蛋白质之活细胞探究

昌增益

北京大学生命科学学院, 蛋白质及植物基因研究国家重点实验室, 北京大学蛋白质科学中心, 北京 100871

E-mail: changzy@pku.edu.cn

2021-08-19 收稿, 2021-09-23 修回, 2021-09-27 接受, 2021-09-30 网络版发表

国家杰出青年科学基金(39725008)、国家自然科学基金(31971189, 31670775, 31470766)、国家重点基础研究发展计划(2012CB917300)和启东-北大生命科学学院创新基金资助

摘要 人类对蛋白质的认识过程经历了200多年, 但大部分研究都在活细胞体之外开展. 随着认识生命之手段和方法的不断更新与改进, 科学家更加关注生命体内发生的事件. 因此, 从体外到体内, 从定性到定量将成为蛋白质研究领域的新趋势. 本文简要总结了活细胞内蛋白质研究的现状以及我们实验室的部分代表性研究成果. 首先介绍已经建立的主要方法体系, 包括绿色荧光蛋白标记法、活细胞核磁共振波谱法、非天然氨基酸标记法等. 之后, 重点介绍非天然氨基酸标记方法体系, 以及我们实验室对这一技术体系进行的改进, 包括: 创建用于在目标蛋白质中引入非天然氨基酸Bpa的LY928大肠杆菌菌株; 建立将非天然氨基酸随机引入目标蛋白质的任何氨基酸残基位点的方法体系; 建立直接在大肠杆菌基因组中对目标蛋白质编码基因进行碱基点突变的方法体系; 开发单块胶可分析384种蛋白质样品的高通量聚丙烯酰胺凝胶电泳技术. 为说明这种通过非天然氨基酸引入而开展活细胞内蛋白质研究体系的有效性, 列举了我们实验室所获得的部分代表性成果: 在研究细菌分裂关键蛋白质FtsZ动态组装机制时意外发现的、存在于停止分裂细胞中的、可逆亚细胞结构——复苏延迟体(regrowth delay body); 揭示细菌ATP合酶复合体在细胞处于不同条件下实施独特的“换挡”活性调节机制; 揭示一种参与细菌外膜蛋白新生肽链转运的超级蛋白质复合体的存在; 揭示抗酸分子伴侣蛋白HdeA在极端酸性细胞条件下通过丧失其三维空间结构而保护包括中性分子伴侣蛋白在内的底物蛋白质的独特机制; 揭示一直被认为具有双功能的DegP蛋白主要发挥蛋白质水解酶功能, 不具明显的分子伴侣功能; 揭示活细胞中小分子热休克蛋白响应温度变化而发生层级性结构和活性变化. 最后, 作者通过提出一系列有待回答的科学问题对在活细胞中探究蛋白质的前景进行了展望.

关键词 非天然氨基酸介导的蛋白质活细胞光交联, 高通量聚丙烯酰胺凝胶电泳, 复苏延迟体, ATP合酶, 蛋白质生成, 抗酸分子伴侣蛋白HdeA

对被称为蛋白质的这类生命物质的认识, 大约开始于1750年. 最初, 它们作为存在于植物与动物组织及体液中的为人类和动物的必须食物成分而被重视^[1]. 那时, 它们之所以被归为同一类生命物质, 是因为它们具有在一定条件下发生聚集这一共同特征(如鸡蛋蛋白加热后、动物乳汁加酸后以及血液加热后都发生肉眼可见的聚集现象). 通过对这类物质的元素组成进行系

统分析之后, 瑞士化学家Jacob Berzelius和荷兰医生Gerard Mulder于1838年将这类物质命名为“protein”, 认为它们是植物为动物所准备的、“占据主要地位的”营养物质^[1,2]. 后来, 随着显微镜的使用, 以及细胞是动、植物基本组成单位这一理论的提出, 人们开始探究细胞内被称为原生质的一类组分的特征, 并逐渐认识到蛋白质是细胞的重要组分. 现在, 人们清楚地认识

引用格式: 昌增益. 蛋白质之活细胞探究. 科学通报, 2021, 66: 4524–4538

Chang Z Y. Exploring proteins in living cells (in Chinese). Chin Sci Bull, 2021, 66: 4524–4538, doi: [10.1360/TB-2021-0845](https://doi.org/10.1360/TB-2021-0845)

到,蛋白质是生命活动的直接执行者,几乎参与生命的所有过程。而且,存在于细胞内的成千上万种结构和功能各异的蛋白质分子皆由基因编码,蛋白质演化是生物演化的关键内容。

就像对所有生命分子的认识过程一样,我们对蛋白质的认识也是从体外开始的。从早期对蛋白质的元素组成和化学构建单位组成的认识,到之后对其共价和非共价结构、不可逆和可逆共价修饰、功能、作用机制、相互作用、生成机制、活性调节等方面的认识,从有序三维空间结构的物理测定到内在无序空间结构的发现,从静态晶体结构到动态溶液结构的认识等,我们从不断深入的对蛋白质这种复杂生命机器的大量体外研究中,获得了目前有关蛋白质的丰富知识。

然而,与体外研究所获得的丰富知识相比,我们对活细胞中蛋白质的认识却相当肤浅,包括对其在细胞内发生的很多过程,如蛋白质的生成(包括跨膜转运、折叠、组装及质量控制等)、每种蛋白质在体内所发挥的功能及其分子机制、蛋白质的动态空间结构,以及蛋白质分子之间的动态相互作用等。这些过程极为精致巧妙,又复杂高效,很难通过体外研究获得深刻认识。尽管人们早已意识到对蛋白质开展活细胞探究很重要,但目前,无论就理论体系还是方法体系而言,要对蛋白质进行活细胞研究仍充满挑战^[3,4]。本文拟结合我们实验室所开展的蛋白质活细胞研究经验和成果,对该领域的部分进展进行总结。

1 蛋白质活细胞探究需建立更为有效的方法体系

生命科学研究的总体发展趋势是:从体外到体内、从定性到定量。然而,与体外蛋白质研究可供利用的多种多样物理化学研究手段相比,可用于对活细胞中(体内)蛋白质进行直接研究的手段仍旧非常有限,这阻碍了深入研究的有效开展。

若要对活细胞中某种目标蛋白质进行观察,就必须使之在细胞中的含量足够高,同时还必须将某种容易被检测的化学探针通过一定的途径引入到目标蛋白质分子上。然而,为了认识目标蛋白质在活细胞中的真实特征,目标蛋白质的含量和时空特征又需尽量与内源性目标蛋白质接近。这样的矛盾也反映了为什么在技术上难以对活细胞中的蛋白质开展研究。

就活细胞中的蛋白质而言,人们希望了解的内容很多。比如,就一种特定目标蛋白质而言,它的新生肽

链如何通过特异的跨膜转运而抵达特定亚细胞部位?之后如何折叠形成其三维结构?这种亚细胞定位是否会随细胞条件的变化而变化?它在细胞中折叠形成的三维结构与我们在体外所观察到的是否完全一致?这种非共价三维结构如何随细胞条件的改变而改变?它在细胞中是否发生可逆或不可逆的共价修饰?它在细胞中与什么蛋白质结合组装成稳定的四级结构或发生瞬时性的相互作用?每种蛋白质分子如何发挥其生物学功能?异常折叠的蛋白质分子或者是遇到胁迫条件时,空间结构遭破坏的蛋白质分子如何通过蛋白质质量控制体系被重新折叠成具有生物活性的分子或被降解成氨基酸?

一般而言,蛋白质分子可以依据环境条件提供的信息而调节自己的行为。它们“知道”自己应完成什么生命任务,如何“见机行事”;它们通过相互协作完成某些重要的生命功能;同时,每个蛋白质分子也都面临“生老病死”的生命周期,在合适的时候需被动或主动地结束自己的“生命”以维持蛋白质群体的“活力”等。

要全面认识蛋白质,就必须在活细胞中对它们进行探究,尽管开展这方面的工作充满挑战。下面先总结开展蛋白质活细胞研究的几种相对较为有效的方法体系。

1.1 被荧光蛋白融合标记的某些目标蛋白质分子在活细胞中的身影可通过荧光显微镜追踪

一个活细胞中可能同时存在成千上万种不同的蛋白质分子,每一种蛋白质分子都具有不同的结构特征,发挥不同的生物学功能。我们需要知道它们各自在细胞的何处“安家”,与细胞中的何种蛋白质发生接触,当细胞的条件发生改变时,它是否会“搬家”,是否会结交“新朋友”,是否有新的蛋白质个体诞生或消亡等。这些都是我们认识目标蛋白质在细胞内所形成的结构及所发挥功能的重要内容。绿色荧光蛋白(green fluorescent protein, GFP)的发现、改造和应用,为探究活细胞中原本不可见的目标蛋白质带来了新契机,提供了新手段。

绿色荧光蛋白由日裔美国科学家Shimomura等人^[5]于20世纪60年代初作为水母发光蛋白(aequorin)的伴随蛋白质被发现。大约30年后,Chalfie等人^[6]发现,这种来自水母的GFP竟然可以在大肠杆菌或线虫细胞中有效重组表达,无需添加任何额外底物或辅助因子即可产生具有荧光的GFP分子。后来的研究表明,在氧气存在的条件下,它依赖自身的催化能力将3个相邻氨基

酸残基($^{65}\text{Ser-Tyr-Gly}^{67}$)经过环化、脱氢等作用形成一种发色基团^[7]。因此,被重组表达于任何类型细胞中的GFP都可在有氧条件下自发产生荧光基团,并在紫光或蓝光照射时发出绿色荧光。而且其荧光基团因位于蛋白质分子内部而被保护,从而相对比较稳定,不易发生淬灭。

后来,华裔美国科学家Tsien(钱永健)^[7]对该蛋白质分子进行了大量的遗传工程改造,制造出各种颜色不同的荧光蛋白质。20世纪90年代,GFP被开发作为一种有效工具用于标示活细胞内特定的目标蛋白质,以观察其细胞定位等特征。通过DNA重组技术将编码GFP的DNA序列与编码目标蛋白质的DNA序列以合适的方式进行遗传融合,经过转录和翻译后,目标蛋白质即被荧光蛋白质“标记”,从而可以通过荧光显微镜等对其在活细胞内的行为实施有效监测^[8]。通过利用GFP的一系列具有不同光学性质的(如产生红色、黄色等荧光)突变体,蛋白质之间所发生的相互作用可以在活细胞中被直接实时表征(如通过显微镜共定位观察或荧光共振能量转移分析等)^[7]。

然而,这种在目标蛋白质分子上融合另外一个蛋白质分子的做法并非理想,因为它可能改变某些目标蛋白质在细胞内的结构和行为特征等。而且,如果被GFP标记的目标蛋白质分子含量低且不集中存在,而是分散存在的话,也会因为荧光强度不够而难以通过荧光显微镜进行有效观察。

1.2 利用改进的核磁共振光谱学技术尝试观察活细胞内特定蛋白质的三维结构

除了利用荧光蛋白标示以观察目标蛋白在活细胞中的定位等工作之外,人们进行的另一方面尝试是通过核磁共振(nuclear magnetic resonance, NMR)光谱学技术直接观察活细胞中蛋白质的三维空间结构。尽管这方面成功的工作还很有限,但也值得提及。为此,除了需对核磁共振光谱学技术的硬件和软件予以改进之外,还有一个基本的条件是,特定目标蛋白质必须以高浓度并被能产生核磁共振信号的特定同位素(如 ^2H 、 ^{13}C 、 ^{15}N 等)所标记的状态存在于活细胞中^[9,10]。

对于像大肠杆菌这样的模式细菌细胞而言,以上对目标蛋白质的要求通过努力也许可以被基本满足^[9,10]。但对于高等真核细胞而言,即使在培养的细胞内也难以实现。有人试图将高浓度、被特定同位素标记的蛋白质注入人体积比较大的卵母细胞中,然后再

利用NMR技术对引入的蛋白质结构分析^[9,10]。但这样获得的蛋白质结构信息,其价值有限,因为这样引入的蛋白质其实是经过体外处理后被放置于细胞中,这更像是一种体外研究体系,并未真实反映天然存在于细胞中的蛋白质的结构。

即使是活的细菌细胞,在核磁共振仪中生存的时间最多就几个小时(真核细胞更短)。而利用核磁共振光谱学技术测定蛋白质结构时一般需花费几天的时间来收集数据。用这样的所谓细胞内核磁共振(in-cell NMR)技术试图测定活细胞中天然存在的蛋白质的三维结构,显然还存在诸多挑战。

1.3 通过遗传方法引入的非天然氨基酸探针使对目标蛋白质在活细胞中的研究更接近真实状况

与上述融合荧光蛋白方法不同的是,引入非天然氨基酸方法只是将目标蛋白质分子中原有的一种特定氨基酸替换为一种大小与之相近的非天然氨基酸,而且是通过遗传学方法引入(而非像利用核磁共振光谱学对细胞中的蛋白质进行结构研究中那样,将标记的目标蛋白质“注射入”细胞中),因此干扰最小,能更加真实地反映活细胞中蛋白质的状态。也就是说,引入了非天然氨基酸的目标蛋白质在细胞内的定位、结构、功能和调节特性与内源野生型蛋白质非常接近,特别是经过一系列改进后,这已经成为一种颇具威力的蛋白质活细胞探究方法体系^[11]。

最初,这种具有某种物理化学特性的非天然氨基酸只能通过固相多肽合成方法被引入肽段中^[12]。而后,Peter Schultz实验室^[13]利用体外转录翻译系统,将拟引入的非天然氨基酸化学连接到抑制型tRNA分子的氨基酸臂上,同时将目标基因中编码拟替换氨基酸残基的密码子改变成终止密码子TAG,就可以有效地将非天然氨基酸引入到特定目标蛋白质的特定氨基酸残基位点。但这样的标记蛋白质只可用于开展体外研究,用途有限。后来,Schultz实验室将该方法体系再次升级,获得了对所引入非天然氨基酸高度特异的所谓“正交”氨基酰-tRNA合成酶和“正交”tRNA(这里“正交”的意思是,它们只识别特定的非天然氨基酸,不识别细胞中的20种天然氨基酸)。这就使得通过遗传学方法将特定非天然氨基酸引入到活细胞中特定目标蛋白质的特定氨基酸残基位置成为可能^[14],因而创建出一种全新的活细胞蛋白质研究路径^[11]。

通过遗传学方法将非天然氨基酸引入特定蛋白质

的流程如图1(a)所示^[11,14]。首先,需将特定目标蛋白质的编码基因中对应于拟替换氨基酸残基的密码子(目标蛋白质的编码基因一般由表达质粒携带)通过DNA核苷酸定点突变技术替换成终止密码子TAG(转录后在信使RNA中对应的碱基序列为UAG)。然后在细胞中同时引入另外一种表达质粒,它既产生可识别终止密码子UAG的特异“正交”tRNA(它携带的反密码子的碱基序列是CUA),也产生一种催化将这种正交tRNA分子与特定非天然氨基酸连接在一起的“正交”氨基酰-tRNA合成酶。这样加入培养基中的非天然氨基酸(如常用的对-苯甲酰基-L-苯丙氨酸, p-benzoyl-L-phenylalanine, 简称为Bpa)在进入这样的细胞后可被引入到目标蛋白质中。

多种具有不同物理化学特性的非天然氨基酸,比如带有荧光基团或像上述Bpa这样的能被紫外光激活产生自由基的非天然氨基酸,可以通过这种方法被引入目标蛋白质分子中。不同类型非天然氨基酸的引入可以揭示活细胞中蛋白质的不同行为特征^[11]。

2 通过非天然氨基酸介导的光交联分析可有效探究活细胞中蛋白质的动态相互作用

通过在目标蛋白质分子中引入可介导光诱发的蛋白质分子之间共价交联的非天然氨基酸,可以有效分析活细胞中所发生的各类蛋白质相互作用(图1(b))。这类非天然氨基酸在通常条件下是“惰性”的,只有当细胞被合适波长的紫外线照射时,才形成化学反应能力

极强的自由基,与邻近的碳原子等反应形成共价键。因此,可以将处于一定条件下的活细胞中目标蛋白质所参与的各类相互作用通过共价交联方式予以固定。此后,细胞可以被破碎,在活细胞中与目标蛋白质发生相互作用的蛋白质可以被鉴定:通过免疫印迹分析或质谱分析^[15]。

激活非天然氨基酸需尽量利用相对比较温和的紫外光(即波长相对较长、能量相对较低的)短时间照射活细胞即可,对细胞的损伤相对较小。比如,通过利用波长为365 nm的紫外光照射细胞几分钟即可激活引入到蛋白质分子中的非天然氨基酸Bpa,不会造成明显的损伤。当然,为了有效捕获短时间内发生的瞬时性蛋白质相互作用事件,我们也可以通过利用高能量紫外光进行短时间(如15 s)照射来激活非天然氨基酸^[16]。

经过多年努力,这种最初在细菌中建立的方法体系,通过获得合适的正交tRNA和正交氨基酰-tRNA合成酶这两种分子,并使它们在真核细胞中有效表达而开始被应用于对真核生物蛋白质的研究之中了。具体而言,这种活细胞蛋白质非天然氨基酸标记方法已经在酵母细胞^[17]、体外培养的哺乳动物细胞系^[18]、甚至是线虫这样的简单动物个体中^[19]得到了初步建立^[11,15,20]。期盼这些方法能被生物学家有效利用,以探究蛋白质分子在活的真核生物细胞内的行为规律、所发挥的功能,以及作用机制等方面,特别是针对那些通过基因组测序所揭示的未知蛋白质而言。无疑这将有有力促进在原核和真核细胞中蛋白质研究从体外(*in vi-*

图1 (网络版彩色)通过遗传学方法在特定目标蛋白质的特定氨基酸残基位点引入非天然氨基酸Bpa(a)及利用光交联非天然氨基酸Bpa分析活细胞中所发生蛋白质相互作用(b)的流程示意图^[11,15]

Figure 1 (Color online) Schematic diagrams illustrating how an unnatural amino acid (e.g., p-benzoyl-L-phenylalanine [Bpa]) is genetically introduced at a selected residue position of a specific target protein A (a), and how the nature of a protein photocrosslinked to a target protein in living cells is characterized through either immunoblotting or mass spectrometric analyses (b)^[11,15]

tro)到体内(*in vivo*)的方向转变。

3 对非天然氨基酸介导活细胞蛋白质光交联方法体系的改进

在多年利用非天然氨基酸介导活细胞蛋白质光交联分析的研究实践中, 我们实验室对其方法体系进行了一定的改进和扩展, 使之更加简便有效, 总结如下。

3.1 制备出可有效进行非天然氨基酸 Bpa 介导的蛋白质光交联分析的LY928大肠杆菌菌株

通过将引入Bpa所需的正交tRNA和正交氨酰-tRNA合成酶的编码基因整合进大肠杆菌基因组的两个合适位点, 并进行组成性表达, 构建了LY928菌株(图2)^[21]。它可以被用于在目标蛋白质中引入Bpa。因为无需将表达这两种关键成分的质粒引入细菌细胞中, 从而大大方便了相关研究工作的开展。

3.2 建立了一套将非天然氨基酸随机引入到目标蛋白质中任何氨基酸残基位点的方法体系

对很多目标蛋白质而言, 在不同活细胞条件下, 究竟哪些残基介导蛋白质相互作用, 一般都并非知晓(除非蛋白质复合物的结构被解析)。依据传统的方法体系(图1), 我们须将目标蛋白质中的所有氨基酸残基逐一替换为非天然氨基酸, 然后逐一分析蛋白质交联结果才可得知。这样的过程相对低效, 特别是对大的蛋白质分子而言工作量极大。为克服这一难题, 我们建立了一

套将目标基因中编码氨基酸的所有碱基三联体随机替换为TAG(从而可引入非天然氨基酸)的方法系统^[21,22], 可以将蛋白质多肽链中的每一个氨基酸残基随机替换为非天然氨基酸。通过对这样随机引入了非天然氨基酸的目标蛋白质变异体库进行交联分析, 可以先筛选出那些可介导特定蛋白质相互作用的那些含非天然氨基酸的蛋白质形式。最后, 通过对相应突变基因DNA进行测序分析, 则可知介导特定相互作用的氨基酸残基的准确位置。这一策略与酶定向演化(*directed evolution*)研究中所采用的先获得具有特定性质的突变体酶, 然后通过DNA测序确定发生突变的具体氨基酸残基这种探索路径高度可比^[23,24]。

3.3 建立了一种直接在大肠杆菌基因组上对目标蛋白质的编码基因进行点突变的方法体系

通过操纵质粒所携带目标基因将非天然氨基酸引入目标蛋白质中的方法也存在一定的缺陷。首先, 这样产生的蛋白质的水平一般都高于内源的野生型蛋白质。其次, 一般都需先将基因组上内源的目标蛋白编码基因删除。再有, 质粒上控制目标蛋白质编码基因表达所用的启动子一般与目标蛋白质的野生型基因不同, 导致目标蛋白质表达的条件(如何时以何种水平表达等)有异于内源性目标蛋白质。我们通过“敲除野生型基因-敲入变异体基因”两步法, 将内源性目标基因中编码特定氨基酸残基的碱基三联体替换为TAG(图2)^[25]。这种引入了非天然氨基酸的目标蛋白质的表达特征(如表

图2 (网络版彩色)基因组中整合了编码正交tRNA和正交氨酰-tRNA合成酶基因的LY928菌株以及通过基因组水平的核苷酸点突变将非天然氨基酸Bpa引入目标蛋白质的示意图^[21,22]

Figure 2 (Color online) Schematic diagram of the genes encoding the orthogonal tRNA and amino acyl-tRNA synthetase integrated at proper positions in the genome of an LY928 bacterium. The schematic shows how an unnatural amino acid, in this case Bpa, is introduced via site-directed mutagenesis at the endogenous target gene in the genome^[21,22]

达的时空条件和水平等)与内源性蛋白质高度接近,故实验结果更能反映细胞中的真实情况。

3.4 建立了一种单块胶可分析 384 个蛋白质样品的高通量聚丙烯酰胺凝胶电泳技术

在对话细胞中某种目标蛋白质的全部或大量氨基酸位点被非天然氨基酸替换后的光交联产物进行分析时, 常需对大量的蛋白质交联样品进行比较分析. 比如, 当对细菌ATP合酶复合体中的 ϵ 亚基在细胞处于不同条件下的结构模式进行分析时, 需将所有含Bpa的 ϵ 亚基变异体细菌放置于不同培养条件下进行处理, 然后需对不同条件下 ϵ 亚基交联产物之间进行半定量比较. 如果利用多块常规电泳胶(每块一般只可分析十几个蛋白质样品)进行分析, 很难对这些数量巨大的细胞样品的交联产物进行半定量比较(因为不同的免疫印迹分析胶的结果之间难免存在实验误差). 为此, 我们经过大量试错摸索, 成功建立了一种高通量聚丙烯酰胺凝胶电泳(high-throughput PAGE)技术. 这种方法的关键在于, 将传统的分析蛋白质的垂直式电泳技术改造为类似于分析DNA的浸没式水平式电泳技术. 这种凝胶体系可以在单块胶上同时分析多达384种蛋白质样品(图3)^[25], 从而可以有效地对大量蛋白质样品中的特定蛋白质带进行半定量比较.

通过这种高通量电泳和免疫印迹技术, 我们能对一种蛋白质在活细胞中所发生特定蛋白质-蛋白质相

图 3 每块胶上可分析384种蛋白质样品的高通量聚丙烯酰胺凝胶电泳及免疫印迹分析代表性结果^[25]

Figure 3 A representative high-throughput polyacrylamide gel electrophoresis immunoblot image showing 384 protein samples analyzed on a single gel^[25]

相互作用进行“自印证”。这可以通过在多个引入了非天然氨基酸的变异体中检测到相同的相互作用模式而实现。当然更为重要的是, 通过这种高通量检测系统我们可以对目标蛋白质在不同条件下所发生的蛋白质相互作用的变化进行有效揭示(第5节)。总之, 该高通量蛋白质样品检测分析体系为认识蛋白质在活细胞内的功能及调节机制等提供了一种高度有效的通用方法体系。

利用上述经过改进的活细胞蛋白质光交联方法体系,我们获得了多项科学发现.下面列举几个代表性例子,以说明这种活细胞蛋白质研究方法体系的有效性.

4 在研究FtsZ这种细菌分裂关键蛋白质之活细胞动态组装机制过程中意外发现复苏延迟体

在一个细菌“母细胞”分裂成两个“子细胞”的细胞质分裂阶段, FtsZ蛋白发挥了关键功能. 它动态组装成一个可以收缩的Z-环^[26]. 针对纯化的FtsZ蛋白, 人们已经在体外通过X射线晶体学解析了其三维分子结构, 初步认识了它如何纵向组装成“原丝纤维”的分子细节^[27]. 但对于这样的“原丝纤维”如何进一步动态组装形成Z-环, 以及Z-环如何收缩等分子过程, 我们仍知之甚少^[26]. 因为这样的动态组装或收缩过程难以在体外重现, 需开展活细胞研究其机制才可能被揭示. 尽管人们利用分辨率不断提高的荧光显微镜对细菌细胞Z-环的动态结构进行了大量观察^[26,28], 但鉴于光学显微镜分辨率的限制, 这样的研究还是难以揭示Z-环形成和收缩的分子机制.

为了探究活细胞中FtsZ蛋白的动态组装机制, 我们构建了上述可在目标蛋白质中有效引入非天然氨基酸的LY928菌株(图2), 建立了在目标蛋白质中随机引入非天然氨基酸的技术体系^[21, 22]。通过非天然氨基酸介导的活细胞蛋白质光交联技术, 分析了在活细胞中FtsZ蛋白的哪些氨基酸残基参与了Z-环形成的自组装过程。除了确认晶体结构所暗示的纵向相互作用中的氨基酸残基参与自组装之外, 更为重要的是, 我们还观察到了FtsZ蛋白质亚基在活细胞中由参与自组装的其他氨基酸残基所组成的其他之前未知的相互作用表面^[21]。我们正在研究这些新鉴定到的氨基酸残基如何参与活细胞中Z-环的组装和收缩过程。

人们对活细胞中Z-环结构特征的荧光显微镜观察, 大多是利用正处于活跃分裂状态(即对数期)的细菌细胞开展的. 我们在进行非天然氨基酸介导的FtsZ蛋白光

交联分析时,自然也主要以正在活跃分裂的细菌细胞为研究对象.但作为“负”对照,我们也分析了停止分裂的静止后期细菌细胞中FtsZ蛋白的光交联情况.结果表明,在停止生长和分裂的细胞中,FtsZ蛋白不再形成原来在活跃分裂细胞中所观察到的自我交联形式,而是与其他未知蛋白质发生特异交联.令我们意外的是,FtsZ蛋白与其他蛋白质交联形成的这些产物以及未发生交联的自由FtsZ蛋白皆出现于细胞被破碎和离心后的沉淀部分,而不是像处于活跃分离的细胞中那样出现在离心后的上清部分.

为了解释上述意外观察到的结果,我们将FtsZ蛋白与绿色荧光蛋白融合后开展了活细胞成像研究.结果表明,在停止分裂的静止后期细胞中,不再形成Z-环的FtsZ蛋白定位于细胞的两极,形成荧光显微镜下明显可见的一种全新亚细胞结构(图4).当这样的细胞被放置于新鲜培养基中时,伴随该结构的解聚,细胞开始复苏;相反,如果该结构被维持不解聚状态,则细胞维持其不分裂状态.因此,我们将这种亚细胞结构命名为复苏延迟体(regrowth-delay body)^[29,30].

通过进一步的系统研究,我们观察到复苏延迟体具有以下特征(图4)^[29,30]: (1) 它在细菌细胞停止生长和分裂一段时间后形成,一般定位于细胞两极; (2) 多种参与细胞生长和分裂的关键蛋白质成分几乎全部被特

图4 复苏延迟体在停止分裂的静止期细菌细胞中形成,在复苏过程中解聚^[29,30].右侧下图为细菌细胞的荧光显微镜照片,其中位于两端的绿色结构为复苏延迟体;含复苏延迟体的细胞表现出休眠细菌的所有抗生素具有耐受性的特征

Figure 4 A cartoon showing how the regrowth-delay body is formed in non-dividing, stationary phase bacterial cells; maintained during the lag phase; and dissolved when the cells resume growth^[29,30]. The green colored structures in the fluorescence photomicrograph are regrowth-delay bodies. Bacterial cells containing those structures are as tolerant of all antibiotics as dormant persister bacterial cells are

异锁定于复苏延迟体中,可能正是因此而阻止了细胞生长分裂; (3) 当细胞被放置于有利的培养条件时,复苏延迟体的存在与否,决定细胞个体处于“延迟期”时间的长短,只有当它们发生解聚并释放其中的蛋白质以重新发挥功能时,细胞才复苏,否则细胞将维持其“休眠”非生长状态; (4) 细菌细胞群体中存在极少数量的个体可维持其复苏延迟体不解聚,处于非生长和非分裂状态,对所有抗生素皆表现出耐受性.这些维持休眠状态的细菌很可能就是近百年前定义的“持留菌(persisters)”,它们在抗生素存在的条件下仍可以存活.

这项意外发现复苏延迟体的成果发表后,论文不久即被Faculty1000学术评价组织推荐.其推荐词(图5)中提及:“这篇文章的重要性怎么说都不过分”.希望推荐人对这项成果的积极评价能被进一步的科学数据所支持.复苏延迟体的发现,至少为我们在一个巨大的细菌群体里鉴定出为数极少的持留菌——从而研究它们的结构和代谢特征——提供了一种有效路径.假如我们能通过某种物理或化学途径有效加快休眠细菌细胞中这种复苏延迟体的解聚,也许能为破解长期困扰医学界的细菌传染性疾病预防的难题提供一种新思路.

5 细菌ATP合酶实施一种独特的“换挡”活性调节机制

作为细胞能量货币ATP的主要制造机器——ATP合酶是存在于所有原核和真核细胞中最为重要、最为复杂和最为神奇的多亚基蛋白质复合体之一.细菌细

It would be hard to over-estimate the importance of this paper that reveals
...
dynamics of persistence. That these serendipitous discoveries were driven by curiosity should serve as a lesson to funding agencies and their reviewers.

图5 Faculty Opinion对复苏延迟体论文推荐词的一部分(<https://facultyopinions.com/prime/734927744>)

Figure 5 Part of the Faculty Opinion paragraph commending the paper that reported the discovery of the regrowth-delay body (<https://facultyopinions.com/prime/734927744>)

胞ATP合酶的亚基组成通常可以表示为 $\alpha_3\beta_3\gamma\delta\epsilon ab_2c_{8-15}$, 其中由 ab_2c_{8-15} 组成的 F_0 部分整合于细胞质膜内, 而由 $\alpha_3\beta_3\gamma\delta\epsilon$ 组成的 F_1 部分则存在于细胞质可溶环境中^[31]. 依据大量的体外研究结果, 人们普遍认为, 它通过一种独特的旋转机制催化ATP分子的产生. 我们对体外所获得的所有研究数据(包括结构数据)进行了理论分析, 提出了一种新的ATP合酶催化模型, 即与按压式伸缩圆珠笔中推杆的直线往复运动驱动凸轮产生连续转动的工作机理十分相似. 该模型的核心内容为: 质子的跨膜转运引起ATP合酶c亚基环的周期性构象改变(即并非在细胞膜中发生快速的旋转运动), 使附着在c亚基环上的主要由 γ 亚基组成的中心杆(“推杆”)产生直线往复运动, γ 亚基继而驱动 $\alpha_3\beta_3$ 六聚体连续转动^[32]. 尽管我们的理论模型解释了传统模型无法解释的观察结果, 但其正确性有待未来更为深入研究结果的检验, 特别是活细胞方面的研究.

ATP合酶蛋白质复合体在活细胞内如何发挥功能? 在它生成的过程中, 多种不同亚基之间如何发生高度有序组装? 它的活性如何响应细胞环境条件的改变而被调节等? 这些都是重要的科学问题, 也是我们正在积极研究的问题. 作为第一个阶段的研究成果, 我们开展的活细胞研究揭示, ATP合酶可能通过改变其 ϵ -亚基的结构(及与其他亚基之间的相互作用)而实施一种催化活性的“换挡”调节机制.

通过非天然氨基酸介导的蛋白质光交联分析, 我们探究了活细胞中ATP合酶 ϵ -亚基与其他亚基之间的相互作用以及它随细胞环境条件改变而发生变化的情况^[25]. 通过上述基因组定点突变技术(图2), 我们将 ϵ -亚基C末端近40个氨基酸残基位点逐一替换成了非天然氨基酸Bpa. 这些通过基因组突变而引入了Bpa的 ϵ -, 其表达的时空特征与内源野生型 ϵ -亚基基本保持一致, 这就维持了ATP合酶不同亚基之间合适的化学计量比, 避免对ATP合酶组装造成任何干扰. 我们将这些含有 ϵ -亚基的近40个Bpa变异体的细胞放置于各种不同环境中, 并分析了365 nm紫外线照射后产生的 ϵ -亚基与ATP合酶中其他亚基之间形成的不同交联产物. 交联结果清楚揭示了 F_1 部分的 ϵ -亚基C末端两种截然不同的相互作用模式: 一种是与 F_0 部分的c亚基发生相互作用的“非插入状态”, 另一种是与 F_1 部分的 α 、 β 亚基发生相互作用的“插入”状态. 我们建立的高通量凝胶电泳免疫印迹分析(图3), 使我们可以通过对多种氨基酸残基位点在大量不同培养条件下所介导的相互作用进行半定量比

较, 从而有效观察到ATP合酶这两种结构状态在不同条件下的存在情况及相互转变情况^[25].

通过活细胞蛋白质相互作用的高通量分析, 以及无法形成“插入”状态的 ϵ 亚基突变体赋予细胞生长及ATP合成特征等现象的分析, 我们揭示了基于 ϵ -亚基结构状态改变的ATP合酶“换挡”活性调控机制^[25]. 这种机制的主要特征可以被概况如下(图6(a)): (1) 当细胞处于适宜生长条件时, ϵ -亚基主要以“非插入”状态存在, ATP合酶处于其“低挡位”活性调节模式; (2) 当细胞处于高度不适宜生长条件时, ϵ -亚基主要以“插入”状态存在, ATP合酶处于其“高挡位”模式; (3) 当细胞处于中、低度不适宜条件时, ATP合酶处于一种“混合挡位”活性调节模式; (4) ϵ -亚基的这种“非插入”和“插入”结构处于一种平衡状态, 随着细胞条件的改变可以发生快速有效的相互转变; (5) 这两种状态的最终目标都是, 即使在差异很大的细胞条件下仍然维持细胞ATP水平处于一种动态平衡状态. 该工作发表后, 很快就有国际同行撰文进行了积极的评价(图6(b))^[33].

6 一种超级蛋白质复合体负责细菌新生外膜蛋白质的跨膜转运

很多蛋白质的新生多肽链在位于细胞质中的核糖体上合成后都要被靶向定位于特异亚细胞部位(如线粒体、细胞质膜等), 这一过程由新生肽链自身所携带的信息决定, 涉及跨膜转运、修饰、折叠、组装等过程, 并最终在特定亚细胞部位形成具有功能的蛋白质形式. 革兰氏阴性细菌的新生外膜蛋白需经由细胞质、细胞质膜(内膜)、膜间质才可抵达外膜, 相对比较复杂^[22,34]. 我们实验室对该过程进行了多年探究, 对这一过程所涉及的质量控制因子, 如DegP蛋白^[35-38]、FkpA蛋白^[39]、SurA蛋白^[38]等开展了生物化学和遗传学方面的研究.

通过利用非天然氨基酸介导的蛋白质光交联所进行的活细胞研究, 我们鉴定到一种横跨细胞质、细胞内膜、膜间质和外膜的超级蛋白质复合体, 它负责 β 桶类外膜蛋白质新生肽链的有效转运(图7). 该复合体中的蛋白质组分包括存在于细胞质中的SecA马达蛋白、存在于内膜中的SecYEG跨膜蛋白质通道、存在于膜间质中的SurA分子伴侣蛋白和PpiD蛋白、存在于外膜中的BamA蛋白等^[22,40]. 我们不清楚如DegP蛋白等是否也存在于该超级蛋白质复合体中. 尽管对这种超级蛋白质复合体的详细结构还不清楚, 但它的存在一方面

图 6 (网络版彩色)细菌ATP合酶基于其 ϵ -亚基结构改变而响应不同环境条件调节自身活性的“换挡”活性调控机制示意图(a)^[25]及国际同行的专文评价(b)^[33]

Figure 6 (Color online) (a) Schematic illustration of how the “gear-shifting” mechanism based on the ATP synthase ϵ -subunit exists in either the “non-inserted” or “inserted” state, and that bacterial ATP synthase modulates its activity in response to a variety of environmental conditions^[25]. (b) The article title and journal comment commending this important discovery^[33]

可以为未折叠新生肽链提供保护, 另一方面也可使这一新生肽链转运过程快速有效. 传统模型认为, 新生肽链被逐步传递穿过不同亚细胞部位^[34]. 这种传统的接力赛式的传递机制, 尽管可能符合直接的思维范式, 但很难想象在活细胞中会利用这样一种低效而且“危险”的传递机制, 因为未折叠好的新生肽链不仅很容易发生聚集, 也很容易被蛋白质水解酶降解.

可以设想的是, 要纯化这样的超级蛋白质复合体以进行详细的体外表征研究, 如进行X射线晶体衍射或冷冻电子显微镜结构分析等, 可能并非易事. 换句话说, 如果不开展活细胞研究, 这样的超级蛋白质复合体可能难以被揭示. 以此类推, 可以设想, 一定还有不少存在于活细胞中、又无法被体外纯化的蛋白质复合体有待被揭示. 这也从一个侧面表明, 开展活细胞蛋白质研究非常必要.

7 抗酸分子伴侣蛋白 HdeA 在极端酸性细胞条件下丧失空间结构后发挥生物学功能且可保护其他分子伴侣蛋白

非天然氨基酸光交联技术的一大优点是可以用于

探究活细胞处于如极端酸性这样的不利条件下所发生的蛋白质-蛋白质相互作用. 如大肠杆菌这样的细菌能在穿过动物胃里的极端酸性(pH大约为1~3)环境后生存下来, 位于其膜间质中的抗酸分子伴侣蛋白HdeA和HdeB发挥了重要作用. 尽管HdeA在中性pH条件下的空间结构通过X射线晶体衍射技术已经被测定^[41,42], 但我们的体外研究表明, 在胃酸条件下, HdeA将丧失其在pH中性条件下所维持的三维结构, 转变成一种基本无序的结构. 然而我们同时也发现, 正是通过这种丧失了三维空间结构的多肽链所展现出的结构特征, HdeA在体外表现出酸性条件下保护底物蛋白质(或客户蛋白)的分子伴侣活性^[43,44]. 通过对高量表达HdeA蛋白的活细菌细胞开展核磁共振光谱学研究表明, 当活细胞处于极端酸性条件时, HdeA的三维空间结构的确丧失, 而当细胞恢复到中性pH条件时, 其三维结构也得到恢复(该部分工作为本人实验室与北京大学金长文教授合作研究结果; 尚未发表). 显然, HdeA这种“拥有三维结构却无生物活性, 丧失三维结构却拥有生物活性”的结构-活性关系是独特的, 不同于绝大多数蛋白质发挥生物学活性的模式.

图7 (网络版彩色)存在于活细胞中的转运外膜蛋白新生肽链的超级蛋白质复合体示意图。超级蛋白质复合体跨越革兰氏阴性细菌的内外膜,从而将转运、运输和插膜等外膜蛋白生成事件在活细胞内进行整合^[22,40]

Figure 7 (Color online) Schematic diagram illustrating the supercomplex that was unveiled by living cell studies and how it is responsible for the translocation of nascent outer membrane proteins. The supercomplex spans the inner and outer membrane of Gram-negative bacteria and integrates nascent outer membrane proteins events, such as the translocation, transport and membrane insertion of nascent outer membrane proteins in living cells^[22,40]

通过使用非天然氨基酸介导的蛋白质光交联技术,我们分析了当大肠杆菌被置于极端酸性条件时,活细胞中的HdeA分子伴侣蛋白所结合底物蛋白质的种类。对交联产物进行质谱鉴定分析发现,HdeA结合的底物蛋白质中包括了如SurA(分子伴侣蛋白)和DegP(蛋白质水解酶)这样的蛋白质质量控制因子(图8)^[45,46]。进一步的体外研究表明,当细菌从极端酸性环境返回中性环境时,这些质量控制因子似乎可以迅速地由HdeA上释放并恢复结构和功能。这种“分子伴侣蛋白保护分子伴侣蛋白”的独特机制似乎也表明,生物演化所产生的不同分子伴侣蛋白家族成员各有专长,在不利条件下可以实现相互保护,这使细菌细胞能在不同胁迫条件下有效生存。此外,细菌在高温下生存所必需的这种质量控制因子DegP蛋白,在细菌抗酸过程中也发挥一定的功能。

8 一直被认为具有双功能的 DegP 蛋白主要发挥蛋白质水解酶功能,分子伴侣功能不明显

DegP蛋白最初通过遗传学方法被鉴定,是细菌在高温下(如44℃)生存所必需的一种蛋白质^[47,48]。研究发现,其同源蛋白质也存在于动、植物体内。有研究依据体外分析结果认为,该蛋白质在较低温度下(如28℃)表

图8 (网络版彩色)细菌抗酸分子伴侣蛋白HdeA在活细胞处于强酸条件下发挥功能的两个特征:丧失三维结构后才发挥分子伴侣功能;可以保护如SurA分子伴侣和DegP蛋白质水解酶这样的在中性pH条件下发挥功能的质量控制因子^[43-46]

Figure 8 (Color online) The acid stress chaperone HdeA exhibits two unique functional features in living cells: It performs its chaperone function only after its tertiary structure becomes largely disordered, and under acidic conditions, it is able to protect the functioning of quality control factors such as the SurA molecular chaperone and DegP protease, each of which normally function under neutral pH conditions^[43-46]

现出分子伴侣活性,而在较高温度时(如44℃)表现为蛋白质水解酶活性,因此被长期描述为一种分子伴侣-蛋白质水解酶双功能蛋白质^[49,50]。应该说明的是,这样的体外分析结果观察到的是蛋白质的“活性”而非“功能”。我们的体外、遗传学和活细胞研究结果表明,存在于革兰氏阴性细菌膜间质中的DegP蛋白主要发挥的是蛋白质水解酶功能,分子伴侣功能不明显(图9)^[36,50]。在高温下,它帮助清除存在于细菌膜间质中异常折叠的细菌外膜蛋白,以阻止异常折叠的外膜蛋白聚集体对细菌细胞膜产生毒性^[36]。

基于对DegP蛋白在极端酸性条件下被抗酸分子伴侣蛋白HdeA保护这一活细胞非天然氨基酸光交联的研究结果^[45,46],我们进一步研究了DegP蛋白在细菌抗酸过程中发挥了何种生物学功能。结果表明,DegP蛋白的确可以帮助经历了酸性条件的细菌细胞在恢复到中性条件时对膜间质中一些底物蛋白质的降解清除^[51]。这似乎通过一种独特机制进行:这样的底物蛋白质在酸性条件下即与同样处于变性状态的DegP蛋白进行某种形式的预结合,当细胞恢复到中性条件时,DegP蛋白恢复其结构并降解所结合的底物蛋白,从而实现其抵抗酸性胁迫的功能。

9 活细胞中小分子热休克蛋白的结构和活性随温度变化而变化

小分子热休克蛋白是一类存在于所有生物种类中

图9 (网络版彩色) DegP蛋白在活细胞中主要发挥蛋白质水解酶功能, 几乎不发挥分子伴侣功能^[36,50]
 Figure 9 (Color online) The DegP protein exhibits protease activity, with hardly any chaperone activity^[36,50]

的分子伴侣蛋白^[52]。我们对该家族蛋白质的多个成员(如结核杆菌中的Hsp16.3、大肠杆菌中的IbpB等)开展的体外研究表明, 这类蛋白质的三级和四级结构对包括热休克在内的胁迫条件具有高度的敏感性和响应能力, 这些细微结构的变化导致其分子伴侣活性明显的改变^[53-59]。我们好奇的是, 这样的体外胁迫条件诱导发生的结构和活性改变是否反映活细胞中的情况?

通过非天然氨基酸介导的蛋白质光交联活细胞分析, 我们观察到, 随着温度的逐步升高(20~50°C), 细菌中的IbpB小分子热休克蛋白参与和底物蛋白质结合的残基数目也逐渐增多, 因此, 分子伴侣活性也逐渐提高(图10(a))^[60]。显然, 这种在活细胞中体现出来的小分子热休克蛋白的层级式激活机制与我们在体外观察到的结果高度吻合。同时, 我们对细菌小分子热休克蛋白在活细胞中所结合的底物蛋白质的种类进行了质谱分析。结果表明, 在IbpB所保护的底物蛋白质中, 与蛋白质翻译过程相关的蛋白质因子相对比较富集^[61]。这一现象背后的生物学意义有待深入认识。

同样重要的是, 我们在体外实验中观察到的小分子热休克蛋白形成寡聚体的亚基之间相互作用的动态性^[54-58], 也通过非天然氨基酸蛋白质交联分析在活细胞内得到了验证^[22,62]。这样的分析结果还表明, 这种分子伴侣蛋白质在活细胞中构象的动态性似乎远超出我们依据体外观察到的结果所想象到的(图10(b))^[22,62]。这些基于活细胞所开展的研究大大加深了我们对蛋白质发挥功能机制的认识。

10 结语

活细胞是蛋白质发挥功能的舞台, 在那里每个蛋白质个体都会历经“生老病死”。体外研究是认识蛋白质分子所必须的非常重要的第一步, 但最终我们还是要认识活细胞中蛋白质的行为, 这无疑将是未来蛋白质研究的方向所在。需探究的方面很多, 这里列出一些代表性的科学问题, 供读者参考。

- (1) 活细胞内蛋白质的生成过程(包括新生肽链的定向转运、折叠、组装等)如何发生?
- (2) 活细胞内蛋白质的质量控制过程(包括异常合成、异常折叠和异常组装形式的蛋白质分子的清除或重折叠及重组装等, 特别是对尚未折叠好的新生肽链与折叠异常的新生肽链之间的区分)如何实现?
- (3) 活细胞内蛋白质的空间结构与体外所观察到的是否完全一致?
- (4) 活细胞内蛋白质之间的可逆动态相互作用及特异识别如何发生?
- (5) 活细胞中特定蛋白质所完成的生物学功能(可能是多种)是什么?
- (6) 活细胞内每种蛋白质发挥生物学功能的动态过程(机制)如何?
- (7) 活细胞中蛋白质的结构(包括共价结构、非共价结构)和活性如何随细胞条件(包括物理、化学和生物的条件)的改变而改变(调节)?
- (8) 活细胞中每种特定蛋白质的寿命差异甚大, 发

图 10 小分子热休克蛋白在活细胞中响应温度变化以层级方式被激活(a)^[60]并具高度的结构动态性(b)^[62]

Figure 10 Small heat shock proteins modulate their activities in a cascade manner while responding to increasing temperatures (a)^[60], and exhibit highly dynamic structures in living cells (b)^[62]

生的原理和机制是什么?

(9) 活细胞中不同蛋白质含量差异甚大, 这是如何发生的?

(10) 如何对活细胞中含量极低、寿命极短的蛋白质开展观察?

在活细胞中对这些蛋白质的不同方面开展研究极

具挑战性. 因为不仅要进行定性的, 也需进行定量的描述; 不仅要认识蛋白质的静态结构, 更需认识其动态机制. 对这样的问题开展有效研究, 还需建立更加有效的技术和理论体系. 深入认识活细胞中的蛋白质是我们理解生命本质的基础, 也是生物医学和生物技术有效实践的重要基础.

致谢 感谢北京大学刘洋和福建师范大学付新苗、王妍帮助绘制图片并参与讨论.

参考文献

- 1 Fruton J S. Proteins, Enzymes and Genes—the Interplay of Chemistry and Biology. New Heaven and London: Yale University Press, 1999
- 2 Tanford C, Reynolds J. Nature's Robots—A History of Proteins. Oxford: Oxford University Press, 2001
- 3 Clark P L. Proteins in the cell. Protein Sci, 2019, 28: 1175–1176
- 4 Ma J, Ge X, Chang Z Y. Protein function studies: History, current status and future trends (in Chinese). Chin Bull Life Sci, 2007, 19: 294–300 [马静, 葛熙, 昌增益. 蛋白质功能研究: 历史, 现状和未来. 生命科学, 2007, 19: 294–300]
- 5 Shimomura O, Johnson F H, Saiga Y. Extraction, purification and properties of aequorin, a bioluminescent protein from the luminous

- hydromedusan, aequorea. *J Cell Comp Physiol*, 1962, 59: 223–239
- 6 Chalfie M, Tu Y, Euskirchen G, et al. Green fluorescent protein as a marker for gene expression. *Science*, 1994, 263: 802–805
- 7 Tsien R Y. The green fluorescent protein. *Annu Rev Biochem*, 1998, 67: 509–544
- 8 Wang S, Hazelrigg T. Implications for *bcd* mRNA localization from spatial distribution of *exu* protein in *Drosophila* oogenesis. *Nature*, 1994, 369: 400–403
- 9 Freedberg D I, Selenko P. Live cell NMR. *Annu Rev Biophys*, 2014, 43: 171–192
- 10 Luchinat E, Banci L. A unique tool for cellular structural biology: In-cell NMR. *J Biol Chem*, 2016, 291: 3776–3784
- 11 Liu C C, Schultz P G. Adding new chemistries to the genetic code. *Annu Rev Biochem*, 2010, 79: 413–444
- 12 Kauer J C, Erickson-Viitanen S, Wolfe H R Jr, et al. p-Benzoyl-L-phenylalanine, a new photoreactive amino acid. Photolabeling of calmodulin with a synthetic calmodulin-binding peptide. *J Biol Chem*, 1986, 261: 10695–10700
- 13 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: 182–188
- 14 Chin J W, Martin A B, King D S, et al. Addition of a photocrosslinking amino acid to the genetic code of *Escherichia coli*. *Proc Natl Acad Sci USA*, 2002, 99: 11020–11024
- 15 Chin J W, Schultz P G. *In vivo* photocrosslinking with unnatural amino acid mutagenesis. *ChemBioChem*, 2002, 3: 1135–1137
- 16 Miyazaki R, Myougo N, Mori H, et al. A photo-cross-linking approach to monitor folding and assembly of newly synthesized proteins in a living cell. *J Biol Chem*, 2018, 293: 677–686
- 17 Chin J W, Cropp T A, Anderson J C, et al. An expanded eukaryotic genetic code. *Science*, 2003, 301: 964–967
- 18 Liu W, Brock A, Chen S, et al. Genetic incorporation of unnatural amino acids into proteins in mammalian cells. *Nat Methods*, 2007, 4: 239–244
- 19 Greiss S, Chin J W. Expanding the genetic code of an animal. *J Am Chem Soc*, 2011, 133: 14196–14199
- 20 Chin J W. Expanding and reprogramming the genetic code of cells and animals. *Annu Rev Biochem*, 2014, 83: 379–408
- 21 Guan F, Yu J, Yu J, et al. Lateral interactions between protofilaments of the bacterial tubulin homolog FtsZ are essential for cell division. *eLife*, 2018, 7: e35578
- 22 Fu X M, Chang Z. Biogenesis, quality control and structural dynamics of proteins as explored in living cells via site-directed photo-crosslinking. *Protein Sci*, 2019, 28: 1194–1209
- 23 Chen K, Arnold F H. Tuning the activity of an enzyme for unusual environments: Sequential random mutagenesis of subtilisin E for catalysis in dimethylformamide. *Proc Natl Acad Sci USA*, 1993, 90: 5618–5622
- 24 Chang Z. The 2018 Nobel Prize in Chemistry: Engineering proteins (enzymes/peptide/antibodies) towards desired properties via the construction of random libraries. *Sci China Life Sci*, 2019, 62: 713–724
- 25 Yang L, Yu J Y, Wang M Y, et al. A high-throughput genetically directed protein crosslinking analysis reveals the physiological relevance of the ATP synthase “inserted” state. *FEBS J*, 2021, 288: 2989–3009
- 26 Barrows J M, Goley E D. FtsZ dynamics in bacterial division: What, how, and why? *Curr Opin Cell Biol*, 2021, 68: 163–172
- 27 Li Y, Hsin J, Zhao L, et al. FtsZ protofilaments use a hinge-opening mechanism for constrictive force generation. *Science*, 2013, 341: 392–395
- 28 Yang X, Lyu Z, Miguel A, et al. GTPase activity-coupled treadmilling of the bacterial tubulin FtsZ organizes septal cell wall synthesis. *Science*, 2017, 355: 744–747
- 29 Yu J, Liu Y, Yin H, et al. Regrowth-delay body as a bacterial subcellular structure marking multidrug-tolerant persisters. *Cell Discov*, 2019, 5: 8; doi.org/10.1038/s41421-019-0080-3
- 30 Yu J Y, Liu Y, Yin H J, et al. Regrowth-delay body as a reversible subcellular structure marking dormant bacterial cells (in Chinese). *Chin J Biochem Mol Biol*, 2019, 35: 349–352 [余家钰, 刘洋, 殷会佳, 等. 复苏延迟体(regrowth-delay body)作为一种标示休眠细菌的可逆亚细胞结构. *中国生物化学与分子生物学报*, 2019, 35: 349–352]
- 31 Walker J E. The ATP synthase: The understood, the uncertain and the unknown. *Biochem Soc Trans*, 2013, 41: 1–16
- 32 Liu J F, Fu X M, Chang Z Y. A reciprocating motion-driven rotation mechanism for the ATP synthase. *Sci China Life Sci*, 2016, 59: 44–48 [刘佳峰, 付新苗, 昌增益. ATP合酶旋转催化的一种新机制. *中国科学: 生命科学*, 2016, 46: 269–273]
- 33 Miranda-Astudillo H, Zarco-Zavala M, Garcia-Trejo J J, et al. Regulation of bacterial ATP synthase activity: A gear-shifting or a pawl-ratchet mechanism? *FEBS J*, 2021, 288: 3159–3163
- 34 Driessen A J, Nouwen N. Protein translocation across the bacterial cytoplasmic membrane. *Annu Rev Biochem*, 2008, 77: 643–667
- 35 Jiang J, Zhang X, Chen Y, et al. Activation of DegP chaperone-protease via formation of large cage-like oligomers upon binding to substrate proteins. *Proc Natl Acad Sci USA*, 2008, 105: 11939–11944
- 36 Ge X, Wang R, Ma J, et al. DegP primarily functions as a protease for the biogenesis of β -barrel outer membrane proteins in the Gram-negative bacterium *Escherichia coli*. *FEBS J*, 2014, 281: 1226–1240
- 37 Li S, Wang R, Li D, et al. Thermal-triggered proteinquake leads to disassembly of DegP hexamer as an imperative activation step. *Sci Rep*, 2014, 4:

4834

- 38 Wu S, Ge X, Lv Z, et al. Interaction between bacterial outer membrane proteins and periplasmic quality control factors: A kinetic partitioning mechanism. *Biochem J*, 2011, 438: 505–511
- 39 Ge X, Lyu Z X, Liu Y, et al. Identification of FkpA as a key quality control factor for the biogenesis of outer membrane proteins under heat shock conditions. *J Bacteriol*, 2014, 196: 672–680
- 40 Wang Y, Wang R, Jin F, et al. A supercomplex spanning the inner and outer membranes mediates the biogenesis of β -barrel outer membrane proteins in bacteria. *J Biol Chem*, 2016, 291: 16720–16729
- 41 Yang F, Gustafson K R, Boyd M R, et al. Crystal structure of *Escherichia coli* HdeA. *Nat Struct Mol Biol*, 1998, 5: 763–764
- 42 Gajiwala K S, Burley S K. HDEA, a periplasmic protein that supports acid resistance in pathogenic enteric bacteria 1 | Edited by P. E. Wright. *J Mol Biol*, 2000, 295: 605–612
- 43 Hong W, Jiao W, Hu J, et al. Periplasmic protein HdeA exhibits chaperone-like activity exclusively within stomach pH range by transforming into disordered conformation. *J Biol Chem*, 2005, 280: 27029–27034
- 44 Wu Y E, Hong W, Liu C, et al. Conserved amphiphilic feature is essential for periplasmic chaperone HdeA to support acid resistance in enteric bacteria. *Biochem J*, 2008, 412: 389–397
- 45 Zhang M, Lin S, Song X, et al. A genetically incorporated crosslinker reveals chaperone cooperation in acid resistance. *Nat Chem Biol*, 2011, 7: 671–677
- 46 Hong W, Wu Y E, Fu X, et al. Chaperone-dependent mechanisms for acid resistance in enteric bacteria. *Trends Microbiol*, 2012, 20: 328–335
- 47 Lipinska B, Fayet O, Baird L, et al. Identification, characterization, and mapping of the *Escherichia coli* *htrA* gene, whose product is essential for bacterial growth only at elevated temperatures. *J Bacteriol*, 1989, 171: 1574–1584
- 48 Strauch K L, Johnson K, Beckwith J. Characterization of *degP*, a gene required for proteolysis in the cell-envelope and essential for growth of *Escherichia coli* at high-temperature. *J Bacteriol*, 1989, 171: 2689–2696
- 49 Spiess C, Beil A, Ehrmann M. A temperature-dependent switch from chaperone to protease in a widely conserved heat shock protein. *Cell*, 1999, 97: 339–347
- 50 Chang Z Y. The function of the DegP (HtrA) protein: Protease versus chaperone. *IUBMB Life*, 2016, 68: 904–907
- 51 Fu X, Wang Y, Shao H, et al. DegP functions as a critical protease for bacterial acid resistance. *FEBS J*, 2018, 285: 3525–3538
- 52 Chang Z Y. What do small heat shock proteins do for bacterial cells. In: Tanguay R M, Hightower L E, eds. *The Big Book on Small Heat Shock Protein*. Berlin: Springer, 2015. 511–525
- 53 Yang H, Huang S, Dai H, et al. The *Mycobacterium tuberculosis* small heat shock protein HSP16.3 exposes hydrophobic surfaces at mild conditions: Conformational flexibility and molecular chaperone activity. *Protein Sci*, 1999, 8: 174–179
- 54 Gu L, Abulimiti A, Li W, et al. Monodisperse Hsp16.3 nonamer exhibits dynamic dissociation and reassociation, with the nonamer dissociation prerequisite for chaperone-like activity. *J Mol Biol*, 2002, 319: 517–526
- 55 Fu X, Liu C, Liu Y, et al. Small heat shock protein Hsp16.3 modulates its chaperone activity by adjusting the rate of oligomeric dissociation. *Biochem Biophys Res Commun*, 2003, 310: 412–420
- 56 Fu X, Chang Z. Temperature-dependent subunit exchange and chaperone-like activities of Hsp16.3, a small heat shock protein from *Mycobacterium tuberculosis*. *Biochem Biophys Res Commun*, 2004, 316: 291–299
- 57 Jiao W, Qian M, Li P, et al. The essential role of the flexible termini in the temperature-responsiveness of the oligomeric state and chaperone-like activity for the polydisperse small heat shock protein IbpB from *Escherichia coli*. *J Mol Biol*, 2005, 347: 871–884
- 58 Fu X, Zhang H, Zhang X, et al. A dual role for the N-terminal region of *Mycobacterium tuberculosis* Hsp16.3 in self-oligomerization and binding denaturing substrate proteins. *J Biol Chem*, 2005, 280: 6337–6348
- 59 Jiao W, Hong W, Li P, et al. The dramatically increased chaperone activity of small heat-shock protein IbpB is retained for an extended period of time after the stress condition is removed. *Biochem J*, 2008, 410: 63–70
- 60 Fu X, Shi X, Yin L, et al. Small heat shock protein IbpB acts as a robust chaperone in living cells by hierarchically activating its multi-type substrate-binding residues. *J Biol Chem*, 2013, 288: 11897–11906
- 61 Fu X, Shi X, Yan L, et al. *In vivo* substrate diversity and preference of small heat shock protein IbpB as revealed by using a genetically incorporated photo-cross-linker. *J Biol Chem*, 2013, 288: 31646–31654
- 62 Fu X, Wang Y, Song X, et al. Subunit interactions as mediated by “non-interface” residues in living cells for multiple homo-oligomeric proteins. *Biochem Biophys Res Commun*, 2019, 512: 100–105

Summary for “蛋白质之活细胞探究”

Exploring proteins in living cells

Zengyi Chang

Protein Science Center, State Key Laboratory of Protein and Plant Gene Research, School of Life Sciences, Peking University, Beijing 100871, China
E-mail: changzy@pku.edu.cn

Proteins have been studied for over 200 years, but much of that work has been performed *in vitro*. This article briefly summarized the current status of protein exploration in living cells, and then presented some of my own representative findings in this field. First, I summarized the major methodologies hitherto developed for studying proteins *in vivo*, including visualizing a target protein by fusing it with the green fluorescent protein; investigating the tertiary structure of the target protein via in-cell nuclear magnetic resonance spectroscopy; and exploring the behaviors of a target protein by genetically labeling it with different types of unnatural amino acids.

Next, I described how my laboratory applied and modified an *in vivo* methodology in which the genetic introduction of a photoactivatable, unnatural amino acid at a selected residue position in a target protein then mediates crosslinking of that protein with other proteins when the cells are irradiated by UV light. By improving that method, my laboratory contributed the following: (1) The creation of a bacterial strain with a genome that has a gene that encodes the orthogonal aminoacyl-tRNA synthetase and one that encodes the tRNA that are both needed to incorporate an unnatural amino acid (i.e., Bpa) into any target protein. (2) The development of a system in which the unnatural amino acid is randomly introduced at each of the target protein's amino acid residue positions and, after detecting a meaningful interaction, DNA sequencing subsequently identifies that unnatural amino acid's exact position. (3) The incorporation of the unnatural amino acid into a target protein by directly modifying the endogenous target gene in the genomic DNA so that the variant protein is produced spatiotemporally identical to that of the wild-type protein. (4) The development of a high-throughput polyacrylamide gel electrophoresis system in which 384 protein samples can be analyzed in a single gel.

Finally, I described a few representative results from my lab by using *in vivo* protein photocrosslinking analysis: (1) The serendipitous discovery of a reversible subcellular structure (regrowth-delay body) that forms when a highly selected group of important proteins assembles only in non-dividing bacterial cells, and then dissolves when the cells resume growth. Their presence thus marks dormant persister bacterial cells that are known to tolerate antibiotics. (2) A revelation of a “gear shifting” mechanism that modulates bacterial ATP synthase activity via adaptation of its ϵ -subunit to the “inserted” or “non-inserted” conformation as the cell responds to a variety of favorable to unfavorable cellular conditions. (3) The identification of a supramolecular protein complex that is composed of proteins present in the cytoplasm, inner membrane, periplasm and outer membrane of Gram-negative bacteria, and that mediates the biogenesis of a group of outer membrane proteins. (4) The unveiling of a unique way that the acid stress chaperone HdeA acts. It exhibits biological activity while its tertiary structure becomes largely disordered, and then it protects other chaperones under acidic conditions. (5) A clarification that the long-designated bifunctional protein DegP acts primarily as a proteinase with little chaperone activity. (6) The discovery that small heat shock proteins are activated in a temperature-responsive, cascade manner that allows them to significantly increase their biological activity at high temperatures. I then conclude with a list of unresolved issues regarding protein exploration in living cells.

***in vivo* protein photocrosslinking mediated by unnatural amino acids, high-throughput PAGE (polyacrylamide gel electrophoresis), regrowth-delay body, ATP synthase, protein biogenesis, acid stress chaperone HdeA**

doi: [10.1360/TB-2021-0845](https://doi.org/10.1360/TB-2021-0845)