

· 药学前沿 ·

基因编码非天然氨基酸技术及其在生物医学领域的应用

朱银雪, 王德祥, 孔影, 陆文捷, 叶慧*, 郝海平**

(中国药科大学药物代谢动力学重点实验室, 南京 210009)

摘要 人体内蛋白质通常由20种天然氨基酸组成。氨基酸的不同排列组合及立体构象是构建丰富多样性的蛋白质的基础。近年来开发的突破性基因密码子拓展技术通过向目标蛋白质中定向引入非天然氨基酸, 能够赋予原有的靶蛋白以共价结合邻近蛋白质、携带荧光基团、模拟蛋白质的翻译后修饰等新的生物学特性。本文从非天然氨基酸的结构和功能出发, 介绍了不同类型的非天然氨基酸分别在调节蛋白质稳定性、研究蛋白质构象、表达水平、定位以及识别和增强蛋白质间亲和作用等方面的用途。此外, 基因编码非天然氨基酸技术在生物医学领域的应用前景, 将为生物药物的设计及优化提供全新的思路和方法。

关键词 非天然氨基酸; 基因密码子拓展技术; 蛋白质; 反应性; 荧光基团; 翻译后修饰; 生物药物

中图分类号 Q78 **文献标志码** A **文章编号** 1000-5048(2022)04-0383-09

doi: 10.11665/j.issn.1000-5048.20220401

引用本文 朱银雪, 王德祥, 孔影, 等. 基因编码非天然氨基酸技术及其在生物医学领域的应用[J]. 中国药科大学学报, 2022, 53(4): 383-391.

Cite this article as: ZHU Yinxue, WANG Dexiang, KONG Ying, *et al.* Genetic incorporation of unnatural amino acids into proteins and its translational application in biomedicine[J]. *J China Pharm Univ*, 2022, 53(4): 383-391.

Genetic incorporation of unnatural amino acids into proteins and its translational application in biomedicine

ZHU Yinxue, WANG Dexiang, KONG Ying, LU Wenjie, YE Hui*, HAO Haiping**

Key Laboratory of Drug Metabolism and Pharmacokinetics, China Pharmaceutical University, Nanjing 210009, China

Abstract Proteins in the human body are usually made of 20 natural amino acids. Through different amino acid combinations and isomerization, proteins of diverse functions are built. An emerging genetic code expansion technology can introduce unnatural amino acids into specific sites of target protein, endowing the protein with new biological characteristics including covalently binding with proximal proteins, carrying fluorescence, and mimicking specific protein post-translational modifications. In this paper, based on the structure and function of unnatural amino acids, the applications of different types of unnatural amino acids in regulating protein's stability, studying protein's conformation, expression level, and localization, and uncovering heretofore unknown protein-protein interactions were reviewed. Besides, genetic code expansion of unnatural amino acids is anticipated to find broad utilities in biomedicine by bringing new ideas and methods to the design and optimization of biologics.

Key words unnatural amino acids; genetic code expansion; proteins; reactivity; fluorescent group; post-translational modification; biologics

This study was supported by the National Natural Science Foundation of China (No. 82173783, No. 81930109, No. 81720108032)

收稿日期 2022-03-12 **通信作者** *Tel: 025-83271191 E-mail: epuyehui@cpu.edu.cn

**Tel: 025-83271179 E-mail: haipinghao@cpu.edu.cn

基金项目 国家自然科学基金资助项目 (No. 82173783, No. 81930109, No. 81720108032); 中央高校基本科研业务费资助项目 (No. 2632022YC03)

and the Fundamental Research Funds for the Central Universities (No. 2632022YC03)

蛋白质是生命活动的基础,参与生命体内运输、催化、调节等多种过程,而这些生物学功能与蛋白质的序列密切相关。为了探究蛋白质的序列与功能的关联,将特定的氨基酸定点突变成其他氨基酸,研究蛋白质相应的功能改变,是研究该问题的经典手段^[1-2]。但是,氨基酸的突变通常局限于20种常用的氨基酸,而天然氨基酸自身携带的功能基团有限,难以满足改变、甚至赋予蛋白质以更加丰富的、多样的生物学功能的需求^[3-4]。

近年来,通过基因密码子拓展技术,向蛋白质中引入非天然氨基酸,赋予目标蛋白以新的生物学功能的研究取得了重要进展^[5]。基因编码的非天然氨基酸种类繁多,携带的官能团包括烯基、卤代烷烃、磺酰基、炔基、叠氮、醌甲基、喹啉基、磷酸基、乙酰基等^[6]。这类化学基团能够通过亲核取代、光激活、点击化学等多种反应赋予蛋白质以新的特性,帮助阐明蛋白质及其特定结构域和位点的理化性质和生物学功能。现有的非天然氨基酸已被编码插入不同类型的活细胞(比如大肠埃希菌、酵母、哺乳细胞等^[7-9])中,可用于增加靶蛋白的

光、热稳定性,定位蛋白质在体内的分布,揭示未知的蛋白质-蛋白质相互作用,研究翻译后修饰对蛋白质功能的调节等^[10-11]。值得一提的是,向目标蛋白质位点特异性地引入含有化学反应性或光激活交联基团的非天然氨基酸能够使得目标蛋白与其相互作用的靶标蛋白在互作区域形成共价键,从而获得与靶标蛋白亲和力更强的类似物,在生物药物研发领域具有极大的转化价值。这是由于相较于传统的小分子,高反应活性的共价生物药物对靶标蛋白具有更高的选择性,能够显著降低脱靶的副作用^[12-13],为具有明确结合靶标的生物药物设计带来了新的机遇。

目前已有超过200种非天然氨基酸能够通过基因密码子拓展技术引入到蛋白质中^[14-15]。本综述根据其结构和功能大致分为4类(图1):化学反应型(1~10)、光交联型(11~16)、荧光标记型(17~19)以及翻译后修饰型(20~21),并举例说明了基因编码的非天然氨基酸如何改变蛋白质的生物学活性及其在生物医学领域中的应用。

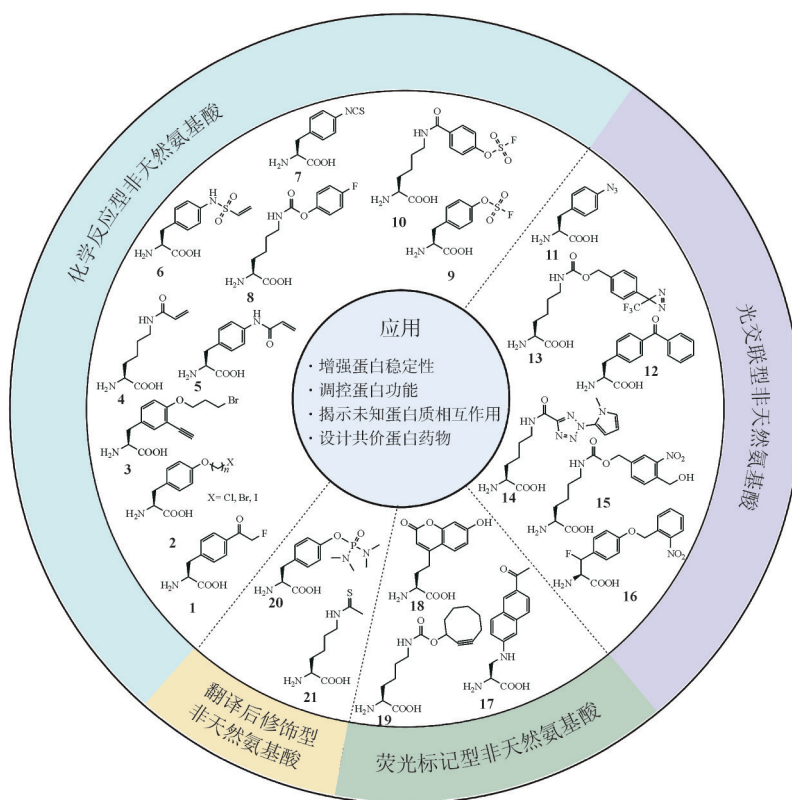


图1 基因编码的非天然氨基酸的分类及结构示意图

1 基因密码子拓展技术

基因密码子扩展技术指的是利用生物正交的氨酰-tRNA合成酶和tRNA分子对,将非天然氨基酸定点插入到蛋白质中的技术。该技术首先通过

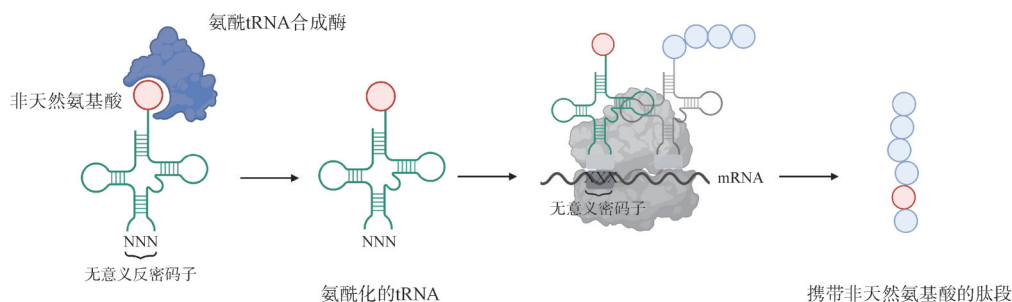


图2 基因密码子拓展技术示意图

2 化学反应型非天然氨基酸

将化学反应型非天然氨基酸引入目标蛋白质能调控目标蛋白与其互作蛋白的亲合力,影响互作介导的信号传递,具有转化为治疗疾病的生物药物的潜力,并且能够用于揭示未知的蛋白质互作关系。常规的化学交联剂多依赖于与靶标蛋白的氨基酸的氨基、巯基等活性基团发生反应。这种非特异性的交联作用会导致目标蛋白产生脱靶并结合其他蛋白的风险;使用邻近激活的非天然氨基酸则能够保证只有当靶蛋白的特定区域与插入了非天然氨基酸的目标蛋白足够接近时,才会触发共价交联反应。这种赋予目标蛋白共价结合靶标蛋白能力的基因编码非天然氨基酸的技术在蛋白质药物研发方面已经展现出巨大的转化潜力。

2.1 靶向半胱氨酸

Xiang等^[16]首次将非天然氨基酸 F_{fact} (*p*-2'-fluoroacetylphenylalanine)(**1**)引入 Z_{SPA} 拟抗体(affibody)蛋白,利用弱亲电性的C原子与邻近的强亲核性半胱氨酸的亲核取代反应(图3),在体外实现对Z蛋白的共价交联。此外, F_{fact} 还被引入哺乳细胞的1型促肾上腺皮质激素释放因子受体(corticotropin-releasing factor receptor type 1, CRF1R)中,用于研究CRF1R与内源性配体尿皮质素蛋白质复合物的拓扑结构,为理解受体的活化机制提供了独特的结构信息,有助于针对CRF1R的药物设计^[17]。

进化出特异性识别非天然氨基酸的氨酰-tRNA合成酶和正交的tRNA,合成出结合了非天然氨基酸的氨酰化tRNA,随后,该氨酰化tRNA识别mRNA上的无义密码子,在目标蛋白质的特定位点插入非天然氨基酸^[5](图2)。

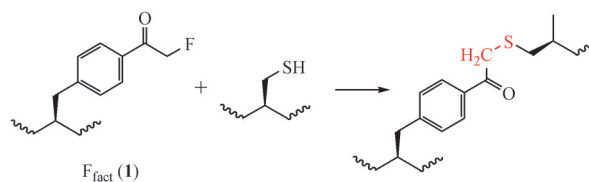


图3 非天然氨基酸 F_{fact} 靶向半胱氨酸的反应示意图

为了得到可调控的半胱氨酸反应性的非天然氨基酸,一系列含有不同卤素原子与不同长度脂肪链的卤代烷类非天然氨基酸 Haloalkane Uaas (unnatural amino acids)(**2**)被设计出来。这些卤代烷非天然氨基酸能够与目标半胱氨酸残基形成距离远超过天然二硫键的共价键,其稳定性也超过天然的二硫键,显著提升了蛋白质的稳定性,为蛋白质的结构研究和蛋白质工程提供了多样性^[18]。Yang等^[19]在卤代烷类非天然氨基酸的基础上引入了生物正交反应性的炔基基团,获得了EB3(**3**)。以Z蛋白和拟抗体为模型,发现通过EB3交联的肽段可借助点击化学反应进行富集,富集后的交联肽段的质谱响应比未富集的质谱响应提高了近30倍,提示EB3有助于提高基于质谱检测的交联蛋白、肽段的灵敏度和准确性。该方法已成功运用于捕获微弱、瞬时以及未知的蛋白质的相互作用,已识别出泛素结合酶(ubiquitin conjugating enzyme E2 D3, UBE2D3)与增殖细胞核抗原(proliferating cell nuclear antigen, PCNA)的弱相互作用,并鉴定出71个与硫氧还蛋白-1(thioredoxin-1, Trx1)具有相互作用而未见报道的蛋白质。

2.2 靶向赖氨酸

部分靶蛋白的结合界面或位点缺乏可供反应的半胱氨酸,为实现共价结合这类靶标蛋白,Furman等^[20]设计了3种能与赖氨酸发生迈克尔加成反应的非天然氨基酸 AcrK (*N*^ε-acryloyl-(*S*)-lysine)、AcrF(*p*-acrylamido-(*S*)-phenylalanine)和VSF(*p*-vinylsulfonamido-(*S*)-phenylalanine)(4~6)。以ErbB2(receptor tyrosine kinase 2)抗原结合片段-受体为例,研究发现当ErbB2抗原结合片段插入了含

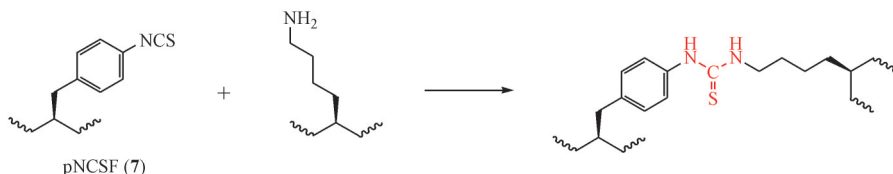


图4 非天然氨基酸pNCSF靶向赖氨酸的反应示意图

2.3 靶向多种氨基酸

鉴于部分靶标蛋白质参与结合的位点不含有半胱氨酸和赖氨酸,因此研究能够靶向多种氨基酸的非天然氨基酸具有重要价值。Xuan等^[22]设计出携带芳基氨基甲酸酯基团的非天然氨基酸FPheK(8),具有交联邻近的赖氨酸、半胱氨酸以及酪氨酸的能力。然而,FPheK需要在碱性条件下与赖氨酸、酪氨酸反应,并且具有中度至重度的细胞毒性,因此主要适用于体外实验。Wang等^[23-24]通过将芳基氟硫酸盐基团引入到酪氨酸和赖氨酸的侧链,设计出基于硫(VI)-氟置换反应与邻近的赖氨酸、组氨酸以及酪氨酸反应的非天然氨基酸FSY(fluorosulfate-L-tyrosine)和FSK(fluorosulfonyloxy-benzoyl-L-lysine)(9~10)。这类非天然氨基酸没有明显的细胞毒性,具有广泛应用到活细胞的潜力。此外,相比于FSY,FSK具有更长且柔韧的侧链,增加了交联邻近蛋白质的可能性。Li等^[25]将FSY引入人类程序性细胞死亡蛋白-1(human programmed cell death protein-1, PD-1),发现PD-1(FSY)能够共价结合人类程序性细胞死亡-配体1(human programmed cell death 1 ligand 1, PD-L1),阻断PD-L1/PD-1的结合和信号通路的激活,逆转肿瘤细胞对T细胞的免疫抑制作用,促进T细胞增殖并增强细胞因子的释放,更加有效地杀伤肿瘤细胞(图5)。该研究提示:应用基因编码的非天然氨基酸技术可以将各种生物药物与靶标蛋白的非共价亲和作用转化为共价结合,增强常规生物药

有乙烯基磺酰胺的VSF后,能够在生理条件下与ErbB2受体的赖氨酸发生高效的交联,进而干扰ErbB2受体的二聚化,诱导抗体依赖的细胞毒性作用,干扰细胞增殖,发挥抑制肿瘤生长的作用。此外,Xuan等^[21]利用芳香基异硫氰酸酯合成了非天然氨基酸pNCSF(*para*-isothiocyanate phenylalanine)(7),可以在温和的条件下与赖氨酸的侧链氨基发生亲核加成反应(图4),亦能高效地形成蛋白质的分子间、分子内的交联。

物的治疗效果。

3 光交联型非天然氨基酸

通过向基因编码的非天然氨基酸引入光激活基团与邻近氨基酸残基产生反应,能够实现光控的交联反应的发生,可用于捕捉瞬时、动态变化的蛋白质相互作用,有助于生物过程的研究和可控的药物设计(图6)。

3.1 残基非选择性的光交联非天然氨基酸

芳基叠氮化物是应用最广泛的光交联剂。Chin等^[26]通过对詹氏甲烷球菌的酪氨酰tRNA合成酶进行突变,筛选出一种新的正交氨酰tRNA合成酶/tRNA对,能选择性地非天然氨基酸AziF(*p*-azido-L-phenylalanine)(11)引入到蛋白质中,获得产率较高的重组蛋白。目前已成功运用于谷胱甘肽-S-转移酶(glutathione *S*-transferase, GST)二聚体拓扑结构的研究。此外,Chin等^[27]将另一种非天然氨基酸pBpa(*p*-benzoyl-L-phenylalanine)(12)基因编码至大肠埃希菌表达的蛋白质中,利用pBpa的二苯甲酮基团共价交联邻近的氨基酸残基,拓展了光交联反应性非天然氨基酸的多样性。Hino等^[28]将pBpa引入生长因子受体结合蛋白2(growth factor receptor-bound protein 2, Grb2)的SH2结构域,并在中国仓鼠卵巢(Chinese hamster ovary, CHO)细胞中进行表达。在波长365 nm紫外光照射下,Grb2-pBpa成功交联了表皮生长因子受体(epidermal growth factor receptor, EGFR)以及内

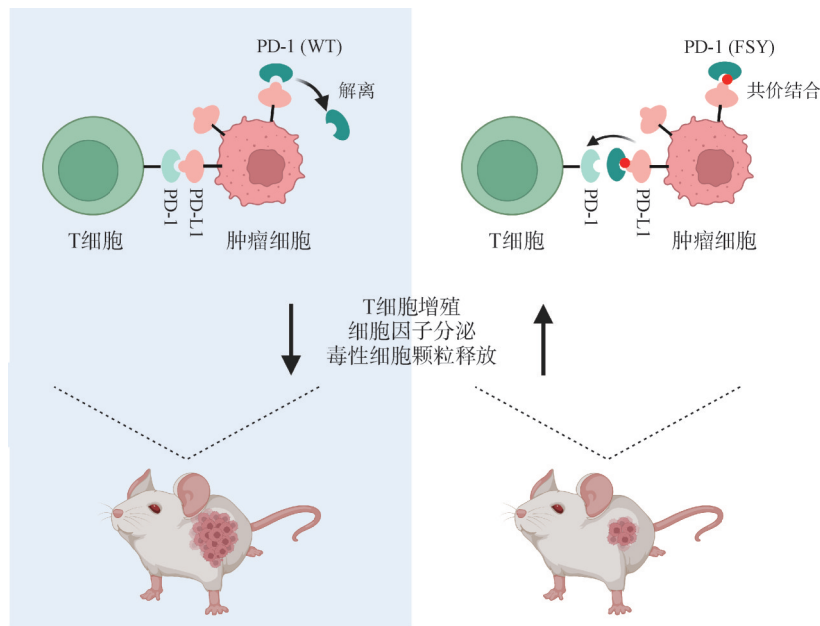


图 5 共价生物药物 PD-1(FSY)抑制肿瘤生长的原理图

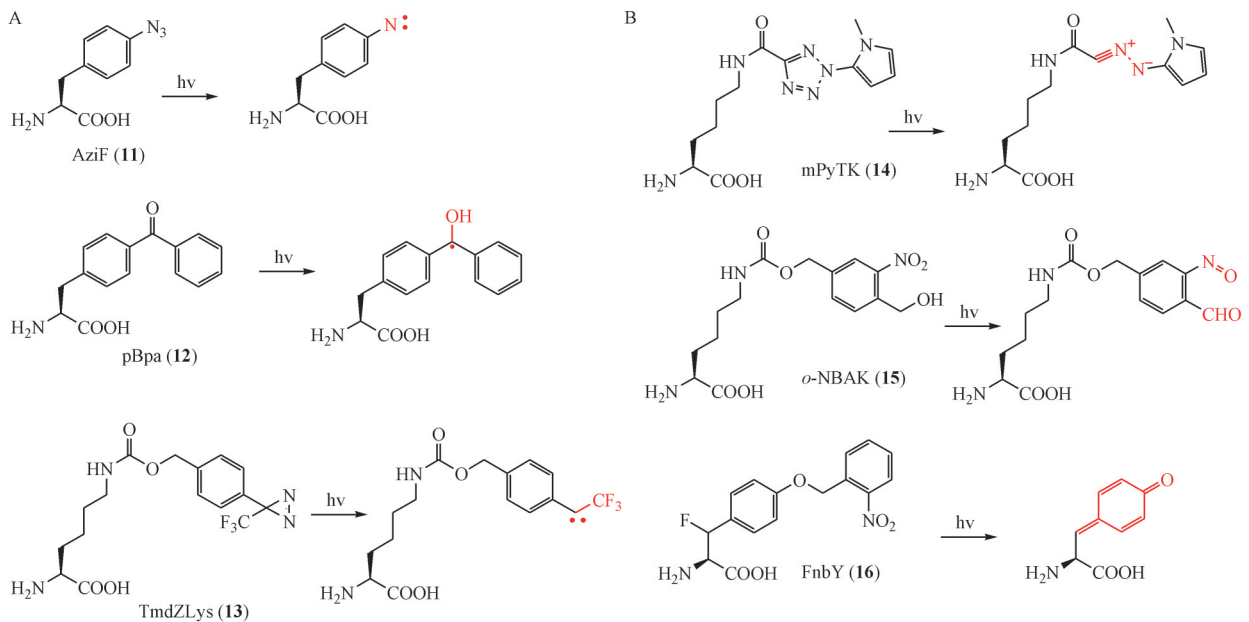


图 6 残基非选择性的(A)和残基选择性的(B)光交联非天然氨基酸的结构和反应示意图

源性酪氨酸激酶受体 ErbB2,充分证明了 pBpa 在研究体外、体内蛋白质间相互作用的价值。

上述两种以苯丙氨酸为模板衍生的非天然氨基酸,其光激活基团与蛋白质的骨架距离较短,只能在较近的范围共价交联相互作用的蛋白质。基于该局限,Yanagisawa 等^[29]设计并合成了一种赖氨酸衍生物 TmdZLys (N^{ϵ} -[(4-(3(trifluoromethyl)-3H-diazirin-3-yl) benzyl) oxy) carbonyl] -L-lysine)

(13),其双吡丙啶基团与 C_{α} 的距离为 15 Å,而 pBpa 与 C_{α} 距离仅为 7.9 Å。以 Grb2 为模型蛋白,发现插入 TmdZLys 后,Grb2 可以共价交联 EGFR 的 103 ~ 106 和 112 位点,而插入 pBpa 的 Grb2 仅能与上述位点形成弱交联甚至无交联。TmdZLys 这种长侧链的非天然氨基酸增加了蛋白质之间交联的机会,有助于识别蛋白质更广泛的结合伴侣。因此,不同长度的交联剂能够相互补充,共同勾勒出蛋

白质相互作用的拓扑结构。

3.2 残基选择性的光交联非天然氨基酸

上述传统的光交联非天然氨基酸对邻近的氨基酸残基不具有选择性,最近发展起来的残基选择性的光交联非天然氨基酸则能与特定氨基酸残基发生光交联反应。Tian等^[30]报道了系列的2-芳基-5-羧基四唑-赖氨酸类似物,其中含有甲基吡咯四唑基团的类似物 mPyTK (*N*-methylpyrroletetrazole-lysine)(**14**) 在紫外光激发后生成的羧脒亚胺可与邻近谷氨酸的羧基反应。因此,当 mPyTK 被引入哺乳细胞的 Grb2 中,能够通过光控瞬时交联其相互作用的蛋白 EGFR。此外, Hu 等^[31]设计合成了邻硝基苯甲醇类赖氨酸衍生物 *o*-NBAK (*o*-nitrobenzyl alcohol derived lysine)(**15**), 发现 *o*-NBAK 经过光异构化生成芳基亚硝基中间体后,能选择性地与邻近的赖氨酸交联,并用于捕获酶与底物蛋白之间的相互作用。

上述两种残基选择性的光交联非天然氨基酸仅能与特定氨基酸残基反应,限制了其靶向蛋白质的范围。Liu 等^[32]设计出 FnbY((*2R*)-2-amino-3-fluoro-3-(4-((2-nitrobenzyl)oxy)phenyl)propanoic acid)(**16**), 通过光激活生成的甲基活性醌可共价交联体内9种天然氨基酸残基。以大肠埃希菌中的 GST 为模型,在 103 位点插入 FnbY 后,发现 GST 除了与其二聚体界面邻近的 107 位赖氨酸反应,还与亲核性的组氨酸、酪氨酸、半胱氨酸、色氨酸、甲硫氨酸、精氨酸、谷氨酰胺以及天冬酰胺反应。与传统的光交联非天然氨基酸生成的反应中间体普遍具有纳秒至微秒级别的半衰期相比^[33-34],甲基活性醌半衰期可达到秒级别^[35-36],大大提高了蛋白质的交联效率。FnbY 的这一特性有助于实现活细胞中目标蛋白的高效共价交联,在化学生物学、生物疗法和蛋白质工程方面具有应用价值^[37]。

这类光交联非天然氨基酸借助瞬时控制以及不同的残基反应性,为研究时间分辨的蛋白质-蛋白质相互作用奠定基础,对理解生物学过程和开展转化蛋白质相互作用为共价蛋白质药物的研究具有推动作用。

4 荧光标记型非天然氨基酸

对蛋白质的荧光标记有助于探索生物大分子在细胞环境中的动态过程,实现细胞内生理过程

的可视化。Chatterjee 等^[38]报道了一种带荧光基团的非天然氨基酸 Anap(3-(6-acetylnaphthalen-2-yl-amino)-2-aminopropanoic acid)(**17**), 可借助常规的荧光显微镜观察基因编码 Anap 的蛋白质的荧光强度,已成功用于定位哺乳细胞中表达的组蛋白 H3。Charbon 等^[39]以香豆素为原型,设计了非天然氨基酸 CouAA (coumarin-derived amino acid)(**18**), 以 FtsZ 蛋白为模型,发现带荧光的 CouAA 能够在不影响 FtsZ 蛋白功能的基础上,确定 FtsZ 蛋白的亚细胞定位。

基因密码子拓展技术除了可以引入本身带有荧光基团的非天然氨基酸,还可以位点特异性地插入携带点击化学基团的非天然氨基酸,再通过点击化学反应引入荧光基团。例如, Jagadish 等^[40]将携带叠氮基团的非天然氨基酸 AziF(**11**) 引入环肽 MCoTI-I 中,利用叠氮与带有荧光基团的二苄基环辛炔(dibenzylcyclooctyne, DBCO)的衍生物进行环加成反应,使得环肽携带荧光,用于后续研究活细胞中环肽与蛋白质的相互作用。Alamudi 等^[41]对荧光探针做了筛选优化,根据探针脂溶性、水溶性以及范德华表面积电荷确定探针的细胞通透性和非特异性结合的强弱,最终选择分别含有环辛炔、叠氮化物基团的 BODIPY 探针 CO-1 和 AzG-1, 分别用于标记活细胞中插入了 AziF(**11**) 和 CoK(*N*- ϵ -(cyclooct-2-yn-1-yloxy)carbonyl)L-lysine)(**19**) 的靶蛋白,并进行活细胞成像。

5 翻译后修饰型非天然氨基酸

翻译后修饰是调控蛋白质功能的关键要素之一,影响着细胞中基因转录、增殖、信号转导、免疫调控等生命过程。因此,翻译后修饰的水平 and 失调与多种疾病密切相关^[42-44]。基因密码子拓展技术通过向目标蛋白质定点的引入翻译后修饰及其模拟物,有助于阐明翻译后修饰诱导的底物蛋白发生的结构和功能的变化。针对最重要的翻译后修饰之一——蛋白质磷酸化, Hoppmann 等^[45]通过基因编码引入 pTyr(phosphotyrosine)的类似物(**20**), 通过酸处理切除保护基团,高效地获得特定位点酪氨酸磷酸化的目标蛋白质。聚丙烯酰胺凝胶电泳以及质谱数据证明了该方法成功地向钙调蛋白和绿色荧光蛋白位点特异性地引入了 pTyr。此外,向泛素中 59 位酪氨酸引入磷酸化后,发现该位

点与 51 位谷氨酸的氢键被破坏,从而影响了 Tyr59-Glu51 环区,引起泛素的构象改变,降低其与 UBE2D3 结合的能力,揭示了泛素的 59 位酪氨酸磷酸化在泛素化过程中的负性调节作用。

由于体内存在多种移除翻译后修饰的酶, Venkat 等^[46]设计出一种有效模拟赖氨酸乙酰化、同时不能被去修饰酶识别的乙酰化赖氨酸类似物——硫代乙酰化赖氨酸 TAcK (*N*^ε-thioacetyl-L-lysine)(**21**)。以苹果酸脱氢酶为模型,研究人员发现插入 TAcK 的苹果酸脱氢酶可以被乙酰化赖氨酸的特异性抗体所识别,并且其酶活性与乙酰化形式相似。同时,该修饰不会被去乙酰化酶移除。

目前,赖氨酸泛素化、脂酰化、丝氨酸磷酸化等修饰皆通过基因密码子拓展技术被设计和引入到目标蛋白质中^[47-50],成为揭示翻译后修饰调控蛋白质的有力工具。但是,复杂的翻译后修饰的合成和插入仍旧充满挑战,最著名的例子是糖基化修饰。糖基化是最重要和复杂的蛋白质翻译后修饰之一,也是影响抗体药物的稳定性、安全性和生物活性的重要因素。我们期待在不久的将来能够实现定点引入糖基化修饰,生成均一性修饰的蛋白质,用于生物药物的药学性质的理性设计和优化。

6 结 论

非天然氨基酸通过引入反应性基团、荧光基团和翻译后修饰基团等,赋予了蛋白质新的特性。其中,化学反应型和光交联型的非天然氨基酸借助反应性官能团与邻近氨基酸残基发生共价结合,可用于靶向目标蛋白结合的靶标蛋白,不仅具有发现目标蛋白的新互作伴侣、阐明生物学过程的潜力,还能够用于构建具有共价结合靶标蛋白能力的目标蛋白的衍生物,从而转化开发出新的生物药物。荧光标记型和翻译后修饰型非天然氨基酸则为阐明蛋白质的功能提供了新的手段。根据研究的科学问题,多种非天然氨基酸被不断的优化、改进和创新。我们期待基因密码子拓展技术与不同性质的非天然氨基酸的融合能够用于设计出丰富多样的生物大分子,为生物疗法、生物学研究和蛋白质工程等多个领域带来重要的突破。

References

- [1] Humer D, Spadiut O. Improving the performance of horseradish peroxidase by site-directed mutagenesis [J]. *Int J Mol Sci*, 2019, **20**(4):916.
- [2] Lv KM, Shao WY, Pedroso MM, et al. Enhancing the catalytic activity of a GH5 processive endoglucanase from *Bacillus subtilis* BS-5 by site-directed mutagenesis [J]. *Int J Biol Macromol*, 2021, **168**:442-452.
- [3] Doering JA, Lee SH, Kristiansen K, et al. *In silico* site-directed mutagenesis informs species-specific predictions of chemical susceptibility derived from the Sequence Alignment to Predict Across Species Susceptibility (SeqAPASS) tool [J]. *Toxicol Sci*, 2018, **166**(1):131-145.
- [4] Drienovská I, Roelfes G. Expanding the enzyme universe with genetically encoded unnatural amino acids [J]. *Nat Catal*, 2020, **3**(3):193-202.
- [5] Dumas A, Lercher L, Spicer CD, et al. Designing logical codon reassignment - Expanding the chemistry in biology [J]. *Chem Sci*, 2015, **6**(1):50-69.
- [6] Cao L, Wang L. New covalent bonding ability for proteins [J]. *Protein Sci*, 2022, **31**(2):312-322.
- [7] Zambaldo C, Koh M, Nasertorabi F, et al. An orthogonal seryl-tRNA synthetase/tRNA pair for noncanonical amino acid mutagenesis in *Escherichia coli* [J]. *Bioorg Med Chem*, 2020, **28**(20):115662.
- [8] Hu LM, Qin XW, Huang YJ, et al. Thermophilic pyrrolysyl-tRNA synthetase mutants for enhanced mammalian genetic code expansion [J]. *ACS Synth Biol*, 2020, **9**(10):2723-2736.
- [9] Shao SD, Koh M, Schultz PG. Expanding the genetic code of the human hematopoietic system [J]. *Proc Natl Acad Sci U S A*, 2020, **117**(16):8845-8849.
- [10] Nguyen TA, Cigler M, Lang K. Expanding the genetic code to study protein-protein interactions [J]. *Angew Chem Int Ed Engl*, 2018, **57**(44):14350-14361.
- [11] Chung CZ, Amikura K, Söll D. Using genetic code expansion for protein biochemical studies [J]. *Front Bioeng Biotechnol*, 2020, **8**:1233.
- [12] Bedard PL, Hyman DM, Davids MS, et al. Small molecules, big impact: 20 years of targeted therapy in oncology [J]. *Lancet*, 2020, **395**(10229):1078-1088.
- [13] Wang NX, Wang L. Genetically encoding latent bioreactive amino acids and the development of covalent protein drugs [J]. *Curr Opin Chem Biol*, 2022, **66**:102106.
- [14] Zhu HQ, Tang XL, Zheng RC, et al. Recent advancements in enzyme engineering via site-specific incorporation of unnatural amino acids [J]. *World J Microbiol Biotechnol*, 2021, **37**(12):213.

- [15] Smolskaya S, Andreev YA. Site-specific incorporation of unnatural amino acids into *Escherichia coli* recombinant protein: methodology development and recent achievement[J]. *Biomolecules*, 2019, **9**(7):255.
- [16] Xiang Z, Ren HY, Hu YS, *et al.* Adding an unnatural covalent bond to proteins through proximity-enhanced bioreactivity[J]. *Nat Methods*, 2013, **10**(9):885-888.
- [17] Coin I, Katritch V, Sun TT, *et al.* Genetically encoded chemical probes in cells reveal the binding path of urocortin-I to CRF class B GPCR[J]. *Cell*, 2013, **155**(6):1258-1269.
- [18] Xiang Z, Lacey VK, Ren HY, *et al.* Proximity-enabled protein crosslinking through genetically encoding haloalkane unnatural amino acids[J]. *Angew Chem Int Ed Engl*, 2014, **53**(8):2190-2193.
- [19] Yang B, Tang SB, Ma C, *et al.* Spontaneous and specific chemical cross-linking in live cells to capture and identify protein interactions[J]. *Nat Commun*, 2017, **8**(1):2240.
- [20] Furman JL, Kang MC, Choi S, *et al.* A genetically encoded azamichael acceptor for covalent cross-linking of protein-receptor complexes[J]. *J Am Chem Soc*, 2014, **136**(23):8411-8417.
- [21] Xuan WM, Li J, Luo XZ, *et al.* Genetic incorporation of a reactive isothiocyanate group into proteins[J]. *Angew Chem Int Ed Engl*, 2016, **55**(34):10065-10068.
- [22] Xuan WM, Shao SD, Schultz PG. Protein crosslinking by genetically encoded noncanonical amino acids with reactive aryl carbamate side chains[J]. *Angew Chem Int Ed Engl*, 2017, **56**(18):5096-5100.
- [23] Wang NX, Yang B, Fu CY, *et al.* Genetically encoding fluorosulfate-L-tyrosine to react with lysine, histidine, and tyrosine via SuFEx in proteins *in vivo*[J]. *J Am Chem Soc*, 2018, **140**(15):4995-4999.
- [24] Liu J, Cao L, Klauser PC, *et al.* A genetically encoded fluorosulfonyloxybenzoyl-L-lysine for expansive covalent bonding of proteins via SuFEx chemistry[J]. *J Am Chem Soc*, 2021, **143**(27):10341-10351.
- [25] Li QK, Chen Q, Klauser PC, *et al.* Developing covalent protein drugs via proximity-enabled reactive therapeutics [J]. *Cell*, 2020, **182**(1):85-97.e16.
- [26] Chin JW, Santoro SW, Martin AB, *et al.* Addition of *p*-azido-L-phenylalanine to the genetic code of *Escherichia coli*[J]. *J Am Chem Soc*, 2002, **124**(31):9026-9027.
- [27] Chin JW, Martin AB, King DS, *et al.* Addition of a photocrosslinking amino acid to the genetic code of *Escherichia coli*[J]. *Proc Natl Acad Sci USA*, 2002, **99**(17):11020-11024.
- [28] Hino N, Okazaki Y, Kobayashi T, *et al.* Protein photo-cross-linking in mammalian cells by site-specific incorporation of a photo-reactive amino acid[J]. *Nat Methods*, 2005, **2**(3):201-206.
- [29] Yanagisawa T, Hino N, Iraha F, *et al.* Wide-range protein photocrosslinking achieved by a genetically encoded *N*^ε-(benzyloxycarbonyl)lysine derivative with a diazirinyl moiety[J]. *Mol Biosyst*, 2012, **8**(4):1131-1135.
- [30] Tian YL, Jacinto MP, Zeng Y, *et al.* Genetically encoded 2-aryl-5-carboxytetrazoles for site-selective protein photo-cross-linking [J]. *J Am Chem Soc*, 2017, **139**(17):6078-6081.
- [31] Hu W, Yuan Y, Wang CH, *et al.* Genetically encoded residue-selective photo-crosslinker to capture protein-protein interactions in living cells[J]. *Chem*, 2019, **5**(11):2955-2968.
- [32] Liu J, Li SS, Aslam NA, *et al.* Genetically encoding photocaged quinone methide to multitarget protein residues covalently *in vivo*[J]. *J Am Chem Soc*, 2019, **141**(24):9458-9462.
- [33] Row RD, Nguyen SS, Ferreira AJ, *et al.* Chemically triggered crosslinking with bioorthogonal cyclopropenones [J]. *ChemComm*, 2020, **56**(74):10883-10886.
- [34] Roy A, Barman S, Padhan J, *et al.* Engineering an acetyllysine reader with a photocrosslinking amino acid for interactome profiling[J]. *ChemComm*, 2021, **57**(77):9866-9869.
- [35] Zhu XH, Akiyama T, Yokoyama T, *et al.* Stereoselective formation of β -O-4 structures mimicking softwood lignin biosynthesis: effects of solvent and the structures of quinone methide lignin models[J]. *J Agric Food Chem*, 2019, **67**(25):6950-6961.
- [36] Ito S, Sugumaran M, Wakamatsu K. Chemical reactivities of *ortho*-quinones produced in living organisms: fate of quinonoid products formed by tyrosinase and phenoloxidase action on phenols and catechols[J]. *Int J Mol Sci*, 2020, **21**(17):6080.
- [37] Liu J, Cheng RJ, Van Eps N, *et al.* Genetically encoded quinone methides enabling rapid, site-specific, and photocontrolled protein modification with amine reagents [J]. *J Am Chem Soc*, 2020, **142**(40):17057-17068.
- [38] Chatterjee A, Guo JT, Lee HS, *et al.* A genetically encoded fluorescent probe in mammalian cells [J]. *J Am Chem Soc*, 2013, **135**(34):12540-12543.
- [39] Charbon G, Brustad E, Scott KA, *et al.* Subcellular protein localization by using a genetically encoded fluorescent amino acid [J]. *Chem Bio Chem*, 2011, **12**(12):1818-1821.
- [40] Jagadish K, Borra R, Lacey V, *et al.* Expression of fluorescent cyclotides using protein trans-splicing for easy monitoring of cyclotide-protein interactions [J]. *Angew Chem Int Ed Engl*, 2013, **52**(11):3126-3131.
- [41] Alamudi SH, Satapathy R, Kim J, *et al.* Development of background-free tame fluorescent probes for intracellular live cell imaging[J]. *Nat Commun*, 2016, **7**:11964.
- [42] Macek B, Forchhammer K, Hardouin J, *et al.* Protein post-translational modifications in bacteria[J]. *Nat Rev Microbiol*, 2019, **17**(11):651-664.
- [43] Zhou JH, Zhao SW, Dunker AK. Intrinsically disordered proteins link alternative splicing and post-translational modifications to complex cell signaling and regulation [J]. *J Mol Biol*, 2018, **430**(16):2342-2359.
- [44] Müller MM. Post-translational modifications of protein backbones: unique functions, mechanisms, and challenges [J]. *Bio-*

- chemistry, 2018, **57**(2): 177-185.
- [45] Hoppmann C, Wong A, Yang B, *et al.* Site-specific incorporation of phosphotyrosine using an expanded genetic code [J]. *Nat Chem Biol*, 2017, **13**(8): 842-844.
- [46] Venkat S, Nannapaneni DT, Gregory C, *et al.* Genetically encoding thioacetyl-lysine as a non-deacetyltable analog of lysine acetylation in *Escherichia coli* [J]. *FEBS Open Bio*, 2017, **7**(11): 1805-1814.
- [47] Luo XZ, Fu GS, Wang RE, *et al.* Genetically encoding phosphotyrosine and its nonhydrolyzable analog in bacteria [J]. *Nat Chem Biol*, 2017, **13**(8): 845-849.
- [48] Fottner M, Brunner AD, Bittl V, *et al.* Site-specific ubiquitylation and SUMOylation using genetic-code expansion and sortase [J]. *Nat Chem Biol*, 2019, **15**(3): 276-284.
- [49] Beránek V, Reinkemeier CD, Zhang MS, *et al.* Genetically encoded protein phosphorylation in mammalian cells [J]. *Cell Chem Biol*, 2018, **25**(9): 1067-1074.e5.
- [50] Fu CY, Chen Q, Zheng F, *et al.* Genetically encoding a lipidated amino acid for extension of protein half-life *in vivo* [J]. *Angew Chem Int Ed Engl*, 2019, **58**(5): 1392-1396.

· 校园新闻 ·

郝海平/叶慧团队发表有关环状亚胺离子示踪技术揭示人类蛋白质组存在丰富乳酰化修饰谱的研究论文

2022年6月27日, *Nature Methods* 期刊在线发表了我校郝海平/叶慧团队的最新研究工作成果(Cyclic immonium ion of lactyllysine reveals widespread lactylation in the human proteome)。2019级博士研究生皖宁和2018级硕士研究生王念为本论文的共同第一作者, 我校叶慧研究员、郝海平教授和南京中医药大学王南溪教授为本文的共同通讯作者, 中国药科大学为本文的第一通讯单位。该工作的主要合作者包括王广基院士、邵畅博士和余思琴、张汉卿、王德祥、陆文捷、孔影、王鑫淼、吕浪浪等; 我校药科院江苏省药物代谢动力学重点实验室对该工作的开展给予了大力支持。

该工作针对乳酸是否可以共价修饰非组蛋白进而发挥生物学效应的科学问题, 提出在公共的人类蛋白质组深度测序数据中搜索乳酰化修饰的新底物蛋白的策略。然而, 由于从非富集的蛋白质组数据中检索修饰的假阳性率极高, 若能发现修饰特异性的特征离子则能通过谱图筛选, 显著降低鉴定的假阳性率, 揭示真实修饰位点和靶蛋白, 指导后续的生物验证。基于此需求, 该团队通过合成和研究模型乳酰化肽段的谱图, 首次发现了携带乳酰化修饰赖氨酸的多肽在碰撞室中经过二级断裂会形成链状亚胺离子, 该离子经过脱氨形成次生碎片——环状亚胺离子。进一步通过分析生物样本中富集的大量乳酰化阳性肽段, 再以近十万条人类蛋白质组的非修饰肽段谱图作为阴性对照, 确证了环状亚胺离子指征乳酰化修饰的灵敏度和特异性, 提出该离子作为判定数据库检索的乳酰化真实与否的金标准。

基于该诊断离子策略, 研究者从现有的非富集、大规模人类蛋白质组数据资源中挖掘了全新的、丰富的乳酰化修饰底物蛋白和位点的信息, 特别是从2020年 *Nature Methods* 发表的 Meltome Atlas 库的人类细胞蛋白质组热稳定性数据中, 发现乳酰化修饰高度富集在糖酵解通路。其中, 代谢酶 ALDOA 的乳酰化修饰存在于多种人类肿瘤细胞系且修饰占比高, 引发了乳酰化修饰调节代谢酶活性等功能, 进而调控糖酵解通路的猜想。

郝海平、叶慧团队进一步联合王南溪课题组, 利用先进的化学生物学技术——基因密码子扩展技术, 首次实现向靶蛋白 ALDOA 定点引入乳酰化修饰, 并发现修饰后酶活性显著降低, 揭示了乳酸蓄积后, 能通过共价修饰糖酵解通路的上游代谢酶, 抑制糖酵解活跃度的反馈调节机制, 对生物化学领域现有的“终产物抑制”的调控模式进行了补充。

综上, 该研究表明乳酰化是广泛存在于人类组织和细胞中的一种非组蛋白特异性的、具有生物学功能的翻译后修饰。该示踪技术的应用能够助力在未来发现更多样的乳酰化修饰蛋白, 并且揭示乳酰化修饰的动态变化与乳酸紊乱驱动炎症、肿瘤等重大慢性疾病发生发展的关联, 提示新的治疗靶点。该工作得到了国家重点研发计划、国家自然科学基金面上项目、中国药科大学兴药学者计划、天然药物活性组分与药效国家重点实验室自主研究课题、中央高校基本科研业务费等资助。

(科技处)