

· 综 述 ·

遗传密码扩充技术及其在蛋白质功能研究及标记成像中的应用

李哲¹, 凌虹^{1,2}

1 哈尔滨医科大学 微生物学教研室, 黑龙江 哈尔滨 150081

2 哈尔滨医科大学 免疫学教研室, 黑龙江 哈尔滨 150081

李哲, 凌虹. 遗传密码扩充技术及其在蛋白质功能研究及标记成像中的应用. 生物工程学报, 2022, 38(2): 620-631.

LI Z, LING H. Genetic code expansion and its application in characterization and imaging of proteins. Chin J Biotech, 2022, 38(2): 620-631.

摘 要: 遗传密码扩充 (genetic code expansion, GCE) 技术利用终止密码子将非天然氨基酸掺入到蛋白质中, 再结合点击反应对蛋白质实现定点标记。相较于荧光蛋白、标签抗体等其他标记工具, 该技术在蛋白标记中使用的化合物分子较小、对蛋白空间结构影响较小, 且能通过点击反应实现蛋白分子与染料分子 1:1 的化学计量比, 从而能够依据荧光强度对蛋白质定量。因此, 在活细胞单分子追踪和超分辨率显微成像等需要细胞长时间暴露在高激光功率下的研究中, GCE 技术具有极大的优势。同时, 该技术也为提高活细胞成像过程中的定位精度和分子计数准确度奠定了基础。文中旨在总结近年来 GCE 技术在蛋白质研究中的应用进展, 特别是在蛋白质标记成像方面的应用进展。

关键词: 遗传密码扩充; 非天然氨基酸; 活细胞成像; 蛋白质

Genetic code expansion and its application in characterization and imaging of proteins

LI Zhe¹, LING Hong^{1,2}

1 Department of Microbiology, Harbin Medical University, Harbin 150081, Heilongjiang, China

2 Department of Immunology, Harbin Medical University, Harbin 150081, Heilongjiang, China

Abstract: Genetic code expansion (GCE) allows the incorporation of unnatural amino acids into

Received: March 25, 2021; Accepted: June 2, 2021

Supported by: National Natural Science Foundation of China (81772190)

Corresponding author: LING Hong. Tel/Fax: +86-451-87503179; E-mail: lingh@ems.hrbmu.edu.cn

基金项目: 国家自然科学基金 (81772190)

proteins via using stop codons. GCE may achieve site-specific labeling of proteins in combination with the click reaction. Compared with other labeling tools such as fluorescent proteins and tagged antibodies, the compound molecules used in protein labeling by GCE technology are smaller, and therefore, may less interfere the conformational structure of proteins. In addition, through click reaction, GCE allows a 1:1 stoichiometric ratio of the target protein molecule and the fluorescent dye, and the protein can be quantified based on the fluorescence intensity. Thus, GCE technology has great advantages in the researches that require the exposition of living cells under high laser power for longer time, for example, in the context of single molecule tracing and super-resolution microscopic imaging. Meanwhile, this technology lays the foundation for improving the accuracy of positioning and molecule counting in the imaging process of living cells. This review summarized the GCE technology and its recent applications in functionally characterizing, labeling and imaging of proteins.

Keywords: genetic code expansion; unnatural amino acids; live cell imaging; proteins

近年来, 研究分子功能及分子之间的相互作用成为生物分子研究的一个重要方向, 如何在活细胞甚至活体状态下对这些分子进行长时间示踪成像成为一大难点与挑战; 同时, 不同类型的显微成像技术应运而生^[1]。新成像技术的出现开辟了生物学研究的一个新领域, 即在单细胞水平上理解细胞变化过程和分子动力学改变^[2]。然而, 当前常用的荧光蛋白、标签抗体等标记及检测工具由于分子量过大、光物理特性等限制, 仍无法满足更为严苛精确的示踪要求^[3]。因此, 亟需既能减少荧光标记复合物对蛋白的外源干扰又能实现精确示踪蛋白的标记技术。

遗传密码扩充 (genetic code expansion, GCE) 是指在生物正交对——氨酰化 tRNA 合成酶 (aminoacyl-tRNA synthetase, aaRS)/转运 RNA (tRNA) (即 aaRS/tRNA 对) 的帮助下, 使无义密码子 (通常为琥珀密码子 UAG) 执行有义编码功能, 并在相应位点引入非天然氨基酸 (unnatural amino acid, UAA), 由此增加遗传密码的多样性。GCE 技术依据上述原理, 利用基因工程技术将目的蛋白中的氨基酸密码子突变为无义密码子, 并外源引入 UAA 及其特异 aaRS/tRNA 对; 在 aaRS 催化下, tRNA 携带

UAA 与无义密码子进行特异性识别, 将 UAA 插入到肽链相应位置, 实现多肽链的顺利通读。其中, 外源引入 aaRS/tRNA 使琥珀密码子 UAG 执行有义密码子功能, 进而导致多肽链顺利通读, UAG 原本的终止翻译功能被抑制的过程被称为琥珀抑制^[4] (图 1)。

GCE 技术结合生物正交点击反应能够实现精准定位到氨基酸的特异性荧光标记。用于氨基酸标记的点击反应主要有 3 类 (图 2A): (1) 应变促进的逆电子需求狄尔斯-阿尔德环加成 (strain-promoted inverse-electron-demand Diels-Alder cycloaddition, SPIEDAC) 反应^[5]; (2) 应变促进叠氮化物-炔环加成 (strain promoted azide-alkyne cycloaddition, SPACC) 反应^[5]; (3) 铜催化的叠氮化物-炔环加成 (Cu(I)-catalyzed azide-alkyne cycloaddition, CuAAC) 反应^[6]。其中, 侧链为环烯烃的 UAA 可以与四嗪 (tetrazine, Tet) 基团发生 SPIEDAC 反应, 侧链为环炔烃的 UAA 可以与叠氮基-荧光染料发生 SPACC 或者 CuAAC 反应。由此, UAA 可以与携带 Tet 或者叠氮基的荧光染料高度特异结合, 最终实现对蛋白中特定氨基酸的精准标记。机体内的天然氨基酸不具备环状结构, 则不能发生上述

点击反应, 由此保证了 UAA 的特定荧光标记。

与其他蛋白质标记技术相比, GCE 标记技术具备以下关键优势: (1) 荧光团改造的灵活性较高, 凡是能够通过化合反应生成携带四嗪的荧光团, 均可以通过 SPIEDAC 反应快速整合到掺入 UAA 的蛋白中^[3]; (2) 能够最大化降低荧光标记物对目标蛋白的干扰。当前, GCE 技术已被应用于蛋白质的构象、功能及作用机制研究、胞内标记成像等方面^[7-8], 为生物分子功能及分子之间相互作用的深入研究提供了极有价值的手段。下文将简要综述 GCE 技术的优化及其在蛋白质标记示踪研究中的应用。

1 GCE 技术的核心要素及优化

GCE 技术有以下 3 个核心要素 (图 1): (1) 容易被目标生物体吸收、代谢稳定、无毒的 UAA。为赋予蛋白不同功能, UAA 由人工合成并在侧链基团上携带了不同性质的官能团。目前, 通过 GCE 技术已有 150 多种 UAA 可用于掺入到目的蛋白中^[9]。 (2) 引入到目的基因中的无义密

码子。通常, 首选琥珀密码子 UAG 作为无义密码子, 有少数研究也选择四联密码子, 将目的氨基酸的密码子突变为四联密码子 (如 AGGA), 同时, 构建相应反密码子 tRNA (如 tRNA_{UCCU}), 目的在于降低引入终止密码子所致突变出现回复的可能性^[10]。 (3) 由 aaRS/tRNA 对组成 UAA 正交翻译系统 (orthogonal translation system, OTS)。OTS 的正交性表现为: OTS 中的 tRNA、aaRS 组分不会与胞内原有的任何一类内源性氨基酸、aaRSs 或 tRNAs 发生相互作用, 彼此之间不相互干扰。同样, 内源性的 aaRSs、tRNAs 也不会与 OTS 中的 aaRS、tRNA 以及 UAA 发生相互作用^[11]。

OTS 的正交性可以通过筛选来得到提高。最初人们发现, 原核、真核生物的翻译系统难以识别古菌的 aaRS/tRNA 对, 因此, 古菌中琥珀密码子的 aaRS/tRNA 对就为 OTS 提供了初始水平的正交性^[12]。随后, 通过对 tRNA 进行突变改造以及连续轮次的正 (负) 筛选来提高 OTS 正交性^[11]。最后, 对 aaRS 的底物结合域

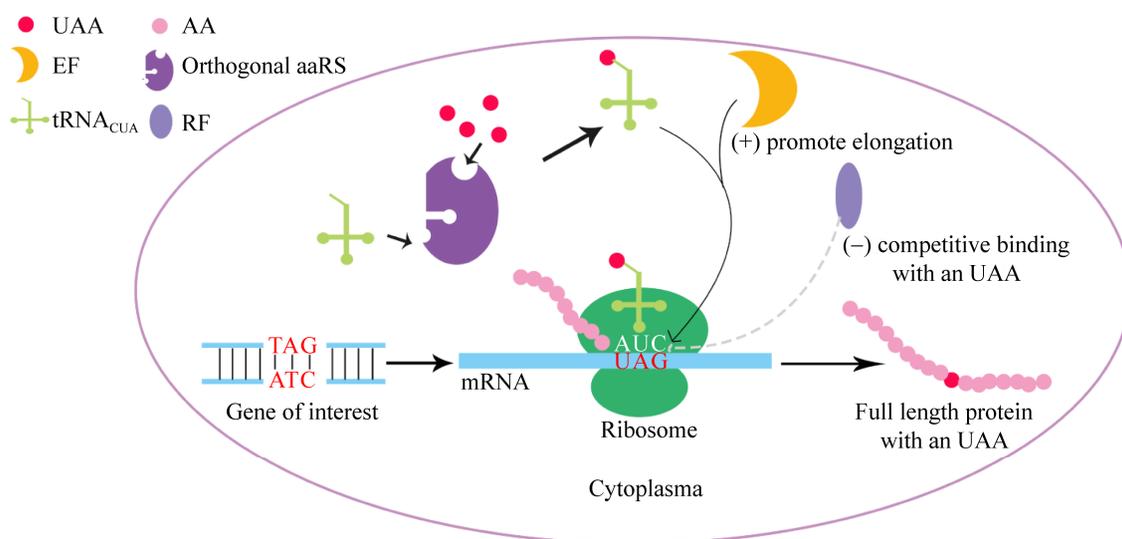


图 1 遗传密码扩充技术原理

Figure 1 The principle of genetic code expansion. UAA: unnatural amino acid. EF: elongation factor. RF: release factor.

进行定点突变,并选择能够与目的 UAA 良好结合的突变体^[11]。通过上述工作将会得到高效选择的 OTS,保证 UAA 与正交 tRNA 的结合。

高效表达掺入 UAA 的蛋白依赖于优化的 3 个核心要素,其中优化措施包括改进编码 OTS 的表达载体、增强 aaRS 的工作效率、优化正交延伸因子 (elongation factor, EF)、降低释放因子 (release factor, RF) 1 的影响等^[13]。

除去上述优化条件,影响蛋白表达的实验变量还包括:(1) 选择的蛋白质^[14];(2) 翻译蛋白所用的细胞类型^[15];(3) 胞浆中 UAA 的浓度^[15-16];(4) 在 mRNA 内引入琥珀密码子位置的上下游序列^[17-18]。

当前,尚未充分了解这些实验变量对于蛋白表达效率影响的程度,但一些研究结果表明,高表达水平的正交 tRNA 与 aaRS 的胞浆定位、胞内高浓度的 UAA 能够促进蛋白的生产^[14]。

2 GCE 技术在蛋白质功能研究中的应用

2.1 蛋白功能研究

研究蛋白质的功能及作用机制,往往需要精细化到某一功能区甚至是某一残基。但很多蛋白分子的某些功能区或小分子蛋白对蛋白修饰较为敏感,不易耐受外源标记对其的改变,因而研究其功能遇到极大障碍。

近年来,研究者们应用 GCE 技术,把 UAA 引入到蛋白质的不同位点,将对蛋白分子的干扰降到最小,同时实现了对其特定功能位点的标记,创造出更多的蛋白质变体,为研究蛋白的功能及作用机制提供了更为直观有利的方法。Lacoursiere 等在研究泛素功能中,利用 GCE 技术构建了所有可能的乙酰化泛素 (acetylated ubiquitin, acUb) 变体^[19]。在将这些

变体用于泛素化实验时,发现了 acUb 能与 E3 连接酶结合,并证明这些用 GCE 技术得到的 acUb 变体能用于发现、识别潜在泛素化模式的研究^[19]。这项研究展示了 GCE 技术在蛋白质改造中的灵活应用,即通过制备出所有可能的 acUb 变体分析其下游发生的反应,进而进行对比研究,发现潜在的泛素化模式。

不过,作者是在已知泛素仅有 7 个氨基酸残基会发生乙酰化的前提下,构建了所有 acUb 变体。若在乙酰化的残基数未知或者残基数过多的情况下,构建变体并逐一研究下游反应的工作量和难度将大大增加。因而,类似研究方法只适于需要探讨的蛋白分子功能位点及可能的变体较少的情况。通过引入特定 UAA 到蛋白的单一残基研究某些位点或区域的功能,是一种新的、精准研究手段,特别对不易耐受外源标记的蛋白分子有更重要的意义。

2.2 蛋白功能改造

目前,采用化学修饰手段可以灵活改造 UAA 的侧链基团,使其具备不同的功能。因此,通过 GCE 技术引入研究所需特定 UAA,对一些蛋白进行改造,可以赋予其新的功能及应用。已知细胞内诸如内体、溶酶体等细胞器内部的环境是酸性的^[20];而大多数荧光蛋白以及由荧光蛋白衍生出的 pH 传感器却仅在中性或微碱性环境下发出荧光,在酸性环境中荧光会淬灭^[21]。因此,Wang 等将含有喹啉基团的一种非天然氨基酸 Qui 掺入到增强型绿色荧光蛋白 (enhanced green fluorescent protein, EGFP) 的生色团中,研制出能在酸性微环境下产生明亮荧光的酸增亮荧光蛋白 (acid-brightening fluorescent protein, abFP)^[22]。当 abFP 暴露于酸性环境时,Qui 发生质子化产生阳离子生色团增强荧光,在生理 pH 值或高于生理 pH 值条件下则几乎不会发出荧光^[22]。同时,他们还建立了在大肠杆菌和哺

乳动物细胞中表达 abFP 的方法, 并成功应用在 δ 阿片受体-abFP 融合蛋白在哺乳细胞的内吞作用研究中, 证明了 abFP 在酸性环境中应用的可行性。abFP 的出现打破了荧光蛋白酸性环境荧光成像的应用局限。

通过对荧光蛋白关键生色团进行改造, 改变了荧光生成方式。这种策略同样适用于其他荧光蛋白, 通过掺入不同 UAA 赋予荧光蛋白新的颜色或功能。增加荧光蛋白的多样性, 也为荧光成像提供了更多的选择, 例如对多种蛋白同时成像等。最近, Roman-Arocho 等用特定 UAA 替代荧光蛋白生色团中的 Tyr66 残基, 研制出基于荧光蛋白的传感器, 进一步拓展了该蛋白的功能及应用^[23]。

通过 GCE 技术改造蛋白质, 本质上是通过改造 UAA 的理化性质来实现功能的改变。迄今, 已有 150 多种 UAA 可被掺入到蛋白质中^[9]。对这些 UAA 的侧链基团进行改造, 获得具备不同功能的 UAA 并应用于蛋白新功能的开发, 已经显露出在蛋白类药物研制、蛋白质疫苗设计、探针设计、药物递送等方面的应用前景^[11,24]。

3 GCE 技术在蛋白质标记成像中的应用

基于 GCE 技术的标记方法已实现了对各种活细胞内成分 (包括溶酶体、内质网、细胞膜等)、受体的标记和成像^[7]。其中, GCE 技术成功应用于无法使用蛋白质标签标记的小蛋白和不易耐受外源标记物的大分子蛋白的研究中。例如, Peng 和 Hang 首次报道了利用 GCE 技术对仅有 137 个氨基酸的干扰素诱导跨膜蛋白 3 (interferon-induced transmembrane protein 3, IFITM3) 标记成像^[25]; Sakin 等将 GCE 技术首

次应用于动态记录人免疫缺陷病毒 (human immunodeficiency virus, HIV) 包膜糖蛋白的迁移运动^[16]。

3.1 应用 GCE 技术标记成像的优势

传统的标记技术依赖于荧光蛋白或荧光抗体。通常, 荧光蛋白通过基因工程被直接引入到目的蛋白中, 因而在进行分子示踪时有高度特异性。但与荧光染料相比, 荧光蛋白分子量较大 (荧光染料约 0.5 nm, 荧光蛋白约 4.2 nm, 荧光抗体大于 10 nm) (图 2B)^[26]、光稳定性更差, 因此, 并不是适宜的标记探针。荧光染料可在较低的激光强度下进行长时成像, 且小分子荧光染料的定位测量精度更高、对蛋白质的结构及生理特性干扰更小; 但荧光染料要通过特定的复杂方法才能附着到蛋白分子上, 缺乏分子特异性。

GCE 标记方法结合了上述两类标记探针的优势, 弥补了不足。基于 GCE 技术的荧光标记依赖于 Tet 与 UAA 侧链环烯烃或环炔烃之间发生的点击反应, 将 Tet-荧光染料掺入到蛋白分子中实现荧光标记。

由于生物体内的 20 种天然氨基酸不具备 UAA 的侧链基团, 所以, GCE 标记系统中 Tet-荧光染料与 UAA 之间的点击反应保证了标记的更高特异性。此外, 与其他荧光标记系统相比, GCE 标记具有以下更为关键的优势: (1) 由于荧光染料比荧光蛋白小一个数量级, 因此对目的蛋白的结构与功能的干扰最小^[16,25]; (2) 荧光染料通常比荧光蛋白更亮, 更具光稳定性, 能够在较低的激光功率和较高的时间分辨率下进行活细胞成像^[3]; (3) 荧光染料种类多、可适用于包括近红外光在内的全光谱, 即能够在对细胞光毒性较小的长波长下成像, 比荧光蛋白在此波长下的光物理性质更好^[27]; (4) 对点击反应可以进行化学计量 (Tet-荧光染料以 1:1

的比例结合 UAA), 从而能够基于异源表达系统中的荧光强度对蛋白质进行定量 (图 2A)^[28]。以上优势使 GCE 标记突破了荧光蛋白标记的应用限制, 特别是在单分子及超分辨率显微镜成像等需要细胞长时间暴露在高激光功率下的应用中可能是划时代的。典型的实例是 Sakin 等构建的琥珀突变型 Env 并成功可视化了这一蛋白^[16]。琥珀突变型 Env 结构改变微小, 能够最大化接近天然状态的 Env。作者证明了琥珀突变后的 Env 在质膜定位、糖基化程度、假病毒组装、感染性和融合活性等方面与野生型有相同的特性与功能^[16]。进而, 作者用受激发射损耗 (stimulated emission depletion, STED) 显微技术成功可视化了琥珀突变型 Env, 发现了在病毒组装位点处所形成的环形 Env 簇。作者通过荧光漂白恢复 (fluorescence recovery after photo bleaching, FRAP) 技术测量出 Env 向细胞膜迁移的迁移速率; 同时他们发现, 与先前用免疫荧光方法所得结果相比, GCE 标记方法最

大限度地减少了由额外质量引起的迁移率测量中的伪影^[16]。

3.2 GCE 技术用于蛋白质标记成像

3.2.1 GCE 技术与超分辨率显微技术的联合应用

基于 GCE 技术的上述优势, 研究者们逐渐将其与超分辨率显微技术如结构光照明 (structured illumination microscopy, SIM)、单分子定位 (single molecule localization microscopy, SMLM) 和随机光学重建显微 (stochastic optical reconstruction microscopy, STORM)、STED 显微术等联合应用, 极大地提高了标记的荧光强度及分辨率^[29-31]。

Schwartz 等发现, 在 GCE 技术标记微管蛋白的 SIM 成像中, 可以分辨相隔距离小于 100 nm 的纤维^[32] (图 3A)。Serfling 等使用 GCE 结合单粒子跟踪技术及图像相关光谱法标记 G 蛋白偶联受体发现, GCE 标记与荧光蛋白标记产生的荧光信号强度呈线性相关^[28]; 而且, 采用单粒

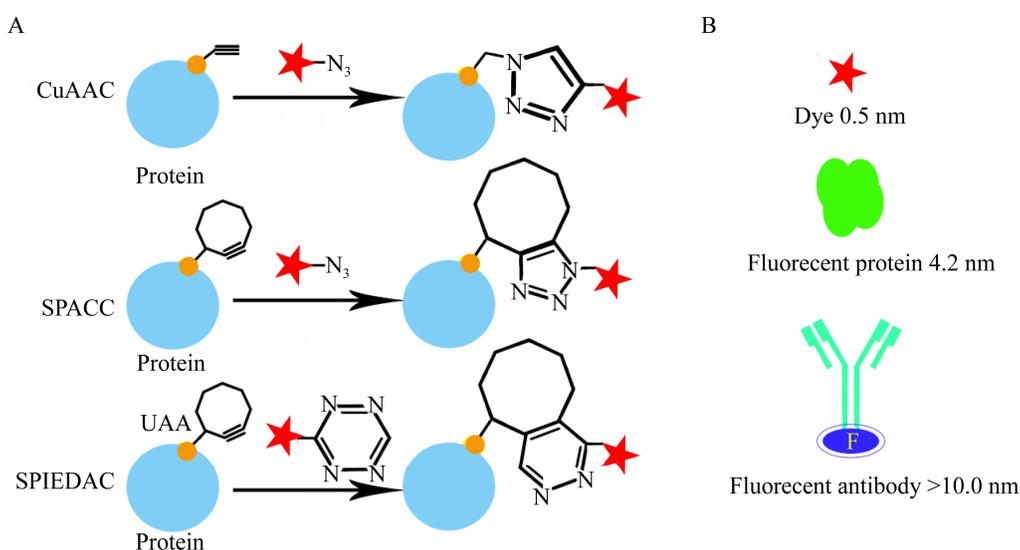


图 2 GCE 标记方法的特点 A: GCE 技术所用的点击反应类型; B: 荧光标记物大小

Figure 2 The characteristic of click chemistry GCE labeling. (A) The type of click reactions used in GCE. (B) The scale of fluorescent labeling.

子跟踪技术测量蛋白扩散率的实验结果显示, GCE 标记技术获得了比荧光蛋白标记更多且更长的蛋白扩散轨迹^[26]。Neubert 等把 UAA 引入 N-甲基-D-天冬氨酸受体 (N-methyl-D-aspartate receptor, NMDAR), 用 STORM 成像发现, GCE 标记产生的荧光密度明显高于荧光抗体标记, 证实了 GCE 标记用于蛋白成像的效果更佳^[33] (图 3B)。

上述研究结果充分证明, 联合 GCE 技术不但能够提高超分辨率显微成像的分辨率, 同时, 保证甚至增强了对蛋白的标记效果; 而 1:1 的化学计量比特点, 也为提高成像过程中的定位精度和分子计数奠定了良好基础。

3.2.2 GCE 技术与荧光共振能量转移系统的联合应用

荧光共振能量转移 (fluorescence resonance energy transfer, FRET) 是指两个荧光发色基团在足够靠近时, 供体分子吸收一定频率的光子后被激发到更高的电子能态, 在该电子回到基态前, 通过偶极子相互作用, 实现了能量向邻近的受体分子转移 (即发生能量共振转移)^[34]。其中, 发生能量转移的供体和受体组成了 FRET 对。通过 GCE 技术可将 UAA 掺入到 FRET 对中, 与荧光蛋白构成新的 FRET 系统。例如 Park 等构建了由 Hsp70-YFP 融合蛋白和掺入了 ANAP (一种荧光 UAA) 的 Bax 蛋白组成的 FRET

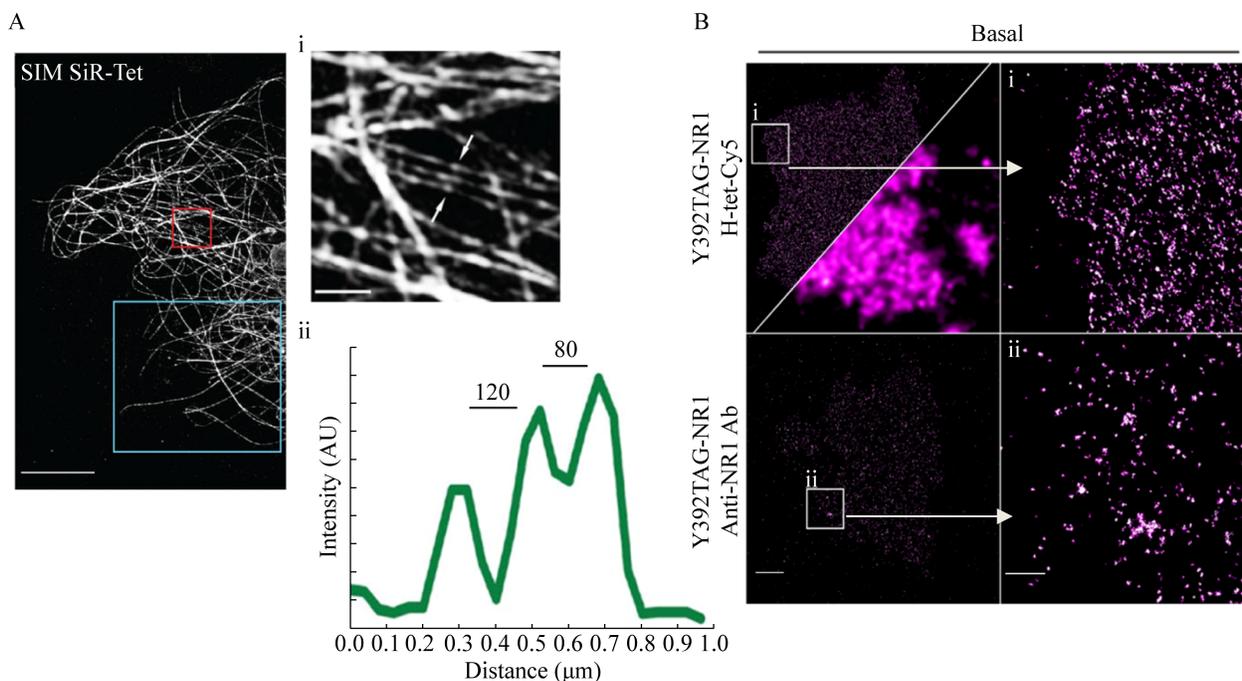


图 3 GCE 技术与超分辨率显微技术联合应用实例

Figure 3 Examples of the application of GCE technology and super-resolution microscopy. (A) SIM imaging of UAA-tubulin labeled with SiR-Tet. Scale bar, $10 \mu\text{m}$ ^[32]. i) Enlarged image of the red rectangle in the SIM image. ii) Intensity profile drawn between the arrows in enlarged image i. Scale bar= $1 \mu\text{m}$. (B) STORM image of UAA-NMDAR in basal regions of the cell^[33]. Bioorthogonal click-labeling with H-tet-Cy5 in NMDAR (upper panel). Antibody labeling with antibody anti-NR1 and Alexa Fluor 647 as secondary antibody was performed (lower panel). The lower triangle in the upper left panel shows a diffraction-limited wide-field image. Scale bar: $2.5 \mu\text{m}$ (overview images) and $0.5 \mu\text{m}$ (enlarged images i and ii). NR1: NMDAR subunit.

对, 探究 Bax 在凋亡过程中从与 Hsp70 的复合物中解离的因素^[35]。FRET 系统可以实时监测蛋白质的运动和相互作用, 但很少有 FRET 对不破坏蛋白质折叠和功能。因此, Jones 等又提出利用 GCE 技术构建分子量最小的 FRET 对^[36]。他们分别将甲氧基香豆素马来酰亚胺 (methoxycoumarin maleimide, Mcm-Mal) 和吖啶基丙氨酸 (acridonylalanine, Acd) (一种荧光 UAA) 掺入到蛋白质中, 构建半胱氨酸/Acd 双突变蛋白组成 FRET 对。其中, Acd 作为受体, Mcm-Mal 作为供体通过与半胱氨酸反应发出荧光。这种标记策略提供了一个构建简单方便的 FRET 对, 且能够实现能量转移的距离范围达到 1.5–4.0 nm, 是监测大多数分子运动的理想选择^[36]。此外, Mcm/Acd 还可与色氨酸结合 (色氨酸具有荧光特性), 通过三色荧光监测多种蛋白质的分子运动及分子间的相互作用^[36]。

常规 FRET 技术用荧光蛋白、荧光素等荧光物质作为探针, 往往容易干扰蛋白结构、影响功能; 同时, 过大的标记分子且离蛋白主链距离相对远, 也为 FRET 测量增加了不确定性^[36]。而 Mcm-Mal 和 Acd 分子与天然氨基酸大小基本一致, 避免了对目的蛋白的干扰作用, 同时, Mcm 吸收光谱峰值与 Acd 吸收光谱最小值重合, 避免了 Acd 的直接激发, 进一步提高了 FRET 的精准测量。因此, Mcm/Acd 对是一个较为适宜的 FRET 对。该构建策略体现了 GCE 技术应用于 FRET 系统中的潜力。通过改造 UAA 侧链基团, 使其具备荧光, 由此解决荧光团干扰蛋白结构或者测量不准确的问题。随着技术的不断优化改进, UAA 将会在 FRET 系统中得到广泛应用, 成为研究蛋白质胞内动态变化、构象改变、亲和力分析, 以及与其他蛋白质相互作用的机制中极为有力的研究手段与工具。

3.2.3 GCE 双色标记策略

除了用单个 UAA 进行点击标记的标记方法外, 近年研究提出了在不同实验需求下的 GCE 双色标记策略, 可以实现同时标记同一细胞内的不同蛋白。

2018 年, Saal 等首先报道了双色标记系统。作者分别将两种 UAA 整合到两个细胞中的不同目的蛋白上, 然后用灭活的乙型副流感病毒将细胞融合, 实现融合细胞的双色 STED 成像^[29] (图 4A)。但由于该方法的成像仅限于在固定细胞以及融合细胞中的蛋白质对, 因此在细胞生物学中的适用性受到限制。继而, 有几个研究团队相继报道了单个细胞内双色标记的方法。他们采用两种荧光特性不同的 Tet-荧光染料对同一 UAA 进行竞争性标记, 构建单分子荧光共振能量转移 (single molecular fluorescence resonance energy transfer, smFRET) 成像系统并用不同的成像方式同时从单个细胞获取数据。目前, 该方法已被用于蛋白质寡聚化研究^[26,28,37]。

Meineke 等开创了基于两套相互正交的 GCE 表达及标记体系从而标记活细胞内同一蛋白不同位点的双色标记方法^[38] (图 4B)。作者优化出两组 aaRS/tRNA 正交对, 分别识别琥珀密码子和赭色密码子 (UAA), 依次编码轴向反式-环辛-2-烯-赖氨酸 (axial trans-cyclooct-2-ene-Llysine, TCO*K) 和 N-炔丙基-L-赖氨酸 (N-propargyl-L-lysine, ProK)。TCO*K 通过 SPIEDAC 点击反应被四嗪-荧光染料标记, ProK 通过 CuAAC 点击反应被叠氮基-荧光染料标记。

双无义密码子的成功抑制取决于两大条件: (1) 两类无义密码子各自所属的 OTS 具备正交性; (2) 两套 OTS 之间相互正交。其中, aaRS 负责与 UAA 特异性识别, tRNA 负责与无义密码子特异性识别, 而 aaRS 与 tRNA 之间

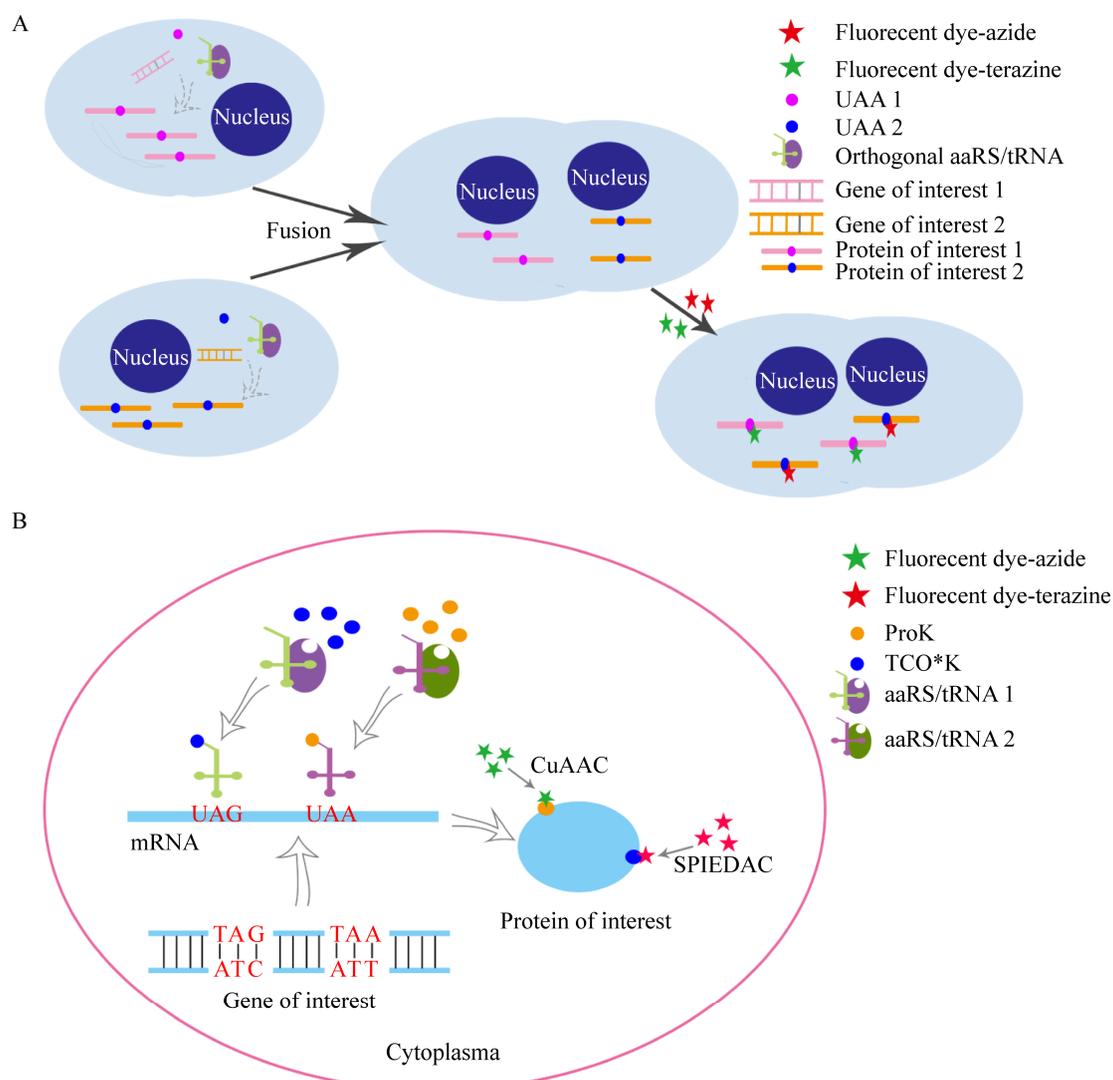


图4 GCE 双色标记原理 A: Saal K 团队建立的双色标记方法; B: Meineke B 团队建立的双色标记方法

Figure 4 The principle of GCE dual-color labeling. (A) The dual-color labeling method established by Saal K team. (B) The dual-color labeling method established by Meineke B team.

的催化同样也需要高度特异。由上述列举的种种条件可以看出, 双无义密码子的表达标记要求极为严苛。因而, 当前 GCE 双色标记策略往往选择用两种染料标记同一 UAA, 或者一种无义密码子掺入两种 UAA 再随机生成具有双重标记的分子^[37,39]。Meineke 等所创建的标记方法打破了双色标记的局限性, 展示了两套完全独

立的表达标记体系在同一蛋白表达中共用的可能^[38]。这为蛋白研究提供了新的视角与思路, 尤其是对蛋白构象改变、组装迁移或转运、免疫学研究等。但是, 双无义密码子抑制效率低下, 不易表达出掺入两种 UAA 的全长蛋白。而且, 细胞实验涉及的质粒、UAA 的添加、荧光染料的添加等因素较为复杂。因此, 该双色标

记方法应用难度较大。

3.3 利用 GCE 技术研发荧光探针工具

3.3.1 GCE 标签的构建

由于 GCE 技术对蛋白的成功标记依赖于 OTS 的正交性, 以及选择蛋白中合适的标记位点等因素。为了构造出既能保证表达量同时又不干扰蛋白功能的 UAA 蛋白, 往往需要对目的蛋白进行大量烦琐的突变位点筛选实验。因此, Segal 等着手设计一个 GCE 标签, 使用流感病毒血凝素 (hemagglutinin, HA) 表位作为骨架优化出含有 14 个残基的 N 末端标签, 将 UAA 掺入到标签 C 端^[7]。通过标记细胞内不同结构及细胞间隔, 包括细胞膜、溶酶体、过氧化物酶体和内质网等, 证明了这个 GCE 标签用于标记哺乳动物活细胞中蛋白质、细胞器的可行性。

该方法的最大优势在于省略了烦琐的位点筛选步骤, 可直接用 GCE 标签标记细胞蛋白进行活细胞成像, 节省了很多时间及经济成本。相较于其他标签抗体而言, GCE 标签能够进一步降低对蛋白结构的干扰。但 GCE 标签只能连接在蛋白质 N 端, 这降低了标记的灵活性, 限制了蛋白标记的选择范围。因此, GCE 标签对于示踪目的蛋白位置的意义更大, 对蛋白具体功能区的成像或研究而言并不是一个适宜的选择。

3.3.2 携带 UAA 的荧光蛋白传感器的构建

除了构建独立的 GCE 标签标记蛋白质外, 当前也有研究者利用将特定功能 UAA 直接引入荧光蛋白修饰其理化性质, 构建新型荧光探针来满足不同研究需求, 例如构建荧光蛋白传感器对生物体内生成的物质进行检测及追踪等^[40]。

Roman-Arocho 等提出了对荧光蛋白生色团的保护-去保护策略, 并设计了 UAA 荧光蛋白传感器^[23]。荧光蛋白的发光依赖于其蛋白内部的生色团 Ser65-Tyr66-Gly67, 其中 Tyr66 会在脱质子 (酚盐) 状态和质子化 (羟酚基) 状态

之间转化从而发出荧光^[41]。当 Tyr66 的羟基被封闭, 则无法发生状态转化, 继而不能形成关键的酚盐阴离子, 最终导致了荧光缺失。作者用 UAA 替代 Tyr66, 使酚盐阴离子缺失, 生色团无法发出荧光, 即为生色团的保护策略。然而, 在 UAA 与所要检测的离子之间发生特定反应后, 会形成酚盐阴离子, 生色团继而发出明亮荧光, 即为生色团的去保护策略^[23]。当前, 这类传感器设计策略已应用于如汞离子等小分子代谢物的检测与定量分析^[23,42-43]。

GCE 技术与荧光蛋白的结合, 为构建基因编码类型传感器检测活细胞中的小分子代谢物提供了有力的工具, 也扩展了荧光蛋白传感器在体内体外所能检测的分析物范围^[23]。上述荧光蛋白传感器的设计是基于氧化还原反应原理, UAA 替换了生色团的关键残基 Tyr66, 阻止了酚盐的形成。这改变了荧光蛋白的光谱性质, 导致荧光的消失。但当 UAA 与检测物发生反应时, 生色团会重新生成, 继而发出明亮的荧光^[23], 携带 UAA 的荧光蛋白传感器的构建, 在研究活细胞中离子转运和运输机制方面具有潜在的应用价值。但该传感器仍存在不足, UAA 掺入水平较低^[23], 这可能会导致在检测过程中真正发挥检测作用的荧光蛋白较少, 影响了检测的准确性。因此, 在荧光蛋白传感器的设计应用中, 还应大力优化 UAA 的掺入效率。

4 总结与展望

当前, 在蛋白质的构象、功能及作用机制研究、胞内标记成像等方面, GCE 技术已有应用, 该技术为生物分子功能及分子之间相互作用的深入研究提供了开创性的手段。GCE 技术与新的显微成像技术、定量技术联合应用, 为动态研究蛋白质功能、细胞成分的功能、设计荧光探针、研发疫苗等诸多领域提供了极具价

值的支撑。

理论上, UAA 可以被整合到蛋白质的任何位置, 荧光染料与 UAA 发生点击反应即可实现对蛋白质的标记。然而实际上, 在 GCE 技术及应用的每一个环节都还有大量的问题没有解决。其中, 由于 UAA 掺入蛋白效率低导致蛋白产量不高, 即改造蛋白在任何体系中的顺利合成是一个关键问题。针对这一问题, 未来应不断优化 OTS, 研制新的正交对, 提高 aaRS/tRNA 正交对在原核表达系统及真核表达系统中的正交性, 提高其与表达系统内释放因子竞争结合的能力。同时, 还需优化 UAA, 不断改进侧链基团的特性, 增强其与 tRNA 结合的能力等, 从根本上解决 UAA 掺入效率低下的问题。而且, 在标记实践中, 为保证蛋白加工折叠后掺入的 UAA 保持与荧光染料发生反应的能力, 仍需要不断筛选出新的、具有特定功能的 UAA。尽管仍存在局限与明显不足, 但随着技术的不断改进与完善, GCE 技术将会是蛋白质功能及活细胞成像研究的强大工具。

致谢: 感谢本课题组徐鑫钰同学协助绘制本文图片。

REFERENCES

- [1] Wen L, Fan ZC, Mikulski Z, et al. Imaging of the immune system—towards a subcellular and molecular understanding. *J Cell Sci*, 2020, 133(5): S234922. DOI: 10.1242/jcs.234922.
- [2] Specht EA, Braselmann E, Palmer AE. A critical and comparative review of fluorescent tools for live-cell imaging. *Annu Rev Physiol*, 2017, 79(1): 93-117.
- [3] Elia N. Using unnatural amino acids to selectively label proteins for cellular imaging: a cell biologist viewpoint. *Febs J*, 2021, 288(4): 1107-1117.
- [4] Brabham R, Fascione MA. Pyrrolysine amber stop-Codon suppression: development and applications. *Chembiochem*, 2017, 18(20): 1973-1983.
- [5] Nikić I, Kang JH, Girona GE, et al. Labeling proteins on live mammalian cells using click chemistry. *Nat Protoc*, 2015, 10(5): 780-791.
- [6] Yang Y, Lin SX, Lin W, et al. Ligand-assisted dual-site click labeling of EGFR on living cells. *Chem Bio Chem*, 2014, 15(12): 1738-1743.
- [7] Segal I, Nachmias D, Arbely E, et al. A straightforward approach for bioorthogonal labeling of proteins and organelles in live mammalian cells, using a short peptide tag. *bioRxiv*, 2019. DOI: 10.1101/708545.
- [8] Beliu G, Kurz AJ, Kuhlemann AC, et al. Bioorthogonal labeling with tetrazine-dyes for super-resolution microscopy. *Commun Biol*, 2019, 2: 261.
- [9] Young DD, Schultz PG. Playing with the molecules of life. *ACS Chem Biol*, 2018, 13(4): 854-870.
- [10] Niu W, Schultz P, Guo JT. An expanded genetic code in mammalian cells with a functional quadruplet *Codon*. *ACS Chem Biol*, 2013, 8(7): 1640-1645.
- [11] Fok JA, Mayer C. Genetic-code-expansion strategies for vaccine development. *Chembiochem*, 2020, 21(23): 3291-3300.
- [12] Liu CC, Schultz PG. Adding new chemistries to the genetic code. *Annu Rev Biochem*, 2010, 79(1): 413-444.
- [13] Smolskaya S, Andreev Y. Site-specific incorporation of unnatural amino acids into *Escherichia coli* recombinant protein: methodology development and recent achievement. *Biomolecules*, 2019, 9(7): 255.
- [14] Müller T, Sakin V, Müller B. A spotlight on viruses—application of click chemistry to visualize virus-cell interactions. *Molecules*, 2019, 24(3): 481.
- [15] Jakob L, Gust A, Grohmann D. Evaluation and optimisation of unnatural amino acid incorporation and bioorthogonal bioconjugation for site-specific fluorescent labelling of proteins expressed in mammalian cells. *Biochem Biophys Rep*, 2019, 17: 1-9.
- [16] Sakin V, Hanne J, Dunder J, et al. A versatile tool for live-cell imaging and super-resolution nanoscopy studies of HIV-1 env distribution and mobility. *Cell Chem Biol*, 2017, 24(5): 635-645.e5.
- [17] Pott M, Schmidt MJ, Summerer D. Evolved sequence contexts for highly efficient amber suppression with noncanonical amino acids. *ACS Chem Biol*, 2014, 9(12): 2815-2822.
- [18] Xu H, Wang Y, Lu J, et al. Re-exploration of the *Codon* context effect on amber *Codon*-guided incorporation of noncanonical amino acids in *Escherichia coli* by the blue-white screening assay. *Chembiochem*, 2016, 17(13): 1250-1256.
- [19] Lacoursiere RE, O'Donoghue P, Shaw GS. Programmed ubiquitin acetylation using genetic code

- expansion reveals altered ubiquitination patterns. *FEBS Lett*, 2020, 594(7): 1226-1234.
- [20] Casey JR, Grinstein S, Orlowski J. Sensors and regulators of intracellular pH. *Nat Rev Mol Cell Biol*, 2010, 11(1): 50-61.
- [21] Lin MZ, Schnitzer MJ. Genetically encoded indicators of neuronal activity. *Nat Neurosci*, 2016, 19(9): 1142-1153.
- [22] Wang N, Wang L. Acid-brightening fluorescent protein (abFP) for imaging acidic vesicles and organelles. *Methods Enzymol*, 2020, 639: 167-189.
- [23] Roman-Arocho GM, Shang X, Niu W, et al. Design of fluorescent protein-based sensors through a general protection-deprotection strategy. *Methods Enzymol*, 2020, 640: 63-82.
- [24] Li Q, Chen Q, Klauser PC, et al. Developing covalent protein drugs via proximity-enabled reactive therapeutics. *Cell*, 2020, 182(1): 85-97.e16.
- [25] Peng T, Hang HC. Site-specific bioorthogonal labeling for fluorescence imaging of intracellular proteins in living cells. *J Am Chem Soc*, 2016, 138(43): 14423-14433.
- [26] König AI, Sorkin R, Alon A, et al. Live cell single molecule tracking and localization microscopy of bioorthogonally labeled plasma membrane proteins. *bioRxiv*, 2020. DOI: 10.1101/660118.
- [27] Grimm JB, English BP, Choi H, et al. Bright photoactivatable fluorophores for single-molecule imaging. *Nat Methods*, 2016, 13(12): 985-988.
- [28] Serfling R, Seidel L, Bock A, et al. Quantitative single-residue bioorthogonal labeling of G protein-coupled receptors in live cells. *ACS Chem Biol*, 2019, 14(6): 1141-1149.
- [29] Saal KA, Richter F, Rehling P, et al. Combined use of unnatural amino acids enables dual-color super-resolution imaging of proteins via click chemistry. *ACS Nano*, 2018, 12(12): 12247-12254.
- [30] Uttamapinant C, Howe JD, Lang K, et al. Genetic code expansion enables live-cell and super-resolution imaging of site-specifically labeled cellular proteins. *J Am Chem Soc*, 2015, 137(14): 4602-4605.
- [31] Nikić I, Lemke EA. Genetic code expansion enabled site-specific dual-color protein labeling: superresolution microscopy and beyond. *Curr Opin Chem Biol*, 2015, 28: 164-173.
- [32] Schvartz T, Aloush N, Goliand I, et al. Direct fluorescent-dye labeling of α -tubulin in mammalian cells for live cell and superresolution imaging. *Mol Biol Cell*, 2017, 28(21): 2747-2756.
- [33] Neubert F, Beliu G, Terpitz U, et al. Bioorthogonal click chemistry enables site-specific fluorescence labeling of functional NMDA receptors for super-resolution imaging. *Angew Chem Int Ed Engl*, 2018, 57(50): 16364-16369.
- [34] Bajar BT, Wang ES, Zhang S, et al. A guide to fluorescent protein FRET pairs. *Sensors (Basel)*, 2016, 16(9): E1488.
- [35] Park SH, Ko W, Park SH, et al. Evaluation of the interaction between bax and Hsp70 in cells by using a FRET system consisting of a fluorescent amino acid and YFP as a FRET pair. *ChemBioChem*, 2020, 21(1/2): 59-63.
- [36] Jones CM, Venkatesh Y, Petersson EJ. Protein labeling for FRET with methoxycoumarin and acridonylalanine. *Methods Enzymol*, 2020, 639: 37-69.
- [37] Das DK, Govindan R, Nikić-Spiegel I, et al. Direct visualization of the conformational dynamics of single influenza hemagglutinin trimers. *Cell*, 2018, 174(4): 926-937.e12.
- [38] Meineke B, Heimgärtner J, Eirich J, et al. Site-specific incorporation of two ncAAs for two-color bioorthogonal labeling and crosslinking of proteins on live mammalian cells. *Cell Rep*, 2020, 31(12): 107811.
- [39] Lu M, Ma X, Castillo-Menendez LR, et al. Associating HIV-1 envelope glycoprotein structures with states on the virus observed by smFRET. *Nature*, 2019, 568(7752): 415-419.
- [40] Zhang Y, Du Y, Li M, et al. Activity-based genetically encoded fluorescent and luminescent probes for detecting formaldehyde in living cells. *Angew Chem Int Ed Engl*, 2020, 59(38): 16352-16356.
- [41] 储军, 施华, 杨杰, 等. 分子与细胞事件的光学可视化——解读 2008 年诺贝尔化学奖. *生物化学与生物物理进展*, 2008, 35(10): 1104-1111. Chu J, Shi H, Yang J, et al. Optical visualization of molecular and cellular events: to decode 2008 Nobel prize in chemistry. *Prog Biochem Biophys*, 2008, 35(10): 1104-1111 (in Chinese).
- [42] Ai HW. Fluorescent-protein-based probes: general principles and practices. *Anal Bioanal Chem*, 2015, 407(1): 9-15.
- [43] Liu X, Li J, Hu C, et al. Significant expansion of the fluorescent protein chromophore through the genetic incorporation of a metal-chelating unnatural amino acid. *Angew Chem Int Ed Engl*, 2013, 52(18): 4805-4809.

(本文责编 郝丽芳)