刍议质谱技术发现新型蛋白质翻译后修饰研究在化学与生物学教学中的应用

王志鹏^{1,2,3,*}, 张璇⁴, 车子良⁵, 蒋振雄⁶, 马新雨²

¹ Division of Genetics, Department of Medicine, Brigham and Women's Hospital, Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, MA 02115, USA.

² Department of Chemistry, Texas A&M University, College Station, TX 77840, USA.

³ 清华大学化学系, 北京 100084

⁴ 清华大学生命学院, 北京 100084

⁵ Department of Mathematics, Harvard University, MA 02138, USA.

⁶ Department of Biology, Texas A&M University, College Station, TX 77840, USA.

摘要: 随着质谱技术的不断发展, 质谱技术已经被广泛应用于化学、材料、生命科学等众多领域, 特别是其在生物化学与化学生物学领域的应用极大扩展了研究人员对于蛋白质翻译后修饰的认知。本文从基本的质谱原理入手, 着重探讨了质谱技术在以组蛋白为模式蛋白的赖氨酸翻译后修饰的发现过程, 并展示如何将基础有机化学和生物化学知识融入前沿科学研究。本文通过梳理不同的研究逻辑下质谱技术在蛋白翻译后修饰的研究中发挥的作用, 旨在引导学生建立起将基础教学内容与实际应用相联系的意识及思维模式。

关键词: 质谱技术; 串联质谱; 翻译后修饰; 生物化学

中图分类号: G64; O6

Discussion on the Application of Novel Protein Post-Translational Modification Discovered by Mass Spectrometry in Chemistry and Biology Education

Zhipeng A. Wang^{1,2,3,*}, Xuan Zhang⁴, Ziliang Che⁵, Zhenxiong Jiang⁶, Xinyu Ma²

¹ Division of Genetics, Department of Medicine, Brigham and Women's Hospital, Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, MA 02115, USA.

² Department of Chemistry, Texas A&M University, College Station, TX 77840, USA.

³ Department of Chemistry, Tsinghua University, Beijing 100084, China.

⁴ School of Life Sciences, Tsinghua University, Beijing 100084, China.

⁵ Department of Mathematics, Harvard University, MA 02138, USA.

⁶ Department of Biology, Texas A&M University, College Station, TX 77840, USA.

Abstract: Rapid development of mass spectrum technology (MS) brings about its wide applications in many fields such as chemistry, material science and life sciences, and its application in biochemistry and chemical biology have been greatly expanding the knowledge of novel protein post-translation modification (PTM). This manuscript starts

from the basic principles of mass spectrum and later elaborates the application of MS in discovering lysine post-translation modifications over histones as model proteins, showing a path incorporating fundamental organic chemistry and biochemistry knowledge into cutting-edge researches. By analyzing roles of mass spectrum throughout post-translation modification studies under different research logistics, this manuscript was aimed to help establishing a way of thinking that associates fundamental chemistry and biology education with practical applications for students.

Key Words: Mass spectrum technology; MS/MS; PTM; Biochemistry

质谱技术在最近几十年间得以突飞猛进的发展与应用,尤其是串联质谱的发展对于化学和生命科学等众多分支学科起到了极大的推动作用。可以说一系列组学的创立就是以质谱为基础的,例如蛋白质组学(Proteomics)、表观遗传组学(Epigenomics)和代谢组学(Metabolomics)等^[1]。其中,基于串联质谱的蛋白质翻译后修饰及表观遗传组学是后基因组时代的重要研究领域。

翻译后修饰是表观遗传学的核心组成之一,在体内的胞质蛋白、膜蛋白和核蛋白上都大量存在。尤其是以组蛋白为代表的核蛋白,其上的各类修饰可能改变 DNA 分子在组蛋白异源八聚体上缠绕的紧密程度,进而直接影响染色体结构。同时,这些翻译后修饰还可以影响 mRNA 转录酶复合体 II 及各种其他转录因子(Transcriptional factor)的结合,进而改变基因的转录活性^[2]。值得一提的是,所有的蛋白质翻译后修饰均涉及三类蛋白质或酶类,包括“写录者(Writer)”,“读取者(Reader)”和“擦除者(Eraser)”。这三者精密配合,共同调控翻译后修饰^[3]。

质谱本身作为分析化学技术对于蛋白质翻译后修饰的各个研究领域均有重要作用。而新型蛋白质翻译后修饰的发现和验证则可作为极佳的教学素材用于化学、生物学和医学等交叉学科的基础教学中,可以融入例如有机化学、生物化学、酶学、化学生物学、分子生物学、遗传学、生物物理学等一系列二级学科中。而对于新型蛋白质翻译后修饰这一过程的探索发现极大地体现了化学发现新结构这一永恒的主题,颇有侦探推理“抽丝剥茧”的意味,在将深刻的前沿生物化学知识融入基础化学的同时可以充分调动学生的兴趣^[4]。并让学生从这些研究中看出基础有机化学和分析化学知识对于前沿生命科学研究仍然具有的重要意义。对培养新时代的创造性跨学科人才有重要作用。本文将从质谱技术原理入手,以若干新型蛋白质翻译后修饰的发现研究为典例,按照科学哲学的逻辑为基础,探讨结合质谱技术在化学与生物学教学中的意义与重要性,以期对相关领域的教学和研究有助^[5]。

1 质谱技术简述

质谱技术本身的知识属于分析化学或仪器分析。但可以划归入基础有机化学、生物化学和化学生物学等二级学科中。在基础教学尤其是非分析化学课程教学中有必要首先讲解现代质谱技术的发展与应用。

1.1 质谱概要

质谱作为一种基于分子量的分析化学工具,在过去的 20 年内发生了很多重大的革新。以电喷雾质谱(ESI)^[6]和基质辅助激光离子化(MALDI)^[7]为代表的新型离子化手段的出现使得蛋白质分析成为了可能。伴随硬件的发展,质谱的检测器也实现了多样化,发展出飞行时间(TOF)、四级杆(Quadrupole)、离子阱(Ion traps)、轨道离子阱(Orbitraps)和离子回旋共振(ICR)五大类,并使用傅里叶变换(FT)进行数据处理。这些检测器不仅可以生成被测物的质荷比信号(m/z),还可以得出不同荷质比的相对强度,进而生成最终的质谱图谱。对于蛋白质这些大分子量生物分子,其可以同时带有若干电荷,故而人们发展了对应的软件进行特定峰值的划归计算,即去卷积化(Deconvolution)。

当被分析的生物大分子处于复杂的混合物状态时,我们需要进行液质联用(HPLC-MS)。经由 HPLC 先对样品进行纯化分离后,再送入质谱分析。这样的混合物状态可以直接来自细胞裂解液或

者提纯的细胞器或细胞核,还可以是人为进行蛋白质酶解后所得的多肽片段混合物。值得注意的是,在使用 C18 反相柱色谱进行分离纯化的同时,多肽也被富集到具有一定百分比有机溶剂的水相中,适合进一步分析。

1.2 串联质谱概要

串联质谱(Tandem MS 或 MS/MS)是一种具有分析蛋白质序列这一强有力手段的质谱衍生技术。在检测荷质比后,具有特定荷质比的前体离子(Precursor ion)将被电磁场分离,并与惰性气体分子进行碰撞而裂解形成子代离子(Daughter ion)。子代离子进而被下一级质谱检测。除了与稀有气体分子的碰撞而导致解离(CID)外,电子转移(ETD)、电子捕获(ECD)等也能导致前体离子的解离。

解离涉及到前体离子化学键的断裂,故而键能较弱的共价键易于断裂。由于解离过程包含有前体离子的结构信息,故而串联质谱可以反映前体离子的结构。这对于多肽尤其重要,虽然多肽所裂解产生的子代离子可能有多种不同的模式,但都具有逐个氨基酸的信息^[8]。故而,将同种前体离子检测到的所有子代离子的荷质比合并作图可得串联质谱图谱。将串联质谱图谱与同种前体离子可以产生所有的理论子代离子相互对比即可确定多肽的序列。这一过程通过特定的计算机软件算法得出,之后则需人工进行峰归属划分^[9]。随着检测需求的增加,相关的配套软件也进行了不断的升级和改进。为了进一步提升软件对于肽和翻译后修饰(PTM)识别的准确性和灵敏度,赵英明课题组^[10]近来开发了 PTMap 软件。该软件包含了峰值选择、不精确误差调整和 PTM 的精确定位等功能。而对于不匹配的峰,该软件还能自动执行基于手动验证规则的肽段鉴定,提高了后期修正的效率。

1.3 质谱对翻译后修饰研究的应用

质谱对于翻译后修饰领域的发展具有极为关键的作用,可以说翻译后修饰的发现、功能研究都直接或间接与质谱相关。可以分为主要几个方面。首先,质谱技术可以用于发现和确定新的翻译后修饰模式,这种定性研究是后续研究的基础。其次,借助于质谱技术,我们可以对某种翻译后修饰进行定量研究,以确定某种特定的翻译后修饰在细胞生理或病理过程中的变化情况。这种定量研究同样是后续研究的基础。

之后三个应用都是基于前两个完成的。第三,对于某种翻译后修饰定量研究的结论,我们可以追本溯源寻找其对应的“写录者”和“擦除者”,并进行进一步的生物化学和生物物理学研究。第四,同样是基于某种翻译后修饰定量研究的数据,我们可以搜寻其上游的“写录者”和下游的“读取者”以明确其在细胞信号传递通路中的作用。最后,基于单一翻译后修饰的定性、定量数据,我们可以进一步研究蛋白质翻译后修饰之间的交叉感应(Crosstalk)的情况。本文将主要针对质谱与串联质谱技术对于新型翻译后修饰的发现进行分析。从方法论出发,并列举若干近年来的主要突破性进展。

2 质谱技术发现新型翻译后修饰之总论

质谱的进步对新型翻译后修饰的发现和确认可以算一次技术上的重大革新。这里可以简单分为三大步骤。值得一提的是,人们通常选用组蛋白作为发现新型翻译后修饰并研究翻译后修饰的理想目标。一方面是组蛋白上存在丰富的翻译后修饰,而这些翻译后修饰也有重要的生物调控功能。另一方面,组蛋白相比其他胞质蛋白如代谢酶类具有更加丰富的量并易于进行生物化学和化学生物学处理。

2.1 分子量确认

传统的以组蛋白为典型代表的翻译后修饰的搜寻可以归纳为偶然的意外发现。而现代的精确质谱法对于新型翻译后修饰的发现则形成了系统化的流程。高灵敏度的 HPLC-MS/MS 是检测出低丰度新型翻译后修饰的先决条件。而这一分析过程一般是在将目标蛋白质通过胰蛋白酶酶切形成的多肽片段上进行的。在未知理论分子量变化(ΔM_r)的情况下,一级质谱算法可以实现无差别的全序列筛查。由于从分子量无法直接推断序列及结构的信息,故而来自一级质谱的数据并不能直接确定修饰的更多结构信息。即便如此,人们也可以利用高精度的 ΔM_r 数值与原子量的精确数值进行对比而计

算出新型翻译后修饰的分子式。进而通过串联质谱确定该翻译后修饰所在的氨基酸位置。这可以确定一些可以修饰多种氨基酸残基的翻译后修饰的具体位置，例如赖氨酸和精氨酸都可以烷基化，而丝氨酸、苏氨酸、酪氨酸和组氨酸都可以磷酸化。

2.2 结构确定

在确定分子式后，我们需要对新型翻译后修饰进行可能的异构体排查以确定其结构式。理想情况下，我们直接将对应的翻译后修饰蛋白或多肽进行核磁共振(NMR)分析即可，但是一方面我们很难提取到足够的匀质的翻译后修饰蛋白，另一方面 NMR 分析可能过于复杂。故而，我们需要用穷举法列出所有的可能翻译后修饰结构。一般说来翻译后修饰的分子式不会过于复杂，通常在 10 个碳原子以内，使得可能的结构式也不会过于繁杂。我们进一步合成对应含有新型翻译后修饰的多肽进行质谱和串联质谱研究，并与天然多肽的图谱进行对比。利用一些细小的“指纹图谱”一般能够确定具体的结构。

2.3 最终确认

而对于确定结构的新型翻译后修饰，我们还需要进行验证。在传统方法中，我们需要用同位素标记或化学方法进行验证。例如，组蛋白乙酰化的确定就来源于将含有 ^{14}C 的醋酸钠加入细胞培养基并立即检查出组蛋白同位素标记^[11]，并通过氨基酸分析得以最终确认^[12]。这种方法不仅费时费力，不够精确，且难以做到具体定位与定量^[13]。相比之下，质谱法可以直接获取修饰前后的精确荷质比，而特异性抗体可以直接对翻译后修饰进行免疫沉淀或蛋白质印记分析进行确认。最后再用其他化学和生物化学手段对新型翻译后修饰进行分析。

3 质谱技术发现新型翻译后修饰之分论

在最近的数十年内，人们发现并解析了大量的新型翻译后修饰。这些翻译后修饰不仅存在于组蛋白上，还存在于各类胞质蛋白和膜蛋白上。下面分类列举若干典型案例着重对发现过程进行讨论。这部分主要引用芝加哥大学赵英明课题组的近期研究成果，该课题组主要致力于利用串联质谱技术对蛋白质的新型翻译后修饰进行定性定量研究。

3.1 执果索因：赖氨酸三种二酰化

赖氨酸酰基化是一大类翻译后修饰。最初人们对其的认识仅仅是乙酰、丙酰、丁酰等简单的修饰模式^[14]。直到二酰化的发现极大地突破了人类的认知。2012 年，赵英明课题组^[15]在人、小鼠、果蝇和酵母细胞的组蛋白上发现赖氨酸的丁二酰化(Succinylation)与丙二酰化(Malonylation)。与典型的新型翻译后修饰发现的模式不同，由于丁二酰基在细胞内作为代谢中间体广泛存在，尤其是在线粒体三羧酸循环中。该课题组之前已经在非组蛋白上鉴定到赖氨酸丁二酰化，进而预测丁二酰化赖氨酸亦存在于组蛋白的前提下对其进行蛋白质免疫印迹的检定。之后用 HPLC-MS/MS 的手段检测并分析出组蛋白上丁二酰化的修饰位点，并最终合成修饰后的多肽进行验证。基于以上两种二酰化翻译后修饰，赵英明课题组^[16]又于 2014 年发现了蛋白质的赖氨酸戊二酰化(Glutarylation)。考虑到结构的相似性，戊二酰基可以视为丙二酰基、丁二酰基的第三种同系物。该课题组同样从免疫印迹法入手进而用串联质谱进行分析，发现并验证了该翻译后修饰在人体蛋白质组中广泛存在。而重要的代谢酶氨基甲酰磷酸合成酶 1 (CPS1)的戊二酰化修饰可以被组蛋白去乙酰化酶 Sirtuin5 脱除。这三种二酰化赖氨酸的发现极大地拓宽了人们对于翻译后修饰的认识，不仅消除了赖氨酸在天然状态下原本的正电荷，还附加了一个额外的负电荷，故而极大地影响了对应的底物蛋白质的结构功能与调控^[17]。

3.2 上下求索：赖氨酸苯甲酰化

前三种二酰化赖氨酸的发现可以算是“有的放矢”，即在已经根据细胞代谢过程或其他信息猜想特定结构存在的前提下进行的检测与验证。其本身的过程较为直截了当，并不太符合前文所述的“发现-猜想-验证”的模式。

新型翻译后修饰层出不穷。直到 2018 年,赵英明课题组^[18]还进一步发现了组蛋白上赖氨酸的苯甲酰化(Benzoylation)。该课题组在将胰蛋白酶处理后的 HepG2 细胞组蛋白提取物进行串联质谱分析过程中发现其 H2B 的第五位赖氨酸(K5)上带有+104.0268 Da 的增加。经过精细分子量比对,得出可能的修饰分子式为 C_7H_4O 。考虑到其带有的一个自由价(单电子),且取代赖氨酸时需要脱去一个氢原子,划归后的修饰分子式为 C_7H_6O 。此分子式的不饱和度是 5,推断只能是“苯甲醛”(图 1)。故而得出该修饰为苯甲酰基修饰。在后续验证中,合成的多肽和天然多肽,及二者的混合物均展现相似的特征串联质谱图谱。该修饰是由于服用了作为美国食品药品监督管理局(FDA)批准作为药物及食品防腐剂的苯甲酸钠而促进产生的,并可以进一步作为组蛋白翻译后修饰参与生理调控。

图 1 赖氨酸苯甲酰化的结构发现

从这里可以看出,要想准确确定修饰的分子式,其前提条件是精确测定的分子量增量。类似的还有赖氨酸的甲酰化的发现。虽然对于甲酰化和二甲基化均为+28 Da,但是二者相距 0.0364 Da。Wisniewski 课题组^[19]通过串联质谱精确测定的分子量变化进而确定之,而甲酰基作为最简单的一个碳原子的酰基化,只有唯一一种可能结构。而这一修饰的来源可能是 DNA 的氧化损伤^[20]。

3.3 左右逢源: 赖氨酸羟基异丁酰化(Hydroxyisobutyrylation)

如果说苯甲酰化的发现是一帆风顺,而另外一种修饰的发现过程则一波三折。由于细胞内物质与能量代谢中可能产生各种各样的活性酰基,这些酰基可以以酶促反应或非酶促反应的方式对蛋白质赖氨酸和氮端氨基进行修饰。故而存在各种可能的新型修饰结构。

早在 2010 年,赵英明课题组^[21]就开始对新型蛋白质翻译后修饰进行发现和研究。他们首先将胰蛋白酶处理后的小鼠睾丸细胞组蛋白提取物用 HPLC-MS/MS 分离鉴定,发现组蛋白 H4 第 77 位赖氨酸(K77)上带有未知的+86.0354 Da 的增加。根据精细分子量比对,得出可能的修饰分子式为 $C_4H_7O_2$,再确定划归后的修饰分子式 $C_4H_8O_2$ 。与前例中“苯甲醛”不同,单纯从这一分子式的化学角度看似乎具有太多结构可能性,但接下来需要整合各种信息进行系统性推理并加以验证。

首先, $C_4H_8O_2$ 的不饱和度是 1,结合赖氨酸修饰模式,最大可能是被酰基占用。再结合人体细胞代谢产物中一般不具有醚的结构,故而剩下的结构为羟基丙基结构。考虑其碳链的构造异构体,具有五种可能。首先是直链丙基上的伯碳取代形成羟基正丙基对应 4-羟基丁酰(图 2A)、2 位仲碳取代形成的 2-羟基丙基对应 3-羟基丁酰(图 2B)、3 位仲碳取代形成的 1-羟基丙基对应的 2-羟基丁酰(图 2C);其次是支链异丙基的伯碳取代形成的 2-羟基异丙基对应的 3-羟基异丁酰(图 2D)和仲碳取代形成的 1-羟基异丙基对应的 2-羟基异丁酰(图 2E)。

图 2 赖氨酸羟基丁酰化可能结构(1)

为了进一步确认新型翻译后修饰的结构, Zhao 等^[22]利用多肽固相合成产生含有以上五种修饰的多肽, 进行 HPLC-MS/MS 分析并与天然对应多肽进行对比。最终发现含有 2-羟基异丁酰(图 2E)的多肽与天然修饰多肽具有一致的色谱保留时间及相同的串联质谱指纹图谱。故而得出结论新型翻译后修饰为赖氨酸的 2-羟基异丁酰化(K_{hib})。该修饰被证明广泛存在于细胞质组中。

3.4 前呼后应: 赖氨酸 β -羟基丁酰化(β -hydroxybutyrylation)

如果说羟基异丁酰化的发现结合了基础有机化学同分异构体的知识, 那么基础生物化学知识在另一种修饰的发现过程中同样起到了重要作用。这一典例进一步说明新型蛋白质翻译后修饰的发现集合了多种交叉学科的知识, 是极佳的教学素材。

无独有偶, 2016 年赵英明课题组^[23]用 HPLC-MS/MS 对胰蛋白酶处理后的 HEK293 细胞组蛋白提取物进行分析, 在组蛋白 H3 的第 18 位赖氨酸(K18)上发现同样的未知+86.0376 Da 的增加。这种修饰虽然与 2-羟基异丁酰化具有相同的分子式, 但结构并不相同。由于生物体系都具有特定的手性选择, 考虑到手性中心的不同, 需要将该列表拓展至 8 种, 即(*R/S*)-3-羟基丁酰(图 3B1/2)、(*R/S*)-2-羟基丁酰(图 3C1/2)、(*R/S*)-3-羟基异丁酰(图 3D1/2)(图 3)。

图 3 赖氨酸羟基丁酰化可能结构(2)

为了验证结构, 该课题组合成了分别带有以上 8 种可能的修饰模式的多肽, 并进行 HPLC-MS/MS 的鉴定。对于非对映异构体, 其具有不同的化学结构, 即便在非手性的 HPLC 柱色谱上也可根据不同的保留时间进行区分。可以将可能性缩小至 3-羟基丁基(β -羟基丁基)即的两种异构体(图 3B1/2)。

这里又需要借助于生物化学的分析。考虑到代谢中 β -羟基丁酸(图 4I)是由乙酰乙酸(图 4G)还原而来, 而后者又由乙酰辅酶 A (Ac-SCoA, 图 4F)缩合形成。 β -羟基丁酸、乙酰乙酸和其非酶促脱羧产物丙酮(图 4H)这三者共同构成“酮体”(Ketone body)。其中, β -羟基丁酸是乙酰乙酸在 3-羟基丁酸脱氢酶催化下还原形成, 主要产物为 *R* 构象而非 *S* 构象。其之的可能代谢路线是形成-羟基丁酰辅酶 A (图 4J), 并最终酰化赖氨酸。故而可以推测修饰结构为 *R*- β -羟基丁酰化。合成的含有 *R*- β -羟基丁酰化赖氨酸的多肽片段和天然片段具有一致的保留时间和相同的串联质谱指纹图谱, 故而结构(K_{Bhb})得以最终确认。

图 4 酮体与 β -羟基丁酸的代谢

此外, 2011 年 Zhao 等^[24]由组蛋白 H2B 第 5 位赖氨酸上发现+68.0230 Da, 结合分子式的可能性和细胞代谢中间体分析确定为(Crotonylation); 2019 年由赵英明课题组^[25]利用类似方式发现的赖氨酸乳酰化(Lactylation), 也是由+72.021 Da 的分子量增加入手, 结合分子式的有机化学分析与代谢过程的生物化学分析进行确定验证。

3.5 为果设因: 赖氨酸 4-氧代-2-壬酰化修饰

在前文所列举的若干案例中, 都是直接用液相色谱-串联质谱联用直接从正常的培养细胞中提取蛋白质进行分析。而另外一些特别的情况则是让细胞在人为控制的情况下生长进而进行类似分析。一个典型案例是 4-氧代-2-壬酰化(4-ONEylation)的发现。由细胞代谢过程可知, 长链脂质在氧化代谢过程中会产生多种 α,β -不饱和醛类, 例如 4-羟基-2-壬醛(4-HNE)和 4-氧代-2-壬醛(4-ONE)等(图 5)。从有机化学角度看这些 α,β -不饱和醛类可以作为极好的迈克尔加成受体(Michael acceptor)并被细胞内的亲核基团进攻, 例如蛋白质上广泛存在的赖氨酸、半胱氨酸和组氨酸等。其中, 由于 4-氧代-2-壬醛存在第 2 个羰基, 故而理论上可以形成稳定的 1,2-加成产物。

2014 年, Marnett 课题组^[26]分析了在含有炔基官能化的 4-HNE 和 4-ONE 的培养基中生长的细胞, 发现这两类代谢产物会共价修饰组蛋白上的赖氨酸和组氨酸, 代表性产物为 4-氧代-2-壬酰化赖氨酸(图 5N)。而组蛋白 H3 第 27 位赖氨酸(K27)的 4-氧代-2-壬酰化修饰甚至会抑制核小体的折叠。这种新型赖氨酸翻译后修饰同样被质谱和串联质谱中的分子量之改变所证实。而赖氨酸的 4-羟基壬醛化修饰(4-HNEylation)也被证明以多种不同结构(图 5K/L)存在, 且与一些重要代谢疾病相关^[27]。

图 5 可能的赖氨酸 4-ONEylation 与 4-HNEylation 修饰结构

4 讨论

对于翻译后修饰, 其生物化学性质与生物学功能是最为重要的。各类翻译后修饰对于调控底物蛋白质结构功能至关重要, 进而可以影响细胞的整体结构、代谢、功能与活动^[28]。而质谱技术对新型翻译后修饰研究的贡献还远不止发现。例如, 将新型翻译后修饰在全蛋白质组进行定位与定量研究、系统地预测可以对蛋白功能产生干扰的修饰位点^[29]等。

4.1 教学素材选取

本文着重质谱技术对于辅助新型蛋白质翻译后修饰的发现过程的讨论, 并未涉及质谱技术对其进一步的定性定量研究的相关内容, 难免有失偏颇。但这部分从教学的趣味性上看是难得的素材。在前沿科学研究中融入了基本的有机化学与生物化学知识, 极为生动地给学生们展示了基础教学内容可以被实实在在用于科研发现中去。

同时,对应的案例教学可以融入五种经典科学研究中可以遵循的研究逻辑:“执果索因”即先有结构后寻觅;“上下求索”即发现导致唯一下游结果;“左右逢源”即发现导致不唯一结果,需要利用穷举法进行排除试错;“前呼后应”即与之前的结果相关,但需要借助其他各个学科知识进行分析;“为果设因”即为了探寻假说而认为创造可能情况。

4.2 化学生物学应用

串联质谱对于蛋白质翻译后修饰的研究的应用与贡献还远不止于此。一方面,近年来发现的新型翻译后修饰远不止于此。单纯以赖氨酸为例,在酶促或非酶促条件下可能的修饰有几十种,例如通过蛋白质组分析发现,糖酵解中间体 1,3-双磷酸甘油酸酯与蛋白质中的特定赖氨酸残基的反应不需要酶催化即可形成 3-磷酸甘油基赖氨酸(3-phosphoglyceroylation)的过程^[30]。这里只是与质谱相关的少数典例。

另一方面,由于在研究特定蛋白质翻译后修饰蛋白质时,我们需要首先获取足够量的目标蛋白。而生物学家却难以直接使用生物工程手段进行表达^[31]。故而对于化学生物学家,制备含有特定修饰的蛋白质是一项挑战^[32]。目前已经有非天然氨基酸引入法^[33]、蛋白质全合成与半合成法^[34]、生物正交反应法^[35]等一系列合成手段。在合成后,仍然需要用串联质谱进行分析与验证。

5 结语

随着蛋白质翻译后修饰的逐步发现和深入研究,质谱技术也显示出其在生命科学领域中广泛存在的应用价值。尤其是伴随着一系列串联质谱技术的发展,质谱能提供一系列蛋白质等生物大分子系列、结构和修饰等信息。其中极高的分子量精度更是极大促进了新型翻译后修饰的发现。

本文首先对质谱特别是串联质谱进行了概述。其次,对质谱技术在发现新型翻译后修饰方面发挥的作用,包括分子量的确认、结构确定及最终的验证等方面进行了论述。随后,我们以五个典型案例为例,包括赖氨酸三种二酰化、赖氨酸苯甲酰化、赖氨酸羟基异丁酰化、赖氨酸 β -羟基丁酰化和赖氨酸 4-氧代-2-壬酰化修饰,展示了不同逻辑指导下质谱技术如何在新型蛋白质翻译后修饰发现过程中发挥作用。通过上述对于质谱技术应用案例的研究,笔者为学生生动展示了基础教学内容与科学研究之间的联系。在培养学生学习基础化学和生命科学兴趣^[36]的同时,为教学研究中解析研究逻辑方面的内容提供了良好的范例,并应用于蛋白质翻译后修饰的化学生物学研究中去^[37]。这对于培养新型创新性交叉学科人才极为重要^[38]。

致谢: 感谢Texas A&M University博士生Erol Can Vatansever参与讨论和提出的建议。

参 考 文 献

- [1] Arnaudo, A. M.; Molden, R. C.; Garcia, B. A. *Crit. Rev. Biochem. Mol. Biol.* **2011**, *46* (4), 284.
- [2] Strahl, B. D.; Allis, C. D. *Nature* **2000**, *403* (6765), 41.
- [3] Filippakopoulos, P.; Knapp, S. *Nat. Rev. Drug Discov.* **2014**, *13* (5), 337.
- [4] 田杰, 王志鹏, 马新雨, 蒋振雄, 许兵. *大学化学*, **2019**, *34* (2), 81.
- [5] 蒋振雄, 王鹏, 马新雨, 田杰, 张军, 姜汉杰, 石瑾慧, 王志鹏. *大学化学*, **2019**, *34* (5), 80.
- [6] Fenn, J. B.; Mann, M.; Meng, C. K.; Wong, S. F. *Science* **1989**, *246* (4926), 64.
- [7] Tanaka, K.; Waki, H.; Ido, Y.; Akita, S.; Yoshida, Y.; Yoshida, T.; Matsuo, T. *Rapid Commun. Mass Spectrom* **1988**, *2*, 151.
- [8] Steen, H.; Mann, M. *Nat. Rev. Mol. Cell Biol.* **2004**, *5* (9), 699.
- [9] 王志鹏, 王鑫, 刘伟. *化学教育*, **2017**, *38* (12), 1.
- [10] Chen, Y.; Chen, W.; Cobb, M. H.; Zhao, Y. *Proc. Natl. Acad. Sci. USA* **2009**, *106* (3), 761.
- [11] Allfrey, V. G.; Faulkner, R.; Mirsky, A. E. *Proc. Natl. Acad. Sci. USA* **1964**, *51* (5), 786.

- [12] Gershey, E. L.; Vidali, G.; Allfrey, V. G. *J. Biol. Chem.* **1968**, 243 (19), 5018.
- [13] Huang, H.; Lin, S.; Garcia, B. A.; Zhao, Y. *Chem. Rev.* **2015**, 115 (6), 2376.
- [14] Chen, Y.; Sprung, R.; Tang, Y.; Ball, H.; Sangras, B.; Kim, S. C.; Falck, J. R.; Peng, J.; Gu, W.; Zhao, Y. *Mol. Cell Proteomics* **2007**, 6 (5), 812.
- [15] Xie, Z.; Dai, J.; Dai, L.; Tan, M.; Cheng, Z.; Wu, Y.; Boeke, J. D.; Zhao, Y. *Mol. Cell Proteomics* **2012**, 11 (5), 100.
- [16] Tan, M.; Peng, C.; Anderson, K. A.; Chhoy, P.; Xie, Z.; Dai, L.; Park, J.; Chen, Y.; Huang, H.; Zhang, Y.; *et al.* *Cell Metab.* **2014**, 19 (4), 605.
- [17] 王志鹏, 程农壹, 梁妍钰, 马新雨. 大学化学, **2017**, 32 (12), 41.
- [18] Huang, H.; Zhang, D.; Wang, Y.; Perez-Neut, M.; Han, Z.; Zheng, Y. G.; Hao, Q.; Zhao, Y. *Nat. Commun.* **2018**, 9 (1), 3374.
- [19] Wisniewski, J. R.; Zougman, A.; Mann, M. *Nucleic Acids Res.* **2008**, 36 (2), 570.
- [20] Jiang, T.; Zhou, X.; Taghizadeh, K.; Dong, M.; Dedon, P. C. *Proc. Natl. Acad. Sci. US* **2007**, 104 (1), 60.
- [21] Sabari, B. R.; Zhang, D.; Allis, C. D.; Zhao, Y. *Nat. Rev. Mol. Cell Biol.* **2017**, 18 (2), 90.
- [22] Dai, L.; Peng, C.; Montellier, E.; Lu, Z.; Chen, Y.; Ishii, H.; Debernardi, A.; Buchou, T.; Rousseaux, S.; Jin, F.; *et al.* *Nat. Chem. Biol.* **2014**, 10 (5), 365.
- [23] Xie, Z.; Zhang, D.; Chung, D.; Tang, Z.; Huang, H.; Dai, L.; Qi, S.; Li, J.; Colak, G.; Chen, Y.; *et al.* *Mol. Cell* **2016**, 62 (2), 194.
- [24] Tan, M.; Luo, H.; Lee, S.; Jin, F.; Yang, J. S.; Montellier, E.; Buchou, T.; Cheng, Z.; Rousseaux, S.; Rajagopal, N.; *et al.* *Cell* **2011**, 146 (6), 1016.
- [25] Zhang, D.; Tang, Z.; Huang, H.; Zhou, G.; Cui, C.; Weng, Y.; Liu, W.; Kim, S.; Lee, S.; Perez-Neut, M.; *et al.* *Nature* **2019**, 574 (7779), 575.
- [26] Galligan, J. J.; Rose, K. L.; Beavers, W. N.; Hill, S.; Tallman, K. A.; Tansey, W. P.; Marnett, L. J. *J. Am. Chem. Soc.* **2014**, 136 (34), 11864.
- [27] Castro, J. P.; Jung, T.; Grune, T.; Siems, W. *Free Radic Biol. Med.* **2017**, 111, 309.
- [28] Sabari, B. R.; Zhang, D.; Allis, C. D.; Zhao, Y. *Nat. Rev. Mol. Cell Biol.* **2017**, 18 (2), 90.
- [29] Huang, J. X.; Lee, G.; Cavanaugh, K. E.; Chang, J. W.; Gardel, M. L.; Moellering, R. E. *Nat. Methods* **2019**, 16 (9), 894.
- [30] Moellering, R. E.; Cravatt, B. F. *Science* **2013**, 341 (6145), 549.
- [31] Wang, Z. P.; Wang, Y. H.; Chu, G. C.; Shi, J.; Li, Y. M. *Curr. Org. Synth.* **2015**, 12 (2), 150.
- [32] 田杰, 马新雨, 王志鹏. 生命的化学, **2018**, 38 (5), 5.
- [33] 王志鹏, 李曼, 程农壹, 王田. 化学教育, **2018**, 39 (18), 9.
- [34] 王志鹏, 李曼, 李琚, 刘志华, 李影, 郑基深. 有机化学, **2018**, 38 (9), 2400.
- [35] 王志鹏, 程农壹, 李曼, 王田. 大学化学, **2018**, 33 (3), 48.
- [36] 赵力维, 邓利娟, 范华军. 化学教育, **2016**, 37 (20), 67.
- [37] Wang, Z. A.; Millard, C. J.; Lin, C.-L.; Gurnett, J. E.; Wu, M.; Lee, K.; Fairall, L.; Schwabe, J. W. R.; Cole, P. A. *elife* **2020**, 9, e57663.
- [38] 王志鹏, 程农壹, 马新雨, 田杰. 化学教育, **2019**, 40 (10), 14.