



遗传密码子扩展技术在蛋白质及多肽类药物中的应用*

高晓威 韦思平** 王 钦**

(西南医科大学药学院, 泸州 646000)

摘要 通过遗传密码子扩展技术位点特异性插入非天然氨基酸 (noncanonical amino acids, ncAAs) 可在原子水平上对蛋白质的结构与功能进行操控。目前该技术能够向包括高等动植物在内的各种生命体中插入 200 多种 ncAAs, 已被广泛应用于生物医药领域。凭借能够在蛋白质中定点引入可控生物正交化学官能团的独特优势, 该技术不仅可以用于蛋白质及多肽药物的研发, 提高蛋白质及多肽药物的质量与疗效, 而且可以为一些人类重大疾病的预防和治疗提供开创性解决方案。本文将重点关注遗传密码子扩展技术的前沿进展及其在各类抗体、细胞因子以及抗菌肽等蛋白质及多肽类药物中的应用, 同时也对其衍生的新型生物治疗手段进行简单阐述。

关键词 遗传密码子扩展技术, 非天然氨基酸, 抗体, 抗菌肽, 疫苗, 共价交联活性蛋白药物

中图分类号 Q51

DOI: 10.16476/j.pibb.2021.0332

蛋白质参与包括细胞骨架、生物催化、信号转导、转录调控、物质运输等在内的所有生物学过程, 其几乎全部由 20 种天然氨基酸 (canonical amino acids, cAAs) 组成^[1]。蛋白质功能的多样性由 20 种 cAAs 的排列组合及侧链间的相互作用共同决定。在自然界中通过对蛋白质进行翻译后修饰 (如甲基化、乙酰化、磷酸化、糖基化等), 可向蛋白质中引入新的化学基团进一步丰富蛋白质的功能; 同时, 一些蛋白质只有在辅酶存在或者结合金属离子的情况下才能发挥正常的生物学功能^[2], 这表明 20 种 cAAs 提供的有限化学基团在某些情况下并不能够完全满足蛋白质功能的需求, 而引入新的化学基团有望丰富蛋白质的物理化学性质, 产生新的生物学功能。

非天然氨基酸 (noncanonical amino acids, ncAAs) 种类繁多, 其结构与性质的多样性远超 cAAs。通过插入 ncAAs 来提高蛋白质或者多肽的生物化学特性已见诸许多报道^[3]。鉴于 ncAAs 在蛋白质研究和应用中的巨大潜力, 研究人员目前已开发出多种向蛋白质或多肽中插入 ncAAs 的技术。其中固相肽合成 (solid-phase peptide synthesis) 可

用于合成不超过 50 个氨基酸残基的含 ncAAs 多肽; 通过内含肽 (intein) 技术将合成的短肽共价相连可获得较大分子质量的含 ncAAs 蛋白质^[4]。无细胞翻译系统 (cell-free translation systems) 亦可用于合成含 ncAAs 的蛋白质, 该技术的关键在于首先要在体外制备能被核糖体识别的装载有 ncAAs 的氨酰 tRNA^[5]。研究表明, 有些与 cAAs 结构类似的 ncAAs 能够在胞内被相应的氨酰 tRNA 合成酶 (aminoacyl tRNA synthetase, aaRS) 识别并装载到特定 tRNA 上^[6]。在某种 cAA 营养缺陷型菌株的培养基中加入该 cAA 的结构类似物 (ncAA), 菌体内的翻译系统可将其错误地插入到整个蛋白质组中^[7]。该方法有可能扰乱菌体内一些关键蛋白质的功能, 存在细胞毒性, 降低目标蛋白产量。上述方法尽管均可有效制备含 ncAAs 的蛋白质或多肽,

* 中国博士后科学基金 (2020M683364) 和四川省科技计划项目 (2020YJ0129, 2019JDTD0016) 资助。

** 通讯联系人。

韦思平 Tel: 0830-3162291, E-mail: swei1225@swmu.edu.cn

王钦 Tel: 0830-3162291, E-mail: wq_ring@hotmail.com

收稿日期: 2021-10-28, 接受日期: 2021-11-29

但都具有各自的局限性,比如非位点特异性、不可持续且生产成本高、只能应用于cAAs结构类似物等。遗传密码子扩展技术(genetic code expansion)能够有效克服上述局限性,实现在生物体内向蛋白质或多肽中位点特异性插入各种ncAAs^[2]。

遗传密码子扩展技术首次由美国Scripps研究所Peter G Schultz教授领导的研究组建立起来^[8]。研究人员目前已可利用该技术将200多种含有不同功能侧链基团的ncAAs位点特异性地插入到各类蛋白质中^[2]。遗传密码子扩展技术已被广泛应用于生物医药、蛋白质工程、合成生物学、结构生物学等诸多生物学领域,为一些重大生命科学问题的研究与解决提供了极大便利。本文将重点关注遗传密码子扩展技术的前沿进展及其在各类抗体、细胞因子以及抗菌肽等蛋白质及多肽类药物中的应用,同时也对其衍生的新型生物治疗手段进行简单阐述。

1 遗传密码子扩展技术

1.1 原理

遗传密码子扩展技术是指通过在生物体内引入外源正交aaRS/tRNA对,借助特殊密码子(如UAG、UAGA等)向目的蛋白中位点特异性插入ncAAs(图1)^[8]。目前应用较为广泛的外源aaRS/tRNA对包括来源于古生菌詹氏甲烷球菌

(*Methanococcus jannaschii*)的酪氨酰aaRS/tRNA对(*MjTyrRS/tRNA*)、巴氏甲烷八叠球菌(*Methanosarcina barkeri*)的吡咯赖氨酰aaRS/tRNA对(*MbPylRS/tRNA*),以及来源于大肠杆菌(*E. coli*)的酪氨酰aaRS/tRNA对(*EcTyrRS/tRNA*)和亮氨酰aaRS/tRNA对(*EcLeuRS/tRNA*)(表1)^[9]。其中*MjTyrRS/tRNA*对可在原核生物中应用,*EcTyrRS/tRNA*对和*EcLeuRS/tRNA*对可在真核生物中应用,而*MbPylRS/tRNA*对在原核生物和真核生物中均可应用。以野生型外源aaRS/tRNA对为基础,通过构建突变体库并在宿主细胞中进行基于特定ncAAs的正负筛选,获得与内源性aaRS/tRNA对和cAAs不存在交叉反应且能特异性识别ncAAs的突变体是遗传密码子扩展技术的关键所在^[8]。遗传密码子扩展技术首次在*E. coli*中获得成功,目前其应用已实现从单细胞到多细胞、从低等到高等物种的跨越^[10]。研究人员已经能够利用遗传密码子扩展技术在乳酸乳球菌、结核分支杆菌、枯草芽孢杆菌、酵母菌、委内瑞拉链霉菌、哺乳动物细胞、人造血干细胞、线虫、果蝇、斑马鱼、拟南芥、小鼠等物种中定点插入ncAAs(表1)。尽管遗传密码子扩展技术应用潜力巨大,但其自身仍有多方面的局限性需要突破才能实现大规模工业应用。

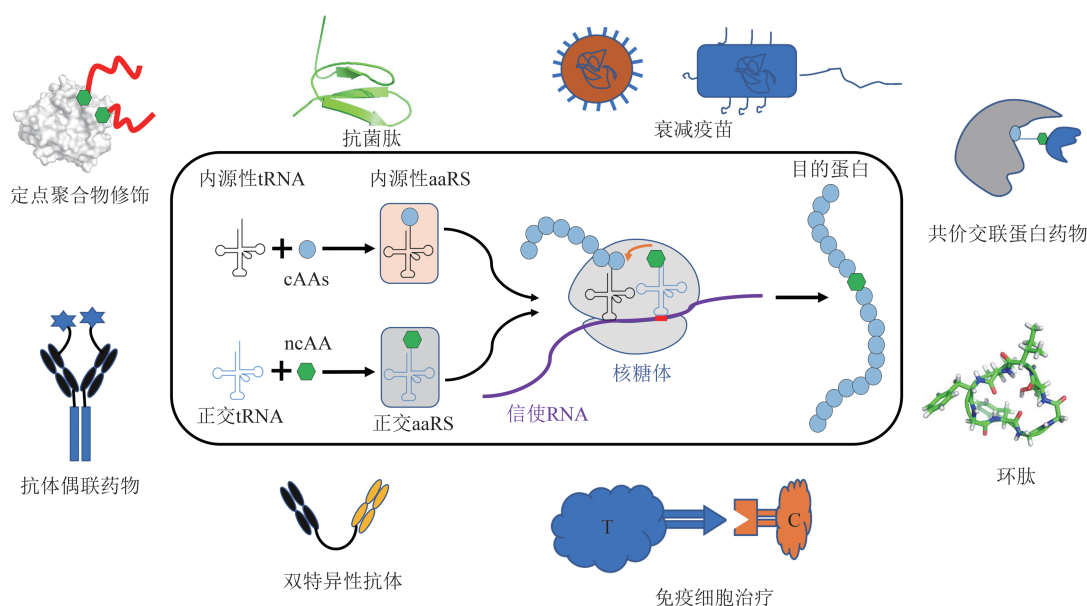


Fig. 1 Principles of genetic code expansion and its applications in protein and peptide related drugs

图1 遗传密码子扩展技术原理及其在蛋白质、多肽相关药物中的应用

Table 1 Orthogonal tRNA/synthetase pairs used in different hosts
表1 不同物种中使用的外源正交aaRS/tRNA对

宿主	正交aaRS/tRNA 对	来源	参考文献
原核细胞			
大肠杆菌	DhPylRS/tRNAPyl CUA	<i>Desulfitobacterium hafniense</i>	[11]
	MaTyrRS/tRNATyr CUA	<i>M. acetivorans</i>	[12]
	MbPylRS/tRNAPyl CUA	<i>M. barkeri</i>	[13]
	MjTyrRS/tRNATyr CUA	<i>M. jannaschii</i>	[8]
	MmPylRS/tRNAPyl CUA	<i>M. mazei</i>	[14]
	MtLeuRS/ HstRNA ^{Leu} CUA/UCCU/UCA	<i>Methanobacterium thermoautotrophicum</i> / <i>Halobacterium</i> sp. NRC-1	[15]
	PhProRS/ AftRNA ^{Pro} CUA/AGGG/CUAG	<i>Pyrococcus horikoshii</i> / <i>Archaeoglobus fulgidus</i>	[16]
	PhGluRS/_tRNA ^{Glu} CUA	<i>P. horikoshii</i> / Consensus tRNA from archaeal tRNA ^{Glu} sequences	[17]
	PhLysRS/_tRNA ^{Lys} UCCU	<i>P. horikoshii</i> /Consensus tRNA from archaeal tRNA ^{Lys} sequences	[18]
	ScGlnRS/tRNAGln CUA	<i>Saccharomyces cerevisiae</i>	[19]
	ScTyrRS/tRNATyr CUA	<i>S. cerevisiae</i>	[20]
	ScTyrRS/EctRNA ^{fMet} 2	<i>S. cerevisiae</i> / <i>E. coli</i>	[21]
	ScAspRS/tRNA ^{Asp} CUA	<i>S. cerevisiae</i>	[22]
	ScTrpRS/tRNA ^{Trp} CUA	<i>S. cerevisiae</i>	[23]
	MjTyrRS/tRNATyr CUA	<i>M. jannaschii</i>	[24]
	MbPylRS/ MmtRNA ^{Pyl} CUA	<i>M. barkeri</i> / <i>M. mazei</i>	[25]
	MbPylRS/tRNAPyl CUA	<i>M. barkeri</i>	[26]
	MbPylRS/tRNAPyl CUA	<i>M. barkeri</i>	[26]
委内瑞拉链霉菌	MjTyrRS/tRNATyr CUA	<i>M. jannaschii</i>	[27]
真核细胞			
酵母	EcGlnRS/人启动因子tRNA	<i>E. coli</i> / <i>Human</i>	[21]
	EcTyrRS/tRNATyr CUA	<i>E. coli</i>	[28]
	EcLeuRS/tRNA ^{Leu} CUA	<i>E. coli</i>	[29]
	MbPylRS/tRNAPyl CUA	<i>M. barkeri</i>	[30]
哺乳动物细胞	BsTrpRS/tRNA ^{Trp} CUA	<i>B. subtilis</i>	[31]
	EcTyrRS/tRNATyr CUA	<i>E. coli</i>	[32]
	EcGluRS/tRNAGlu CUA	<i>E. coli</i>	[33]
	EcTyrRS/BstRNA ^{Tyr} CUA	<i>E. coli</i> / <i>Bacillus stearothermophilus</i>	[34]
	EcLeuRS/tRNA ^{Leu} CUA	<i>E. coli</i>	[35]
	MmPylRS/tRNAPyl CUA	<i>M. mazei</i>	[36]
秀丽隐杆线虫	MmPylRS/tRNAPyl CUA	<i>M. mazei</i>	[37]
斑马鱼胚胎	MbPylRS/tRNAPyl CUA	<i>M. barkeri</i>	[38]
拟南芥	MbPylRS/tRNAPyl CUA	<i>Methanosarcina barkeri</i>	[39]
果蝇	MmPylRS/tRNAPyl CUA	<i>M. mazei</i>	[40]

1.2 提高ncAAs整合效率

目前在遗传密码子扩展技术中使用频率最高的密码子为琥珀密码子 (UAG), 该密码子在蛋白质翻译过程中被释放因子 1 (release factor 1, RF1) 所识别, 介导新生多肽链的翻译终止^[41]。正交 aaRS/tRNA 对需要在蛋白质翻译的过程中与 RF1 竞争性识别 mRNA 上的 UAG 密码子才能获得含

ncAAs 的全长蛋白质, 因此含 ncAAs 蛋白质的表达量一般只有野生型的 10%~20%^[10]。为了解决这一问题, 研究人员尝试了采用不同的方法敲除 *RF1* 基因, 避免 RF1 对 UAG 密码子识别导致翻译提前终止。通过将 *E. coli* 中 7 个必需基因 (*coaD*、*had*、*hemA*、*mreC*、*murF*、*lola* 和 *lpxK*) 的终止密码子 UAG 替换为 UAA, Sakamoto 课题组^[42] 成功敲除

了 *RF1* 基因。通过将框内 UAG 自调节元件 (in-frame UAG autoregulation element) 失活并将 *RF2* 中的 T246 突变为 A, Wang 课题组^[43] 同样也成功敲除了 *RF1* 基因。利用 *RF1* 敲除菌株作为表达宿主, 含 ncAAs 蛋白质的产量, 尤其是多位点插入 ncAAs 蛋白质的产量, 得到了显著提高。但值得注意的是, 这两种 *RF1* 敲除菌株的生长速率都明显下降。利用多元自动化基因组编辑技术 (multiplex automated genome engineering) 和基因组接合组装技术 (conjunctive assembly genome engineering), Isaacs 课题组^[44] 将 *E. coli* MG1655 中所有已知 UAG 密码子全部替换成了 UAA 并将 *RF1* 基因进行了敲除, 成功构建了菌株 *E. coli* C321.ΔA。在菌株 *E. coli* C321.ΔA 中, UAG 已不再作为终止密码子发挥功能而是成为了一个空白密码子。以 *E. coli* C321.ΔA 作为表达宿主, 可将含 ncAAs 蛋白质的产量提高至与野生型相当。我们近期从转录组和蛋白质组学的角度分析了 UAG 密码子重排对 *E. coli* 生理特性的影响, 并通过向基因组中整合 T7 RNA 聚合酶系统进一步提高了 *E. coli* C321.ΔA 菌株含 ncAAs 蛋白质的表达能力^[45]。

由于遗传密码子扩展技术使用的正交 aaRS/tRNA 对均为外源性的, 其在宿主细胞中的活性往往偏低。通过增加正交 aaRS/tRNA 对在细胞内的基因拷贝数和调整启动子强度, 提高正交 aaRS/tRNA 对的表达量, 可在一定程度上提高含 ncAAs 蛋白质的产量^[46]。Isaacs 课题组^[47] 在 *E. coli* 中利用多元自动化基因组编辑技术对正交 aaRSs 进行了定向进化, 显著提高了酶活性和特异性, 将含 ncAAs 蛋白质的产量提高了 25 倍。他们利用进化后的 aaRS/tRNA 对获得了最多含有 30 个 ncAAs 插入的蛋白质。Liu 课题组^[48] 利用噬菌体辅助的连续进化技术 (phage assisted continuous evolution) 分别对 PylRS 和 TyrRS 进行定向进化, 显著提高了酶的催化效率和底物特异性。类似地, Söll 课题组^[49] 利用噬菌体辅助的非连续进化技术 (phage assisted noncontinuous evolution) 对 PylRS 进行了定向改造, 获得了活性和特异性显著提高的突变体。通过不同策略提高 ncAAs 在宿主细胞内的浓度 (如过表达宿主细胞的氨基酸转运酶、将 ncAAs 以二肽的形式添加至培养基等), 亦可增加含 ncAAs 蛋白质的产量^[50]。氨酰化的 ncAA-tRNA 需要延伸因子 (elongation factor thermo-unstable, Ef-tu) 辅助才能进入核糖体参与蛋白质的合成。然而 ncAA-

tRNA 并不是 Ef-tu 的天然底物, 这导致二者的识别和结合效率偏低。通过对 Ef-tu 进行工程化改造提高其对非天然底物的亲和力可有效提高含 ncAAs 蛋白质的产量^[51]。

1.3 同一多肽链中整合不同 ncAAs

遗传密码子扩展技术一般只能向蛋白质中插入单种 ncAA, 这主要是由两方面的原因造成的: a. 缺乏相互正交的外源 aaRS/tRNA 对; b. 缺乏介导不同 ncAAs 插入的密码子。前期研究表明, *MjTyrRS/tRNA* 对与 *PylRS/tRNA* 对相互正交, 可用于向同一多肽链中插入两种不同 ncAAs^[18]。Chatterjee 课题组^[52] 和 Chin 课题组^[53] 分别采用不同的方法获得了三重正交 aaRS/tRNA 对, 实现了向同一多肽链中插入 3 种不同 ncAAs。从头设计产生多重相互正交的 aaRS/tRNA 对亦获得了一定突破^[54]。最近, Lin 课题组^[55] 通过蛋白质工程将 *MbPylRS/tRNA* 对中赋予其正交性的结构元件转移至 *E. coli* HisRS/tRNA 对中, 获得了新的正交嵌合 aaRS/tRNA 对, 为正交 aaRS/tRNA 对的构建提供了新的思路。要想实现向同一多肽链中插入多种 ncAAs, 除需要多重相互正交的 aaRS/tRNA 对之外, 还需要更多介导 ncAAs 插入的密码子。受菌株 *E. coli* C321.ΔA 的启发, 目前多个研究小组正通过不同的基因组工程技术将 *E. coli* 基因组中编码特定 cAAs 的密码子进行同义替换, 从而实现对 64 个三碱基密码子进行压缩和重排, 产生更多的空白密码子用于 ncAAs 的插入^[56-57]。截止目前, 在 *E. coli* 基因组中已有 7 个三碱基密码子实现了同义替换, 并被用于介导 ncAAs 的插入^[58]。通过设计四碱基密码子介导 ncAAs 的插入亦见诸多报道, 将四碱基密码子广泛应用于遗传密码子扩展技术可有效解决介导 ncAAs 插入密码子数量不足的问题, 然而宿主细胞中的翻译系统对四碱基密码子的识别效率较低且无法避免对四碱基密码子中前三个碱基的错误识别。Chin 课题组^[59] 通过对核糖体进行工程化改造获得了正交核糖体 Ribo-Q, 显著提高了含 ncAAs 蛋白的产量, 同时也为工程化改造获得有效解码四碱基密码子的正交核糖体提供了研究思路。通过向 *E. coli* 基因组中引入非天然碱基 (X, Y) 获得自然界中不存在的密码子, 并将其用于遗传密码子扩展技术介导 ncAAs 的插入同样也获得了成功^[60]。

1.4 原位合成 ncAAs

现阶段遗传密码子扩展技术主要依赖于人为外源添加 ncAAs 才能获得含相应 ncAAs 的蛋白质。

若某些 ncAAs 价格昂贵则势必会增加含 ncAAs 蛋白质的生产成本。通过蛋白质工程 (protein engineering) 或代谢通路工程 (metabolic pathway engineering) 改造宿主细胞原位合成 ncAAs, 并将其直接用于蛋白质合成可有效解决这一问题。Xiao 课题组^[61]通过引入来源于委内瑞拉链霉菌 (*Streptomyces venezuelae*) 的代谢通路实现了在 *E. coli* 中有效合成对氨基苯丙氨酸 (*p*-aminophenylalanine, pAF), 利用遗传密码子扩展技术将其插入到单链抗体中与荧光探针偶联可用于细胞成像。该课题组还通过引入来源于野油菜黄单胞菌 (*Xanthomonas campestris*) 的苯丙氨酸 4-羟基化酶 (phenylalanine 4-hydroxylase) 和辅酶再生系统成功实现了在 *E. coli* 中合成 5-羟基色氨酸 (5-hydroxytryptophan, 5-HTP)^[62]。Liu 课题组^[63]通过工程化改造 *E. coli* 中 Cys 合成路径并向培养基中补加各种廉价芳香硫醇, 实现了合成 50 种 ncAAs。同样地, L-磷酸苏氨酸 (O-phospho-L-threonine)、L-二羟基苯丙氨酸 (3,4-dihydroxy-L-phenylalanine) 和 S-烯丙基 L-半胱氨酸 (S-allyl-L-cysteine) 等 ncAAs 均实现了在 *E. coli* 中生物合成并直接用于遗传密码子扩展技术生产含 ncAAs 的蛋白质^[64-66]。随着越来越多的新颖酶分子以及代谢通路被发掘, 更多的 ncAAs 将能够被生物合成并直接用于遗传密码子扩展技术。随着遗传密码子扩展技术在上述各个环节的瓶颈问题得以突破, 该技术的应用也日趋多样化, 有望为下一代蛋白质及多肽类药物的开发提供技术支持。

2 在蛋白质及多肽类药物中的应用

2.1 抗体偶联药物

抗体是体内获得性免疫应答的重要组成部分, 对抗原具有优异的识别特异性和亲和力。将抗体与其他毒性小分子或者治疗性药物通过一定方式连接到一起, 制备成抗体偶联药物 (antibody-drug conjugates, ADCs), 可有效提高药物的靶向性, 实现向肿瘤等病灶组织精确递药, 降低药物毒副作用 (图 1)。目前有超过 175 种 ADCs 正在开展临床试验, 超过 10 种 ADCs 已成功上市^[67-68]。尽管 ADCs 在包括癌症等人类重大疾病的治疗中表现出了巨大的应用潜力, 但是其安全性和有效性一直受到制备技术的制约。传统 ADCs 主要依赖于抗体分子中的 Cys 或 Lys 残基与亲电子基团反应形成共价交联实现药物分子的装载。由于抗体分子中 Cys 和

Lys 残基数量较多, 往往无法与配体基团形成均一的化学修饰, 造成药物与抗体的偶联比率 (drug-to-antibody ratios, DAR) 和偶联位点不均一、不固定。通过还原抗体分子内部的二硫键暴露 Cys 残基进行化学修饰则会影响抗体分子的稳定性^[69]。因此传统方法制备的 ADCs 是不稳定的混合物, 其可能同时含有未装载药物的抗体以及过量装载药物的抗体。未装载药物的抗体会与装载药物的抗体竞争性结合目标抗原, 减弱 ADCs 的疗效, 而过量装载药物的抗体则会导致 ADCs 发生沉淀、毒性增加、血清稳定性降低等^[70]。传统方法制备的 ADCs 在体内的有效性、安全性、免疫原性、半衰期、亲和力和药物动力学等性质可能会受到不同生产批次的影响。实现抗体和药物定点偶联, 制备均一的 ADCs 有望解决这一问题。Mallet 课题组^[71]通过定点突变引入反应性 Cys 并减少抗体分子中非必需 Cys 残基的数量 (the engineered thio monoclonal antibody, THIOMAB) 首次制备出了均一的 ADCs。该 ADCs (THIOMAB antibody-drug conjugates) 与非均一 ADCs 相比显示出了更高的剂量耐受性和血清稳定性, 有效提高了药物的治疗窗浓度范围。然而抗体中引入的 Cys 会与其他含巯基物质或者抗体分子间形成二硫键从而影响偶联, 同时许多基于 Cys 形成的加合物在生理条件下并不稳定, 这限制了 THIOMAB 技术在 ADCs 中的应用。通过蛋白质标签 (protein tagging) 和酶催化等技术同样也实现了均一化 ADCs 的制备, 但考虑到抗体上偶联位点的选择对 ADCs 的有效性和安全性具有重要影响, 而这些方法对抗体上的修饰位点选择性有限, 可能会影响对 ADCs 的进一步优化^[72]。

利用遗传密码子扩展技术位点特异性插入 ncAAs 制备均一化 ADCs 具有以下明显优势: a. 遗传密码子扩展技术可实现在抗体分子中任意一个或者多个位点插入 ncAAs, 实现药物偶联位点的自由选择; b. 借助 ncAA 侧链上的生物正交化学官能团 (如叠氮基、炔基、酮基、卤代基等) 不仅可以建立起抗体与药物分子之间稳定、高效、高特异性的定点偶联, 而且还可以使偶联反应多样化, 允许更多不同化学结构的药物用于 ADCs 的制备 (图 2)^[73]。Schultz 研究组^[74]利用遗传密码子扩展技术向曲妥珠单抗 (trastuzumab) 中引入 L 乙酰基苯丙氨酸 (4-acetyl-L-phenylalanine, pAcF), 使其与烷氧基胺修饰的药物 auristatin F 形成肟键制备成功

了DAR为2的均一化ADCs。与非均一化ADCs相比,该ADCs在小鼠肿瘤移植模型中具有消除Her2阳性肿瘤的能力,并且展现出了偶联效率高、安全性好、血清清除率低等优点。随后该方法被成功用于制备治疗前列腺癌、肾癌以及癌症转移等疾病的均一化ADCs。Chin课题组^[75]利用遗传密码子扩展技术将含有环丙烯基团的Lys衍生物CypK插入到曲妥珠单抗中,随后通过Diels-Alder反应实现了与药物auristatin E的定点偶联,显著提高了偶联反应的速率,提高了均一化ADCs的产量。Sapra研究组^[76]将正交*EcTyrRS/tRNA*对基因整合到中国仓鼠卵巢癌细胞(Chinese hamster ovary, CHO)基因组中,构建了稳定的重组细胞系4E2,将含ncAAs抗体产量提升至1 g/L,为均一化ADC的大规模生产提供了极大便利。

凭借ncAAs结构与化学性质的多样性,遗传密码子扩展技术不仅可以实现ADCs的均一化制备,而且还可以极大丰富ADCs的种类与功能。随着众多含有不同生物正交化学官能团的ncAAs被用于遗传密码子扩展技术,越来越多在生理条件下稳定的偶联反应,如叠氮-炔环加成(azide-alkyne cycloaddition)、点击化学(click chemistry)等被用于制备均一化ADCs^[77]。通过遗传密码子扩展技术引入两种不同的ncAAs既可以在单个抗体分子中装载两种具有协同作用的药物提高ADCs的效力,又可以在单个抗体中实现同时装载药物分子和荧光探针用于ADCs在组织细胞中的荧光成像。Schultz研究组^[78]通过遗传密码子扩展技术分别将pAcF和叠氮化赖氨酸类似物(azido-lysine analog, pAzK)插入到anti-Her2 IgG抗体的重链和轻链上,成功制备了能同时偶联毒性药物分子和荧光探针的均一化ADCs。该课题组还通过遗传密码子扩展技术将pAcF位点特异性插入到anti-Her2 IgG抗体中与小干扰RNA(small interfering RNAs, siRNAs)偶联,成功制备出了能够沉默肿瘤细胞内基因表达的ADCs药物,并且他们发现偶联位点对该ADC分子的活性影响较大^[79]。Guo课题组^[80]利用遗传密码子扩展技术将含叠氮基团ncAA定点引入到利妥昔单抗(rituximab)中,通过与双功能臂DIBO-DOTA发生click反应成功制备出了有放射性核素⁶⁴Cu标记的ADCs,其具有肿瘤成像和放射性免疫治疗的应用前景。除了用于抗击肿瘤外,目前由遗传密码子扩展技术制备的均一化ADCs还被用于其他疾病的治疗。Schultz研究组^[81]通过遗传密码

子扩展技术将免疫抑制剂——达沙替尼(dasatinib)定点偶联到能够与造血干细胞(hematopoietic stem cells)表面抗原CXCR4高效结合的人源化抗体HLCX上,制备的ADCs能够有效避免达沙替尼靶向性不足而产生的毒副作用,可用于治疗T细胞活化引起的免疫紊乱。利用遗传密码子扩展技术将甲状腺激素受体激动剂与靶向肝脏的抗体进行定点偶联可用于非酒精性脂肪肝炎的治疗^[82]。此外,相关ADCs还被应用于治疗其他炎症及代谢性疾病^[2]。目前基于ncAAs偶联构建的ADCs数量约占临床试验ADCs的4%,我们认为凭借遗传密码子扩展技术的优势未来将有更多基于ncAAs的ADCs被应用^[67]。

2.2 双特异性抗体

双特异性抗体(bispecific antibody)是基因工程抗体的一种,具有两个不同的抗原结合位点,可发挥类似于免疫突触(immunological synapse)的功能。双特异性抗体能够同时结合靶细胞(通常为肿瘤细胞)和功能细胞(通常为T细胞),增强对靶细胞尤其是免疫原性较低的肿瘤细胞的杀伤作用(图2)。双特异性抗体结构独特、功能多样,已被用于病毒性传染病的酶联免疫检测、体内组织细胞成像、肿瘤免疫治疗以及穿透血脑屏障治疗神经退行性疾病等领域^[82]。目前大量双特异性抗体药物仍处于临床试验阶段,至少有3种双特异性抗体药物已获批上市。2009年由欧盟批准上市的双特异性抗体——卡妥索单抗(catumaxomab)可用于治疗EpCAM阳性肿瘤导致的恶性腹水^[83]。2014年由美国批准上市的博纳吐单抗(blinatumomab)是首个具有双特异性T细胞接合器(bispecific T cell engager)结构的双特异性抗体,可同时结合T细胞的CD3受体和淋巴瘤细胞的CD19受体,用于治疗前体B细胞急性淋巴细胞白血病^[84]。该抗体药物于2015年在欧盟批准上市。由罗氏公司(Roche)开发的双特异性抗体——艾美赛珠单抗(emicizumab)于2017年在美国批准上市,并于2018年获得FDA认证。该抗体药物是首个非凝血因子类血友病药物,可用于预防体内不含VIII因子抑制物的血友病频繁出血,为血友病的治疗提供了新的解决方案^[85]。虽然与普通抗体相比,双特异性抗体表现出了优异的灵敏度、特异性以及多靶点靶向能力,但其面临着表达量低、容易聚集、血清半衰期短、副作用强等缺点^[85]。

传统双特异性抗体的构建方法主要依赖于将两

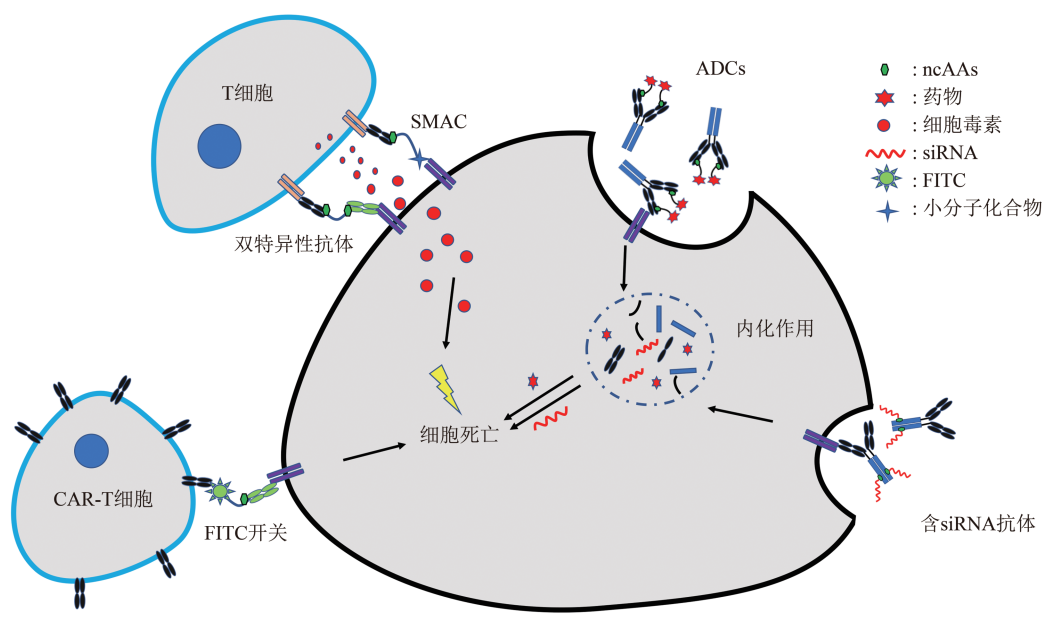


Fig. 2 Applications of genetic code expansion in antibody-related drugs
图2 遗传密码子扩展技术在抗体类相关药物中的应用

个抗原结合区的编码基因进行融合表达，然而这将限制二者的空间排列并可能对稳定性、有效性、选择性产生影响^[86]。利用遗传密码子扩展技术在两个抗原结合区中引入ncAAs，通过正交反应与不同长度的聚乙二醇（polyethylene glycol, PEG）连接臂相连可对二者在空间位置和距离上进行调控。Schultz 研究组^[87]利用遗传密码子扩展技术将pAcF 定点插入到 anti-CD3 和 anti-Her2 抗体的 Fab 区，随后与一端为氨基、另一端为叠氮或环辛烷基团的 PEG 连接臂通过两步反应进行共价连接，成功制备了 anti-HER2/anti-CD3 Fab 双特异性抗体。只需皮摩尔级浓度的 anti-HER2/anti-CD3 Fab 即可引导 T 细胞对 Her2⁺ 阳性肿瘤细胞进行选择性杀伤，该抗体同样对低表达 Her2⁺ 的肿瘤细胞显示出了较强的抑制作用。利用相同的策略，Schultz 研究组^[88]还成功制备出了双特异性抗体 anti-CLL1/anti-CD3，用于急性髓系白血病的治疗。Young 研究组^[89]向 anti-BCMA 和 anti-CD3 抗体 Fab 区中定点引入 pAcF 并利用 Diels-Alder 反应实现了二者之间的高效交联，成功制备的双特异性抗体 BiFab-BCMA，能够介导 T 细胞识别成熟 B 细胞抗原，特异性杀伤恶性多发性骨髓瘤细胞，为多发性骨髓瘤的耐药性问题提供了新的治疗手段。

鉴于 Watson-Crick 碱基配对具有高度特异性，利用寡核苷酸链作为连接臂进行双特异性抗体的快

速构建同样取得了成功。Smider 研究组^[90]利用遗传密码子扩展技术将 pAcF 位点特异性地插入到 anti-Her2 或 anti-CD20 抗体片段的 Fab 区并与一端经过氨基修饰的寡核苷酸链或者肽核苷酸链进行正交反应，随后与携带有 anti-CD3 Fab 区的反义寡核苷酸链或肽核苷酸链进行互补配对形成完整连接臂，成功制备的双特异性抗体分别用于乳腺癌和淋巴瘤的治疗。利用该策略可以快速制备双特异性抗体，甚至可以制备出异源多价特异性抗体用于提高抗体药物的有效性。

基于 ncAAs 的正交反应可以突破通过传统基因融合方式产生双特异性抗体的限制，将具有靶向作用的小分子与抗体片段 Fab 连接到一起，制备成均一的小分子/Fab 双特异性抗体（small-molecule-antibody conjugate, SMAC）（图2）。Zn²⁺ 依赖金属蛋白酶 PSMA 在前列腺癌细胞表面大量表达而在正常组织中的表达水平极低，是理想的前列腺癌分子标志物。DUPA（2-[3-(1,3-dicarboxy propyl)-ureido] pentanedioic）是 PSMA 的高效抑制剂，通过与 PSMA 结合抑制其酶活性^[91]。Schultz 研究组^[92]利用遗传密码子扩展技术将 pAcF 定点引入到 anti-CD3 抗体片段 Fab 中并将其与 DUPA 通过连接臂相连，制备成了 DUPA/anti-CD3 双特异性抗体。DUPA/anti-CD3 不仅在体外对前列腺癌细胞展现出了高效地选择性杀伤作用，而且在小鼠肿瘤移

植模型中也显现出了良好的安全性和延长的血清半衰期。叶酸受体 (folate receptor) 在许多炎症细胞和肿瘤细胞 (如卵巢癌和非小细胞肺癌等) 表面过度表达, 其对叶酸具有极强的亲和力。该研究组采用相同的策略成功制备了叶酸分子/anti-CD3 Fab 双特异性抗体并将其用于靶向叶酸受体高表达的癌症治疗, 同样获得了较好的实验结果^[93]。

2.3 基于蛋白质开关控制的细胞治疗

与传统的放化疗相比, 免疫细胞疗法具有靶向性好、毒副作用小等特点。嵌合抗原受体T细胞免疫疗法 (chimeric antigen receptor T-cell immunotherapy, CAR-T) 在癌症治疗中表现出了巨大潜力, 工程化T细胞表面的scFv抗体能够靶向识别体内肿瘤组织, 特异性杀伤肿瘤细胞。目前该疗法已被FDA批准用于急性淋巴细胞白血病的治疗^[94]。然而工程化T细胞在体内的不可控性可能导致严重后果, 包括引起细胞因子释放综合症 (cytokine release syndrome) 以及长效B细胞发育不全等^[95]。为了解决这一问题, Young 课题组^[96]和Cao课题组^[97]设计了一种基于遗传密码子扩展技术的蛋白质“开关分子”, 实现了在体内对工程化T细胞活性的调控。该“开关分子”由异硫氰酸荧光素 (fluorescein isothiocyanate, FITC) 和靶向肿瘤细胞的小抗体分子组成, 二者通过基于ncAAs的正交反应共价相连 (图2)。工程化T细胞在体内不再直接靶向肿瘤细胞, 而是特异性识别蛋白质“开关分子”中的FITC, 随后由蛋白质“开关分子”介导其对肿瘤细胞的识别和杀伤。该方法能够使人们通过控制蛋白质“开关分子”在体内的浓度来精细调控工程化T细胞的活性, 避免了CAR-T在体内的不可控性。同时只需对“开关分子”中的小分子抗体进行简单替换, 即可实现将工程化T细胞靶向不同肿瘤细胞, 避免了对工程化T细胞进行重复改造。

通过诱导剂控制治疗性蛋白质的表达可用于多种疾病的治疗。目前研究人员主要通过诱导剂控制转录水平的应答来实现对治疗性蛋白质表达的调控。这种调控方式响应速度慢, 并不能够完全满足对疾病治疗的需求。Liu 课题组^[98]近期开发出了一种基于遗传密码子扩展技术的治疗开关系统 (ncAA triggered therapeutic switch, NATS), 实现了在翻译水平上对治疗性蛋白质表达的控制, 并成功将其用于小鼠糖尿病的治疗。他们将遗传密码子扩展系统以及含有UAG终止密码子的胰岛素基因

同时整合进工程化细胞中, 随后将该工程化细胞移植到小鼠体内, 通过饲喂含有相应ncAA的饲料实现了对工程化细胞中胰岛素表达的精准调控, 有效降低了小鼠的血糖水平。

2.4 具有共价交联活性的蛋白质药物

与靶点存在共价交联机制的药物一般具有选择性高、药效持久、药代动力学性质优良等特点, 因此共价交联药物的开发一直备受关注。由于小分子化合物药物易于通过化学合成或修饰添加各种反应官能团, 一段时间以来共价交联药物的研发主要聚焦于一些小分子化合物。一般小分子共价交联药物会与靶点蛋白的口袋或活性中心特异性结合并与其其中的Lys或Cys残基侧链发生交联形成共价键^[99]。值得注意的是, 小分子共价交联药物并不适合用于干扰蛋白质与蛋白质之间的相互作用 (protein-protein interactions, PPIs), 这是因为PPIs往往发生在面积高达数百埃的界面上, 远超小分子共价交联药物的作用能力。PPIs与诸多生命过程密切相关, 包括信号转导、分化调控以及大分子合成等。PPIs紊乱会导致多种疾病的发生^[100]。以干扰PPIs为目标设计药物已被证实是治疗一些炎症疾病的有效手段。受抗原-抗体特异性相互作用启发, 研究人员认为以同为大分子的多肽或蛋白质为基础设计药物有望成为干扰PPIs, 治疗相关疾病的有效手段。

与小分子共价交联药物类似, 若蛋白质或多肽类药物能够与其靶点共价连接, 有望提高药物的选择性并增强药效。然而20种cAAs仅Cys能够在没有酶催化的情况下形成二硫键。二硫键在生理条件下易受还原势的影响, 不适合用于共价交联蛋白药物的开发^[101]。通过向蛋白质或多肽中引入含有正交反应官能团的ncAAs有望解决这一问题。Pellecchia 研究组^[102]利用固相肽合成将Lys-丙烯酸酰胺插入到靶向E3泛素连接酶Siah的短肽BI-107D1中, 发现具有共价交联活性的BI-107D1肽能够更好地抑制Siah的活性, 有望用于由Siah诱发的癌症的治疗。采用相同的策略, BIM肽被设计成为共价交联肽靶向致癌蛋白Bcl2A1, 表现出了极好的位点选择性^[103]。在人类癌症中肿瘤蛋白Mdm2和Mdm4会与肿瘤抑制蛋白p53发生相互作用导致其失活。Wang 课题组^[104]将含芳基磺酰氟的氨基酸 (aryl sulfonyl fluoride, Ar-SO₂F) 插入到靶向肿瘤蛋白Mdm2和Mdm4的短肽SAHp53-8中制成了共价交联肽, 成功干扰了p53与Mdm2/4之间的相互

作用。他们还发现该共价肽对Mdm4的选择性明显高于Mdm2, 而野生型SAHp53-8对二者的亲和力是相同的。随后又有靶向不同PPIs的共价交联肽被设计合成出来, 其优点也得到越来越多的验证。

由于固相肽合成只能用于合成相对较短的多肽, 这导致该方法不能够用于大分子共价交联蛋白药物的开发, 而遗传密码子扩展技术有望解决这一问题。为避免共价交联蛋白药物中ncAAs与胞内其他蛋白质组分发生无区别的交联反应干扰细胞功能, Wang课题组^[105-108]率先提出了邻近触发化学共价交联 (proximity-triggered chemical crosslinking) 的概念并合成了一系列新型ncAAs, 包括含卤代烷烃的EB3、含全氟苯的F-PSCaa、含氟乙酰胺的FAcK以及含芳基氨基甲酸酯的FPheK等。这些ncAAs在生理条件下不会与胞内低浓度游离的cAAs或蛋白质发生交联, 只有含ncAAs的共价交联蛋白药物与靶蛋白成功建立起稳定的相互作用, ncAA以适当方向接近其靶点cAA时才会激发共价交联反应的发生。该课题组最近利用遗传密码子扩展技术开发出了第一种共价交联蛋白药物用于癌症的治疗^[109]。人程序性死亡蛋白-1 (PD-1) 是位于T细胞膜表面可调控T细胞活性的转膜受体。PD-L1可与PD-1发生相互作用抑制T细胞增

殖活化, 导致肿瘤特异性T细胞衰竭和凋亡, 其往往在不同肿瘤细胞表面过度表达 (图3)。使用外源PD-1竞争性阻断体内T细胞上PD-1分子与肿瘤细胞上PD-L1分子间的相互作用可用于癌症免疫治疗。Wang课题组将氟磺酸-L-酪氨酸 (4-flourosulfate-L-tyrosine, FSY) 通过遗传密码子扩展技术定点引入到PD-1中制备成了可发生共价交联的蛋白药物PD-1 (FSY)。FSY可在PD-1 (FSY) 与PD-L1发生相互作用时与PD-L1上特定的His残基发生共价交联 (图3)。他们发现, 不管是体内还是体外, PD-1 (FSY) 都比野生型PD-1表现出了更好的抗肿瘤活性^[109]。利用蛋白水解靶向嵌合体 (proteolysis-targeting chimera methodology, PROTAC) 降解肿瘤源性蛋白可用于癌症的治疗。Chen课题组^[110]通过遗传密码子扩展技术向anti-PD-L1纳米抗体中引入FYS, 构建了能够与PD-L1形成共价交联蛋白药物GlueTAC。GlueTAC能够与肿瘤细胞表面的PD-L1形成共价交联, 促进PD-L1的内吞和在溶酶体中的降解, 有效缓解PD-L1造成的肿瘤细胞免疫逃逸。共价交联蛋白药物概念的提出以及遗传密码子扩展技术的应用将蛋白药物的设计与开发提升到了新的高度。

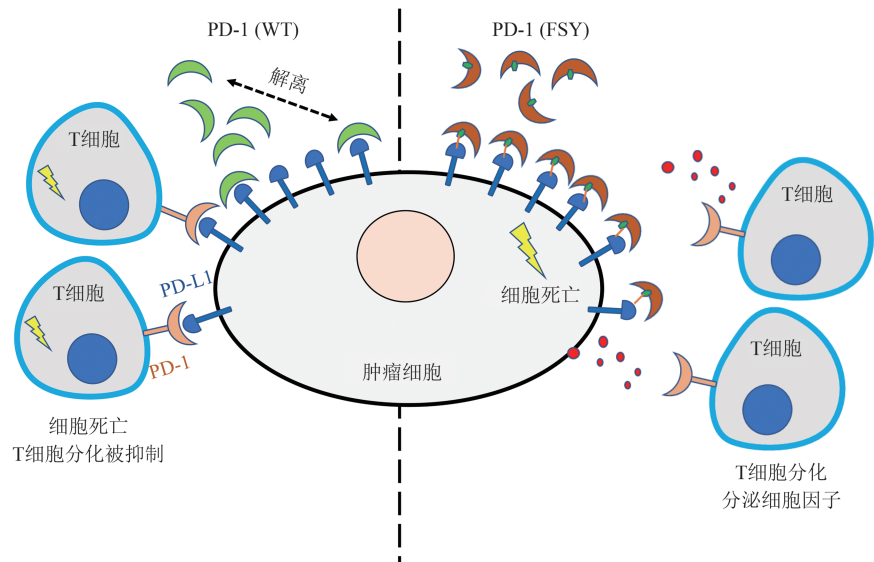


Fig. 3 The mechanisms of PD-1 (FSY) in cancer treatment
图3 共价蛋白药物PD-1 (FSY) 治疗肿瘤的原理示意图

2.5 治疗性蛋白疫苗

遗传密码子扩展技术为治疗性疫苗和预防性疫

苗的开发提供了开创性解决方案, 有望在疫苗领域引起一场新的技术革命。一般人体内免疫系统不会

对自身蛋白质产生免疫应答的现象称之为免疫耐受, 自身免疫耐受可有效防止自身免疫疾病的发生^[111]。然而在某些情况下打破自身免疫耐受, 引导免疫系统靶向自身蛋白质有助于包括癌症、慢性炎症以及骨质疏松等多种疾病的治疗^[112-113]。通过向自身蛋白质中引入新的抗原表位或者进行大量化学修饰可在一定程度上打破自身免疫耐受激活体内免疫应答, 但同时有可能对目标蛋白的折叠产生影响^[114]。机体在受到病毒感染以及发生炎症时会通过氧化应激反应引发自身某些蛋白质中的 Tyr 残基发生翻译后修饰形成硝基酪氨酸 (nitrotyrosine) 和磺基酪氨酸 (sulfotyrosine), 目前这一过程被证实与机体打破免疫耐受, 诱发自身免疫疾病相关。随后研究人员发现硝基和磺基在体内具有较强免疫原性, 是打破自身免疫耐受的关键所在。因此, 一些含硝基或者磺基的 ncAAs, 包括 *p*-nitro-L-phenylalanine (*p*NO₂F)、3-nitro-L-tyrosine (3-NO₂Y) 和 *p*-sulfo-L-phenylalanine (*p*SO₃F) 被认为可用于治疗性蛋白疫苗的开发 (图4)^[115-116]。

肿瘤坏死因子- α (TNF- α) 是在机体炎症发生过程中扮演重要角色的多功能细胞因子, 被认为与多种炎症疾病相关, 包括克罗恩病和类风湿性关节炎等。以 TNF- α 为靶点, 打破其自身免疫耐受, 诱导免疫系统产生中和 TNF- α 抗体被认为可用于多种炎症疾病的治疗。Smider 课题组^[117-118] 利用遗传密码子扩展技术将小鼠 mTNF- α 中第 11 位 Lys 或第 86 位 Tyr 残基定点突变成 *p*NO₂Phe 后, mTNF- α 产生了新的抗原决定簇, 有效打破了自身免疫耐受, 产生了能够同时识别突变体及野生型 mTNF- α 的高效价抗体。他们发现 *p*NO₂Phe 在 mTNF- α 中的突变位置对能否打破自身免疫耐受至关重要。为了明确 *p*NO₂Phe⁸⁶-mTNF- α 的作用, 他们分别将接种了野生型 mTNF- α 和突变体 *p*NO₂Phe⁸⁶-mTNF- α 的小鼠进行脂多糖 (LPS) 处理, 发现免疫接种 *p*NO₂Phe⁸⁶-mTNF- α 后可有效降低小鼠的死亡率。RANKL 是 TNF 家族中的细胞因子, 其与能够分解骨骼组织的噬骨细胞激活有关, 是治疗骨质疏松症和类风湿性关节炎等疾病的潜在靶点^[119]。Tao 课题组^[120] 通过遗传密码子扩展技术将 RANKL 中第 234 位和第 240 位的 Tyr 残基突变为 *p*NO₂Phe 后, 有效打破了其自身免疫耐受, 产生了能够同时特异性识别野生型和突变体 RANKL 且在体内滞留时间长的高效价抗体。他们随后还明确了免疫接种 *p*NO₂Phe²³⁴-mRANKL 后, 小鼠的骨质疏松和类风

湿性关节炎症明显减轻。其他自身蛋白质如补体系统组分 5 α 、蛋白前体转化酶 subtilisin/kexin 9 等同样可以通过类似的策略打破自身免疫耐受^[117]。通过遗传密码子扩展技术插入免疫原性强的 ncAAs 不仅可以用于打破自身免疫耐受, 同时还有望提高一些外源蛋白的免疫原性, 用于一些感染性疾病的预防和治疗。

2.6 预防性衰减疫苗

目前人们主要通过接种灭活或者减毒疫苗来预防一些传染病。灭活疫苗中的病原体在体内无法复制繁殖, 往往需要多次免疫接种才能达到理想的预防效果; 而减毒活疫苗在体内能够正常复制繁殖, 只需接种一次就可以获得较强的免疫应答, 但存在减毒病原体通过回复突变重新获得致病性的使用安全问题。将病原微生物中关键蛋白进行基于遗传密码子扩展技术的 ncAAs 突变, 可构建出严格依赖 ncAAs 生长的营养缺陷型突变株。这些突变株由于 ncAAs 的缺乏而不能在人体内正常复制繁殖, 会逐渐消亡 (图4)^[67]。将该技术用于疫苗领域具有诸多优势: a. 突变株与野生型病原体的结构完全一致, 具有与野生型一样的免疫原性, 能够激发宿主发生较强免疫应答; b. 在病原微生物必需基因中插入多个 ncAAs 可有效降低发生回复突变重新获得致病性的概率, 提高安全性; c. 与获得减毒突变株的传统方法相比, ncAAs 营养缺陷型突变株构建方法简单且适用范围较广。

目前遗传密码子扩展技术已被应用于病毒衰减疫苗的研发。ncAAs 营养缺陷型突变病毒株只能够在相应 ncAAs 存在的情况下, 在含有正交 aaRS/tRNA 对的细胞系中生长繁殖, 在人体正常细胞中则不能复制繁殖 (图4)。Guo 课题组^[121] 率先将遗传密码子扩展技术用于 HIV-1 ncAA 营养缺陷型突变株的构建。他们向 HIV-1 不同基因中引入 UAG 突变, 随后利用含有正交 aaRS/tRNA 对的转基因细胞系对 HIV-1 病毒进行诱导表达, 以期能获得具有感染活性的 ncAA 营养缺陷型突变株。通过一系列筛选, 他们发现将 HIV-1 蛋白酶第 59 位 Tyr 残基或将 Gag 蛋白第 36 位 Trp 残基和第 127 位 Gln 残基突变为 *p*AzPhe 可保留病毒的自组装能力和感染性。然而 UAG 回复突变为其他有义密码子的风险限制了该 ncAAs 营养缺陷型突变株的进一步应用。为了解决这一问题, 该课题组将四碱基密码子 (UAGA) 替换了 UAG 密码子用于介导 ncAAs 的插入, 大幅降低了 ncAAs 营养缺陷型突变株发生回复

突变的频率^[122]。除此之外, 他们还将编码正交 aaRS/tRNA 对的基因整合到 HIV-1 基因组中, 其可在感染过程中与病毒基因组一起插入到机体正常细胞中, 随后可通过控制 ncAAs 的有无来调控 HIV-1 在体内的传代次数, 达到增强免疫的效果^[123]。Zhou 课题组^[124]利用遗传密码子扩展技术向流感病毒 (influenza virus) NP 蛋白中引入多个 UAG 突变, 成功构建了低逃逸率的流感病毒 ncAAs 营养缺陷型突变株。通过多种动物模型实验, 他们发现获得的流感病毒 ncAAs 营养缺陷型突变株与商品化的疫苗相比具有更强的免疫原性, 可激发宿主免疫系统发生更强、更广泛的免疫应答。

利用遗传密码子扩展技术构建严格依赖 ncAAs 生长的致病菌同样获得了成功。与病毒相比, 由于病原菌具有独立的大分子合成系统, 其逃逸频率,

即发生回复突变的频率相对更高。目前研究人员已采用多种策略有效解决了这一问题。Isaacs 课题组^[125]通过向 *E. coli* 22 个必需基因中同时引入多个 UAG 突变, 大幅降低了逃逸频率。Church 课题组^[126]通过计算生物学和基于 ncAAs 的筛选, 获得了代谢过程严格依赖 ncAAs 的 *E. coli* 菌株, 有效避免了由基因突变或者水平基因转移造成的逃逸。Schultz 课题组^[127-129]通过利用定点突变建立起严格依赖 ncAAs 的关键酶活性位点、关键蛋白质间相互作用以及关键蛋白质修饰/去修饰的方式同样大幅降低了逃逸效率。目前严格依赖 ncAAs 的结核分枝杆菌 (*Mycobacterium tuberculosis*) 已被成功构建, 其作为活体疫苗预防肺结核的研究也正在进行中^[24]。

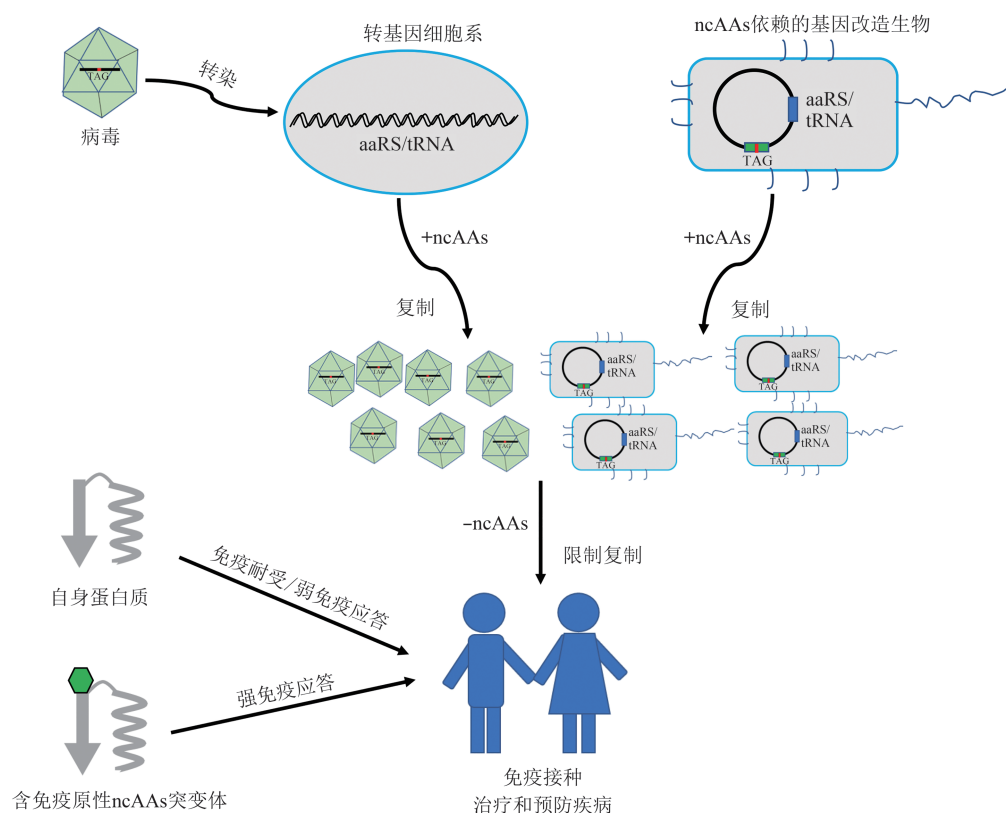


Fig. 4 Applications of genetic code expansion in the vaccine development

图4 遗传密码子扩展技术在疫苗中的应用

2.7 治疗性细胞因子的长效化

蛋白质类药物一般都具有分子质量高、半衰期短、免疫原性强等特点。通过对蛋白药物进行共价聚乙二醇修饰 (PEGylation) 可有效提高蛋白药物

的稳定性以及在体内的半衰期^[130]。过去一般采用对蛋白质 N 端氨基、C 端羧基、Lys 的 ϵ -氨基或者 Cys 的巯基进行修饰。然而对蛋白质末端修饰可能会影响其活性, 对内部的 Lys 和 Cys 残基进行不加

选择地修饰会导致产物不均一,无法获得药效稳定的产物并可能影响蛋白药物的免疫原性^[131]。通过遗传密码子扩展技术定点引入ncAAs不仅可以获得均一的PEG修饰产物,还可以明确不同修饰位点对治疗性蛋白性质的影响从而获得药效更加优良的修饰蛋白。Kimmel研究组^[132]率先利用遗传密码子扩展技术向人生长激素(human growth hormone, hGH)不同位点引入pAcF,通过与含氨基修饰的PEG分子形成酰胺键构建了一系列在不同位点实现均一化PEG修饰的hGH,并对它们的性质进行了比较分析。其中Y35pAcF-hGH在大鼠模型中与野生型hGH活性相当并且具有更优良的药代动力学性质,极大提高了hGH在体内的半衰期。随后Y35pAcF-hGH被命名为ARX201,用于成人生长激素缺乏症(AGHD)的临床治疗并取得了理想效果,有助于大幅降低hGH的给药频率。Zhou研究组^[133]在上述研究的基础上利用相同的策略对hGH进行了多位点PEG修饰,获得的最终突变体Y35/G131/K145-NEAK与Y35pAcF相比具有更低的免疫原性和更优良的药代动力学性质。目前人成纤维细胞生长因子21(FGF21)、牛粒细胞集落刺激因子(GCSF)、干扰素 α 2b(IFN- α 2b)、生长激素受体拮抗剂(GHR antagonists)、白介素-4(interleukin-4, IL-4)等均采用遗传密码子扩展技术进行了定点PEG修饰,在保留了其生物学功能的同时大幅提高了血清半衰期^[72]。PEG虽然能够有效延长治疗性蛋白在体内的半衰期,但研究表明PEG黏附脂质体会激活体内补体系统产生中和PEG抗体,导致二次给药时PEG修饰蛋白的血清半衰期缩短,因此有必要寻找性质更加优良并可用于蛋白质修饰的聚合物^[134]。Meinel研究组^[135]通过遗传密码子扩展技术向IL-4中定点引入含炔丙基的赖氨酸类似物(propargyl-L-lysine, Plk)并将其用于聚2-恶唑啉(poly 2-oxazoline, Pox)聚合物修饰,结果表明修饰后的IL-4生物学功能、热稳定性和化学稳定性均未受到影响。

将生长因子、细胞因子以及其他调控因子固定到病灶周围可在保证治疗疾病的同时有效缓解这些治疗性蛋白存在的全身毒副作用。利用蛋白质分子内部的Lys或Cys残基固定到介质表面会导致产物不均一、空间排列混乱,影响其药代动力学性质以及免疫原性^[136]。通过遗传密码子扩展技术将治疗性蛋白位点特异性地固定到载体上,可制备空间排列单一的均一化产物。Meinel研究组^[137]利用遗传

密码子扩展技术将Plk位点特异性插入到成纤维细胞生长因子2(FGF2)中,通过铜催化的叠氮炔环加成反应将其连接到琼脂糖颗粒上,制备的固定化Plk-FGF-2可高效促进其周围人骨肉瘤细胞的增殖。他们采用相同的策略将IL-4固定到琼脂糖颗粒上,发现固定后的IL-4仍具有生物学功能,可有效促进巨噬细胞的分化^[138]。该研究组还通过遗传密码子扩展技术将Plk定点插入到胰岛素样生长因子I(IGF-1)中并将其固定到琼脂糖颗粒上,他们发现固定化plk-IGF-I比游离IGF-1更能有效促进C2C12细胞肌管形成^[139]。Nickel课题组^[140]利用基于ncAAs的点击化学反应将骨形态发生蛋白2(BMP2)固定到含有活化叠氮基的琼脂糖微珠上,发现BMP2-K3Plk具有促进其周围细胞向成骨细胞系分化的潜力。上述研究表明利用遗传密码子扩展技术制备固定化蛋白因子有望用于伤口愈合、骨缺陷、类风湿性关节炎以及动脉硬化症等疾病的治疗。

2.8 抗菌肽

随着抗生素滥用导致的细菌耐药性问题日趋严重,开发下一代高效广谱抗菌药物已迫在眉睫。抗菌肽(antimicrobial peptide)是一类具有抑菌杀菌功能、带正电荷的两亲性小分子肽,通常由11~50个氨基酸残基组成,广泛存在于微生物以及动植物体内,是宿主天然免疫的重要组成部分。抗菌肽种类和数量繁多,目前已发现有3 000多种,被认为是抗生素药物的理想替代品^[141]。一般在生理条件下,抗菌肽的稳定性及可溶性较差,易受蛋白酶水解,这限制了其在生物医药行业中的应用。已有大量研究表明,通过蛋白质工程可有效改良抗菌肽的性质,提高其应用价值^[142]。一些抗菌肽以前体的形式被核糖体翻译出来,随后通过翻译后加工修饰形成有活性的形态,翻译后修饰的过程往往涉及一些ncAAs的形成。受此启发,研究人员认为对抗菌肽进行基于ncAAs的改造有望获得稳定性、特异性、半衰期、抗菌谱等性质显著提高的突变体^[143]。

目前有关利用遗传密码子扩展技术对抗菌肽进行基于ncAAs改良的研究才刚刚起步。Donk课题组^[144]通过遗传密码子扩展技术向抗菌肽lactacin 481和nukacin ISK-1核心区与前导肽之间定点引入 α -羟基酸(alpha hydroxy acid),实现了利用酸碱处理去除前导肽,避免了特异性蛋白酶的使用。Schultz研究组^[145]将含有 α 氯乙酰胺的ncAAs通过

遗传密码子扩展技术插入乳链菌肽 A (nisin A) 中形成了由硫醚键组成的新型环状拓扑结构。Link 研究组^[146]利用遗传密码子扩展技术将 4 种 ncAAs 引入抗菌肽 MccJ25 结构中, 丰富了该抗菌肽的化学多样性。Süssmuth 研究组^[147]同样通过遗传密码子扩展技术将 5 种 ncAAs 插入抗菌肽 capistrin 中, 利用 Hoveyda-Grubbs 催化剂修饰 ncAAs 的侧链基团成功引入了新的化学结构。Young 课题组^[25]利用遗传密码子扩展技术向抗菌肽 thiocillin 中引入了一些含有生物物理探针 (荧光探针、光交联基团) 的 ncAAs, 在提高 thiocillin 化学多样性的同时为探测其抗菌机理提供了便利。Donk 课题组^[148]将多种 ncAAs 定点引入到抗菌肽 lacticin 481 中获得了抑菌活性显著提高的突变体。尽管多项研究表明插入 ncAAs 可能会导致抗菌肽的生物学活性下降, 但该领域的先驱们普遍认为向抗菌肽分子内部引入 ncAAs 必将丰富抗菌肽的化学多样性, 有望开发出针对新型病原体的杀伤能力。

2.9 病毒载体外壳蛋白修饰

腺病毒 (adeno-associated virus, AAV) 隶属于细小病毒科, 其基因组为线状 DNA 分子。AAV 无明显致病性, 转导效率高, 是基因治疗中常用的病毒载体。然而其在体内靶向性差, 免疫原性强, 血清半衰期短等特点限制了其应用^[149]。相关学者通过基于遗传密码子扩展技术的多种策略在一定程度上解决了上述问题。Zhou 课题组^[150]利用遗传密码子扩展技术将 Nε-2-azidoethoxycarbonyl-L-lysine (NAEK) 定点插入到衣壳蛋白 VP1 中, 构建了一系列含有定点 PEG 修饰的 AAV 突变体并从中获得了血清稳定性提高了 2.4 倍的修饰病毒颗粒。Chatterjee 课题组^[151]利用遗传密码子扩展技术将靶向整合素的环肽 RGD (cyclic-RGD peptide) 定点锚定到 AAV 衣壳蛋白上, 显著提高了 AAV 对高表达 α_vβ₃ 整合素受体细胞的识别和转染。Zhou 课题组^[152]通过遗传密码扩展技术将荧光探针连接到病毒载体上, 实现了对病毒颗粒感染细胞及在胞内运输的可视化监测。最近, Chatterjee 课题组^[153]利用遗传密码子扩展技术实现了对 AAV 载体感染性的光调控, 为未来在使用过程中对 AAV 载体的时空调控打下了坚实基础。他们将 AAV2 衣壳蛋白 VP1 中参与宿主细胞表面硫酸乙酰肝素蛋白多糖 (heparan sulfate proteoglycan, HSPG) 结合的关键氨基酸残基 (R585, R588) 进行了 ncAAs 突变,

新引入的赖氨酸衍生物 NBK 由于正电荷屏蔽作用及空间位阻效应可干扰 VP1 与 HSPG 的结合, 从而抑制 AAV2 的感染性。通过紫外光照射可使 NBK 脱去保护基团转变为 Lys, 使 AAV2 恢复感染活性。

3 其他应用

环肽一般结构紧密, 与靶点结合亲和力高并能耐受蛋白酶的降解, 其在蛋白质抑制剂开发方面具有重要应用。Schultz 研究组^[154]将遗传密码子扩展技术用于 HIV-1 蛋白酶抑制剂环肽的开发, 通过多轮进化筛选, 他们发现将对苯甲酰苯丙氨酸 (4-benzoyl-L-phenylalanine, pBzF) 定点引入到环肽 G12 中, 可诱使 G12 与 HIV 蛋白酶发生共价交联, 显著抑制其活性。Kawakamide 研究组^[155]同样借助于遗传密码子扩展技术获得了能够抑制 TNF-α 功能的环肽分子。Smider 研究组^[156]利用噬菌体展示技术和遗传密码子扩展技术获得了对 HIV gp120 蛋白亲和力显著提高的抗体分子, 并且发现 ncAAs 在其中扮演的角色具有不可替代性。目前 CRISPR/Cas 已被广泛应用于基因编辑领域, 在未来有望用于人类一些重大遗传病以及癌症的治疗, 然而其潜在的脱靶效应仍待解决。Deiters 研究组^[157]利用遗传密码子扩展技术向 Cas9 中定点插入光保护的赖氨酸 (PCK) 抑制其活性, 随后通过紫外光照使 PCK 脱去光保护基团形成 Lys 恢复活性, 实现了在时间和空间上对 Cas9 活性的调控。Liu 课题组^[158]通过将 4-(2-azidoethoxy)-L-phenylalanine (AeF) 定点插入到 Cas12 蛋白中并与相应含有二苯并环辛炔修饰的 crRNA 形成共价复合物, 显著提高了 Cas12-crRNA 的基因编辑能力, 并成功将其用于 CAR-T 细胞的制备。此外, 遗传密码子扩展技术还被广泛应用于一些病原微生物致病机理的研究。Chen 课题组^[159]利用遗传密码子扩展技术在一些肠道致病菌中 (如 *E. coli*, *Shigella*, *Salmonella*) 引入一系列光交联 (photo crosslinking) ncAAs, 研究了这些致病菌在感染人体过程中抵御胃酸及分泌毒素的分子机制。Müller 课题组^[160]利用遗传密码子扩展技术以及荧光手段研究了 HIV 包膜糖蛋白在病毒出芽过程中的作用。Shin 课题组^[161]利用遗传密码子扩展技术对 β-N-乙酰己糖氨基酶进行甘露糖-6-磷酸盐 (M6Ps) 修饰并将其靶向细胞内的溶酶体, 成功用于缺乏内源性 β-N-乙酰己糖氨基酶细胞疾病的治疗。

4 总结与展望

本文综述了遗传密码子扩展技术近年来的研究进展及其在蛋白质及多肽等生物技术药物中的应用。虽然研究人员目前已经能够通过遗传密码子扩展技术将 200 多种 ncAAs 位点特异性插入到蛋白质中,但是这些 ncAAs 在结构上大多是对 Tyr 或 Lys 结构的延伸,将更多不同化学结构的 ncAAs 用于遗传密码子扩展技术有望扩展该技术的能力。由于真核生物基因组庞大且复杂,现阶段通过密码子重分配提高遗传密码子扩展技术在真核生物中的 ncAA 整合效率、增加含 ncAAs 蛋白质的产量仍无法获得突破。由于密码子和相互正交 RS/tRNA 对的匮乏,向除 *E. coli* 之外的物种中同时插入两个以上不同的 ncAAs 依然具有困难。将遗传密码子扩展技术应用于更多物种,包括一些特殊致病菌以及高等多细胞生物等仍面临巨大挑战。然而随着合成生物学、化学生物学、分子生物学以及生物化学等学科的不断发展与积累并指导遗传密码子扩展技术的改良与突破,人类最终有望能够在细胞中从头合成全部由 ncAAs 组成的非天然聚合物 (non-canonical polymers)。随着计算化学以及计算生物学技术的兴起,通过计算模拟辅助遗传密码子扩展技术用于蛋白质及多肽药物的设计将大大提高成功率并有效节约生产成本。随着遗传密码子扩展技术不断成熟与发展,未来其将为诸多生物学问题提供开创性解决方案。

参 考 文 献

- [1] Polcarpo C, Ambrogelly A, Berube A, *et al.* An aminoacyl-tRNA synthetase that specifically activates pyrrolysine. *Proc Natl Acad Sci USA*, 2004, **101**(34): 12450-12454
- [2] Xiao H, Schultz P G. At the interface of chemical and biological synthesis: an expanded genetic code. *Cold Spring Harb Perspect Biol*, 2016, **8**(9): a023945
- [3] Pagar A D, Patil M, Flood D T, *et al.* Recent advances in biocatalysis with chemical modification and expanded amino acid alphabet. *Chem Rev*, 2021, **121**(10): 6173-6245
- [4] Kimmerlin T, Seebach D. '100 years of peptide synthesis': ligation methods for peptide and protein synthesis with applications to beta-peptide assemblies. *J Pept Res*, 2005, **65**(2): 229-260
- [5] Noren C J, Anthony-Cahill S J, Griffith M C, *et al.* A general method for site-specific incorporation of unnatural amino acids into proteins. *Science*, 1989, **244**(4901): 182-188
- [6] Hendrickson T L, Cr  cy-Lagard V D, Schimmel P. Incorporation of nonnatural amino acids into proteins. *Annu Rev Bioche*, 2004, **73**(1): 147-176
- [7] Ibba M, Hennecke H. Relaxing the substrate specificity of an aminoacyl-tRNA synthetase allows *in vitro* and *in vivo* synthesis of proteins containing unnatural amino acids. *FEBS Lett*, 1995, **364**(3): 272-275
- [8] Wang L, Brock A, Herberich B, *et al.* Expanding the genetic code of *Escherichia coli*. *Science*, 2001, **292**(5516): 498-500
- [9] Davis L, Chin J W. Designer proteins: applications of genetic code expansion in cell biology. *Nat Rev Mol Cell Biol*, 2012, **13**(3): 168-182
- [10] Chin J W. Expanding and reprogramming the genetic code of cells and animals. *Annu Rev Bioche*, 2014, **83**(1): 379-408
- [11] Nozawa, Kayo, *et al.* Pyrrolysyl-tRNA synthetase-tRNA Pyl structure reveals the molecular basis of orthogonality. *Nature*, 2009, **457**(7233): 1163-1167
- [12] Ikedaoku A, Ohno S, Hibino Y, *et al.* A simple system for expression of proteins containing 3-azidotyrosine at a pre-determined site in *Escherichia coli*. *J Biochem*, 2013, **153**(3): 317-326
- [13] Neumann H, Peakchew S Y, Chin J W. Genetically encoding N (epsilon)-acetyllysine in recombinant proteins. *Nat Chem Biol*, 2008, **4**(4): 232-234
- [14] Wang Y S, Russell W K, Wang Z, *et al.* The *de novo* engineering of pyrrolysyl-tRNA synthetase for genetic incorporation of L-phenylalanine and its derivatives. *Mol Biosyst*, 2011, **7**(3): 714-717
- [15] Anderson J C, Schultz P G. Adaptation of an orthogonal archaeal leucyl-tRNA and synthetase pair for four-base, amber, and opal suppression. *Biochemistry*, 2003, **42**(32): 9598-9608
- [16] Chatterjee A, Xiao H, Schultz P G. Evolution of multiple, mutually orthogonal prolyl-tRNA synthetase/tRNA pairs for unnatural amino acid mutagenesis in *Escherichia coli*. *Proc Natl Acad Sci USA*, 2012, **109**(37): 14841-14846
- [17] Santoro S W, Anderson J C, Lakshman V, *et al.* An archaeobacteria-derived glutamyl-tRNA synthetase and tRNA pair for unnatural amino acid mutagenesis of proteins in *Escherichia coli*. *Nucleic Acids Res*, 2003, **31**(23): 6700-6709
- [18] Anderson J C, Wu N, Santoro S W, *et al.* An expanded genetic code with a functional quadruplet codon. *Proc Natl Acad Sci USA*, 2004, **101**(20): 7566-7571
- [19] Liu D R, Schultz P G. Progress toward the evolution of an organism with an expanded genetic code. *Proc Natl Acad Sci USA*, 1999, **96**(9): 4780-4785
- [20] Ohno S, Yokogawa T, Fujii I, *et al.* Co-expression of yeast amber suppressor tRNA^{Tyr} and tyrosyl-tRNA synthetase in *Escherichia coli*: possibility to expand the genetic code. *J Biochem*, 1998, **124**(6): 1065-1068
- [21] Kowal A K, Kohrer C, Rajbhandary U L. Twenty-first aminoacyl-tRNA synthetase-suppressor tRNA pairs for possible use in site-specific incorporation of amino acid analogues into proteins in eukaryotes and in eubacteria. *Proc Natl Acad Sci USA*, 2001, **98**(5): 2268-2273
- [22] Pastrnak M, Magliery T J, Schultz P G. A new orthogonal suppressor tRNA/aminoacyl-tRNA synthetase pair for evolving an

- organism with an expanded genetic code. *Helvetica Chimica Acta*, 2000, **83**(9): 2277-2286
- [23] Hughes R A, Ellington A D. Rational design of an orthogonal tryptophanyl nonsense suppressor tRNA. *Nucleic Acids Res*, 2010, **38**(19): 6813-6830
- [24] Wang F, Robbins S, Guo J, *et al.* Genetic incorporation of unnatural amino acids into proteins in *Mycobacterium tuberculosis*. *PLoS One*, 2010, **5**(2): e9354
- [25] Luo X, Zambaldo C, Liu T, *et al.* Recombinant thiopeptides containing noncanonical amino acids. *Proc Natl Acad Sci USA*, 2016, **113**(13): 3615-3620
- [26] Li J, Lin S, Wang J, *et al.* Ligand-free palladium-mediated site-specific protein labeling inside gram-negative bacterial pathogens. *J Am Chem Soc*, 2013, **135**(19): 7330-7338
- [27] He J, Van Treeck B, Nguyen H B, *et al.* Development of an unnatural amino acid incorporation system in the actinobacterial natural product producer *Streptomyces venezuelae* ATCC 15439. *ACS Synth Biol*, 2016, **5**(2): 125-132
- [28] Chin J W, Cropp T A, Chu S, *et al.* Progress toward an expanded eukaryotic genetic code. *Chem Bio*, 2003, **10**(6): 511-519
- [29] Wu N, Deiters A, Cropp T A, *et al.* A genetically encoded photocaged amino acid. *J Am Chem Soc*, 2004, **126**(44): 14306-14307
- [30] Hancock S M, Uprety R, Deiters A, *et al.* Expanding the genetic code of yeast for incorporation of diverse unnatural amino acids via a pyrrolysyl-tRNA synthetase/tRNA pair. *J Am Chem Soc*, 2010, **132**(42): 14819-14824
- [31] Zhang Z, Alfonta L, Tian F, *et al.* Selective incorporation of 5-hydroxytryptophan into proteins in mammalian cells. *Proc Natl Acad Sci USA*, 2004, **101**(24): 8882-8887
- [32] Wang W, Takimoto J K, Louie G V, *et al.* Genetically encoding unnatural amino acids for cellular and neuronal studies. *Nat Neurosci*, 2007, **10**(8): 1063-1072
- [33] Drabkin H J, Park H J, Rajbhandary U L. Amber suppression in mammalian cells dependent upon expression of an *Escherichia coli* aminoacyl-tRNA synthetase gene. *Mol Cell Biol*, 1996, **16**(3): 907-913
- [34] Sakamoto K, Hayashi A, Sakamoto A, *et al.* Site-specific incorporation of an unnatural amino acid into proteins in mammalian cells. *Nucleic Acids Res*, 2002, **30**(21): 4692-4699
- [35] Liu C C, Schultz P G. Adding new chemistries to the genetic code. *Annu Rev Biochem*, 2010, **79**: 413-444
- [36] Chen P R, Groff D, Guo J, *et al.* A facile system for encoding unnatural amino acids in mammalian cells. *Angew Chem Int Ed*, 2009, **48**(22): 4052-4055
- [37] Greiss S, Chin J W. Expanding the genetic code of an animal. *J Am Chem Soc*, 2011, **133**(36): 14196-14199
- [38] Liu J, Hemphill J, Samanta S, *et al.* Genetic code expansion in zebrafish embryos and its application to optical control of cell signaling. *J Am Chem Soc*, 2017, **139**(27): 9100-9103
- [39] Li F, Zhang H, Sun Y, *et al.* Expanding the genetic code for photoclick chemistry in *E. coli*, mammalian cells, and *A. thaliana*. *Angew Chem Int Ed*, 2013, **52**(37): 9700-9704
- [40] Bianco A, Townsley F M, Greiss S, *et al.* Expanding the genetic code of *Drosophila melanogaster*. *Nat Chem Biol*, 2012, **8**(9): 748-750
- [41] Ryden M, Murphy J, Martin R, *et al.* Mapping and complementation studies of the gene for release factor 1. *J Bacteriol*, 1986, **168**(3): 1066-1069
- [42] Mukai T, Hayashi A, Iraha F, *et al.* Codon reassignment in the *Escherichia coli* genetic code. *Nucleic Acids Res*, 2010, **38**(22): 8188-8195
- [43] Johnson D B, Xu J, Shen Z, *et al.* RF1 knockout allows ribosomal incorporation of unnatural amino acids at multiple sites. *Nat Chem Biol*, 2011, **7**(11): 779-786
- [44] Lajoie M J, Rovner A J, Goodman D B, *et al.* Genomically recoded organisms expand biological functions. *Science*, 2013, **342**(6156): 357-360
- [45] Yi H, Zhang J, Ke F, *et al.* Comparative analyses of the transcriptome and proteome of *Escherichia coli* C321. Δ A and further improving its noncanonical amino acids containing protein expression ability by integration of T7 RNA polymerase. *Front Microbiol*, 2021, **12**: 744284
- [46] Young T S, Ahmad I, Yin J A, *et al.* An enhanced system for unnatural amino acid mutagenesis in *E. coli*. *J Mol Biol*, 2010, **395**(2): 361-374
- [47] Amiram M, Haimovich A D, Fan C, *et al.* Evolution of translation machinery in recoded bacteria enables multi-site incorporation of nonstandard amino acids. *Nat Biotechnol*, 2015, **33**(12): 1272-1279
- [48] Bryson D I, Fan C, Guo L T, *et al.* Continuous directed evolution of aminoacyl-tRNA synthetases. *Nat Chem Biol*, 2017, **13**(12): 1253-1260
- [49] Suzuki T, Miller C, Guo L T, *et al.* Crystal structures reveal an elusive functional domain of pyrrolysyl-tRNA synthetase. *Nat Chem Biol*, 2017, **13**(12): 1261-1266
- [50] Luo X, Fu G, Wang R E, *et al.* Genetically encoding phosphotyrosine and its nonhydrolyzable analog in bacteria. *Nat Chem Biol*, 2017, **13**(8): 845-849
- [51] Park H S, Hohn M J, Umehara T, *et al.* Expanding the genetic code of *Escherichia coli* with phosphoserine. *Science*, 2011, **333**(6046): 1151-1154
- [52] Italia J S, Addy P S, Erickson S B, *et al.* Mutually orthogonal nonsense-suppression systems and conjugation chemistries for precise protein labeling at up to three distinct sites. *J Am Chem Soc*, 2019, **141**(15): 6204-6212
- [53] Dunkelmann D L, Willis J C W, Beattie A T, *et al.* Engineered triply orthogonal pyrrolysyl-tRNA synthetase/tRNA pairs enable the genetic encoding of three distinct non-canonical amino acids. *Nat Chem*, 2020, **12**(6): 535-544
- [54] Neumann H, Slusarczyk A L, Chin J W. *De novo* generation of mutually orthogonal aminoacyl-tRNA synthetase/tRNA pairs. *J Am Chem Soc*, 2010, **132**(7): 2142-2144
- [55] Ding W, Zhao H, Chen Y, *et al.* Chimeric design of pyrrolysyl-tRNA synthetase/tRNA pairs and canonical synthetase/tRNA pairs for genetic code expansion. *Nat Commun*, 2020, **11**(1): 3154

- [56] Wang K, Fredens J, Brunner S F, *et al.* Defining synonymous codon compression schemes by genome recoding. *Nature*, 2016, **539**(7627): 59-64
- [57] Ostrov N, Landon M, Guell M, *et al.* Design, synthesis, and testing toward a 57-codon genome. *Science*, 2016, **353**(6301): 819-822
- [58] Robertson W E, Funke L F H, De La Torre D, *et al.* Sense codon reassignment enables viral resistance and encoded polymer synthesis. *Science*, 2021, **372**(6546): 1057-1062
- [59] Neumann H, Wang K, Davis L, *et al.* Encoding multiple unnatural amino acids *via* evolution of a quadruplet-decoding ribosome. *Nature*, 2010, **464**(7287): 441-444
- [60] Zhang Y, Ptacin J L, Fischer E C, *et al.* A semi-synthetic organism that stores and retrieves increased genetic information. *Nature*, 2017, **551**(7682): 644-647
- [61] Chen Y, Loredó A, Gordon A, *et al.* A noncanonical amino acid-based relay system for site-specific protein labeling. *Chem Commun*, 2018, **54**(52): 7187-7190
- [62] Chen Y, Tang J, Wang L, *et al.* Creation of bacterial cells with 5-hydroxytryptophan as a 21(st) amino acid building block. *Chem*, 2020, **6**(10): 2717-2727
- [63] Wang Y, Chen X, Cai W, *et al.* Expanding the structural diversity of protein building blocks with noncanonical amino acids biosynthesized from aromatic thiols. *Angew Chem Int Ed*, 2021, **60**(18): 10040-10048
- [64] Zhang M S, Brunner S F, Huguenin-Dezot N, *et al.* Biosynthesis and genetic encoding of phosphothreonine through parallel selection and deep sequencing. *Nat Methods*, 2017, **14**(7): 729-736
- [65] Kim S, Sung B H, Kim S C, *et al.* Genetic incorporation of L-dihydroxyphenylalanine (DOPA) biosynthesized by a tyrosine phenol-lyase. *Chem Comm*, 2018, **54**(24): 3002-3005
- [66] Exner M P, Kuenzl T, To T M, *et al.* Design of S-Allylcysteine *in situ* production and incorporation based on a novel pyrrolysyl-tRNA synthetase variant. *Chembiochem*, 2017, **18**(1): 85-90
- [67] Rezhdo A, Islam M, Huang M, *et al.* Future prospects for noncanonical amino acids in biological therapeutics. *Curr Opin Biotech*, 2019, **60**: 168-178
- [68] Wu K L, Yu C, Lee C, *et al.* Precision modification of native antibodies. *Bioconjug Chem*, 2021, **32**(9): 1947-1959
- [69] Sievers E L, Senter P D. Antibody-drug conjugates in cancer therapy. *Annu Rev Med*, 2013, **64**: 15-29
- [70] Sochaj A M, Świdorska K W, Otlewski J. Current methods for the synthesis of homogeneous antibody-drug conjugates. *Biotechnol Adv*, 2015, **33**(6): 775-784
- [71] Junutula J R, Raab H, Clark S, *et al.* Site-specific conjugation of a cytotoxic drug to an antibody improves the therapeutic index. *Nat Biotechnol*, 2008, **26**(8): 925-932
- [72] Huang Y, Liu T. Therapeutic applications of genetic code expansion. *Synth Syst Biotechnol*, 2018, **3**(3): 150-158
- [73] Kim C H, Axup J Y, Schultz P G. Protein conjugation with genetically encoded unnatural amino acids. *Curr Opin Chem Biol*, 2013, **17**(3): 412-419
- [74] Axup J Y, Bajjuri K M, Ritland M, *et al.* Synthesis of site-specific antibody-drug conjugates using unnatural amino acids. *Proc Natl Acad Sci USA*, 2012, **109**(40): 16101-16106
- [75] Oller-Salvia B, Kym G, Chin J W. Rapid and efficient generation of stable antibody-drug conjugates *via* an encoded cyclopropene and an inverse-electron-demand diels-alder reaction. *Angew Chem Int Ed*, 2018, **57**(11): 2831-2834
- [76] Tian F, Lu Y, Manibusan A, *et al.* A general approach to site-specific antibody drug conjugates. *Proc Natl Acad Sci USA*, 2014, **111**(5): 1766-1771
- [77] Maruani A, Smith M E, Miranda E, *et al.* A plug-and-play approach to antibody-based therapeutics *via* a chemoselective dual click strategy. *Nat Commun*, 2015, **6**: 6645
- [78] Xiao H, Chatterjee A, Choi S H, *et al.* Genetic incorporation of multiple unnatural amino acids into proteins in mammalian cells. *Angew Chem Int Ed*, 2013, **52**(52): 14080-14083
- [79] Lu H, Wang D, Kazane S, *et al.* Site-specific antibody-polymer conjugates for siRNA delivery. *J Am Chem Soc*, 2013, **135**(37): 13885-13891
- [80] Wu Y, Zhu H, Zhang B, *et al.* Synthesis of site-specific radiolabeled antibodies for radioimmunotherapy *via* genetic code expansion. *Bioconjug Chem*, 2016, **27**(10): 2460-2468
- [81] Wang R E, Liu T, Wang Y, *et al.* An immunosuppressive antibody-drug conjugate. *J Am Chem Soc*, 2015, **137**(9): 3229-3232
- [82] 王璐瑶, 魏振华, 熊伟佳, 等. 双特异性抗体构建在肿瘤临床治疗中的应用. *生物工程学报*, 2021, **37**(2): 513-529
- Wang L Y, Wei Z H, Xiong W J, *et al.* *Chinese Journal of Biotechnology*, 2021, **37**(2): 513-529
- [83] Holmes D. Buy buy bispecific antibodies. *Nat Rev Drug Discov*, 2011, **10**(11): 798-800
- [84] Przepiorka D, Ko C W, Deisseroth A, *et al.* FDA approval: blinatumomab. *Clin Cancer Res*, 2015, **21**(18): 4035-4039
- [85] Scott L J, Kim E S. Emticizumab-kxwh: first global approval. *Drugs*, 2018, **78**(2): 269-274
- [86] Brinkmann U, Kontermann R E. The making of bispecific antibodies. *MAbs*, 2017, **9**(2): 182-212
- [87] Kim C H, Axup J Y, Dubrovskaya A, *et al.* Synthesis of bispecific antibodies using genetically encoded unnatural amino acids. *J Am Chem Soc*, 2012, **134**(24): 9918-9921
- [88] Lu H, Zhou Q, Deshmukh V, *et al.* Targeting human C-type lectin-like molecule - 1 (CLL1) with a bispecific antibody for immunotherapy of acute myeloid leukemia. *Angew Chem Int Ed*, 2014, **126**(37): 9999-10003
- [89] Ramadoss N S, Schulman A D, Choi S-H, *et al.* An anti-B cell maturation antigen bispecific antibody for multiple myeloma. *J Am Chem Soc*, 2015, **137**(16): 5288-5291
- [90] Kazane S A, Axup J Y, Kim C H, *et al.* Self-assembled antibody multimers through peptide nucleic acid conjugation. *J Am Chem Soc*, 2013, **135**(1): 340-346
- [91] Kularatne S A, Venkatesh C, Santhapuram H K, *et al.* Synthesis and biological analysis of prostate-specific membrane antigen-targeted anticancer prodrugs. *J Med Chem*, 2010, **53**(21): 7767-7777
- [92] Kim C H, Axup J Y, Lawson B R, *et al.* Bispecific small molecule-antibody conjugate targeting prostate cancer. *Proc Natl Acad Sci*

- USA, 2013, **110**(44): 17796-17801
- [93] Kularatne S A, Deshmukh V, Gymnopoulos M, *et al.* Recruiting cytotoxic T cells to folate-receptor-positive cancer cells. *Angew Chem Int Ed*, 2013, **52**(46): 12101-12104
- [94] Chen Y Y. Increasing T cell versatility with SUPRA CARs. *Cell*, 2018, **173**(6): 1316-1317
- [95] Maude S L, Frey N, Shaw P A, *et al.* Chimeric antigen receptor T cells for sustained remissions in leukemia. *N Engl J Med*, 2014, **371**(16): 1507-1517
- [96] Cao Y, Rodgers D T, Du J, *et al.* Design of switchable chimeric antigen receptor T cells targeting breast cancer. *Angew Chem Int Ed*, 2016, **55**(26): 7520-7524
- [97] Ma J S, Kim J Y, Kazane S A, *et al.* Versatile strategy for controlling the specificity and activity of engineered T cells. *Proc Natl Acad Sci USA*, 2016, **113**(4): 450-458
- [98] Chen C, Yu G, Huang Y, *et al.* Genetic-code-expanded cell-based therapy for treating diabetes in mice. *Nat Chem Bio*, 2022, **18**(1): 41-55 (DOI: 10.1038/s41589-021-00899-z)
- [99] Potashman M H, Duggan M E. Covalent modifiers: an orthogonal approach to drug design. *J Med Chem*, 2009, **52**(5): 1231-1246
- [100] Nguyen T A, Cigler M, Lang K. Expanding the genetic code to study protein-protein interactions. *Angew Chem Int Ed*, 2018, **57**(44): 14350-14361
- [101] Sevier C S, Kaiser C A. Formation and transfer of disulphide bonds in living cells. *Nat Rev Mol Cell Biol*, 2002, **3**(11): 836-847
- [102] Stebbins J L, Santelli E, Feng Y, *et al.* Structure-based design of covalent Siah inhibitors. *Chem Biol*, 2013, **20**(8): 973-982
- [103] De Araujo A D, Lim J, Good A C, *et al.* Electrophilic helical peptides that bond covalently, irreversibly, and selectively in a protein-protein interaction site. *ACS Med Chem Lett*, 2017, **8**(1): 22-26
- [104] Hoppmann C, Wang L. Proximity-enabled bioreactivity to generate covalent peptide inhibitors of p53-Mdm4. *Chem Comm*, 2016, **52**(29): 5140-5143
- [105] Xiang Z, Ren H, Hu Y S, *et al.* Adding an unnatural covalent bond to proteins through proximity-enhanced bioreactivity. *Nat Methods*, 2013, **10**(9): 885-888
- [106] Xiang Z, Lacey V K, Ren H, *et al.* Proximity-enabled protein crosslinking through genetically encoding haloalkane unnatural amino acids. *Angew Chem Int Ed*, 2014, **53**(8): 2190-2193
- [107] Chen X H, Xiang Z, Hu Y S, *et al.* Genetically encoding an electrophilic amino acid for protein stapling and covalent binding to native receptors. *ACS Chem Biol*, 2014, **9**(9): 1956-1961
- [108] Kobayashi T, Hoppmann C, Yang B, *et al.* Using protein-confined proximity to determine chemical reactivity. *J Am Chem Soc*, 2016, **138**(45): 14832-14835
- [109] Li Q, Chen Q, Klauser P C, *et al.* Developing covalent protein drugs *via* proximity-enabled reactive therapeutics. *Cell*, 2020, **182**(1): 85-97
- [110] Zhang H, Han Y, Yang Y, *et al.* Covalently engineered nanobody chimeras for targeted membrane protein degradation. *J Am Chem Soc*, 2021, **143**(40): 16377-16382
- [111] Wing K, Sakaguchi S. Regulatory T cells exert checks and balances on self tolerance and autoimmunity. *Nat Immunol*, 2010, **11**(1): 7-13
- [112] Fok J A, Mayer C. Genetic-code-expansion strategies for vaccine development. *Chembiochem*, 2020, **21**(23): 3291-3300
- [113] Waldmann T A. Immunotherapy: past, present and future. *Nat Medicine*, 2003, **9**(3): 269-277
- [114] Kalliolias G D, Ivashkiv L B. TNF biology, pathogenic mechanisms and emerging therapeutic strategies. *Nat Rev Rheumatol*, 2016, **12**(1): 49-62
- [115] Doyle H A, Mamula M J. Autoantigenesis: the evolution of protein modifications in autoimmune disease. *Curr Opin Immunol*, 2012, **24**(1): 112-118
- [116] Jones L H. Recent advances in the molecular design of synthetic vaccines. *Nat Chem*, 2015, **7**(12): 952-960
- [117] Grünewald J, Hunt G S, Dong L, *et al.* Mechanistic studies of the immunochemical termination of self-tolerance with unnatural amino acids. *Proc Natl Acad Sci USA*, 2009, **106**(11): 4337-4342
- [118] Grünewald J, Tsao M L, Perera R, *et al.* Immunochemical termination of self-tolerance. *Proc Natl Acad Sci USA*, 2008, **105**(32): 11276-11280
- [119] Ono T, Hayashi M, Sasaki F, *et al.* RANKL biology: bone metabolism, the immune system, and beyond. *Inflamm Regen*, 2020, **40**:2
- [120] Li F, Li H, Zhai Q, *et al.* A new vaccine targeting RANKL, prepared by incorporation of an unnatural Amino acid into RANKL, prevents OVX-induced bone loss in mice. *Biochem Biophys Res Commun*, 2018, **499**(3): 648-654
- [121] Wang N, Li Y, Niu W, *et al.* Construction of a live-attenuated HIV-1 vaccine through genetic code expansion. *Angew Chem Int Ed*, 2014, **53**(19): 4867-4871
- [122] Chen Y, Wan Y, Wang N, *et al.* Controlling the replication of a genomically recoded HIV-1 with a functional quadruplet codon in mammalian cells. *ACS Synth Biol*, 2018, **7**(6): 1612-1617
- [123] Yuan Z, Wang N, Kang G, *et al.* Controlling multicycle replication of live-attenuated HIV-1 using an unnatural genetic switch. *ACS Synth Biol*, 2017, **6**(4): 721-731
- [124] Si L, Xu H, Zhou X, *et al.* Generation of influenza A viruses as live but replication-incompetent virus vaccines. *Science*, 2016, **354**(6316): 1170-1173
- [125] Rovner A J, Haimovich A D, Katz S R, *et al.* Recoded organisms engineered to depend on synthetic amino acids. *Nature*, 2015, **518**(7537): 89-93
- [126] Mandell D J, Lajoie M J, Mee M T, *et al.* Biocontainment of genetically modified organisms by synthetic protein design. *Nature*, 2015, **518**(7537): 55-60
- [127] Xuan W, Schultz P G. A strategy for creating organisms dependent on noncanonical amino acids. *Angew Chem Int Ed*, 2017, **56**(31): 9170-9173
- [128] Gan F, Liu R, Wang F, *et al.* Functional replacement of histidine in proteins to generate noncanonical amino acid dependent organisms. *J Am Chem Soc*, 2018, **140**(11): 3829-3832
- [129] Koh M, Nasertorabi F, Han G W, *et al.* Generation of an orthogonal protein-protein interface with a noncanonical amino acid. *J Am*

- Chem Soc, 2017, **139**(16): 5728-5731
- [130] Nischan N, Hackenberger C P. Site-specific PEGylation of proteins: recent developments. *J Org Chem*, 2014, **79**(22): 10727-10733
- [131] Wals K, Ovaa H. Unnatural amino acid incorporation in *E. coli*: current and future applications in the design of therapeutic proteins. *Front Chem*, 2014, **2**:15
- [132] Cho H, Daniel T, Buechler Y J, *et al.* Optimized clinical performance of growth hormone with an expanded genetic code. *Proc Natl Acad Sci USA*, 2011, **108**(22): 9060-9065
- [133] Wu L, Chen J, Wu Y, *et al.* Precise and combinatorial PEGylation generates a low-immunogenic and stable form of human growth hormone. *J Control Release*, 2017, **249**:84-93
- [134] Sroda K, Rydlewski J, Langner M, *et al.* Repeated injections of PEG-PE liposomes generate anti-PEG antibodies. *Cell Mol Biol Lett*, 2005, **10**(1): 37-47
- [135] Lühmann T, Schmidt M, Leiske M N, *et al.* Site-specific POxylation of interleukin-4. *ACS Biomater Sci Eng*, 2017, **3**(3): 304-312
- [136] Franco E J, Hofstetter H, Hofstetter O. A comparative evaluation of random and site-specific immobilization techniques for the preparation of antibody-based chiral stationary phases. *J Sep Sci*, 2006, **29**(10): 1458-1469
- [137] Lühmann T, Jones G, Gutmann M, *et al.* Bio-orthogonal immobilization of fibroblast growth factor 2 for spatial controlled cell proliferation. *ACS Biomater Sci Eng*, 2015, **1**(9): 740-746
- [138] Lühmann T, Spieler V, Werner V, *et al.* Interleukin-4-clicked surfaces drive M2 macrophage polarization. *Chembiochem*, 2016, **17**(22): 2123-2128
- [139] Wu F, Braun A, Lühmann T, *et al.* Site-specific conjugated insulin-like growth factor-I for anabolic therapy. *ACS Biomater Sci Eng*, 2018, **4**(3): 819-825
- [140] Siverino C, Tabisz B, Lühmann T, *et al.* Site-directed immobilization of bone morphogenetic protein 2 to solid surfaces by click chemistry. *J Vis Exp*, 2018, **133**: 56616
- [141] Budisa N. Expanded genetic code for the engineering of ribosomally synthesized and post-translationally modified peptide natural products (RiPPs). *Curr Opin Biotechnol*, 2013, **24**(4): 591-598
- [142] Wang G. Post-translational modifications of natural antimicrobial peptides and strategies for peptide engineering. *Curr Biotechnol*, 2012, **1**(1): 72-79
- [143] Baumann T, Nickling J H, Bartholomae M, *et al.* Prospects of *in vivo* incorporation of non-canonical amino acids for the chemical diversification of antimicrobial peptides. *Front Microbiol*, 2017, **8**:124
- [144] Bindman N A, Bobeica S C, Liu W R, *et al.* Facile removal of leader peptides from lanthipeptides by incorporation of a hydroxy acid. *J Am Chem Soc*, 2015, **137**(22): 6975-6978
- [145] Zambaldo C, Luo X, Mehta A P, *et al.* Recombinant macrocyclic lanthipeptides incorporating non-canonical amino acids. *J Am Chem Soc*, 2017, **139**(34): 11646-11649
- [146] Piscotta F J, Tharp J M, Liu W R, *et al.* Expanding the chemical diversity of lasso peptide MccJ25 with genetically encoded noncanonical amino acids. *Chem Commun*, 2015, **51**(2): 409-412
- [147] Al Toma R S, Kuthning A, Exner M P, *et al.* Site-directed and global incorporation of orthogonal and isostructural noncanonical amino acids into the ribosomal lasso peptide capistrin. *Chembiochem*, 2015, **16**(3): 503-509
- [148] Kakkar N, Perez J G, Liu W R, *et al.* Incorporation of nonproteinogenic amino acids in class I and II lantibiotics. *ACS Chem Biol*, 2018, **13**(4): 951-957
- [149] Daya S, Berns K I. Gene therapy using adeno-associated virus vectors. *Clin Microbiol Rev*, 2008, **21**(4): 583-593
- [150] Yao T, Zhou X, Zhang C, *et al.* Site-specific PEGylated adeno-associated viruses with increased serum stability and reduced immunogenicity. *Molecules*, 2017, **22**(7): 1155
- [151] Kelemen R E, Mukherjee R, Cao X, *et al.* A precise chemical strategy to alter the receptor specificity of the adeno-associated virus. *Angew Chem Int Ed*, 2016, **55**(36): 10645-10649
- [152] Zhang C, Zhou X, Yao T, *et al.* Precision fluorescent labeling of an adeno-associated virus vector to monitor the viral infection pathway. *Biotechnol J*, 2018, **13**(4): e1700374
- [153] Erickson S B, Mukherjee R, Kelemen R E, *et al.* Precise photoremovable perturbation of a virus-host interaction. *Angew Chem Int Ed*, 2017, **56**(15): 4234-4237
- [154] Young T S, Young D D, Ahmad I, *et al.* Evolution of cyclic peptide protease inhibitors. *Proc Natl Acad Sci USA*, 2011, **108**(27): 11052-11056
- [155] Tsukamoto K, Ando T, Fuji D, *et al.* *In vitro* display evolution of the PURE system-expressed TNF α -binding unnatural cyclic peptide containing an N-methyl-D-amino acid. *Biochem Biophys Res Commun*, 2021, **534**:519-525
- [156] Liu C C, Mack A V, Tsao M L, *et al.* Protein evolution with an expanded genetic code. *Proc Natl Acad Sci USA*, 2008, **105**(46): 17688-17693
- [157] Hemphill J, Borchardt E K, Brown K, *et al.* Optical control of CRISPR/Cas9 gene editing. *J Am Chem Soc*, 2015, **137**(17): 5642-5645
- [158] Ling X, Chang L, Chen H, *et al.* Improving the efficiency of CRISPR-Cas12a-based genome editing with site-specific covalent Cas12a-crRNA conjugates. *Mol Cell*, 2021, **81**(22): 4747-4756
- [159] Lin S, Zhang Z, Xu H, *et al.* Site-specific incorporation of photo-cross-linker and bioorthogonal amino acids into enteric bacterial pathogens. *J Am Chem Soc*, 2011, **133**(50): 20581-20587
- [160] 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
- [161] Hyun J Y, Kim S, Lee H S, *et al.* A glycoengineered enzyme with multiple mannose-6-phosphates is internalized into diseased cells to restore its activity in lysosomes. *Cell Chem Biol*, 2018, **25**(10): 1255-1267

Applications of Genetic Code Expansion in Protein and Peptide Drugs^{*}

GAO Xiao-Wei, WEI Si-Ping^{**}, WANG Qin^{**}

(School of Pharmacy, Southwest Medical University, Luzhou 646000, China)

Abstract Site-specific introduction of noncanonical amino acids (ncAAs) by genetic codon expansion enables manipulation of protein structure and function at the atomic level. More than 200 ncAAs have been incorporated into proteins by this method in different organisms, including plants and animals, and its application has been expanded to the development of new generation biotherapeutics. With the ability to introduce biorthogonal reactive groups into proteins site-specifically, genetic code expansion could not only improve the safety and efficacy of protein and peptide drugs, but also provide a pioneering solution for the prevention and treatment of human diseases. Here, we reviewed recent advances of genetic codon expansion and focused on its applications in protein and peptide drugs. Using genetic codon expansion, homogeneous antibody-drug conjugates with improved pharmacokinetic property could be constructed. Bispecific antibodies with flexible PEG-linkers constructed by ncAAs-based biorthogonal reaction could overcome the spatial arrangement limitation existed in the traditional methods, which generated the bispecific antibodies mainly relying on genetic fusion. Development of genetic codon expansion-based protein switch could improve the efficiency of cell therapy. Site-specifically modified cytokines with PEG and immobilized them on solid surface by genetic codon expansion could also improve their pharmacological properties. Incorporation of immunogenic ncAAs in proteins could elicit immune responses and be used to develop therapeutic protein vaccines. We believe that genetic code expansion could be widely used in the next generation protein and peptide drugs.

Key words genetic code expansion, noncanonical amino acids, antibody, antimicrobial peptides, vaccines, covalent protein drugs

DOI: 10.16476/j.pibb.2021.0332

^{*} This work was supported by grants from the China Postdoctoral Science Foundation (2020M683364) and the Department of Science and Technology of Sichuan Province (2020YJ0129, 2019JDTD0016).

^{**} Corresponding author.

WEI Si-Ping. Tel: 86-830-3162291, E-mail: swei1225@swmu.edu.cn

WANG Qin. Tel: 86-830-3162291, E-mail: wq_ring@hotmail.com

Received: October 28, 2021 Accepted: November 29, 2021