体外合成生物学: 无细胞蛋白合成系统研究进展

刘阳[†], 郭小翠[†], 耿金慧, 焦毅, 韩金鹏, 张志昆, 周晓, 仰大勇^{*}

天津大学化工学院, 系统生物工程教育部重点实验室, 天津化学化工协同创新中心, 天津 300350

[†] 同等贡献

^{*} 联系人, E-mail: dayong.yang@tju.edu.cn

2017-04-18 收稿, 2017-08-03 修回, 2017-08-31 接受, 2017-09-27 网络版发表

摘要 合成生物学是近年来融合生物、化学和工程等学科的新兴交叉研究领域, 推动人类由理解生命到创造生命的革新. 体外合成生物学不依赖生命体, 在试管等媒介中研究生命活动和规律, 是合成生物学重要的前沿领域. 作为体外合成生物学最重要的研究平台, 无细胞蛋白合成系统利用外源DNA或mRNA直接在体外合成目标蛋白质, 不依赖于细胞结构, 突破传统生物方法的局限, 具有简便、快速、可控性强和绿色经济等优点, 同时为理解生命进化和人工创造生命系统提供了重要的研究模型. 本文综述了无细胞蛋白合成系统的发展历程、组成、类型、优势等, 并对其在高通量蛋白质和药物合成方面的应用进行了总结.

关键词 合成生物学, 无细胞蛋白合成, 高通量蛋白合成, 蛋白药物

合成生物学是近年来融合生物、化学和工程等学科的新兴交叉研究领域. 2010年, *Science*和*Nature*杂志评出的重大科学突破均包含合成生物学. 2014年, 美国国防部确定了未来6大颠覆性基础研究领域, 合成生物学位居其中. 合成生物学的核心任务是设计构建新的生物分子、生物系统和生物机器, 打破非生命化学物质和生命体之间的界线, 推动人类由理解生命到创造生命的革新^[1-4]. 细胞是生命体的基本结构和功能单位, 是生化反应发生的场所, 也被称为“细胞工厂”, 可用于生产蛋白质药物、酶制剂和生物能源等产品, 在化学品先进制造和生物经济的革命性发展等方面发挥着巨大推动作用^[5]. 通过基因组学、系统生物学和高速计算等技术解析微生物的基因、蛋白、网络与代谢过程, 已实现在分子、细胞和生态系统等尺度上全方位、多层次地认识和改造细胞, 实现细胞工厂的设计构建. 然而, 细胞内生化反应网络复杂, 人工改造易对系统产生不利或未知影响, 导致细胞活性降低甚至丧失, 限制细胞工厂的应

用. 因此, 如何突破细胞工厂的局限是生物合成所面临的巨大挑战. 针对细胞工厂的局限, 无细胞蛋白合成系统(cell-free protein synthesis, CFPS)提供了很好的解决方案. CFPS作为一种体外生命模拟体系, 以外源DNA或mRNA为模板, 利用细胞提取物中的蛋白合成机器、蛋白折叠因子及其他相关酶系, 通过添加氨基酸、聚合酶和能量等物质, 在体外完成蛋白合成翻译及后修饰过程^[6,7]. 如图1所示, CFPS实现蛋白质的简洁高效合成, 避免细胞系统蛋白合成过程中繁琐的基因克隆和细胞培养操作, 在极短时间内合成大量蛋白质, 并满足目前蛋白质结构和功能组学的研究要求. 相较于细胞系统, CFPS的可控性和可操作性更强; 可在蛋白合成过程中, 随时对系统进行人为干预, 调控蛋白表达. 而且, CFPS与现代生物技术系统相容性高, 很容易实现蛋白质的高通量筛选和工业化生产.

CFPS的发展历程大致经历了7个阶段, 如图2所示. 1958年, Zamecnik课题组^[8]首次证明从细胞中提

引用格式: 刘阳, 郭小翠, 耿金慧, 等. 体外合成生物学: 无细胞蛋白合成系统研究进展. 科学通报, 2017, 62: 3851-3860

Liu Y, Guo X C, Geng J H, et al. *In vitro* synthetic biology: Cell-free protein synthesis (in Chinese). Chin Sci Bull, 2017, 62: 3851-3860, doi: 10.1360/N972017-00419

图 1 (网络版彩色)从细胞工厂到无细胞蛋白质合成系统的演进

Figure 1 (Color online) Evolution of the protein synthesis from cell factory system to CFPS

图 2 (网络版彩色)无细胞蛋白质合成系统的发展历程

Figure 2 (Color online) The development history of CFPS

取的翻译机器能够调节体外蛋白质合成,不依赖于完整的细胞结构. 20 世纪 60 年代, Nirenberg 和 Matthaei^[9]利用大肠杆菌 CFPS, 破译了编码 20 种氨基酸的 64 种密码子, 取得突破性进展, 对遗传密码子的发现起到至关重要的作用, 被认为是 CFPS 开创性的工作. 当时的 CFPS 持续时间短、产量低、稳定性差, 严重抑制其发展和应用. 1973 年, Zubay^[10]建立了基于 S30 大肠杆菌提取物的 CFPS, 减少合成背景肽链

的内源 DNA 或 RNA 含量, 同时使用敲除了核酸水解酶和内源蛋白酶的菌种, 直接提高了外源蛋白的表达产量. 由于大肠杆菌提取物的获取简单易行, 且反应活性较高, Zubay 的系统逐渐被广泛使用, 并发展成为一种标准程序. 该系统只需 1~2 天即可合成所需蛋白质, 而传统的体内重组蛋白质合成周期长, 需要经过细胞收集、破碎、纯化等繁琐步骤, 获得蛋白质至少需要 1~2 周. 1988 年, Spirin 等人^[11]通过在 CFPS 中

连续添加反应所需要的能量和底物,并将反应产物或副产物透析出反应体系,使反应体系保持动态平衡,大大延长了CFPS的反应时间,提高了蛋白产量和稳定性,使得CFPS研究有了突飞猛进的发展. 2001年,罗氏(Roche)公司开发出商业化的CFPS. 同年, Shimizu等人^[12]通过添加纯化的各种转录及翻译因子、必需底物和相应的酶、tRNA等,构建了CFPS的“PURE”系统,之后由New England Biolabs公司商业化生产. “PURE”系统不同于提取物系统,由转录翻译所必需组分纯化后构成,所有添加物完全已知并浓度可控. 2009年和2013年Luo课题组^[13,14]开发了DNA和黏土两种水凝胶CFPS系统,大大提高了蛋白质的表达产量. DNA水凝胶CFPS系统相较于液相CFPS体系,蛋白质产量提高了约300倍.

1 无细胞蛋白合成系统

CFPS包含转录、翻译、蛋白折叠和能量代谢四大模块^[15],合成基本原理与体内合成系统相似^[16]:转录,即以目标蛋白质质粒DNA或线性PCR产物为模板,在RNA聚合酶及转录因子等组分作用下,转录出相应的mRNA;翻译,以mRNA为模板,利用体系中的核糖体、调控因子、氨基酸底物、tRNA、能源物质等翻译合成蛋白质.

CFPS克服了细胞系统在蛋白表达和优化纯化过程中的不足,在应用过程中展现出独特优势,主要优点包括如下3个方面: (1) 表达快速,反应周期短; (2) 反应操作简便,体积微小,模板形式多样^[17]; (3) 反应过程可控,产物多样,与其他生物技术兼容性好^[18~22].

CFPS主要分为两种系统,一种是“PURE”系统,即将转录翻译所必需组分进行纯化,添加入CFPS反应体系,所有添加物完全已知并浓度可控. 另一种是我们在本文主要介绍的细胞提取物系统,通过处理并破碎细胞,去除细胞内的基因组DNA、不溶物等,获得含有转录翻译、能量代谢等所需的必要生化组分,以此为基础进行CFPS. 如图3所示,将CFPS的生产方式、提取物种类、生产蛋白质种类及其研究领域进行了总结.

1.1 细胞提取物来源

理论上,几乎所有物种的细胞提取物都能满足构建CFPS蛋白表达的基本需求^[23],但不同物种来源

的CFPS表现差异很大. 细胞提取物根据来源不同可分为原核细胞提取物和真核细胞提取物两大类^[24],其中原核细胞提取物包括大肠杆菌提取物(*E. coli* extract)和古细菌提取物(archaeal extract)等;真核细胞提取物包括原生生物提取物(protozoan extract)、酵母提取物(yeast extract)、烟草细胞提取物(tobacco BY-2 extract)、麦胚提取物(wheat germ extract)、昆虫细胞提取物(insect cell extract)、兔网织红细胞提取物(rabbit reticulocyte lysate)、仓鼠卵巢细胞提取物(Chinese hamster ovary/CHO cells)和人类细胞系提取物(human cell lines, 例如HEK293细胞和HeLa细胞)等^[25]. 在此仅介绍4种应用相对广泛的细胞提取物并简要比较其优缺点.

(i) 大肠杆菌提取物(*E. coli* extract). 大肠杆菌CFPS目前应用最为广泛,也是研究较早的无细胞系统. 1961年,Matthaei和Nirenberg^[26]就已建立较稳定的大肠杆菌无细胞系统,但该系统的能量供应体系不完善,一旦蛋白合成所需的底物和能量消耗完或产生的副产物积累到一定浓度,将抑制蛋白的合成,停止转录翻译过程. 2007年, Kim等人^[27]将果糖-1,6-二磷酸代替葡萄糖-6-磷酸作为能量来源,优化了大肠杆菌CFPS的能量供应体系,将蛋白产率提高将近14倍,同时大大降低了试剂成本. 至今,大肠杆菌CFPS已被广泛应用于各类蛋白的生产及研究. 如Menezes等人^[28]利用该系统实现了出血性金属蛋白酶HF3前体及其非催化域的重组表达. Zhuge课题组^[29]在大肠杆菌CFPS中实现了新型抗菌肽ABP-dHC-cecropin A的高效生产. Kim课题组^[30]利用大肠杆菌CFPS合成磷脂酶,与大肠杆菌胞内表达相比,合成产量提高100多倍. 大肠杆菌CFPS具有合成产量和效率高、生产周期短、成本低等特点,成为目前比较成熟和广泛应用的系统.

(ii) 麦胚提取物(wheat germ extract). 麦胚提取物的原料是小麦胚芽,是面粉加工过程中的一种副产物. 麦胚CFPS作为一种基于真核细胞的CFPS,可对目标蛋白进行翻译后修饰. 1973年, Roberts和Paterson^[31]利用麦胚CFPS翻译了烟草花叶病毒RNA和兔球蛋白RNA. 在不断优化麦胚CFPS的研究过程中, Hunter等人^[32]发现多胺对麦胚CFPS中氨基酸的转运过程具有促进作用. Harbers^[33]将透析法引入麦胚CFPS中,有效清除了一些对蛋白合成具有抑制作用的副产物,大大延长蛋白合成的时间,提高蛋白产

图 3 (网络版彩色)无细胞生物合成系统概述: 生产方式、提取物种类、生产蛋白质种类和研究应用领域

Figure 3 (Color online) Overview of CFPS system: production formats, category of cell extracts, category of proteins produced and application fields

量. 此外, 通过构建特定的表达质粒pEU, 优化mRNA的5'和3'非翻译序列, 强化翻译过程, 使得麦胚CFPS成为一种先进的真核无细胞系统^[34,35].

(iii) 昆虫细胞提取物(insect cell extract). 昆虫细胞CFPS的主要优势是可以在不添加酶和辅因子的条件下进行转录后修饰. 用于制备提取物的昆虫细胞主要有草地夜蛾(*Spodoptera frugiperda* 21, Sf21)和粉纹夜蛾(*Trichoplusia ni*)等. 其中最常用的是Sf21, 该细胞提取物系统已被商业化制成试剂盒(Invitrogen). 昆虫细胞CFPS广泛应用于杆状病毒表达系统^[36]. 研究发现在5'端引入一段来源于杆状病毒多角体蛋白质的非翻译序列可增强系统的翻译过程, 而在3'端引入一段非翻译序列可保证其翻译效率^[37].

(iv) 兔网织红细胞提取物(rabbit reticulocyte lysate). 1958年, Schweet等人^[38]将放射性氨基酸作

为标记嵌入血红蛋白中, 首次证明了兔网织红细胞提取物可用于建立CFPS. 1976年, Pelham和Jackson^[39]发现兔网织红细胞CFPS能够翻译不同来源、不同长度的mRNA, 且翻译产物保真性较强, 是哺乳细胞无细胞系统中最有效的一种. 最初兔网织红细胞的CFPS只应用于研究真核生物的蛋白翻译过程, 经过对该系统的不断完善, 其应用范围越来越广, 涵盖蛋白微阵列技术^[40]、蛋白分子相互作用研究及蛋白筛选^[41]等领域.

这4种细胞提取物发展至今, 均有其各自的优势, 但仍存在技术问题或应用局限性, 比如产量和成本平衡控制、制备工艺简化、基因操作可行性、翻译后修饰能力. 在表1中, 我们简要比较这4种细胞提取物的优缺点.

1.2 无细胞蛋白生产方式

CFPS的生产方式逐渐演进, 主要包括以下5种.

(i) 一锅式(batch). 一锅式无细胞生产方式是将蛋白合成所需的各种能量和底物等一次性加入到系统中. 1966年, Nirenberg等人^[42]利用该反应模式研究RNA密码子并进行相关蛋白的合成. 在此反应系统中, 所有的物质和能量不能进行中间补充, 随着反应的进行, 底物和能量不断被消耗, 致使蛋白合成的持续时间很短, 蛋白质产量往往只有数μg/mL. 另外, 反应中代谢产物的积累也会影响蛋白质的合成效率.

(ii) 连续流加式(continuous-flow). 为克服一锅式无细胞生产方式的不足, Alakhov课题组^[11]提出了连续流加式无细胞反应. 该生产方式在反应过程中, 需要不断添加反应所需的物质和能量, 并定时清除反应副产物. 将这种生产方式应用于大肠杆菌和麦胚CFPS, 反应时间可延长至20 h, 且蛋白产量大大提高. 但是, 连续流加操作的繁琐性和装置的复杂

表 1 四种细胞提取物系统的优缺点
Table 1 Pros and cons of four types of cell extracts

细胞提取物	优点	缺点
大肠杆菌	蛋白产量高、提取物制备简便、遗传背景清晰, 基因操作简便、能量供应成本低	缺少翻译后修饰手段
麦胚	合成真核蛋白范围广、蛋白产量高	产量低、制备耗时长且复杂、基因操作困难
昆虫细胞	细胞易破碎且快、可进行真核特异性翻译后修饰	制备耗时长且成本高、基因操作困难
兔网织红细胞	细胞易破碎且快、可进行真核特异性翻译后修饰	蛋白产量低、动物组织处理复杂、蛋白合成范围窄、基因操作困难、背景mRNA多

性严重限制了该模式的实用性。

(iii) 连续交换式(continuous-exchange). 连续交换式无细胞生产方式也称作半连续流加方式,是针对连续流加无细胞生产方式实用性不强的问题而发展的一种简化改进模式。1996年, Kim和Choi^[43]利用该生产方式,产量相对一锅式生产方式提高了将近10倍。连续交换体系由反应液和补充液组成,并通过一层具有选择透过性的半透膜隔开。反应液部分是CFPS的核心,是基因模板进行转录和翻译的场所,合成的大分子目的蛋白在反应液中富集,反应的副产物如无机小分子通过半透膜扩散至补充液,降低其对目标蛋白合成的抑制作用。补充液部分的能量和底物小分子不断地通过半透膜为反应液部分补充物质和能量,使整个系统处于持续工作状态。

(iv) 双层相式(bilayer). 双层相式无细胞生产方式是连续交换式的简化。在双层相式系统中,补充液只是简单地覆盖在反应液上。Endo课题组^[34]发展了该生产方式,反应时间较一锅式延长了10倍,而且产量上升到了毫克级别。更重要的是,该生产方式适合于蛋白的高通量研究和生产。近年来,越来越多的新型无细胞蛋白表达方式不断涌现,例如利用中空纤维体系^[44,45],或者利用微流控芯片^[46]进行无细胞蛋白合成等。

(v) 水凝胶系统(hydrogel). 将水凝胶应用于CFPS源于对细胞类胶体特性的模拟^[47,48]。美国康奈尔大学的Luo课题组在这方面开展了系列探索性研究。DNA水凝胶作为一种新型的生物医用高分子水凝胶既具备水凝胶的一般特性,也被赋予DNA分子的独特生物特性,包括优异的分子识别特性、特异编码、分子结构完全可控、生物相容性和生物可降解性等^[49,50]。2009年, Luo课题组^[13,51]首次将DNA水凝胶系统应用于CFPS,与液相合成方法相比,蛋白质产量提高了300倍。2013年, Luo课题组^[14]研发了一种成本低廉的黏土水凝胶用于CFPS,研究表明DNA分子倾向于附着在黏土表面,抑制了体系中核酸降解酶的活性,且增强了转录和翻译过程。

2 合成重要蛋白质

蛋白质是生物体结构与功能的重要组成成分,是生命活动的主要承担者。蛋白质功能的多样性决定蛋白质应用的广泛性,越来越多的蛋白产品应用于各个领域,如蛋白药物、蛋白保健品、蛋白食品和

蛋白催化酶等。传统蛋白产品的生产方式主要包括从动物内脏或血液中分离提取、工程菌表达和化学合成。其中工程菌表达虽然实现了蛋白产品的工业化生产,但在合成过程中由于缺乏蛋白质折叠的辅助因子或胞内环境不适,无法形成正确的次级键,导致目标蛋白难以正确折叠而形成包涵体,且复性率一般都很低,有些蛋白甚至具有细胞毒性,致使细胞生长抑制或死亡。对于这些特殊的蛋白产品,使用工程菌表达无法实现工业化生产。但CFPS不依赖于完整的细胞结构,可通过对体系微环境和参与蛋白合成各因子的精确调控优化蛋白合成,克服工程菌表达的局限性。以下介绍3类重要蛋白的无细胞合成。

(i) 细胞毒性蛋白。细胞毒性蛋白是导致细胞无法正常生长甚至死亡的一类蛋白。常见的细胞毒性蛋白包括抗菌肽和膜蛋白。抗菌肽是生物体产生的一类具有抗菌活性的多肽物质,在细菌感染与治疗领域具有十分广泛的应用前景。其分子结构的两亲性使其在病原细胞膜上穿孔而形成离子通道,破坏细胞膜结构,引起胞内水溶性物质大量渗出,最终导致细胞死亡,起到抗菌效果。天然的抗菌肽虽然来源广泛,但在生物体内含量很低,提取、分离、纯化难度很大,且抗菌肽的抗菌性限制了其使用工程菌表达。CFPS不依赖于细胞结构,允许最大限度的人为调控,可高效表达抗菌肽,对于开发抗菌肽等新型药物具有不可替代的作用^[19,52]。

膜蛋白作为重要的功能蛋白,在细胞间接触、表面识别、信号转导和物质运输等方面都具有重要作用,是理想的药物靶点。但膜蛋白的体内过量表达会造成生物毒性和聚集沉淀,后续需添加表面活性剂或脂质体进行重新折叠复性,创造疏水环境。探索膜蛋白包涵体复性的条件耗时长,且操作繁琐,成功率低,是膜蛋白通过工程菌重组表达的瓶颈之一。相对于重组表达,CFPS突破了蛋白质合成与细胞活性的关系,无需考虑细胞毒性问题。同时,CFPS是条件完全可控的开放体系,能够人工提供膜蛋白正确折叠的疏水环境,可成为膜蛋白合成领域中具有应用前景的研究平台^[53,54]。例如, Wu和Swartz^[55]利用经过纯化的大肠杆菌内膜囊,成功在体外合成整合膜蛋白四环素泵(TetA)和甘露醇透性酶(MtlA),产量分别达到 (570 ± 50) $\mu\text{g/mL}$ 和 (130 ± 30) $\mu\text{g/mL}$,是体内表达量的400倍。

(ii) 易错折叠蛋白。蛋白的功能依赖于翻译后

的正确折叠. 工程菌表达蛋白时, 若胞内环境和蛋白正确折叠环境不一致, 会导致蛋白的错误折叠而功能丧失. 利用CFPS合成易错折叠蛋白, 可根据蛋白的性质优化系统环境, 引导蛋白质向正确方向进行折叠. 例如, Murray课题组^[56]将大肠杆菌提取物CFPS与活氧液体携带系统结合, 通过控制体系的氧化还原电势, 大大提高含有二硫键蛋白的产量和折叠正确率. 另外, 在蛋白折叠过程中, 蛋白分子会处于最低能量态被限制在某个状态, 富含分子伴侣的CFPS可以防止蛋白在折叠过程中形成聚集体或无活性结构, 提高折叠正确率, 并增加蛋白的可溶性表达^[57]. 浙江大学徐志南课题组^[58]利用CFPS, 通过额外添加脂类等疏水物质, 辅助通道蛋白AqpZ的正确折叠, 大大提高了AqpZ的产量, 达900 µg/mL.

(iii) 易聚集蛋白. 工程菌表达蛋白时, 蛋白的过量表达容易聚集并形成包涵体, 因此需要后续复性过程. 通过稀释或透析、外加分子伴侣或小分子添加剂, 如表面活性剂和脂质体等, 进行辅助复性, 过程繁琐. 并且, 蛋白复性的分子机理尚不明确, 不同类型的表面活性剂对于蛋白复性效果不同, 无规律可循, 增加了此类蛋白的生产成本, 限制了工业化生产. CFPS作为一种开放可控体系, 可以人工调节体系的环境, 例如, 添加表面活性剂或人工脂质体、改变pH和离子浓度、控制氧化还原电势、改变蛋白的表面带电量等, 从而达到抑制蛋白聚集的效果. Spirin课题组^[59]利用大肠杆菌CFPS, 通过添加蛋白二硫键异构酶(PDI)、分子伴侣DnaK和DnaJ、还原性和氧化性谷胱甘肽, 实现了50%(约8 mg/mL)scFv抗体的活性复苏. 此外, 蛋白生成速率过快, 无充足时间进行折叠, 也会形成蛋白聚集体, 而CFPS连续流加式和连续交换式的生产方式可在合成过程中将表达的目标蛋白移出反应体系, 降低其浓度, 避免由蛋白大量积累导致的蛋白聚集^[11].

3 蛋白高通量表达和筛选

蛋白的高通量表达和筛选是功能蛋白优化的重要研究手段, 然而细胞内蛋白表达周期长、体系封闭等特点在很大程度上限制了蛋白的高通量表达和筛选. CFPS作为一种开放体系, 有效地解决了这一问题. Sawasaki等人^[34]使用麦胚CFPS, 通过96孔板高通量筛选目标蛋白基因, 成功实现表达系统的扩大化. Brizuela课题组^[60]利用96孔板高通量检测并合成63种

绿脓杆菌蛋白, 其中51种蛋白可以成功表达, 部分蛋白可在变性条件下进行纯化, 实现单人劳动力在4 h内完成蛋白表达和纯化操作, 高效快速的体外蛋白表达纯化为大规模筛选克隆蛋白质提供了一个有效途径. Nomura课题组^[61]通过高通量CFPS合成了13364个人源蛋白, 其中77%的蛋白被证明具有生物活性, 甚至在无还原性的麦胚CFPS中合成了具有活性的含二硫键的细胞因子. Endo 和Sawasaki^[62]通过在96孔板进行蛋白高通量转录, 并在24-微量透析杯中进行翻译, 优化麦胚CFPS, 大大提高蛋白产量. 这一系列的研究证明, CFPS的高通量表达可以实现系统在基因层次和表达层次上的优化, 在筛选高性能目标蛋白和蛋白质组学研究方面同样具有巨大的应用价值, 为蛋白质工程的进一步发展提供了一个更加灵活高效的研究平台.

4 合成蛋白类药物

理论上, 蛋白类药物几乎可以用于所有发病机制清楚、病理学靶点明确的疾病治疗. 基于对蛋白质结构和功能认识的不断积累, 蛋白类药物合成领域逐渐发展成熟, 使其逐渐成为21世纪重要的预防、诊断和治疗用药, 在治疗肿瘤、自身免疫性疾病、代谢性遗传病、各种老年病及退行性疾病等方面具有重要作用. 目前, 在国际市场上, 新药研发的难度越来越大, 使得高效重组蛋白药物的开发和蛋白类药物如抗体、疫苗等产品的改性升级倍受重视. 但重组蛋白对细胞毒性较大, 易被胞内蛋白酶降解, 大大限制了基于细胞的体内表达系统在蛋白类药物开发和改性方面的应用. CFPS的开放性特点使其在解决以上蛋白质药物开发难题方面显示出巨大优势. 随着对细胞提取物、生产方式和代谢网络等方面的优化, CFPS已经突破基础研究的应用范围, 逐渐应用于蛋白药物的改性、高通量筛选和工业化生产, 在生物制药业上游研发和下游生产纯化两方面都展现出优于细胞系统的巨大潜力, 为解决蛋白质药物研发难题提供了有效的解决方案.

蛋白类药物的改造和药性优化的重要手段之一是在蛋白中引入非天然氨基酸, 用于提供额外基团, 以便对药物蛋白进行化学修饰, 赋予蛋白质新的物理化学及生物学特性, 拓展对蛋白质操作的空间. 非天然氨基酸引入技术赋予了蛋白质药物优良的药理特性, 如延长体内半衰期, 提高药效等. 然而将非天

然氨基酸掺入蛋白质仍是一项挑战,其主要障碍是促进酰胺tRNA合成酶与非天然氨基酸的识别。CFPS允许便捷地加入非天然氨基酸相关底物,实现非天然氨基酸的定点插入。例如,Pepinsky等人^[64]将 β 干扰素中蛋氨酸替换成非天然类似物并对其特定位点进行PEG修饰,使得产物在治疗多发性硬化症时在体内的存留时间延长。Schultz课题组^[65]利用詹氏甲烷球菌(*Methanococcus jannaschii*)与大肠杆菌Trna/aaRSs对的正交特性,研发出一种新的非天然氨基酸修饰蛋白的方法,为后续非天然氨基酸添加到多肽链的设计和操作奠定了基础。Szostak课题组^[66]向合成系统中添加各种非天然氨基酸,在多肽中掺入了50多种酶催化的氨酰化类似物,优化后甚至可以将高达13种不同的非天然氨基酸掺入一条多肽链中,实现对蛋白类药物的目的性改造。Swartz课题组^[67]利用CFPS将带有炔基的非天然氨基酸引入鞭毛蛋白中,并将此鞭毛蛋白负载于乙肝病毒的类病毒颗粒表面,其对TLR5的刺激效应提高了近10倍。

5 总结与展望

相较于基于细胞的生物合成系统,CFPS不依赖于细胞结构,突破传统生物技术的局限,在蛋白质合成方面具有独特的优势。同时,CFPS也为探索生命起源和进化的化学基础提供了很好的实验模型,是合成生物学的前沿方向^[68]。世界各国科研工作者已经开展了一些基础性研究工作,深化了对CFPS的认知,为其工业化奠定了一定的理论基础。虽然商业化

无细胞蛋白质表达的试剂盒为大学等科研单位提供了实验原料,但用户群相对其他成熟的生化试剂非常小众,因此价格昂贵;且批次间蛋白活性存在一定差异。此外,无细胞生物合成领域在国际范围内尚处初级阶段,大量的基础研究亟待开展,与药物合成相关的应用研究也需要进一步推进。在国内,无细胞生物合成领域逐渐受到科技管理部门和科学家的重视,一些前沿研究工作正在开展。目前,CFPS面临的直接挑战是缺乏一个能够合成具有生物活性蛋白质的通用平台,在反应体系中高效整合所需的生物代谢途径及网络,将体内和体外生物系统的复杂度差距缩小。以大肠杆菌为基础的原核细胞CFPS发展较为成熟,蛋白产量和生产成本都得到了较好的控制,但是糖基化等翻译后修饰手段的缺失导致其在医药工业中的应用仍然受限。而真核细胞CFPS表达量较低,在一锅法生产方式中达不到mg/mL的级别,远远不能满足产业化要求,甚至某些人源蛋白质的产量低于体内重组合成产量。真核CFPS的内部调控等机制也远远比原核复杂,研究难度相应增大。当前CFPS仍着重于单个基因的合成研究,在CFPS中实现多基因的功能性表达可以研究一系列相关蛋白的相互作用、构建代谢通路和构建药物靶点高通量筛选模型,仍存在较大挑战。结合中国的国情和特色,无细胞生物合成在军民融合的蛋白质药物合成方面将大有作为,未来可发展成为引领制药行业,甚至促使行业颠覆发展的高新生物技术。

参考文献

- Smanski M J, Zhou H, Claesen J, et al. Synthetic biology to access and expand nature's chemical diversity. *Nat Rev Microbiol*, 2016, 14: 135–149
- Church G M, Elowitz M B, Smolke C D, et al. Realizing the potential of synthetic biology. *Nat Rev Mol Cell Biol*, 2014, 15: 289–294
- Cameron D E, Bashor C J, Collins J J. A brief history of synthetic biology. *Nat Rev Microbiol*, 2014, 12: 381–390
- Wang Y H, Wei K Y, Smolke C D. Synthetic biology: Advancing the design of diverse genetic systems. *Annu Rev Chem Biomol*, 2013, 4: 69–102
- Nielsen J, Larsson C, Van M A, et al. Metabolic engineering of yeast for production of fuels and chemicals. *Curr Opin Biotechnol*, 2013, 24: 398–404
- Hodgman C E, Jewett M C. Cell-free synthetic biology: Thinking outside the cell. *Metabolic Eng*, 2012, 14: 261–269
- Perez J G, Stark J C, Jewett M C. Cell-free synthetic biology: Engineering beyond the cell. *Cold Spring Harb Perspect Biol*, 2016, 8: a023853
- Hoagland M B, Stephenson M L, Scott J F, et al. A soluble ribonucleic acid intermediate in protein synthesis. *J Biol Chem*, 1958, 231: 241–257
- Nirenberg M W, Matthaei J H. The dependence of cell-free protein synthesis in *E. coli* upon naturally occurring or synthetic polynucleotides. *Proc Natl Acad Sci USA*, 1961, 47: 1588–1602

- 10 Zubay G. *In vitro* synthesis of protein in microbial systems. *Annu Rev Genet*, 1973, 7: 267–287
- 11 Spirin A S, Baranov V I, Ryabova L A, et al. A continuous cell-free translation system capable of producing polypeptides in high yield. *Science*, 1988, 242: 1162–1164
- 12 Shimizu Y, Inoue A, Tomari Y, et al. Cell-free translation reconstituted with purified components. *Nat Biotechnol*, 2001, 19: 751–755
- 13 Park N, Um S H, Funabashi H, et al. A cell-free protein-producing gel. *Nat Mater*, 2009, 8: 432–437
- 14 Yang D, Peng S, Hartman M R, et al. Enhanced transcription and translation in clay hydrogel and implications for early life evolution. *Sci Rep*, 2013, 3: 3165
- 15 Rosenblum G, Cooperman B S. Engine out of the chassis: Cell-free protein synthesis and its uses. *FEBS Lett*, 2014, 588: 261–268
- 16 Whittaker J W. Cell-free protein synthesis: The state of the art. *Biotechnol Lett*, 2013, 35: 143–152
- 17 Ahn J H, Chu H S, Kim T W, et al. Cell-free synthesis of recombinant proteins from PCR-amplified genes at a comparable productivity to that of plasmid-based reactions. *Biochem Biophys Res Commun*, 2005, 338: 1346–1352
- 18 Focke P J, Hein C, Hoffmann B, et al. Combining *in vitro* folding with cell free protein synthesis for membrane protein expression. *Biochemistry*, 2016, 55: 4212–4219
- 19 Martemyanov K A, Shirokov V A, Kurnasov O V, et al. Cell-free production of biologically active polypeptides: Application to the synthesis of antibacterial peptide cecropin. *Protein Expr Purif*, 2001, 21: 456–461
- 20 Berrier C, Park K H, Abes S, et al. Cell-free synthesis of a functional ion channel in the absence of a membrane and in the presence of detergent. *Biochemistry*, 2004, 43: 12585–12591
- 21 Kim D M, Swartz J R. Efficient production of a bioactive, multiple disulfide-bonded protein using modified extracts of *Escherichia coli*. *Biotechnol Bioeng*, 2004, 85: 122–129
- 22 Gan R, Yamanaka Y, Kojima T, et al. Microbeads display of proteins using emulsion pcr and cell-free protein synthesis. *Biotechnol Prog*, 2008, 24: 1107–1114
- 23 Katzen F, Chang G, Kudlicki W. The past, present and future of cell-free protein synthesis. *Trends Biotechnol*, 2005, 23: 150–156
- 24 Zemella A, Thoring L, Hoffmeister C, et al. Cell-free protein synthesis: Pros and cons of prokaryotic and eukaryotic systems. *Chembiochem*, 2015, 16: 2420–2431
- 25 Quast R B, Mrusek D, Hoffmeister C, et al. Cotranslational incorporation of non-standard amino acids using cell-free protein synthesis. *FEBS Lett*, 2015, 589: 1703–1712
- 26 Matthaei H, Nirenberg M W. The dependence of cell-free protein synthesis in *E. coli* upon RNA prepared from ribosomes. *Biochem Biophys Res Commun*, 1961, 4: 404–408
- 27 Kim T W, Keum J W, Oh I S, et al. An economical and highly productive cell-free protein synthesis system utilizing fructose-1,6-bisphosphate as an energy source. *J Biotechnol*, 2007, 130: 389–393
- 28 Menezes M C, Imbert L, Kitano E S, et al. Recombinant expression of the precursor of the hemorrhagic metalloproteinase HF3 and its non-catalytic domains using a cell-free synthesis system. *Amino Acids*, 2016, 48: 2205–2214
- 29 Zhang J, Movahedi A, Wei Z, et al. High-level SUMO-mediated fusion expression of ABP-dHC-cecropin A from multiple joined genes in *Escherichia coli*. *Anal Biochem*, 2016, 509: 15–23
- 30 Lim H J, Park Y J, Jang Y J, et al. Cell-free synthesis of functional phospholipase A1 from *Serratia* sp. *Biotechnol Biofuels*, 2016, 9: 159
- 31 Roberts B E, Paterson B M. Efficient translation of tobacco mosaic virus RNA and rabbit globin 9S RNA in a cell-free system from commercial wheat germ. *Proc Natl Acad Sci USA*, 1973, 70: 2330–2334
- 32 Hunter A R, Farrell P J, Jackson R J, et al. The role of polyamines in cell-free protein synthesis in the wheat-germ system. *Eur J Biochem*, 1977, 75: 149–157
- 33 Harbers M. Wheat germ systems for cell-free protein expression. *FEBS Lett*, 2014, 588: 2762–2773
- 34 Sawasaki T, Ogasawara T, Morishita R, et al. A cell-free protein synthesis system for high-throughput proteomics. *Proc Natl Acad Sci USA*, 2002, 99: 14652–14657
- 35 Fogeron M L, Jirasko V, Penzel S, et al. Cell-free expression, purification, and membrane reconstitution for NMR studies of the non-structural protein 4B from hepatitis C virus. *J Biomol NMR*, 2016, 65: 87–98
- 36 Tarui H, Murata M, Tani I, et al. Establishment and characterization of cell-free translation/glycosylation in insect cell (*Spodoptera frugiperda* 21) extract prepared with high pressure treatment. *Appl Microbiol Biotechnol*, 2001, 55: 446–453
- 37 Smith G E, Summers M D, Fraser M J. Production of human beta interferon in insect cells infected with a baculovirus expression vector. *Mol Cell Biol*, 1983, 3: 2156–2165
- 38 Schweet R, Lamfrom H, Allen E. The synthesis of hemoglobin in a cell-free system. *Proc Natl Acad Sci USA*, 1958, 44: 1029–1035
- 39 Pelham H R, Jackson R J. An efficient mRNA-dependent translation system from reticulocyte lysates. *Eur J Biochem*, 1976, 67: 247–256
- 40 Oleinikov A V, Gray M D, Zhao J, et al. Self-assembling protein arrays using electronic semiconductor microchips and *in vitro* translation. *J*

Proteome Res, 2003, 2: 313–319

- 41 Shao J, Irwin A, Hartson S D, et al. Functional dissection of cdc37: Characterization of domain structure and amino acid residues critical for protein kinase binding. *Biochemistry*, 2003, 42: 12577–12588
- 42 Nirenberg M, Caskey T, Marshall R, et al. The RNA code and protein synthesis. *Cold Spring Harb Symp Quant Biol*, 1966, 31: 11–24
- 43 Kim D M, Choi C Y. A semicontinuous prokaryotic coupled transcription/translation system using a dialysis membrane. *Biotechnol Prog*, 1996, 12: 645–649
- 44 Yamamoto Y, Sugimoto S, Shen X C, et al. In-place condensation of reaction mixture using hollow fiber membrane reactors improved productivity of cell-free protein synthesis. *Biochem Eng J*, 1999, 3: 151–155
- 45 Nakano H, Shinbata T, Okumura R, et al. Efficient coupled transcription/translation from PCR template by a hollow-fiber membrane bioreactor. *Biotechnol Bioeng*, 1999, 64: 194–199
- 46 Hahn G H, Asthana A, Kim D M, et al. A continuous-exchange cell-free protein synthesis system fabricated on a chip. *Anal Biochem*, 2007, 365: 280–282
- 47 Trevors J T, Pollack G H. Hypothesis: The origin of life in a hydrogel environment. *Prog Biophys Mol Biol*, 2005, 89: 1–8
- 48 Moeendarbary E, Valon L, Fritzsche M, et al. The cytoplasm of living cells behaves as a poroelastic material. *Nat Mater*, 2013, 12: 253–261
- 49 Yang D, Campolongo M J, Nhi Tran T N, et al. Novel DNA materials and their applications. *Wires Nanomed Nanobiol*, 2010, 2: 648–669
- 50 Yang D, Hartman M R, Derrien T L, et al. DNA materials: Bridging nanotechnology and biotechnology. *Acc Chem Res*, 2014, 47: 1902
- 51 Kahn J S, Ruiz R C H, Sureka S, et al. DNA microgels as a platform for cell-free protein expression and display. *Biomacromolecules*, 2016, 17: 2019–2026
- 52 Martemyanov K A, Spirin A S, Gudkov A T. Direct expression of PCR products in a cell-free transcription/translation system: Synthesis of antibacterial peptide cecropin. *FEBS Lett*, 1997, 414: 268–270
- 53 Hein C, Henrich E, Orbán E, et al. Hydrophobic supplements in cell-free systems: Designing artificial environments for membrane proteins. *Eng Life Sci*, 2014, 14: 365–379
- 54 Klammt C, Lohr F, Schäfer B, et al. High level cell-free expression and specific labeling of integral membrane proteins. *Eur J Biochem*, 2004, 271: 568–580
- 55 Wu J J, Swartz J R. High yield cell-free production of integral membrane proteins without refolding or detergents. *Biochim Biophys Acta*, 2008, 1778: 1237–1250
- 56 Zawada J F, Yin G, Steiner A R, et al. Microscale to manufacturing scaleup of cellfree cytokine production—A new approach for shortening protein production development timelines. *Biotechnol Bioeng*, 2011, 108: 1570–1578
- 57 Yin G, Swartz J R. Enhancing multiple disulfide bonded protein folding in a cell-free system. *Biotechnol Bioeng*, 2004, 86: 188–195
- 58 Zhang X, Lian J, Kai L, et al. Enhanced functional expression of aquaporin Z via fusion of in situ cleavable leader peptides in *Escherichia coli* cell-free system. *Enzyme Microb Technol*, 2014, 55: 26
- 59 Ryabova L A, Desplancq D, Spirin A S, et al. Functional antibody production using cell-free translation: Effects of protein disulfide isomerase and chaperones. *Nat Biotechnol*, 1997, 15: 79–84
- 60 Murthy T V S, Wu W L, Qiu Q Q, et al. Bacterial cell-free system for high-throughput protein expression and a comparative analysis of *Escherichia coli* cell-free and whole cell expression systems. *Protein Expr Purif*, 2004, 36: 217–225
- 61 Goshima N, Kawamura Y, Fukumoto A, et al. Human protein factory for converting the transcriptome into an *in vitro*-expressed proteome. *Nat Methods*, 2008, 5: 1011–1017
- 62 Endo Y, Sawasaki T. High-throughput, genome-scale protein production method based on the wheat germ cell-free expression system. *Biotechnol Adv*, 2003, 21: 695–713
- 63 Pepinsky R B, LePage D J, Gill A, et al. Improved pharmacokinetic properties of a polyethylene glycol-modified form of interferon- β -1a with preserved *in vitro* bioactivity. *J Pharmacol Exp Ther*, 2001, 297: 1059–1066
- 64 Noren C I, Anthony-Cahill S I, Griffith M C, et al. A General method for site-specific incorporation of unnatural amino acids. *Science*, 1989, 249: 182
- 65 Hartman M C, Josephson K, Lin C W, et al. An expanded set of amino acid analogs for the ribosomal translation of unnatural peptides. *PLoS One*, 2007, 2: e972
- 66 Lu Y, Welsh J P, Chan W, et al. *Escherichia coli*-based cell free production of flagellin and ordered flagellin display on virus-like particles. *Biotechnol Bioeng*, 2013, 110: 2073–2085
- 67 Jia H, Heymann M, Bernhard F, et al. Cell-free protein synthesis in micro compartments: Building a minimal cell from biobricks. *New Biotechnol*, 2017, doi: 10.1016/j.nbt.2017.06.014

Summary for “体外合成生物学: 无细胞蛋白合成系统研究进展”

In vitro synthetic biology: Cell-free protein synthesis

LIU Yang[†], GUO XiaoCui[†], GENG JinHui, JIAO Yi, HAN JinPeng, ZHANG ZhiKun, ZHOU Xiao & YANG DaYong^{*}

School of Chemical Engineering and Technology, Key Laboratory of Systems Bioengineering (Ministry of Education), Collaborative Innovation Center of Chemical Science and Engineering (Tianjin), Tianjin University, Tianjin 300350, China

[†] Equally contributed to this work

^{*} Corresponding author, E-mail: dayong.yang@tju.edu.cn

Synthetic biology is an emerging interdisciplinary field covering the subjects of biology, chemistry and engineering, which promotes the progress from understanding life to creating living systems. As one of important forefront directions of synthetic biology, *in vitro* synthetic biology (IVSB) focuses on studying the phenomenon and rules of life activities in testing tubes, not relying on living organisms. Cell-free protein synthesis (CFPS) is an effective research platform for IVSB, which utilizes exogenous DNA or mRNA as genetic information for synthesizing protein *in vitro*, instead of traditional protein expression in cells. CFPS possesses superior advantages over cell-based systems including convenience, simplicity, rapidness, excellent controllability and green economy. Moreover, CFPS is a robust research model for exploration of life evolution and creation of artificial living system.

Since 1958, it has been demonstrated for the first time that proteins could be synthesized from cell extract; and afterwards, CFPS has emerged to not only a promising platform for proteomics and pharmaceuticals, but also a significant tool for revelation of chemical basis of cells and exploration of origins of life. Cell extracts contained basic transcription and translation machineries, energy regeneration substrates, amino acids, nucleotides and cofactors. A variety of cell extracts have been created from different cell sources such as *Escherichia coli* (*E.coli*), wheat germ, yeast and rabbit reticulocyte. *E.coli* was the earliest prokaryotic system and most widely used due to its high production yield, short production cycle and low cost. The protein yields of *E. coli* system could reach up to several milligrams per milliliter of reaction. Eukaryotic systems such as wheat germ and rabbit reticulocyte although less productive, provided more robust platform for studying protein functions, in particular for post-translationally modified proteins.

Different production formats of CFPS have been developed and classified according to how the reactions were fed, including batch, continuous-flow, continuous-exchange, bilayer and hydrogel.

Batch was the simplest format, but the key limitation was the short lifetimes and the consequent low yield. This was primarily because of the rapid depletion of the high-energy phosphate pool, leading to accumulation of free phosphates and further inhibition of protein synthesis. Continuous-flow overcame the problem in batch format because of the continuous supply of energy and substrates and the continuous removal of the by-products. As a result, the reaction time could be extended to 20 hours, and the product yield was increased by two orders of magnitude in comparison with batch format. In continuous-exchange format, passive exchange of substrates and by-products further extended the reaction lifetime. However, continuous formats were not easily applicable to high-throughput processes, in which miniaturization and automation were required. Bilayer format was simple, efficient and high-throughput-friendly. Very recently, hydrogel formats have been developed because of the gel niche resembling the physicochemical nature of cells, subsequently leading to high yield.

One of the promising applications of CFPS was to synthesize biologically active proteins. CFPS could produce proteins that were difficult or impossible to express with cell-based systems, such as membrane proteins, cell-toxic proteins, isotope-labeling proteins and proteins with unnatural amino acids incorporated. The open nature and high versatility of the CFPS platform enabled high-throughput tools for genomics and proteomics, allowing for biopharmaceutical applications. Through the high-throughput screening of proteins in CFPS the efficacy of protein drugs could be significantly improved by addition of unnatural amino acids, which provided excellent pharmacological properties and prolonged the half-life of drugs in plasma.

In this review, we therefore summarize the development history, compositions, categories and production formats of CFPS, as well as applications in the synthesis of functional proteins, high throughput synthesis and screening of protein, and production of protein drugs.

synthetic biology, cell-free protein synthesis, high-throughput protein synthesis, protein drugs

doi: 10.1360/N972017-00419