

· 知识介绍 ·

doi: 10.3866/PKU.DXHX201708002

www.dxhx.pku.edu.cn

蛋白质中新型赖氨酸翻译后修饰的结构简介

王志鹏^{1,2,*} 程农壹³ 梁妍钰² 马新雨¹⁽¹⁾Chemistry Department, Texas A&M University, College Station, TX 77840, USA;⁽²⁾清华大学化学系, 北京 100084;⁽³⁾Department of Chemistry, University of Wisconsin-Madison, Madison, WI 53706, USA)

摘要: 近年来各种新型蛋白质赖氨酸翻译后修饰被逐渐发现。这些翻译后修饰在蛋白质组中广泛存在, 具有高度的结构多样性, 并对染色体结构和基因表达具有重要的表观遗传调控作用。本文将从化学的角度对这些新型赖氨酸翻译后修饰的发现与结构特点进行简述, 以期对教学与科研有帮助。

关键词: 赖氨酸; 翻译后修饰; 蛋白质组

中图分类号: G64; O6; Q501

Brief Introduction of Structures and Discovery of Novel Lysine Post-Translational Modifications in Native Proteins

WANG Zhipeng A.^{1,2,*} CHENG Nong-Yi³ LIANG Yan-Yu² MA Xin-Yu¹⁽¹⁾Chemistry Department, Texas A&M University, College Station, TX 77840, USA;⁽²⁾Chemistry Department, Tsinghua University, Beijing 100084, P. R. China;⁽³⁾Department of Chemistry, University of Wisconsin-Madison, Madison, WI 53706, USA)

Abstract: An increasing amount of novel protein lysine post-translational modifications (PTMs) have been gradually discovered. These PTMs with a large structural diversity widely exist in human proteome, which play an important epigenetically regulatory function to the chromatin structure and gene expression. This manuscript will briefly summarize the discovery and structure properties of the lysine PTMs from a chemical point of view, which will further benefit to teaching and research.

Key Words: Lysine, Post-translational modification, Proteome

1 引言

21 世纪人类基因组测序完成标志着我们正式进入“后基因组”时代。然而随着蛋白质翻译后修饰的深入研究, 人们发现蛋白质的复杂结构与精细调控除了受基因本身序列信息的影响外, 还与其上的各类翻译后修饰相关。故而表观遗传学应运而生, 并逐渐成为生命科学的重要研究方向之一。以赖氨酸甲基化为例, 其具有化学结构的多样性, 分别为单甲基化、双甲基化及三甲基化。赖氨酸甲基化作为最为重要的修饰方式之一广泛存在于核蛋白与胞质蛋白中, 其主要作用可以分为两方面。一方面, 赖氨酸甲基化作为表观遗传学的标志, 通过参与染色质结构与功能调控进而影响了 DNA 的复制、转录与修复过程; 另一方面, 核蛋白和胞质蛋白的赖氨酸甲基化修饰往往影响的是蛋白翻译和代

*通讯作者, Email: zhipeng.wang@chem.tamu.edu

谢过程^[1]。虽然这些翻译后修饰在细胞内大量存在,但目前具体的调控与信号传递等问题还并不清楚,也就成为表观遗传学的重要问题之一^[2]。

赖氨酸作为构成蛋白质唯一一个侧链含有 ϵ -氨基的氨基酸,具备伯胺的一系列结构特点与反应性能,故而是生物体内存在翻译后修饰种类最多的一种氨基酸。除了前文提到的甲基化,目前已经发现的修饰模式主要还包括酰基化。近年来,一系列新的酰基化修饰被逐渐发现,这些新型翻译后修饰的发现主要归功于基于串联质谱的蛋白质组学的发展^[3]。

然而,各类新型蛋白质表达后修饰模式的发现及其在生物体内重要的调控作用作为化学生物学前沿知识,在国内高校的化学系或生物系本科教育中鲜有涉及。本文将从有机化学的角度出发,首先分析各类新型赖氨酸酰基化的结构特性,进而介绍其各自的发现过程与特点。最后将简要介绍目前赖氨酸翻译后修饰的生命科学研究进展,以期对相关专业的教学与科研有助。

2 新型赖氨酸翻译后修饰的结构与分类

赖氨酸的酰基化种类繁多,且对细胞的结构与功能都有重要的调控作用。一些翻译后修饰早在化学生物学发展之初即被发现,如乙酰化、生物素化和高级脂肪酰化等修饰。随着串联质谱与蛋白质组学的发展,许多新型赖氨酸酰基化修饰模式被逐渐发现,包括甲酰化、丙二酰化、丁二酰化(琥珀酰化)、戊二酰化、 β -羟基丁酰化等等(图1),也是表观遗传学的研究方向之一^[4]。

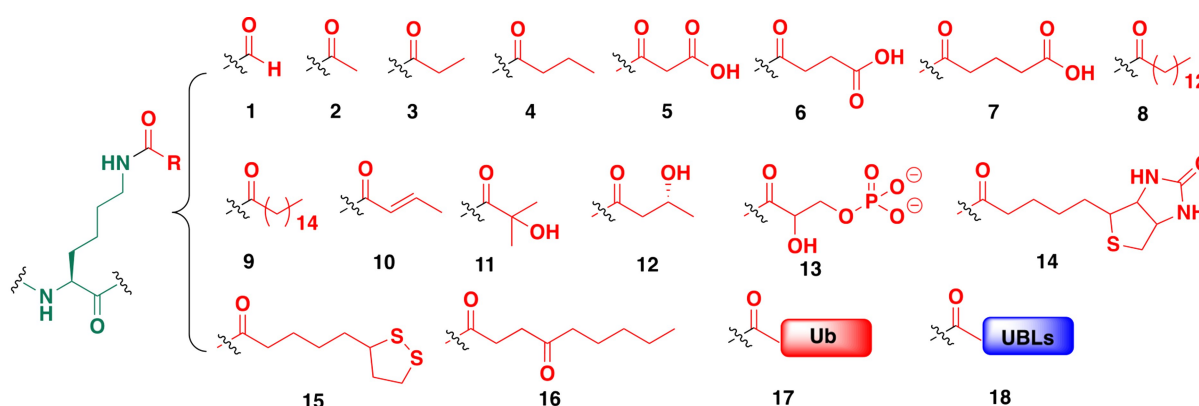


图1 天然蛋白中赖氨酸侧链上各类酰基化的结构

2.1 烷基酰化修饰

最为基本的一类翻译后修饰模式是小分子烷基酰化,例如甲酰化(图2 1)、乙酰化(图2 2)、丙酰化(图2 3)与丁酰化(图2 4)等修饰模式。这些修饰模式是相对较早发现的,例如乙酰化是最早被发现的修饰模式之一。值得注意的是,由于赖氨酸侧链氨基的 pK_a 约为10.5,在pH为7.4的生理条件下带有正电荷。正是由于此,组蛋白上富含赖氨酸,可以与带大量负电荷的DNA相互作用而紧密缠绕。而这些简单的酰基化最为直接的作用就是掩蔽赖氨酸这个自由氨基上的正电荷,进而削弱其与其他负电荷的相互作用(图2 a)。乙酰化由于发现较早,其生物学功能已得到最为广泛的研究,本文不做赘述^[5]。

2.2 阴离子酰化修饰

第二大类酰基化包括三类,即丙二酰化(图2 5)、丁二酰化(琥珀酰化)(图2 6)、戊二酰化(图2 7)三种酰基化修饰。虽然链长度不同,其最大特点是直链的末段带有一个额外的羧基。如上文所述,普通的小分子烷基酰化模式可以掩蔽赖氨酸这个自由氨基上的正电荷成为酰胺。更进一步的是,这一侧链上额外形成新的负电荷,使得对应修饰蛋白与DNA产生更大的排斥作用(图2 b)。

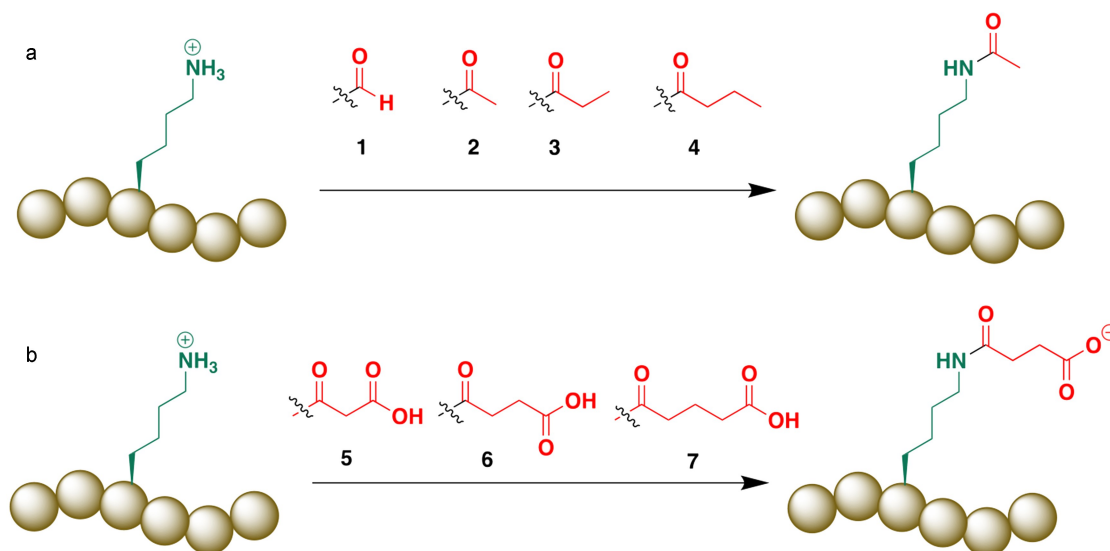


图2 烷基酰化修饰和阴离子酰化修饰对天然蛋白中赖氨酸侧链在生理条件下携带电荷情况的影响
(a) 基本酰基化可以屏蔽原有的正电荷；(b) 三种含有羧基的酰基化可以逆转电荷情况而携带负电荷

这三类修饰中，丙二酰化修饰最早由Zhao课题组于2011年^[6]发现，该课题组于同年发现丁二酰化^[7]。而戊二酰化修饰模式由同一课题组于2014年^[8]发现。蛋白质组学研究发现，这类修饰在真核细胞中广泛存在，且在基因的表达遗传调控与细胞代谢的调控中都有重要作用。例如丙二酰化修饰广泛存在于数百种蛋白的数以千计的赖氨酸侧链上，尤其是线粒体中参与代谢过程中的重要酶类大多都含有丙二酰化修饰^[9]。丙二酰化修饰不仅仅在真核细胞中具有重要的调控作用，其还是一些原核细胞代谢过程的重要调节模式^[10]。

2.3 长链酰化修饰/脂肪酰化修饰

天然蛋白质中存在一类较为特别的酰基化修饰模式，即长烷基链修饰或脂肪酰化修饰(Lipidation)。常见的两类赖氨酸脂肪酰化修饰包括豆蔻酰化(Myristoylation)(图3 8)与棕榈酰化(Palmitoylation)(图3 9)。以前者为例，目前已有若干例子被报道是蛋白质的赖氨酸豆蔻酰化修饰。例如重要的细胞因子肿瘤坏死因子(Tumor Necrosis Factor α , TNF α)^[11]、调控免疫应答与炎症反应的重要细胞因子蛋白白细胞介素1 α (Cytokine Interleukin 1 α , IL-1 α)^[12]等都含有赖氨酸豆蔻酰化修饰。其中，赖氨酸豆蔻酰化修饰(图3 a)与N末端豆蔻酰化修饰(图3 c)共同包含于“N-豆蔻酰化修饰”^[13]，而“N-豆蔻酰化修饰”与半胱氨酸侧链巯基的“S-棕榈酰化修饰”(图3 b)是两类主要的蛋白质脂肪酰化修饰模式^[14]，但是其具体功能与机制尚有待进一步研究^[15]。

脂肪酰化修饰不仅仅与其他酰基化修饰类似地屏蔽了赖氨酸上原有的正电荷，所引入的长链烷基链可以使得蛋白质有效地锚定于细胞膜磷脂双分子层内形成“附着型”膜蛋白。正由于此，蛋白质的脂肪酰化修饰与蛋白质的胞内转运、膜上固着、分泌出胞等过程有紧密联系。虽然这些修饰的形成与下游的信号转导尚无定论，人们已经发现了去脂肪酰化修饰的酶。这类酶能够特异性地催化底物蛋白脱去对应的脂肪酰化修饰^[16]。目前对于此类修饰的研究仍是生物化学与化学生物学的前沿问题^[17]。

2.4 其他小分子酰基化修饰

除了前文所述，还有一些结构相对特殊的酰基化翻译后修饰模式，由于其含有除了羧基以外的官能团，不能归类于前面三类，故而单独列出。由于种类较多且发现较新，大部分还缺乏系统性研究，故而在此略述。

α,β -丁烯酰化修饰(Crotonylation)最早由Zhao课题组^[18]于2011年发现，其结构的显著特点是含

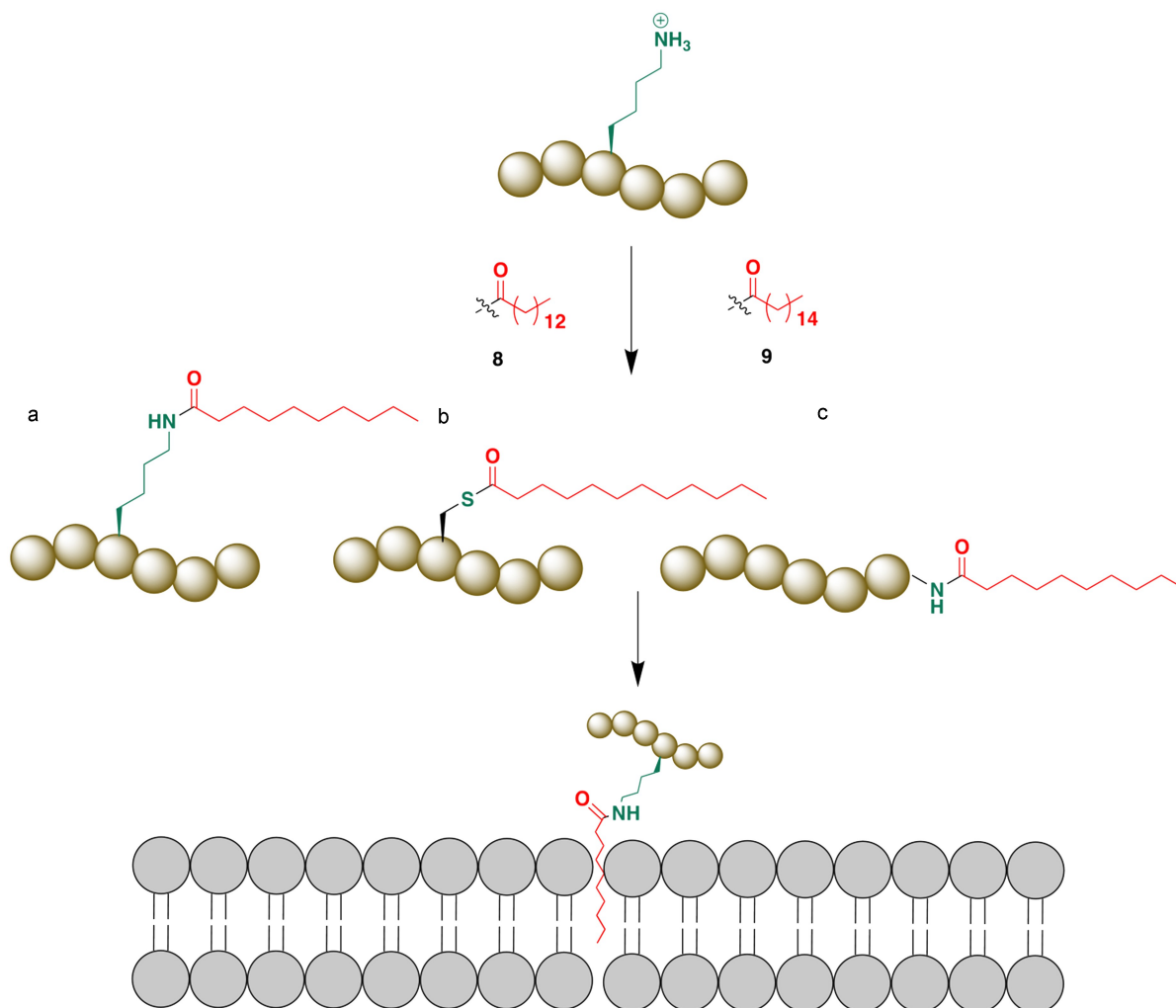


图3 三种脂肪酰化修饰的结构与其对蛋白质分布的影响(其中碳链长度仅为示意)

(a) 赖氨酸豆蔻酰化修饰; (b) 半胱氨酸豆蔻酰化修饰; (c) *N*-末端豆蔻酰化修饰

有 α,β -不饱和结构(图1 10)。这一反式双键平面具有特定的空间构型, 这相对于前文提到的简单烷基酰化修饰有所不同, 从物理有机化学的角度看, 对于所修饰蛋白与其他蛋白或DNA的结合有不同的影响(图4 a)。正由于此, 其异于其他简单酰基化修饰的结构和基因组上的分布都表明其特殊的机制与功能。此翻译后修饰模式已经有一定的研究, 例如其是雄性原始生殖细胞中性染色体上基因激活的一种独特的标记。

2-羟基异丁酰化修饰(2-hydroxyisobutyrylation)是由Zhao课题组^[19]于2014年发现的一种含有羟基的翻译后修饰模式(图1 11)。该修饰模式广泛存在于人与小鼠的组蛋白上, 且其修饰位置与功能都与之前发现的乙酰化修饰及 α,β -反丁烯酰化修饰有所不同。另外一种含有羟基的翻译后修饰模式(*R*)- β -羟基丁酰化修饰(β -hydroxybutyrylation)同样是由Zhao课题组^[20]于2016年发现(图1 12)。这种修饰模式广泛存在于人和小鼠的组蛋白上, 对于增强基因表达启动子有重要作用。而组蛋白上 β -羟基丁酰化修饰与细胞内 β -羟基丁酸物种的浓度水平直接相关。

蛋白质上侧链含有羟基的氨基酸如丝氨酸、苏氨酸和酪氨酸均可被磷酸化修饰。除了单纯含有羟基的酰基化修饰外, 羟基还可以进一步被磷酸化修饰而成为复合型的翻译后修饰模式。2013年由Cravatt课题组^[21]报道了赖氨酸的3-磷酸甘油酰化修饰(3-phosphoglyceroylation)就属于这特殊的一类

(图1 13)。由于3-磷酸甘油酸及其衍生物本来就是糖酵解过程中自然产生的中间代谢产物，故此修饰是直接活性代谢产物发生非酶促反应，且主要存在于对应代谢酶中以调节相关代谢过程。其他还有一些代谢酶类上赖氨酸侧链的辅因子酰基化修饰模式，例如生物素修饰(图1 14)、硫辛酸修饰(图1 15)等等。

4-氧代壬酰化修饰(4-oxononanoylation)是一种烷基链上含有额外羰基的酰基化修饰模式(图1 16)。这种修饰模式最初由Marnett课题组^[22]于2014年发现存在于组蛋白中。从物理有机化学上看，这一烷基链上的羰基具有相对较强的极性与偶极矩，在蛋白-蛋白或蛋白-核酸相互作用中可以与芳香基团形成特殊的羰基与芳环相互作用体系(图4 b)。这一点在去酰基化酶SIRT2与底物的相互作用并脱去此修饰模式的过程中至关重要^[23]。

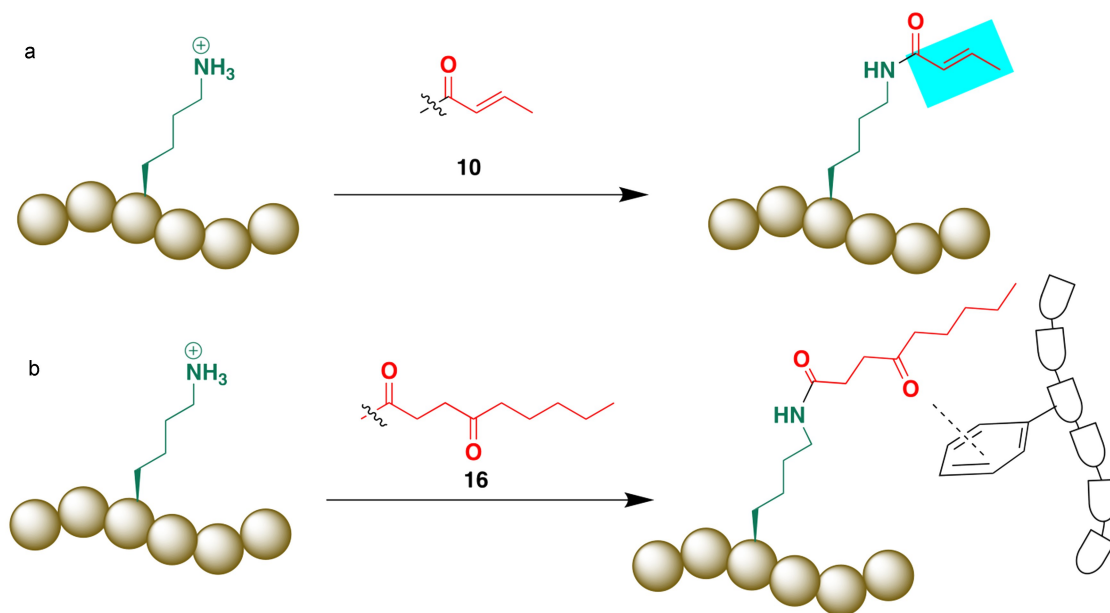


图4 赖氨酸 α,β -反丁烯酰化修饰与4-氧代壬酰化修饰对底物蛋白物理有机化学性质的影响示意图

(a) α,β -反丁烯酰化修饰的平面结构；(b) 4-氧代壬酰化修饰对底物蛋白物理有机化学性质的影响

2.5 小蛋白酰基化修饰

前文所提及的都是小分子酰基化修饰。除此之外还有一些多肽或小蛋白质的修饰模式，最为重要的就是赖氨酸的泛素化修饰(ubiquitination)(图1 17)。如果说前面提到的小分子酰基化修饰是对蛋白质底物结构与功能进行精细的调控，那么泛素化修饰可能是对底物蛋白质的功能甚至是“命运”最为至关重要的决定因素^[24]。泛素是一个含有76个氨基酸且序列高度保守的小型蛋白质，其通过C-末端的甘氨酸与底物蛋白的赖氨酸形成酰胺键键联。而泛素化的蛋白还可以进一步形成多泛素化。由于泛素含有7个赖氨酸，同样是C-末端与赖氨酸侧链形成的异肽键连接，泛素-泛素的连接方式可以多种多样。高度泛素化的底物蛋白会被26S的蛋白酶体所降解，而这一泛素-蛋白酶体系统也参与了众多重要的细胞过程，如细胞分裂、DNA修复等等。人们进一步发现了泛素化修饰的形成与脱去对应的酶系与机制。

除了泛素化修饰外，还有一大类“类泛素蛋白”(ubiquitin-like proteins, UBLs)，也可以通过C-末端形成异肽键与底物蛋白的赖氨酸相键联(图1 18)。这些修饰方式大部分都没有对应的中文翻译，其形成、脱去与信号转导的机制都是前沿的研究问题^[25]。

3 赖氨酸翻译后修饰的后续研究

前文中我们简要介绍了多类新型的赖氨酸翻译后修饰模式的结构和主要特征。本部分简要介绍其对应的研究手段。

3.1 串联质谱与蛋白质组学

串联质谱对于新型翻译后修饰的发现与后续研究都至关重要。除了利用串联质谱发现新型赖氨酸翻译后修饰外,蛋白质组学还可以分析翻译后修饰在蛋白质组的所在位置及其分布情况^[26]。此外,利用化学生物学手段结合蛋白质组学的分析方法也是一种研究翻译后修饰的重要方式,例如利用末端叠氮基团官能化的脂肪酸可以通过点击反应(Click反应)有效钓取并分析具有脂肪酰化的蛋白^[27]。

3.2 蛋白质化学生物学

此外,后续研究需要使用含有均一翻译后修饰的蛋白质,一个前提条件是制备这些含有特定翻译后修饰的均质蛋白。而经典的异源表达的策略无法实现对翻译后修饰的控制。故而,如何高效且便捷地实现天然蛋白中的特定位点的赖氨酸酰基化成为化学生物学家的一项重大挑战^[28]。利用蛋白质全合成和半合成手段,我们可以直接制备含有任何异源结构的蛋白质^[29]。此外,化学生物学家还发展出多种其他策略,如生物正交反应制备翻译后修饰类似物、密码子替代法直接引入非天然氨基酸或引入前提氨基酸进行后续修饰等手段^[30]。通过以上化学方法,我们可以有效地获得具有单一位点翻译后修饰赖氨酸的蛋白质样品,这对于后续的去翻译后修饰酶活性测试、蛋白-蛋白相互作用的钓取与结构生物学研究等生物性质研究极为重要。

4 结论与展望

作为生物化学与化学生物学的交叉前沿,蛋白质的各类翻译后修饰是后基因组时代的一个重要生命科学问题。这些翻译后修饰模式的多元化可以极大地拓宽原有蛋白质的结构与功能多样性与特殊性。同时,复杂的调控网络将细胞内的新陈代谢过程与基因表达建立起交互的桥梁。尤其是随着串联质谱与蛋白质组学的发展,越来越多的新型翻译后修饰模式被逐渐发现,相关问题的研究更加需要合成化学家与分析化学家的参与^[31]。

然而对于本科化学专业教学,传统教材与教学方法基本对本领域没有涉及。本文通过介绍近年来新发展的一些新型赖氨酸翻译后修饰结构及其发现,探索将前沿的蛋白质表观遗传学问题引入相关专业本科生的基础教育中^[32]。这样的教学尝试可以有效提高学生的学习兴趣,并丰富学生的知识广度,并对培养其独立思考与跨学科分析能力有益^[33]。

致谢: 感谢 University of Chicago 博士研究生邱天对本文构思提供的灵感。感谢 Texas A&M University 博士研究生 Erol Can Vatansever 对本文讨论部分提出的意见与建议。

参 考 文 献

- [1] Liu, W.; Wang, Z. A. *Chem. Eur. J.* **2017**, doi: 10.1002/chem.201701655.
- [2] 刘金凤, 王京兰, 钱小红, 蔡 耘. 中国生物化学与分子生物学报, **2007**, 23, 93.
- [3] Chen, B.; Zang, W.; Wang, J.; Huang, Y.; He, Y.; Yan, L.; Liu, J.; Zheng, W. *Chem. Soc. Rev.* **2015**, 44, 5246.
- [4] Wang, Z. A.; Kurra, Y.; Wang, X.; Zeng, Y.; Lee, Y.; Sharma, V.; Lin, H.; Dai, S. Y.; Liu, W. R. *Angew. Chem. Int. Ed.* **2017**, 56, 1643.
- [5] Choudhary, C.; Weinert, B. T.; Nishida, Y.; Verdin, E.; Mann, M. *Nat. Rev. Mol. Cell Biol.* **2014**, 15, 536.
- [6] Peng, C.; Lu, Z.; Xie, Z.; Cheng, Z.; Chen, Y.; Tan, M.; Luo, H.; Zhang, Y.; He, W.; Yang, K.; Zhao, Y. *Mol. Cell. Proteomics* **2011**, 10, M111. 012658.
- [7] Zhang, Z.; Tan, M.; Xie, Z.; Dai, L.; Chen, Y.; Zhao, Y. *Nat. Chem. Biol.* **2011**, 7, 58.

- [8] Tan, M.; Peng, C.; Anderson, K. A.; Chhoy, P.; Xie, Z.; Dai, L.; Park, J.; Chen, Y.; Huang, H.; Zhang, Y.; Ro, J.; Wagner, G. R.; Green, M. F.; Madsen, A. S.; Schmiesing, J.; Peterson, B. S.; Xu, G.; Ilkayeva, O. R.; Muehlbauer, M. J.; Braulke, T.; Muhlhausen, C.; Backos, D. S.; Olsen, C. A.; McGuire, P. J.; Pletcher, S. D.; Lombard, D. B.; Hirschey, M. D.; Zhao, Y. *Cell Metab.* **2014**, *19*, 605.
- [9] Weinert, B. T.; Schölz, C.; Wagner, S. A.; Iesmantavicius, V.; Su, D.; Daniel, J. A.; Choudhary, C. *Cell Reports* **2013**, *4*, 842.
- [10] Yang, M.; Wang, Y.; Chen, Y.; Cheng, Z.; Gu, J.; Deng, J. *Mol. Cell. Proteomics* **2015**, *14*, 796.
- [11] Stevenson, F. T.; Bursten, S. L.; Locksley, R. M.; Lovett, D. H. *J. Exp. Med.* **1992**, *176*, 1053.
- [12] Stevenson, F. T.; Bursten, S. L.; Fanton, C.; Locksley, R. M.; Lovett, D. H. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 7245.
- [13] Thinnon, E.; Serwa, R. A.; Broncel, M.; Brannigan, J. A.; Brassat, U.; Wright, M. H.; Heal, W. P.; Wilkinson, A. J.; Mann, D. J.; Tate, E. W. *Nat. Commun.* **2014**, *5*, 4919.
- [14] Hang, H. C.; Wilson, J. P.; Charron, G. *Acc. Chem. Res.* **2011**, *4*, 699.
- [15] 秦安东, 徐 林. 生命的化学, **2012**, *32*, 332.
- [16] Jiang, H.; Khan, S.; Wang, Y.; Charron, G.; He, B.; Sebastian, C.; Du, J.; Kim, R.; Ge, E.; Mostoslavsky, R.; Hang, H. C.; Hao, Q.; Lin, H. *Nature* **2013**, *496*, 110.
- [17] Wright, M. H.; Clough, B.; Rackham, M. D.; Rangachari, K.; Brannigan, J. A.; Grainger, M.; Moss, D. K.; Bottrill, A. R.; Heal, W. P.; Broncel, M. *Nat. Chem.* **2014**, *6*, 112.
- [18] Tan, M. J.; Luo, H.; Lee, S.; Jin, F. L.; Yang, J. S.; Montellier, E.; Buchou, T.; Cheng, Z. Y.; Rousseaux, S.; Rajagopal, N.; Lu, Z. K.; Ye, Z.; Zhu, Q.; Wysocka, J.; Ye, Y.; Khochbin, S.; Ren, B.; Zhao, Y. M. *Cell* **2011**, *146*, 1015.
- [19] Dai, L.; Peng, C.; Montellier, E.; Lu, Z.; Chen, Y.; Ishii, H.; Debernardi, A.; Buchou, T.; Rousseaux, S.; Jin, F. *Nat. Chem. Biol.* **2014**, *10*, 365.
- [20] Xie, Z.; Zhang, D.; Chung, D.; Tang, Z.; Huang, H.; Dai, L.; Qi, S.; Li, J.; Colak, G.; Chen, Y.; Xia, C.; Peng, C.; Ruan, H.; Kirkey, M.; Wang, D.; Jensen, L. M.; Kwon, O. K.; Lee, S.; Pletcher, S. D.; Tan, M.; Lombard, D. B.; White, K. P.; Zhao, H.; Li, J.; Roeder, R. G.; Yang, X.; Zhao, Y. *Mol. Cell* **2016**, *62*, 194.
- [21] Moellering, R. E.; Cravatt, B. F. *Science* **2013**, *341*, 549.
- [22] 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*, 11864.
- [23] Jin, J.; He, B.; Zhang, X.; Lin, H.; Wang, Y. *J. Am. Chem. Soc.* **2016**, *138*, 12304.
- [24] Mali, S. M.; Singh, S. K.; Eid, E.; Brik, A. *J. Am. Chem. Soc.* **2017**, *139*, 4971.
- [25] van der Veen, A. G.; Ploegh, H. L. *Annu. Rev. Biochem.* **2012**, *81*, 323.
- [26] 王志鹏, 王 鑫, 刘 伟. 化学教育, **2017**, *38* (12), 1.
- [27] Tate, E. W. *J. Chem. Biol.* **2008**, *1*, 17.
- [28] Wang, Z. A.; Zeng, Y.; Kurra, Y.; Wang, X.; Tharp, J. M.; Vatansever, E. C.; Hsu, W. W.; Dai, S.; Fang, X.; Liu, W. R. *Angew. Chem. Int. Ed.* **2017**, *56*, 212.
- [29] Zheng, J.; Tang, S.; Qi, Y.; Wang, Z. P.; Liu, L. *Nat. Protoc.* **2013**, *8*, 2483.
- [30] Wang, Z. A.; Tian, C.; Zheng, J. *RSC Adv.* **2015**, *5*, 107192.
- [31] 王志鹏, 邓 耿. 化学教育, **2017**, *38* (6), 72.
- [32] 王志鹏. 大学化学, **2017**, *32* (6), 86.
- [33] 王志鹏, 冯天师, 崔丽嘉, 石玉洁, 沙耀武. 中国科学: 生命科学, **2013**, *43* (9), 778.