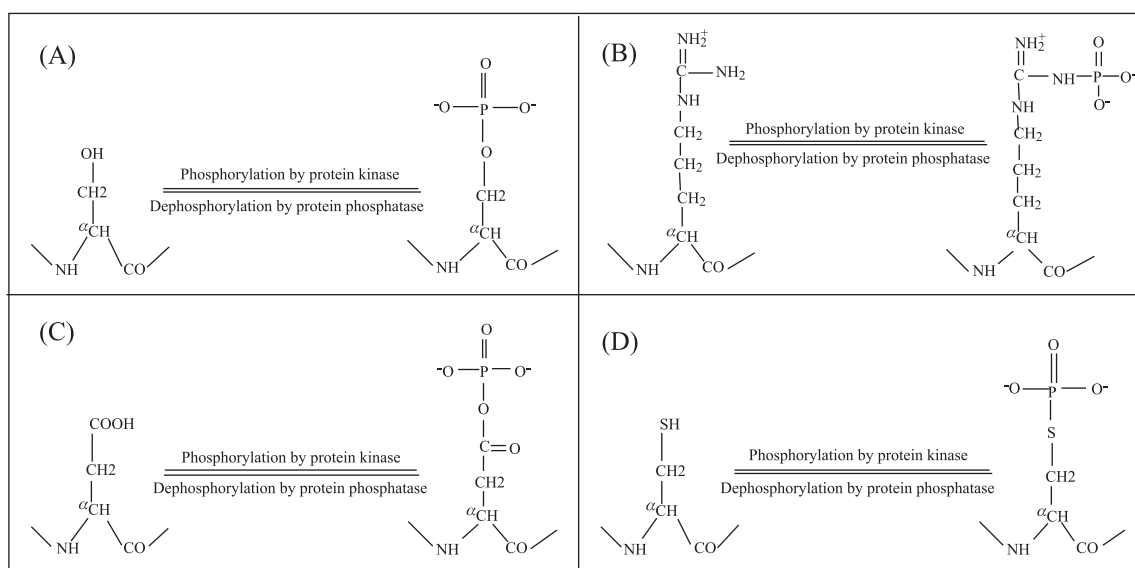
作为研究最广泛以及体内最为常见的蛋白质翻译后修饰——蛋白质磷酸化修饰,在蛋白激酶的作用下,将一个ATP或GTP上 γ 位的磷酸基转移到底物蛋白质的氨基酸残基上,影响了人类细胞中超过1/3的蛋白质的功能。磷酸化发生的位点通常是Ser、Thr、Tyr侧链的羟基或是His、Arg、Lys侧链的氨基,少数发生在Asp和Gln的侧链羧基或Cys的侧链巯基,因此,可将磷酸化蛋白质分为O-磷酸化、N-磷酸化、酰基磷酸化和S-磷酸化四类(图1)^[5]。研究发现,His位点发生磷酸化修饰的频率高于Tyr^[6]。近年来,随着磷酸化肽段富集手段和高精度生物质谱发展,蛋白质磷酸化修饰位点的鉴定从数量和精度上得到了长足发展,然而如何将这此磷酸化蛋白质与催化其磷酸化的蛋白质激酶——对应成为了近年来该领域的热点和难点问题。Tian等^[7]在一项研究中首次从蛋白质组学层面对嗜热四膜虫的蛋白质

磷酸化修饰进行分析,该研究以具有1 069个蛋白质激酶(人类仅为518个)的单细胞真核模式生物——嗜热四膜虫为研究对象,利用TiO₂对磷酸化肽段进行了富集,并利用2D(SCX/RP)-nanoLC-MS/MS技术鉴定得到了1 008个蛋白质的2 238个磷酸化修饰位点。

磷酸化修饰几乎发生在生命活动的各个过程,如细胞膜上的Na⁺-K⁺-ATP酶泵的磷酸化控制钠离子和钾离子的胞内浓度;Src家族Tyr激酶的磷酸化是它激活的“开关”,将使其构象改变从而能识别下游蛋白^[8];在细胞周期调控中,CDKs(cyclin-dependent protein kinases)这一类Ser/Thr激酶是否能激活Cyclins并形成复合物,决定了细胞周期是否阻滞^[7]。磷酸化也是细胞呼吸过程的重要参与者,线粒体中氧化磷酸化的过程需要NADPH氧化酶的磷酸化调控,电子传递链中蛋白分子之间的相互作用由磷酸化调控^[8]。此外,蛋白质降解也需要依赖于ATP参与的磷酸化修饰。

2.2 糖基化

在生物体内,糖类一般并不单独存在,而是连接在蛋白质或脂类分子上分别构成糖蛋白和糖脂,其中以己糖最为多见,包括葡萄糖半乳糖和甘露糖以及它们的一些简单修饰形式,如葡萄糖的 α -羟基被酰化氨基取代生成N-乙酰葡萄糖胺。根据蛋白质



A: O-磷酸化; B: N-磷酸化; C: 酰基磷酸化; D: S-磷酸化。

A: O-phosphorylation; B: N-phosphorylation; C: acetyl-phosphorylation; D: S-phosphorylation.

图1 磷酸化反应过程示意图

Fig.1 Schematic diagram of phosphorylation

被糖类修饰形式的不同把蛋白质糖基化分成N-糖基化、O-糖基化、C-糖基化和糖基磷脂酰肌醇锚定连接四类^[9]。最近几年,对糖基转移酶的研究大幅度增加,主要针对这些酶的三维空间结构与反应机制,在对N-糖基化、O-糖基化的研究中发现,在受体激活过程中存在一些不常见的折叠方式与化学机制^[10]。另外,来自上海交通大学系统生物医学研究院的研究人员首次提出了内质网中可能存在新的蛋白质O-乙酰半乳糖胺(GalNAc)糖基化调控机制,挑战了教科书上关于蛋白质的O-GalNAc糖基化过程只发生在高尔基体中这一传统观念^[11]。N-糖基化是最为常见的糖基化修饰,该修饰是由多糖与蛋白质的Asn残基的酰氨基连接形成,糖链为N-乙酰葡萄糖胺(GlcNAc),发生在特定的氨基酸模序Asn-X-Ser/Thr或Asn-X-Cys上。根据单糖的位置, N-糖基化修饰可以进一步划分为复杂型多糖(图2A)、混合型多糖(图2B)和高甘露糖型多糖(图2C),研究表明,三种类型的糖链具有共同的生物合成起源,即高甘露糖聚合物前体^[12]。

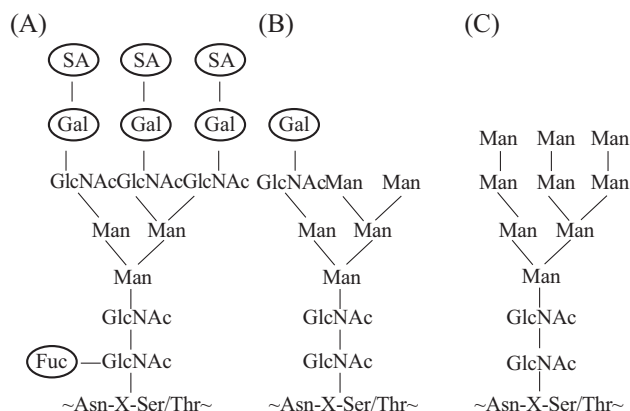


图2 三种不同N-糖基化修饰类型的结构示意图
(根据参考文献[12]改编)

Fig.2 Schematic diagram of three different N-glycosylation type (modified from reference [12])

糖基化是真核生物蛋白质功能调节的基本环节,异常糖基化往往伴随着生理病理过程的异常。有研究发现, N-GlcNAc可以通过Ser或Thr残基,以单糖分子连接到蛋白质上,生成O-GlcNAc。该种修饰可发生在许多与癌变相关的蛋白分子上,如 β 2连环蛋白、p53、pRb家族等。O-GlcNAc化也与糖尿病^[13]、神经退变性疾病^[14]等疾病有关。因此, O-GlcNAc可以作为临床上这些疾病诊断的分子标志物。目前,已经有超过600种蛋白质被鉴定发生了O-GlcNAc化,但

其中只有低于15%的修饰位点得到了鉴定,大多数来自于脑和脊髓组织。O-GlcNAc修饰的重要性及与磷酸化修饰之间的联系,正在成为继磷酸化修饰之后的又一个蛋白质翻译后修饰研究热点^[15]。

2.3 泛素化

泛素是一种含76个氨基酸的多肽,存在于除细菌外的许多不同组织和器官中,具有标记待降解蛋白质的功能。被泛素标记的蛋白质在蛋白酶体中被降解。由泛素控制的蛋白质降解具有重要的生理意义,它不仅能够清除错误的蛋白质,还对细胞周期调控、DNA修复、细胞生长和免疫功能等都有重要的调控作用。

泛素蛋白C末端的Gly通常都经由异肽键与底物蛋白的氨基连接在一起,最常见的连接是与底物蛋白Lys的 ϵ 氨基相连以及与底物蛋白的N末端相连。最近还发现,可以与Cys、Ser和Thr相连^[16]。泛素控制蛋白质水解的整个过程是通过泛素-蛋白酶体系统(ubiquitin proteasome system, UPS)进行的^[17],介导了真核生物体内80%~85%的蛋白质降解。去泛素化酶(deubiquitin, DUB)可以将靶蛋白上的泛素蛋白水解下来。由于DUB酶的存在,泛素化作用只能是暂时的。这种动态修饰过程构成了一个可逆的“开关”,来控制底物蛋白的不同功能状态,调控细胞内的多种生理活动^[18](图3)。

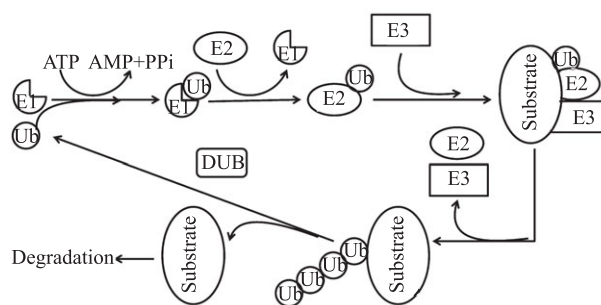


图3 蛋白质的泛素化修饰过程示意图

Fig.3 Schematic diagram of protein ubiquitination process

内质网相关的降解过程能清除错误折叠蛋白的分泌途径,同时介导一些内质网残留蛋白的调控降解过程。研究发现,一种蛋白与一种泛素连接酶之间相互作用的细微增加,都能引发信号底物的降解,一项最新的研究解析了其中的作用机制,指出去泛素化可以作为一种信号放大器,放大信号,从而进行下游调控^[19]。

泛素化修饰途径在人体免疫调节过程中发挥着重要的作用, 其不仅能够通过降解免疫调控蛋白的方式来抑制免疫反应, 也能通过多泛素化修饰作用激活NF- κ B和IRF途径中的激酶来启动一些免疫反应。随着能确定蛋白类型及其修饰状况的质谱检测等高灵敏度检测技术的进步, 我们能够更容易地发现泛素化途径的底物, 并且近年来发现了多种UBD蛋白, 因此我们更容易地弄清楚泛素化修饰途径在机体免疫调控中发挥的真正作用, 这将有助于我们对免疫性疾病的诊疗技术研究^[20]。

细胞内很多受泛素途径修饰的蛋白质都参与控制细胞内肿瘤发生相关生理或病理进程, 比如细胞周期调控、凋亡、受体下调以及基因转录等多种重要的细胞进程, 所以一旦泛素修饰系统出了问题就很有可能引起人体各种肿瘤的发生。因此, UBL系统也是治疗肿瘤的靶点之一, 有望开发出更多有效的抗癌疗法和药物^[21]。

2.4 SUMO化

随着泛素化的研究, 科学家们相继发现了一些类泛素蛋白, 其中小泛素相关修饰物(small ubiquitin-related modifier, SUMO)是最受瞩目的一类。SUMO化与泛素化修饰过程相似, 也是一个多酶参与的反应, 但两个途径参与的酶则完全不同。首先, SUMO化修饰比泛素化修饰多一步成熟化的过程, 即SUMO前体在SUMO蛋白酶(如Ulp1)的作用下, C-端的4个(或多个)氨基酸残基被切除, 生成成熟的SUMO并露出C-端2个Gly残基。接着, 在SUMO化活化酶E1、结合酶E2和连接酶E3的作用下完成SUMO化修饰过程, 并且这个过程是可逆的, 称为去SUMO化^[22-23](图4)。

2012年, 来自德国马普生物化学研究院的研究人员以DNA双链断裂修复作为例, 解析了类泛素蛋白SUMO的作用新机制, 指出蛋白之间的SUMO化修饰稳定相互作用是由DNA结合SUMO连接酶催化, 并由单链DNA开启的, 而且只有整体清除几种修复蛋白的SUMO化修饰, 才会通过大量减慢DNA修复, 影响同源重组途径。因此, 研究人员认为SUMO能协同作用于几种蛋白, 单个修饰叠加促进有效修复, 这也就是说, SUMO化修饰过程也许常常能靶向一组蛋白, 而不是单个蛋白, 局部修饰酶和高度特异性启动则能确保特异性^[24]。

SUMO化修饰不介导蛋白质的降解, 但依然对许多生理过程产生重要作用。例如, 通过调节蛋白质之间的相互作用、影响靶蛋白在细胞内的分布、阻碍泛素蛋白对靶蛋白的共价修饰、提高靶蛋白的稳定性等。此外, SUMO还借助各种方式参与DNA复制和修复以及转录调控过程^[25]。

Akt基因是一种原癌基因, 其表达产物蛋白激酶B(protein kinase B, PKB)具有Ser/Thr激酶活性, 作为磷脂酰肌醇-3-激酶(phosphoinositide-3-kinase, PI3K)信号转导通路下游的信息分子, 已在许多常见肿瘤的研究中发现其过量表达, 而研究人员已证实, SUMO化修饰是一种新型的Akt激活机制, 因此, Akt的SUMO化可能为肿瘤的治疗提供新的靶点^[23]。

2.5 乙酰化

蛋白质的乙酰化修饰影响着细胞生理的各个方面, 主要集中在对染色体结构的影响以及对核内转录调控因子的激活方面, 参与了包括转录调控、信号通路调控、蛋白质稳定性调控、细胞代谢和病原微生物感染调控等多个生理病理过程而被广泛

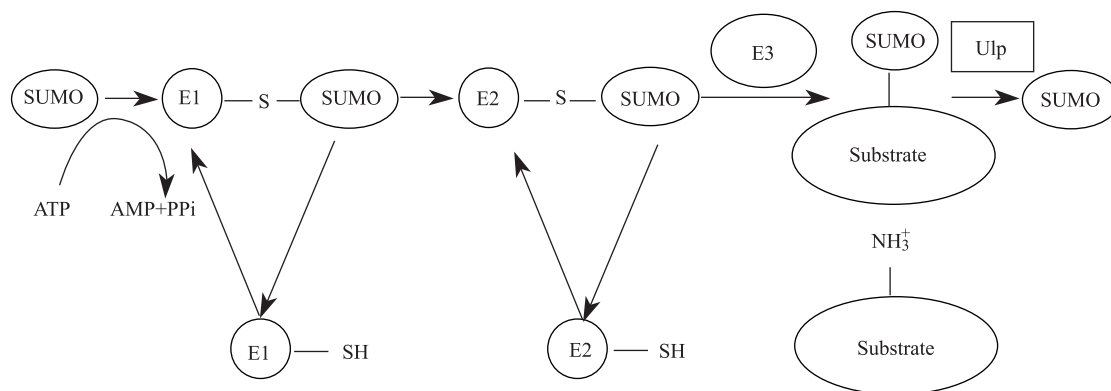


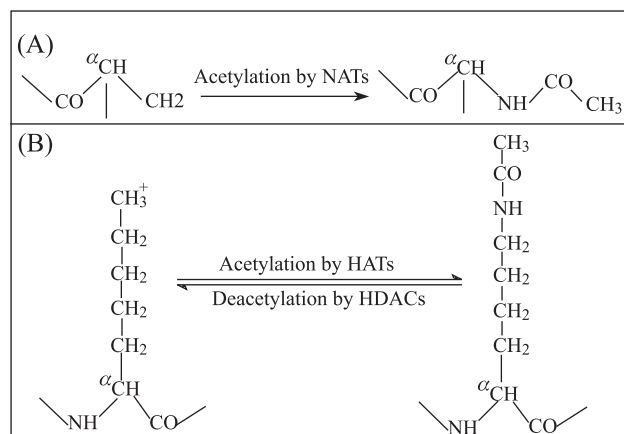
图4 蛋白质的SUMO化修饰过程示意图

Fig.4 Schematic diagram of protein SUMOylation process

关注。2012年, 清华大学生命学院俞立教授课题组在*Science*杂志发表文章, 揭示了蛋白质乙酰化修饰对细胞自噬调控的分子机制, 指出组蛋白乙酰化酶Esa1以及去乙酰化酶Rpd3通过调节自噬发生关键蛋白Atg3的乙酰化水平, 从而实现对自噬过程的动态调控^[26]。

目前文献报道的乙酰化修饰主要有三大类: N-端 α 氨基乙酰化修饰(图5A)、Lys的位氨基乙酰化(图5B)和Ser/Thr的乙酰化修饰。2006年, Mukherjee等^[27]通过研究鼠疫耶尔森菌分泌的毒力因子Yop J对细胞MAPK通路和NF κ B信号通路的阻断作用发现了乙酰化修饰可以发生在磷酸化修饰的Ser位点上, 从而竞争磷酸化修饰来干扰宿主细胞信号转导通路。此类非氨基乙酰化修饰很有可能是一种未被发现的存在于真核生物细胞自身的一种新的修饰类型, 为乙酰化修饰在生理调控领域开拓了新的空间。

科学界早期一般认为, 乙酰化修饰功能主要集中在对细胞染色体结构的影响以及对核内转录调控因子的激活方面。但是, 复旦大学的科研人员通过通量化的蛋白质组研究和不同物种的代谢通路研究发现, 在生理状况下, 存在着大量非细胞核的蛋白被乙酰化修饰。中国科学院上海生命科学院赵国屏院士和复旦大学赵世民领导的科研团队发现, 在沙门氏菌(*Salmonella*)中, 作为对不同碳源的响应, 中心代谢酶被广泛而有差异地可逆乙酰化, 并伴随着细胞生长和代谢流的变化, 可逆代谢酶乙酰化确保细胞可以通过即时感应细胞能量状态和灵活地改变反应速率或方向来响应环境变化^[28-29]。



A: N-端 α 氨基乙酰化; B: 赖氨酸 ϵ 氨基乙酰化。

A: N $^{\alpha}$ -terminal acetylation); B: ϵ -amino group of lysine acetylation.

图5 乙酰化反应过程示意图

Fig.5 Schematic diagram of acetylation

2.6 甲基化

甲基化是指从活性甲基化合物(如S-腺苷甲基硫氨酸)上将甲基催化转移到其他化合物的过程, 底物包括核酸、蛋白、激素和脂类。1964年, Allfrey等^[28]首次报道了Arg甲基化, 在随后的研究中人们发现其在信号转导、转录活化、蛋白质分拣等生命过程中发挥着重要作用。

蛋白质甲基化修饰是在蛋白质甲基转移酶的作用下, 将甲基转移至特定的氨基酸残基。常见的甲基化发生在Lys(图6A)和Arg(图6B)上, 分为组蛋白Lys/Arg甲基化^[29]和非组蛋白Lys/Arg甲基化^[30], 分别由相应的蛋白质Lys甲基化酶和Arg甲基化酶介导反应。甲基化是一种可逆的修饰过程, 去甲基化酶主要有Lys特异性去甲基化酶和包含Jumonji结构域的蛋白质去甲基化酶家族。

蛋白质的甲基化修饰增加了立体阻力, 并且取代了氨基酸残基上的氢, 影响了氢键的形成。蛋白质发生甲基化后, 变换了肽链原来的结构顺序, 可以编码出更多的信息, 从而调控信号分子间和信号分子与目标蛋白之间的相互作用, 进而参与多种生命调控过程, 如转录调控、RNA加工和运输、蛋白翻译、信号转导、DNA修复、蛋白质相互作用、细胞发育与分化等, 并与一些疾病(如肿瘤^[31]、心血管疾病)的发生发展密切相关。

3 蛋白质修饰的协同与动态调控网络

在生物体内, 各种蛋白质的修饰种类繁多, 并且存在“时空特异性”, 即蛋白质修饰类型和程度随着生物发育过程、生存环境与内在状态的变化而不断发生变化, 使蛋白质的修饰成为一个动态变化的过程, 而且往往各种修饰不是孤立存在的, 同一生理或病理过程的发生, 需要各种修饰的蛋白质共同作用, 同一个蛋白质又可以同时拥有一种以上的蛋白质修饰过程, 彼此之间相互影响、相互协调。

阿尔茨海默氏病(Alzheimer's disease, AD)^[32]是一种以记忆力损害和认知障碍为主的中枢神经系统退行性疾病, 在AD中扮演重要角色的淀粉样前体蛋白(amyloid precursor protein, APP)可以被SUMO1和SUMO2修饰, 该修饰可以降低 β -淀粉样蛋白(amyloid- β , A β)的积聚。推测AD患者中可能存在APP代谢过程中的SUMO化失调^[33-34], 但失调的SUMO是否是AD的潜在危险因素以及能

Cdk)^[41]、10号染色体缺失的磷酸酶及张力蛋白同源基因(phosphatase and tensin homolog deleted on chromosome 10, PTEN)^[42]、热休克蛋白90(heat shock protein 90, Hsp90)^[43]等多种蛋白质巯基亚硝基化与AD的病理机制存在诸多广泛的联系。

组蛋白是另一个研究较多的常见的多修饰蛋白质。四种常见组蛋白(H2A、H2B、H3和H4)的游离氨基端具有多个修饰位点,包括乙酰化/去乙酰化、甲基化/去甲基化、泛素化/去泛素化^[44]、磷酸化^[45]以及SUMO化、生物素化等,其最基本的作用是调控基因表达,例如组蛋白甲基化多导致基因沉默,去甲基化则相反;乙酰化一般使转录激活,去乙酰化则相反。这些修饰方式及其作用的发挥并不是相互独立的,很多时候它们通过协同或拮抗来共同发挥作用,形成多亚基复合物,与核小体重修饰复合物(NuRcs,如Swi/Snf、RSC、NURF)相互作用,重修饰染色质^[46]。有证据表明,这些重修饰复合物通过组蛋白尾部由这些复合物所调节的启动子处不同乙酰化方式的因子募集和识别而联合起作用。另外,修饰的发生受细胞生理状态的影响,如DNA复制或转录时常会发生组蛋白的去甲基化与乙酰化,激活基因的表达。

可见,单纯的单一蛋白质、单一修饰的研究难以阐明生理病理机制,只有运用组学的方法,对生理病理状态下的蛋白质修饰动态网络进行研究,才有可能找到问题的答案。

4 蛋白质翻译后修饰常用的研究方法

蛋白质翻译后修饰的常用方法,包括蛋白质分离技术、以物质谱(mass spectrum, MS)为代表的鉴定技术、蛋白质相互作用分析技术以及生物信息学技术等,用于对蛋白质的高级结构、修饰、定位、活性、相互作用以及功能的研究。

4.1 蛋白质的分离技术

双向凝胶电泳技术(two dimensional gel electrophoresis, 2D-PAGE)是常用的蛋白质分离技术,由O'Farrell于1975年创立,用于分离细胞或组织蛋白质粗提物、构建特定组织或细胞的蛋白质“二维参考图谱”,分析特定条件下病理与生理蛋白质组的表达差异。2D-PAGE结合MS分析可以实现较大规模的蛋白质分离和鉴定,特别适用于修饰后蛋白质(如糖基化、磷酸化、脱氨等)的发现和

鉴别。然而该技术也存在很多不足,如对极酸或极碱性蛋白质、疏水性蛋白质、极大或极小蛋白质以及低丰度蛋白质难以有效分离,胶内酶解过程费时、费力,难以与质谱联用实现自动化等^[47]。

双向荧光差异凝胶电泳(two dimensional fluorescence difference gel electrophoresis, 2D-DIGE)是同时分离多个由不同荧光标记的样品,并以荧光标记的样品混合物为内标,对每个蛋白质点和每个差异都进行统计学分析的蛋白质组学凝胶定量代表性技术。这一技术具有以下优点:(1)精确,可以检测到<10%的样品间差异,可信度>95%;(2)标准化,内标最大程度地降低偏差,得出精确的实验结论;(3)高重复性,极佳的定量重复性和准确的量化蛋白表达数据;(4)高效,多个样品在同一块胶上同时分离,减少实验时间;(5)简化,易于使用,可以完全整合到蛋白质组工作流程中^[48]。

高效液相色谱(high performance liquid chromatography, HPLC),基于样品分子在固定相和流动相之间的特殊相互作用而实现样品分离,无须变性处理样品,可实现上样、收集、与质谱联用(LC/MS)、在线分析的自动化。与双向电泳相比,具有快速、分辨率高的优点^[49]。

4.2 蛋白质的鉴定技术

生物质谱技术是蛋白质鉴定的主要技术,通过测定样品的质荷比(m/z)进行高分子物质成分和结构的精确分析,基于电离方法的不同,可分为基质辅助激光解吸电离-飞行时间质谱(matrix assisted-laser desorption ionization-time of flight mass spectrometry, MALDI-TOF-MS)与电喷雾-串联质谱(ESI-MS/MS)^[50-52]。ESI-MS使分析物从溶液相中电离,适合与液相分离手段(如液相色谱和毛细管电泳)联用分析复杂样品。而MALDI是在激光脉冲的激发下,使样品从基质晶体中挥发并离子化,适于分析简单的肽混合物。液质联用(LC/MS)、串联质谱(MS/MS)的应用,不仅可获得蛋白质的肽谱图,更可获得蛋白质氨基酸序列信息和精确的修饰信息。

ETDC(electron-transfer and higher-energy collision dissociation)是2012年Christian等^[53]提出的在串联质谱法中使用的一种新的碎片化方案。通过组合电子转移解离(ETD)和更高的能量碰撞解离(HCD)生成双碎片离子系列,并促进广泛的肽骨架碎片。在电子转移解离步骤之后,所有的离子,包括未反

应的前体离子, 进行碰撞在单一的频谱诱导解离产生b/y-和c/z-型碎片离子。相较于一般的破碎方法, EThcD方案提供了更丰富的MS/MS谱图, 大大增加了肽序列覆盖面, 可以进行确切的肽测序和PTM定位^[54]。

蛋白质基因组学(proteo-genomics)是一种全新的分析方法, 最早由哈佛大学Jaffe等^[55]于2004年提出, 是蛋白质组学与基因组学的新兴交叉, 指运用蛋白质组技术注释基因组, 以弥补以往依赖DNA及RNA序列信息注释基因组的不足。除此之外, 蛋白质基因组学技术还可以进行蛋白质翻译后修饰的研究, 这主要归功于串联质谱技术已经发展成熟, 实现了对蛋白质组的高覆盖, 使得利用串联质谱数据进行基因组注释成为可能, Gupta等^[56]因此成功绘制出了*S. oneidensis* MR-1的蛋白质翻译后修饰全图谱, 包括对蛋白质的化学修饰、信号肽的切割以及N末端甲硫氨酸的切割等, 这项技术的应用将给蛋白质翻译后修饰的研究提供更直观、更全面的信息。

4.3 蛋白质相互作用分析技术

目前, 已建立了包括酵母双杂交系统、免疫共沉淀(Co-IP)、串联亲和纯化(TAP)技术、表面等离子共振技术(surface plasmon resonance, SPR)、荧光共振能量转移(FRET)、GST-Pull-down实验以及蛋白质芯片等多种研究方法, 这些方法各有所长, 为蛋白质相互作用的研究提供了有效的分析方法, 也为蛋白质组学研究奠定了坚实的基础^[57]。表面等离子共振技术(SPR)是目前蛋白质相互作用研究中的新手段, 其利用一种纳米级的薄膜吸附上诱饵蛋白, 当待测蛋白与诱饵蛋白结合后, 薄膜的共振性质会发生改变, 通过检测便可知这两种蛋白的结合情况。SPR技术的优点是不需标记物或染料, 反应过程可

实时监控, 测定快速且安全, 还可用于检测蛋白-核酸及其他生物大分子之间的相互作用^[58]。

4.4 生物信息技术

生物信息学是在生命科学、计算机科学和数学的基础上逐步发展而形成的一门新兴交叉学科, 是以理解各种数据的生物学意义为目的, 运用数学与计算机科学手段进行生物信息的收集、加工、存储、传播、分析与解析的科学。生物信息学虽然以基因组信息学为核心, 但其在蛋白质组学研究中的作用日显突出。通过二维凝胶电泳、生物质谱、蛋白质微阵列等方法所获得的大量数据最终都必须依赖生物信息学的方法和手段才能实现对蛋白质的种类、数量、结构和功能的最后确定, 被称为蛋白质组研究的百科全书。常用的数据库有: Swiss-Prot^[59]、GenBank、PhosphoELM^[60]、Phosphosite^[61]、O-glycbase等^[62](表2), 修饰相关数据库多集中在磷酸化和糖基化, 其中SwissProt是高质量的非冗余蛋白数据库, 同时也有多种翻译后修饰的注释信息, 而PhosphoELM和Phosphosite详细收录了实验验证的磷酸化数据。

5 结论与展望

2007年, Dai和Rasmussen^[63]在干细胞分化研究中首次应用“表观蛋白质组(epi-proteome)”一词来特指组蛋白的CpG甲基化修饰及其形成的特殊图谱, 认为是干细胞分化中基因表达调控的关键。事实上, 除了组蛋白之外, 细胞中绝大多数蛋白质在表达以后都需要经过不同程度的修饰才能发挥相应的功能。这种翻译后修饰过程受到严格的调控, 使得在某一瞬间细胞中蛋白质表现出某种稳定或动态的特定功能。因此我们认为, 这种经过了特定修饰的与

表2 蛋白质翻译后修饰相关数据库
Table 2 Protein post-translational modification databases

数据库	网址	功能
Databases	Websites	Functions
SwissProt	http://www.ebi.ac.uk/swissprot/	Information on post-translational modification of proteins
PhosphoELM	http://phospho.elm.eu.org/	A database of S/T/Y phosphorylation sites
Phosphosite	http://www.phosphosite.org/homeAction.do	A database of phosphorylation sites
PROSITE	http://prosite.expasy.org/	Information on post-translational modification of proteins
HPRD	http://www.hprd.org/	Integrated database
RESID	http://www.biomedsearch.com/sci/RESID-database/0032931901.html	Information on post-translational modification of proteins
O-GlycBase	http://www.cbs.dtu.dk/databases/OGLYCBASE/	A database of O-glycosylation
dbPTM	http://dbptm.mbc.nctu.edu.tw/	Information on post-translational modification of proteins

时间和空间相关联的蛋白质组,已超越了组蛋白的单一甲基化修饰,可以称之为表观蛋白质组。基于这个概念之上的研究很可能形成一个新的组学概念和领域——表观蛋白质组学(epi-proteomics),它可以更客观地反映细胞的各种生理以及病理过程。最新的研究表明,真核细胞中存在的蛋白质修饰过程的大约70%目前还无法解释。可以预见,对经过翻译修饰后的而非“裸”蛋白质及蛋白质组的调控过程进行深入研究将更加关键,能够更深层地揭示蛋白质的多样性和活性状态以及发挥作用的方式和机理,并最终为重大疾病的诊断和治疗奠定基础。

参考文献 (References)

- Rivera CM, Ren B. Mapping human epigenomes. *Cell* 2013; 155(1): 39-55.
- Wilkins MR, Pasquali C, Appel RD, Ou K, Golaz O, Sanchez JC, *et al.* From proteins to proteomes: Large scale protein identification by two-dimensional electrophoresis and amino acid analysis. *Biotechnology (N Y)* 1996; 14(1): 61-5.
- Zhao Y, Jensen ON. Modification-specific proteomics: Strategies for characterization of post-translational modifications using enrichment techniques. *Proteomics* 2009; 9(20): 4632-41.
- Doyle HA, Mamula MJ. Post-translational protein modifications in antigen recognition and autoimmunity. *Trends Immunol* 2001; 22(8): 443-9.
- Rao RS, Moller IM. Large-scale analysis of phosphorylation site occupancy in eukaryotic proteins. *Biochim Biophys Acta* 2012; 1824(3): 405-12.
- Ciesla J, Fraczyk T, Rode W. Phosphorylation of basic amino acid residues in proteins: Important but easily missed. *Acta Biochim Pol* 2011; 58(2): 137-48.
- Tian M, Chen X, Xiong Q, Xiong J, Xiao C, Ge F, *et al.* Phosphoproteomic analysis of protein phosphorylation networks in *Tetrahymena thermophila*, a model single-celled organism. *Mol Cell Proteomics* 2014; 13(2): 503-19.
- Cole PA, Shen K, Qiao Y, Wang D. Protein tyrosine kinases Src and Csk: A tail's tale. *Curr Opin Chem Biol* 2003; 7(5): 580-5.
- Blom N, Sicheritz-Ponten T, Gupta R, Gammeltoft S, Brunak S. Prediction of post-translational glycosylation and phosphorylation of proteins from the amino acid sequence. *Proteomics* 2004; 4(6): 1633-49.
- Hurtado-Guerrero R, Davies GJ. Recent structural and mechanistic insights into post-translational enzymatic glycosylation. *Curr Opin Chem Biol* 2012; 16(5/6): 479-87.
- Li X, Wang J, Li W, Xu Y, Shao D, Xie Y, *et al.* Characterization of ppGalNAc-T18, a member of the vertebrate-specific Y subfamily of UDP-N-acetyl-alpha-D-galactosamine:polypeptide N-acetylgalactosaminyltransferases. *Glycobiology* 2012; 22(5): 602-15.
- Balzarini J. Targeting the glycans of glycoproteins: A novel paradigm for antiviral therapy. *Nat Rev Microbiol* 2007; 5(8): 583-597.
- Wang Z, Park K, Comer F, Hsieh-Wilson LC, Saudek CD, Hart GW. Site-specific GlcNAcylation of human erythrocyte proteins: potential biomarker(s) for diabetes. *Diabetes* 2009; 58(2): 309-17.
- Markkova E, Albahri Z, Valis M. Microheterogeneity of some serum glycoproteins in neurodegenerative diseases. *J Neurol Sci* 2012; 314(1/2): 20-5.
- Hart GW, Slawson C, Ramirez-Correa G, Lagerlof O. Cross talk between O-GlcNAcylation and phosphorylation: roles in signaling, transcription, and chronic disease. *Annu Rev Biochem* 2011; 80: 825-58.
- Ravid T, Hochstrasser M. Diversity of degradation signals in the ubiquitin-proteasome system. *Nat Rev Mol Cell Biol* 2008; 9(9): 679-90.
- Ciechanover A. The ubiquitin proteolytic system: from a vague idea, through basic mechanisms, and onto human diseases and drug targeting. *Neurology* 2006; 66(2 Suppl 1): S7-19.
- Ravid T, Hochstrasser M. Diversity of degradation signals in the ubiquitin-proteasome system. *Nat Rev Mol Cell Biol* 2008; 9(9): 679-90.
- Brodsky JL. Just a trim, please: Refining ER degradation through deubiquitination. *Cell* 2013; 154(3): 479-81.
- Bhoj VG, Chen ZJ. Ubiquitylation in innate and adaptive immunity. *Nature* 2009; 458(7237): 430-7.
- Hoeller D, Dikic I. Targeting the ubiquitin system in cancer therapy. *Nature* 2009; 458(7237): 438-44.
- Wimmer P, Schreiner S, Dobner T. Human pathogens and the host cell SUMOylation system. *J Virol* 2012; 86(2): 642-54.
- Dohmen RJ. SUMO protein modification. *Biochim Biophys Acta* 2004; 1695(1/2/3): 113-31.
- Psakhye I, Jentsch S. Protein group modification and synergy in the SUMO pathway as exemplified in DNA repair. *Cell* 2012; 151(4): 807-20.
- Zhao J. Sumoylation regulates diverse biological processes. *Cell Mol Life Sci* 2007; 64(23): 3017-33.
- Yi C, Ma M, Ran L, Zheng J, Tong J, Zhu J, *et al.* Function and molecular mechanism of acetylation in autophagy regulation. *Science* 2012; 336(6080): 474-7.
- Mukherjee S, Hao YH, Orth K. A newly discovered post-translational modification—the acetylation of serine and threonine residues. *Trends Biochem Sci* 2007; 32(5): 210-6.
- Allfrey VG, Faulkner R, Mirsky AE. Acetylation and methylation of histones and their possible role in the regulation of RNA synthesis. *Proc Natl Acad Sci USA* 1964; 51: 786-94.
- Greer EL, Shi Y. Histone methylation: A dynamic mark in health, disease and inheritance. *Nat Rev Genet* 2012; 13(5): 343-57.
- Xu Y, Ding J, Huang Q, Deng NY. Prediction of protein methylation sites using conditional random field. *Protein Pept Lett* 2013; 20(1): 71-7.
- He Y, Korboukh I, Jin J, Huang J. Targeting protein lysine methylation and demethylation in cancers. *Acta Biochim Biophys Sin (Shanghai)* 2012; 44(1): 70-9.
- Bi X. Alzheimer disease: Update on basic mechanisms. *J Am Osteopath Assoc* 2010; 110(9 Suppl 8): S3-9.
- Li Y, Wang H, Wang S, Quon D, Liu YW, Cordell B. Positive and negative regulation of APP amyloidogenesis by sumoylation. *Proc Natl Acad Sci USA* 2003; 100(1): 259-64.
- Neve RL. A new wrestler in the battle between alpha- and beta-secretases for cleavage of APP. *Trends Neurosci* 2003; 26(9): 479-81.

- 461-3.
- 35 Alonso A, Zaidi T, Novak M, Grundke-Iqbal I, Iqbal K. Hyperphosphorylation induces self-assembly of tau into tangles of paired helical filaments/straight filaments. *Proc Natl Acad Sci USA* 2001; 98(12): 6923-8.
- 36 Dorval V, Fraser PE. Small ubiquitin-like modifier (SUMO) modification of natively unfolded proteins tau and alpha-synuclein. *J Biol Chem* 2006; 281(15): 9919-24.
- 37 Mayeux R, Lee JH, Romas SN, Mayo D, Santana V, Williamson J, *et al.* Chromosome-12 mapping of late-onset Alzheimer disease among Caribbean Hispanics. *Am J Hum Genet* 2002; 70(1): 237-43.
- 38 Cho DH, Nakamura T, Fang J, Cieplak P, Godzik A, Gu Z, *et al.* S-nitrosylation of Drp1 mediates beta-amyloid-related mitochondrial fission and neuronal injury. *Science* 2009; 324(5923): 102-5.
- 39 Padayachee E, Ngqwalala N, Whiteley CG. Association of beta-amyloid peptide fragments with neuronal nitric oxide synthase: Implications in the etiology of Alzheimer's disease. *J Enzyme Inhib Med Chem* 2012; 27(3): 356-64.
- 40 Abrams AJ, Farooq A, Wang G. S-nitrosylation of ApoE in Alzheimer's disease. *Biochemistry* 2011; 50(17): 3405-7.
- 41 Qu J, Nakamura T, Cao G, Holland EA, Mckercher SR, Lipton SA. S-Nitrosylation activates Cdk5 and contributes to synaptic spine loss induced by beta-amyloid peptide. *Proc Natl Acad Sci USA* 2011; 108(34): 14330-5.
- 42 Kwak YD, Ma T, Diao S, Zhang X, Chen Y, Hsu J, *et al.* NO signaling and S-nitrosylation regulate PTEN inhibition in neurodegeneration. *Mol Neurodegener* 2010; 5: 49.
- 43 Martinez-Ruiz A, Villanueva L, Gonzalez DOC, Lopez-Ferrer D, Higuera MA, Tarin C, *et al.* S-nitrosylation of Hsp90 promotes the inhibition of its ATPase and endothelial nitric oxide synthase regulatory activities. *Proc Natl Acad Sci USA* 2005; 102(24): 8525-30.
- 44 Vissers JH, Nicassio F, van Lohuizen M, Di Fiore PP, Citterio E. The many faces of ubiquitinated histone H2A: Insights from the DUBs. *Cell Div* 2008; 3: 8.
- 45 Shimada M, Nakadai T, Fukuda A, Hisatake K. cAMP-response element-binding protein (CREB) controls MSK1-mediated phosphorylation of histone H3 at the c-fos promoter *in vitro*. *J Biol Chem* 2010; 285(13): 9390-401.
- 46 Logie C, Tse C, Hansen JC, Peterson CL. The core histone N-terminal domains are required for multiple rounds of catalytic chromatin remodeling by the SWI/SNF and RSC complexes. *Biochemistry* 1999; 38(8): 2514-22.
- 47 Ong SE, Pandey A. An evaluation of the use of two-dimensional gel electrophoresis in proteomics. *Biomol Eng* 2001; 18(5): 195-205.
- 48 Lilley KS, Dupree P. Methods of quantitative proteomics and their application to plant organelle characterization. *J Exp Bot* 2006; 57(7): 1493-9.
- 49 Romijn EP, Krijgsveld J, Heck AJ. Recent liquid chromatographic-(tandem) mass spectrometric applications in proteomics. *J Chromatogr A* 2003; 1000(1/2): 589-608.
- 50 Ochi H, Horiuchi I, Araki N, Toda T, Araki T, Sato K, *et al.* Proteomic analysis of human brain identifies alpha-enolase as a novel autoantigen in Hashimoto's encephalopathy. *FEBS Lett* 2002; 528(1/2/3): 197-202.
- 51 Gehanne S, Cecconi D, Carboni L, Righetti PG, Domenici E, Hamdan M. Quantitative analysis of two-dimensional gel-separated proteins using isotopically marked alkylating agents and matrix-assisted laser desorption/ionization mass spectrometry. *Rapid Commun Mass Spectrom* 2002; 16(17): 1692-8.
- 52 Chivasa S, Ndimba BK, Simon WJ, Robertson D, Yu XL, Knox JP, *et al.* Proteomic analysis of the Arabidopsis thaliana cell wall. *Electrophoresis* 2002; 23(11): 1754-65.
- 53 Frese CK, Altelaar AF, van den Toorn H, Nolting D, Griep-Raming J, Heck AJ, *et al.* Toward full peptide sequence coverage by dual fragmentation combining electron-transfer and higher-energy collision dissociation tandem mass spectrometry. *Anal Chem* 2012; 84(22): 9668-73.
- 54 Frese CK, Zhou H, Taus T, Altelaar AF, Mechtler K, Heck AJ, *et al.* Unambiguous phosphosite localization using electron-transfer/higher-energy collision dissociation (ETHCD). *J Proteome Res* 2013; 12(3): 1520-5.
- 55 Jaffe JD, Berg HC, Church GM. Proteogenomic mapping as a complementary method to perform genome annotation. *Proteomics* 2004; 4(1): 59-77.
- 56 Gupta N, Tanner S, Jaitly N, Adkins JN, Lipton M, Edwards R, *et al.* Whole proteome analysis of post-translational modifications: applications of mass-spectrometry for proteogenomic annotation. *Genome Res* 2007; 17(9): 1362-77.
- 57 Fenselau C. A review of quantitative methods for proteomic studies. *J Chromatogr B Analyt Technol Biomed Life Sci* 2007; 855(1): 14-20.
- 58 Szunerits S, Maalouli N, Wijaya E, Vilcot JP, Boukherroub R. Recent advances in the development of graphene-based surface plasmon resonance (SPR) interfaces. *Anal Bioanal Chem* 2013; 405(5): 1435-43.
- 59 Khoury GA, Baliban RC, Floudas CA. Proteome-wide post-translational modification statistics: Frequency analysis and curation of the swiss-prot database. *Sci Rep* 2011; 1; doi: 10.1038/srep00090.
- 60 Diella F, Cameron S, Gemund C, Linding R, Via A, Kuster B, *et al.* Phospho.ELM: A database of experimentally verified phosphorylation sites in eukaryotic proteins. *BMC Bioinformatics* 2004; 5: 79.
- 61 Hornbeck PV, Chabra I, Kornhauser JM, Skrzypek E, Zhang B. PhosphoSite: A bioinformatics resource dedicated to physiological protein phosphorylation. *Proteomics* 2004; 4(6): 1551-61.
- 62 Chen C, Huang H, Wu CH. Protein bioinformatics databases and resources. *Methods Mol Biol* 2011; 694: 3-24.
- 63 Dai B, Rasmussen TP. Global epiproteomic signatures distinguish embryonic stem cells from differentiated cells. *Stem Cells* 2007; 25(10): 2567-74.