关于可控组装的一些思考(二)——生命体中的催组装

雷志超², 王忻昌¹, 瞿航¹, 周岑³, 李之豪¹, 杨柳林¹, 曹晓宇¹, 田中群^{1*}

1. 固体表面物理化学国家重点实验室(PCOSS), 厦门大学化学化工学院化学系, 能源材料化学协同创新中心(iChEM), 厦门 361005

2. 北京大学化学与分子工程学院, 软物质科学与工程中心, 高分子化学与物理教育部重点实验室, 北京 100871

3. 闽江学院海洋学院, 福州 350108

*通讯作者, Email: zqtian@xmu.edu.cn

收稿日期: 2020-07-11; 接受日期: 2020-08-27; 网络版发表日期: 2020-11-04

国家自然科学基金(编号: 91427304, 21991130, 21991131)资助项目

摘要 分子组装是构成分子以上层次物质世界的重要方式. 生命体内的分子组装展现了分子组装的最高形式, 在组装基元、组装过程、组装环境、组装体功能等方面都显得精妙复杂. 不同于人工分子组装中最为常见的自组装方式, 生命体内的分子组装具有多路径、多步骤和多级次的复杂性, 常常需要其他分子的参与、协助和调控. 在这些调控方法中, 催组装利用催组剂调控组装过程的速率和路径, 是一种高效且高选择性的调控方法. 本文首次从催组装的视角, 对生命体中的分子伴侣、组蛋白伴侣、别构因子等催组剂及其催组装过程进行了归纳和综述, 这些催组剂在蛋白质折叠、染色质重塑、酶活性调节等过程中发挥着不可替代的作用. 越复杂的组装越需要催组装, 本文在综述这些催组装实例和分析其优势和规律的基础上, 提出了在分子组装研究中, 向生命体系学习而发展催组装的必要性及可能的途径, 以期提升分子组装研究走向更高层次, 并有望为软物质科学和生命科学领域的基础理论和功能应用研究提供新的视角和方法.

关键词 分子组装, 可控组装, 催组装, 催组剂, 分子伴侣

1 引言

分子组装(molecular assembly)指的是在分子间非共价键相互作用下, 若干分子形成某种有序结构聚集体的过程, 是构建分子以上层次的功能材料和生命体系的主要方式^[1-4].

人工分子组装的研究主要集中在超分子化学^[5,6]、高分子组装^[7-9]、表面分子组装^[10-12]等相关领

域. 分子组装研究的主流方向是分子自组装, 它指的是组装基元依靠其之间的相互作用力自发地形成有序结构聚集体的过程^[13,14]. 这些聚集体包括有限基元的组装体, 例如各种主客体体系^[15-17]、拓扑结构组装体^[18]、DNA纳米结构^[19]等; 以及无限基元的组装体, 如超分子聚合物^[20-23]、超分子有机框架^[24-26]、自组装分子单层^[10-12]、DNA三维晶体^[27]等等.

虽然人们已创造出众多分子组装基元, 并对基元

引用格式: Lei ZC, Wang X, Qu H, Zhou C, Li Z, Yang L, Cao X, Tian ZQ. Some thoughts about controllable assembly (II): catassemblies in living organism. *Sci Sin Chim*, 2020, 50: 1781-1800, doi: 10.1360/SSC-2020-0127

间相互作用进行了研究,但目前的人工组装体系仍过于简单,与存在于生命体内的大量复杂组装体系差距较大,尚无法为生命体中的组装现象提供全面清晰的物理化学图景;人工组装体也尚不具备生命体所拥有的复杂精巧的功能。

人们愈发地意识到,分子组装的研究模式的不足是造成这种差距的主要原因^[29-31]。如图1(a)所示,人工分子组装中最为常用的方式是“自组装”,其典型模式为在共价合成组装基元的基础上,通过非共价键相互作用而自发形成组装体。我们认为,分子组装研究的发展,可能长期受到了“自组装”中“自”的观念的限制。这些研究注重面向热力学自发组装过程的组装基元的设计合成,得到的组装体也较为简单(图1(b~g)),而对组装动力学过程、驱动力的精准调控、可重复性等关注较少。

然而并非所有的组装过程都是自发的,复杂的组装往往需要在特定的环境中发生,其过程常常具有多方向性,需施加额外调控。外部调控的引入,将使得组装过程得以发生,且具有更高的可控性。因此,我们在此系列文章的首篇中提出,将分子组装分为自组装和

助组装两种;根据“助”的方式不同,可将助组装进一步分类为共组装、模板助组装、场助组装、催组装等^[29,30]。分子组装研究的进一步发展,需要对现有组装方法进行梳理,需要在自组装之外提出和发展可控组装新方法。

与各种人工分子组装体系相比,生命体展现了更为精巧、复杂、高级的分子组装现象,尤其是生命体中已经发展出的一整套高效、精巧且温和的组装调控手段和相应的分子机制。我们认为,通过辨析生命体中的分子组装和目前所研究的人工分子组装的主要异同,认知并学习生命体中的分子组装调控方式,对于寻找人工分子组装的发展新方向、揭示生命过程与其机制皆具有重要意义。

2 生命体分子组装与化学人工分子组装的差异

我们认为,人工分子组装与生命体分子组装差距巨大,其主要差距在于组装的复杂度和可控程度上,这具体表现在组装基元、组装过程、组装环境和组装体功能等以下四个方面。

2.1 组装基元

人工分子组装基元与生命体内的组装基元的主要差异体现在分子量大小、官能团多少、结构的复杂程度和分子的可控变构程度上,由此造成了各自在组装过程和最终形成的组装体上的巨大差别(图2(a~c))。

生命体中的组装基元往往比人工合成的组装基元分子量更大,官能团数量和种类更多;这意味着其组装基元之间相互作用位点和作用方式更加复杂,进而导致其组装过程的复杂性。尽管各种人工高分子在分子量上与生物大分子相近,但其在官能团的种类和数量上,相较于生物大分子仍然较少。

从结构的复杂程度看,生命体组装基元往往比各种人工组装基元复杂得多。一个生命体组装基元往往具有特定的手性且拥有多个不同的组装位点和作用界面。例如,有些酶分子上既有底物的结合位点,也有激活剂和抑制剂的结合位点。其结构的复杂同样使得生命体组装基元之间的相互作用力和作用方式变得复杂。与此同时,生命体分子的分子内组装也相当复杂,最典型的例子便是蛋白质和RNA分子的折叠。

图1 常见的人工分子组装体系: (a)人工分子组装一般方式为自组装; (b)主客体体系^[28]; (c)分子结^[18]; (d)表面自组装分子单层^[13]; (e)超分子有机框架^[25]; (f)超分子聚合物^[23]; (g)三维DNA晶体^[27] (网络版彩图)

Figure 1 Typical artificial molecular self-assembly: (a) typical method of molecular self-assembly; (b) host guest complexes^[28]; (c) molecular knots^[18]; (d) self-assembled monolayers^[13]; (e) supramolecular organic frameworks^[25]; (f) supramolecular polymers^[23]; (g) 3D DNA crystals^[27] (color online).

图 2 常见的生命体分子组装体系: (a) 血红蛋白^[32]; (b) 蛋白酶体^[33]; (c) 染色质 30 纳米结构^[34]; (d) 常见的病毒组装过程^[35]; (e) 驱动蛋白在微管上的动态组装过程^[36,37] (网络版彩图)

Figure 2 Typical biological molecular assembly: (a) hemoglobin^[32]; (b) proteasome^[33]; (c) 30 nm fiber of chromatin^[34]; (d) assembly of virus^[35]; (e) kinesin moving along microtubule^[36,37] (color online).

生物大分子一般拥有较强的可控变构性, 更适合于多个位点的协同作用. 常见的生物大分子, 如蛋白质、RNA、DNA 都有较强的可控变构能力和响应性. 引起变构的原因多样, 例如, 与其他分子的结合、共价修饰、溶液电解质状态的改变等. 变构与否直接影响了生物大分子的组装行为. 例如酶的别构效应, 结合了激活剂或者抑制剂可以引起酶的变构, 进而影响了其与底物的识别和结合, 从而导致酶活性的改变. 又如核糖开关在结合其他生物分子之后引起分子变构, 进而影响翻译过程^[38,39]. 变构是生命体组装基元和人工组装基元的最重要差别之一, 在变构性的基础上, 生命体组装基元还具有较强的自适应性. 这两者赋予了生命体组装基元更丰富的可调控性和响应性, 不仅使得其组装行为变得更加复杂, 更由此创造了强大的功能.

表观上人工分子基元与生命体分子基元两者的差异体现在分子量大小、官能团多少、分子链结构的复杂程度等, 但更本质的差异是生命体组装基元在组装时所提供信息的量和可控性要远大于人工分子组装基元, 由此造成了两者在组装过程和组装体功能上的巨大差别. 组装基元之间的多位点相互作用不仅可以达到共价键的强度, 更为众多非共价键之间的协同提供

灵活多变的可能性, 由此促进了生命体的多级组装和强大的功能的实现. 目前人们所能获得的最复杂的人工组装体, 都是使用生物大分子, 如 DNA 和蛋白质作为组装基元获得的^[40-44]. 正是由于组装基元的复杂精妙, 方才形成了丰富多彩且功能强大的各类生命体.

2.2 组装过程

生命体中的组装过程除了自组装之外, 也存在着众多的助组装, 其具有多路径、多步骤、非平衡态组装、需要解组装等特点(图 2(d, e)).

生命体的分子组装过程往往是多路径的, 这是由组装基元的结构复杂性所造成的. 例如, 人们早期认为蛋白质的折叠(分子内组装)是自发完成的, 折叠路径较为单一^[45]. 但是, 如图 3 所示, 后续的系统研究表明, 蛋白质的折叠路径是非常复杂的, 很容易发生错误折叠. 目前已经发现了多种蛋白质构象病, 均由蛋白质的错误折叠引发^[46-50]. 例如, 人的 A β 蛋白有两种折叠路径, 一种形成正常的 A β 蛋白, 另一种发生错误折叠, 形成淀粉样的沉淀, 进而引发多种退行性疾病^[51,52]. 很多 RNA 分子也存在着错误折叠的情况^[53,54].

生命体的分子组装过程往往是多步骤的. 以膜蛋白的组装为例, 其折叠是在合成的同时进行的. 对于

图 3 蛋白质折叠的势能面表明其经由复杂多路径形成各种组装产物^[49] (网络版彩图)

Figure 3 The complex folding energy landscape of protein to different assembly products^[49] (color online).

多亚基的膜蛋白, 需要等多个亚基合成并折叠完成后才能进一步组装, 之后在肌动蛋白等多种蛋白的参与和帮助下被转运至细胞膜上, 方可完成最终的组装过程. 整个组装过程需要多种蛋白的参与和协助, 是典型的助组装过程, 其复杂程度是迄今的人工组装过程无法比拟的^[55,56].

生命体是一个远离平衡态的耗散组装体系. 生命体内的很多组装过程是由光能、化学能等所驱动的. 生命体组装基元的可控变构性, 使得其能将ATP等能量分子的结合与组装基元的变构偶联起来, 通过磷酸基团水解释放的化学能来驱动生物分子之间的相互识别和组装^[57]. 例如, 微管蛋白的组装就是在GTP水解的驱动下进行的^[58].

在复杂的生命体分子组装中, 熵扮演着非常重要的角色. 蛋白质等组装基元的表面上往往存在着数量巨大的特定取向的水分子. 在其组装过程中, 这些水分子会被释放到溶液中去, 带来显著的熵变, 其所带来的能量变化远大于组装基元之间相互作用产生的焓变, 故成为生物大分子组装的主要驱动力^[59-62].

在复杂生命体多步骤和多级次的组装中, 解组装也变得非常关键. 通过解组装可以实现组装基元的循环利用以及调整组装体状态等目的, 拥有重要的生物学意义. 例如, 微管蛋白在组装形成微管执行完功能之后, 又会解组装回到细胞基质, 再参与其他微管的形成^[58]. DNA在复制时, 需要先发生解组装, 复制完成之后又会形成新的组装. 生命体巧妙地利用着组装和解组装的协同和反馈过程来实现特定的生命过程.

特别要强调的是, 不可将解组装简单理解为解散所有组装基元而回到组装初始原点, 它往往是在多步骤乃至多级次的组装过程中, 仅仅解开了部分的组装体, 以便进行重构、纠错组装中的偏差, 或者赋予这类组装体新的功能. 从广义而言, 它存在于组装全过程中. 因此, 解组装带给生命体的功能, 丝毫不亚于组装过程, 这显著不同于简单的人工分子组装过程, 也不同于复杂多步骤化学反应中的某个逆过程.

由于组装基元及其相互作用的复杂, 生命体内的分子组装过程要复杂得多, 在分子层次上对这些复杂过程进行研究和理解是非常重要且必要的.

2.3 组装环境

目前人工分子组装的研究中, 大部分使用的是稀

溶液, 且有大量的体系是在有机溶剂中进行的. 反之, 生命体中的分子组装大多在溶质浓度极高的水溶液中进行. 复杂的组装环境使得生命体中的分子组装产生了与稀溶液中迥异的行为 (图4).

细胞内的分子数量巨大、种类繁多. 人类大约有两万多个基因, 在特定细胞中表达的蛋白质数量一般都在数万以上^[63]. 除蛋白质之外, 细胞内还存在着大量的核酸、油脂、各种辅酶因子、游离的氨基酸、核苷酸等分子. 这些大大小小的分子使得细胞内变得十分“拥挤”. 人体质量的70%为水, 即粗略而言, 细胞大约是一个30%质量浓度的浓溶液. 已有研究表明, 细胞中仅蛋白质的浓度就达到了200 g/L. 这使得细胞内存在所谓的大分子拥挤效应^[64-68].

生命体分子组装正是在这样的高浓溶液状况下进行的, 周围的环境对组装行为产生了重要的影响. 拥挤的环境既可能使得分子之间的相互识别和组装变得困难, 也可能在某些特殊情况下加快识别与组装过程.

在拥挤的环境下, 组装基元周围的其他分子很可能会干扰组装基元之间的相互识别和组装. 例如, 当新合成的蛋白质浓度过高时, 新生蛋白质很容易在正确折叠完成前相互干扰, 因其疏水区域的相互作用而

图4 不同环境中分子组装的行为会产生显著差异: (a) 稀溶液中的分子组装过程; (b) 拥挤的环境下的分子组装会减慢; (c) 有膜或者无膜细胞器中分子组装会加快 (网络版彩图)

Figure 4 Molecular assembly in different environments: (a) molecular assembly in dilute solution; (b) molecular assembly will be slowed down in crowded solution; (c) molecular assembly will be accelerated in membranous or non-membranous organelles (color online).

产生蛋白质团聚。

而在某些特殊情况和某类驱动力下, 生命体内的复杂环境会使得分子间的相互识别和结合变得更加容易^[66]。人们已发现, 在各种细胞器中, 有些生物分子之间的识别和结合因为“拥挤”而加快了。近年来, 人们在细胞内发现了在某些特定时刻, 细胞内发生了生物学分相, 形成“无膜细胞器”。在这些特殊的相中, 一些特定的具有协同性的分子在空间上聚集, 这样它们的相互作用也得到了加速。这成为当前生物学的研究热点之一^[69-73], 有关其物理和化学图像、机制的研究方兴未艾, 是生命科学、物理和化学的共同挑战之一。

2.4 组装体功能

各类生命体在进化中不断形成具有更高级功能的复杂组装体。生命体是一个耗散体系, 其活动的维持需要能量、物质和信息的输入, 生命体由此发展出了能量、物质和信息的获取、消耗和存贮的机制和“装置”, 以维持生命活动的运转。例如, 细胞膜这一复杂不对称组装体既可以处理外界的信号刺激, 又可以接受外界物质。又例如, 植物依靠光合作用汲取能量, 因此发展了光合作用系统。又例如动物需要通过摄入植物和微生物来摄取能量, 因此需要运动系统和进食系统。生物体还可以将组装产生的废物等排出体外。生命体在进化中拥有了自修复能力。这些复杂和强大的功能, 都是建立在复杂且精巧的组装之上, 且其之间相互协同, 形成了丰富多彩的生命现象。这是迄今所实现的人工组装体系可望而不可及的。

上述四点分析并阐述了生命组装体的复杂性所在, 表1对此进行了简要的小结。生命体正是在此基础上, 发展出了复杂多级的从分子到细胞器、细胞、器官等的组装。因此, 对于生命体的分子组装特点可归结为, 在复杂环境里, 复杂组装基元经过复杂的组装过程形成拥有复杂功能的组装体。这种复杂性是全方位的, 远远超过人工分子组装体系, 难以用现有的分子组装理论进行解释, 很可能需要从新的、全面的视角给予诠释。

3 越复杂的组装越需要协助与调控

正如生物分子在自然进化中变得越来越复杂和功能越来越强大, 人工分子组装研究的大方向也必然是

表 1 目前人工分子组装与生命体分子组装的主要区别

Table 1 Differences between artificial and biological molecular assembly

	人工分子组装	生命体分子组装
组装基元	较少且简单	分子量大、数量多且复杂
基元变构性	较弱且不可控	高度可控
组装过程	路径较简单	路径多且复杂
能量输入	非耗散为主	耗散为主
熵/焓	以焓为主	熵的影响很大
组装级次	一般为单级、简单协同	多级组装、复杂协同
催组剂	无或简单催组剂	多种不同功能的催组剂
动力学	简单	复杂
组装环境	稀溶液	拥挤
组装体功能	较为单一	多种且强大

从简单走向复杂^[2]。要提升人工分子组装的研究水平, 亟需向具有复杂组装和强大功能的生命体学习, 不断探索和发展用于复杂组装的各种方法。

对于简单组装, 组装基元之间较容易通过自发的相互作用, 通过“自组装”的方式形成组装体。但是对于多步骤、多路径的复杂组装, 其过程的路径和有效势能面都很复杂, 难以自发形成最终的复杂组装体。此外, 细胞内的环境复杂, 组装过程很容易受到周边其他分子的干扰, 这些组装过程似乎更加难以自发完成, 它们究竟如何做到准确高效而不出错呢?

显然, 这些复杂组装需要其他分子的调控和协助, 才能够高效、准确地进行。生命体中存在大量的助组装, 且已经发展出了一整套高效的手段和相应的分子机制, 以确保这些复杂组装得以高效高选择性进行。下面我们将梳理对于一个组装过程可供使用的调控和协助组装的机制, 这对于发展出更具可控性的分子组装方法很有益处。

3.1 分子组装的调控手段分类

总体而言, 要调控一个分子组装的全过程, 可使用的手段依序分为如下四种手段: (1) 组装开始前对组装基元的调控; (2) 组装过程中的化学调控手段; (3) 组装过程中的物理调控手段; (4) 组装结束后对组装体的调控。

组装开始前和结束后的调控往往是利用共价合成的手段对基元或者组装体进行改造。在组装前, 很多组

装基元仍旧处于组装前体状态, 需要使用共价修饰、酶催化等方法改变其结构, 使其变成真正适合组装的基元进而继续组装。这属于我们所分类的助组装中的酶助组装^[74]等。在组装结束后, 许多高级组装体仅依靠非共价键不能稳定存在, 还需要使用共价键对组装体的各连接部分进行加固, 以获得稳固的产物。这是我们之前介绍过的催组联过程(催组装+化学偶联反应)^[29]。这两种手段都是合理利用共价和非共价的多步协同, 从而更好地获得目标产物。

组装过程的化学调控手段, 是指加入组装基元以外的化学物质到体系中, 直接参与或协助组装过程。这些加入的分子可以是溶剂分子, 通过改变溶剂来调控组装过程^[75]; 可以是电解质分子, 通过改变溶液的离子活度来调控组装过程; 可以是酸碱, 通过改变溶液的pH来调控组装过程^[76]; 加入的化学物质还可以是共组装基元, 如模板^[10]、手性诱导剂^[77]等; 也可以加入两亲性分子等形成囊泡, 使某些组装过程能局限在一定的空间内进行。

组装过程中的物理调控手段, 指的是将力、热、光、声等引入到组装体系中。这些物理手段有些为组装过程提供了能量, 如加热、微波等^[78]; 有些则直接作用在组装基元分子上, 促进活性组装基元的结构改变, 如电场^[79]和光场^[80]; 有些则促进了组装的传质过程, 如搅拌、超声^[81]等。

由于生命体的组装大多发生于常温常压和水溶液的环境, 所以, 使用具有调控功能的分子来协助组装过程的温和手段在生命体中最为普遍。例如各种起始因子在时间上调控组装过程; 又如众多的“搬运工”分子在传质上促进组装; 再如分子伴侣等在能量上帮助分子组装越过能垒, 引导其进入正确的组装路径。

3.2 越复杂的组装越需要催组装

在调控和帮助分子组装的助组装手段中, 有一类称为催组装(catassembly), 即通过加入一类称为“催组剂”(catassembler)的物质, 加速组装过程, 或者协助组装基元选择合适的组装路径而获得目标组装体; 催组剂在组装前后不发生变化, 也不存在于最后的组装体中。究其功能而言, 催组剂之于分子组装, 正如催化剂之于化学反应。我们在有关催组装的系列文章的首篇曾较全面比较了催组装与催化反应^[29,30]。

在助组装中, 催组装相较于其他调控手段, 具有以

下多种优势和特点。

第一, 催组装具有高效性。催组剂可以有效地降低组装过程的能垒, 加速组装过程; 催组剂也可以在空间上协助组装基元的有效传质, 运送组装基元到合适的待组装位置。

第二, 催组装具有高选择性。催组剂可以选择特定的组装基元, 帮助其组装。催组剂还可以针对拥有复杂组装路径的组装过程, 调控和选择目标组装路径, 抑制其他路径的组装, 进而得到目标组装体, 提高其选择性。

第三, 催组剂像催化剂一样, 可以反复使用。一个当量的催组剂可以协助多个当量的组装体的形成。催组剂可以在不同时段和不同区间发挥作用, 这样的方式是经济且高效的。

应当指出, 催组剂的种类和功能是非常丰富的: 在时间上, 催组剂可以调控组装过程的快慢, 也可以调控组装过程的起止; 在能量上, 它可以对组装过程的势能面进行调控, 使其进入特定的组装路径, 催组剂还可以将外界能量转化为催组装循环的驱动力; 在空间上, 催组剂可以帮助组装基元的传质, 引导其到达合适的位置。一个复杂的组装过程, 往往会有多种催组剂协同参与到组装的不同阶段。

为了通俗易懂地区分自组装、简单体系催组装和复杂体系催组装, 我们用图5中的三类建桥方式给予说明。自组装可视为在河上自行拼装成浮桥, 将具有多个特定形状的木板块从上游投入河中, 使之顺流而下在河的狭窄之处互相碰撞而不断调整从而实现组装, 最终有望拼装成浮桥整体, 这一过程的时间往往是漫长的。简单催组装体系则是由搬运工(一类催组剂)将所需板块先后运到河的狭窄处, 再由建筑工(另一类催组剂)将其拼装为浮桥。如此拼装浮桥的效率显然会高得多, 且搬运工和建筑工可以离开去搭建更多的浮桥。但是, 若要实现多级且功能强大的组装体——建造一座跨河大桥, 不仅需要各式各样的建桥部件, 而且需要各司其职的建筑工、搬运工和调配人员的全方位协作。总之, 为了实现复杂的多级组装, 需要各式各样的催组剂的参与, 方可从多个层面上协助和调控组装全过程的完成。

由上文所知, 越复杂的组装越需要协助, 在目前各种助组装方法中, 催组装是最为高效和最具选择性的手段之一。因此, 催组装的概念和方法很可能是促进

图 5 几类组装的示意图: (a-c)自组装; (d-f)简单的催组装体系; (g)复杂的催组装体系 (网络版彩图)

Figure 5 Self-assembly and catalytic assembly: (a-c) self-assembly; (d-f) simple catalytic assembly; (g) complex catalytic assembly (color online).

分子组装从简单走向复杂的重要推手, 也很可能成为推进生命体组装从唯象解释走向系统科学理论的途径之一。

通过文献调研可知, 在简单的人工分子组装研究中, 催组装体系较少; 而在复杂的生命体研究中, 催组装的实例则数量众多。扮演催组剂角色的各种大分子、小分子在蛋白质折叠、遗传信息复制和维护、酶活性的调节等重要的生命体活动上发挥着不可替代的作用。这清晰地说明越复杂的组装体系越需要催组装, 越需要多种催组剂的协助。对于复杂的多级组装体, 具有不同功能的催组剂通过某种信号反馈网络, 在各个步骤和时段中发挥特定的作用。我们将在下文中, 对生命体中的催组装进行较详细的总结和讨论, 以期认知复杂体系的催组装, 进而启发人工催组装体系的设计和构建, 并进一步认知生命机制与过程。

4 生命体中的催组装

催组装这一高效、高选择性的组装调控手段在生命体中广泛存在。在生命体组装中, 扮演催组剂角色的既可以是高分子, 也可以是小分子; 催组装的对象既可以是单个基元也可以是分子组装体; 它们的作用机制也是多种多样, 在众多重要的生命体活动方面发挥着不可替代的功能。下面将分类介绍这些催组装过程和相应的催组剂。

4.1 协助分子折叠的分子伴侣

人们在20世纪80年代以后, 陆续发现生物体内存在一类可以帮助其他生物大分子折叠或者组装, 且不存在于最终组装体中的分子, 这些分子被称为分子伴侣(molecular chaperone)^[49,82-84]。“Chaperone”一词原意是指陪伴青少年贵族女性的妇女, 她们的作用主要是帮助青少年贵族女性出入社交场合, 学习各种礼仪, 纠正其错误, 进而成长为合格的贵族妇女。这个词很形象地描述了早期发现的分子伴侣的主要功能, 即帮助分子“成长”为正确构象的蛋白质——进行正确折叠和组装, 以及纠正错误折叠。分子伴侣按照其作用对象的不同, 可以分为两类: 针对分子内组装(折叠)的和针对分子间组装的。早期获得较多重视的分子伴侣大多是帮助蛋白质折叠的。

蛋白质作为生命活动的主要行使者, 其折叠构象的正确与否, 对于其能否正确行使功能非常关键。20世纪60年代, 科学家们发现当生物体的温度过高时, 大量的热休克蛋白(heat shock protein, Hsp)会开始产生^[85-88]。一系列的实验表明, 这些热休克蛋白的存在保证了其他蛋白质的正常折叠。后期的研究则发现, 即使在正常生理条件下, 还是存在一定数量的分子伴侣维持着细胞内蛋白质的折叠。研究的比较充分的蛋白质分子伴侣包括GroEL/GroES家族、Hsp 70家族、Hsp90家族等。这些蛋白质帮助新生肽折叠成具有完整功能的蛋白质, 其作用机制各不相同, 有些依赖于ATP, 有些则不需要, 它们参与到折叠的阶段也不相同^[49]。

图6是被广泛和深入研究的GroEL/GroES家族^[49,84,89-91]。GroEL是一个对称的双桶状分子, 可以在不同的时段在其两端口结合由GroES蛋白组成的“盖子”。在其一端结合上ADP和GroES盖子的状况下, 另

图6 分子伴侣协助蛋白质分子的折叠^[49] (网络版彩图)

Figure 6 The folding of protein assisted by molecular chaperone ^[49] (color online).

一端的桶内侧会保持疏水性和开放。由于新生肽的大量疏水部分还暴露在外, 新生肽很容易进入这一疏水性的桶内。此时该侧GroEL会结合ATP, 导致组成GroEL的7个蛋白亚基发生变构, 将桶内部转变为亲水状态, 同时GroES盖子结合上来。在这个关闭的亲水性的腔体内, 新生肽可以很好地完成其组装过程。当ATP水解成ADP之后, 桶内侧再次变回疏水状态, 不仅驱使GroES盖子离去, 而且会将表面已变得亲水的蛋白质从桶内排出。随着另一新生肽的到来, 这一循环会再次开始。整个过程的能量由ATP水解成ADP提供。

分子伴侣还有另一个重要功能, 可使已经发生了错误折叠的蛋白质先解组装再完成正常折叠。由于各类蛋白在细胞这一拥挤环境内, 很容易与其他分子发生特异或者非特异的相互作用而导致其折叠状态发生改变。蛋白质在折叠错误的情况下甚至会发生团聚形成纤维, 引起阿尔兹海默症、帕金森症等各种蛋白质构象病(图3)。细胞内的分子伴侣起着维持蛋白质正确折叠构象的重要作用^[92]。

值得注意的是, 协助蛋白质折叠的分子伴侣的特异性并不是很高, 往往是一种分子伴侣可以帮助一类蛋白质的正确折叠, 例如, GroEL/GroES可以帮助细胞内约10%种类的蛋白质折叠^[49]。这在细胞内空间和资源有限情况下, 是非常经济和有效的做法, 不可能每个特定蛋白质都配有相应的分子伴侣。这些分子伴侣消耗由ATP供给的化学能, 帮助蛋白质分子回到正确的折叠路径, 完成其分子内组装; 在折叠完成后, 又主动地离开组装体; 它们是很典型的催组剂。

4.2 协助大分子间组装的分子伴侣

分子伴侣(chaperone)不仅可以协助蛋白质和RNA

正确折叠(调控分子内组装), 还可以帮助其他生物大分子完成分子间组装, 例如Laskey等^[93]发现的nucleoplasmin (NAP1)蛋白, 这也是首个报道的分子伴侣。他们发现, 只有在NAP1存在的情况下, DNA才能与组蛋白组装形成染色质。Chaperone一词最早也是由Laskey使用^[93], 之后Ellis等^[82]在总结这两类帮助蛋白质折叠和组装的分子后, 系统地提出了分子伴侣的概念。目前发现的调控组装过程的分子伴侣在多种重要的生物大分子复合物, 如RuBisCO、染色质、核糖体、蛋白酶体等的组装上起到了关键的作用^[94,95]。

植物的RuBisCO复合体是光合作用中一个非常重要的酶, 它由8个长链亚基和4个短链亚基分子组装而成, 分子量高达540 kD。2010年Hayer-Hartl等^[96]发现多种分子伴侣参与了RuBisCO的组装过程, 其中RbcX₂起到了非常关键的作用。RbcX₂先促进了两个长链亚基的组装形成二聚体, 进而形成八聚体; 然后在和短链亚基组装结合形成完整的RuBisCO后, RbcX₂会离开组装体(图7(a))。他们之后又发现, Raf1、Raf2、BSD等分子伴侣也参与了这一过程^[97,98]。

染色质是真核生物遗传信息的载体, 其由DNA和组蛋白组装而成。染色质的基本单元是核小体, 其由约146 bp的DNA缠绕着H2A、H2B、H3和H4组成的组蛋白八聚体而形成^[99], 进而形成更高级结构^[34]。染色质的组装结构在形成之后并非静止不变的, 在行使各种生命功能时, 其结构始终处在动态的组装和解组装过程中。例如, 在染色质的复制和转录过程中, 核小体结构需要解组装, 从而让复制和转录的分子机器通过; 在其通过之后, 核小体需要重新组装。迄今已发现三十多种组蛋白伴侣参与到核小体动态组装过程的调控中^[100]。

FACT(facilitates chromatin transcription)是一种典型的组蛋白伴侣, 它在哺乳动物细胞的转录过程中必不可少, 对于促进RNA聚合酶在染色质上的延伸起着至关重要的作用^[103-106]。李国红等^[101,104]通过单分子磁镊技术发现, 在不存在FACT分子的情况下, 需要使用11 pN的作用力才能使核小体外圈结构解组装, 而且在多次拉伸的情况下核小体结构很难再重新组装; 在FACT结合到核小体的情况下, 只需要7 pN的力就可以使核小体外圈结构解组装, 而且核小体结构很容易重新组装, 可以多次反复拉伸。总而言之, FACT起到了降低解组装的能垒, 并且促进了重新组装的发生, 是

图 7 协助分子组装的分子伴侣: (a)RbcX₂催化RuBisCO的组装^[96]; (b)FACT协助核小体的组装和解组装^[101]; (c)SWR1协助常规核小体的解组装, 促进H2A.Z核小体的组装, INO80则起到相反的作用^[102] (网络版彩图)

Figure 7 The assembly and disassembly of molecules assisted by molecular chaperones: (a) The assembly of RuBisCO complex assisted by RbcX₂ [96]; (b) FACT can assist the disassembly and assembly of nucleosome during transcription [101]; (c) SWR1 can assist the disassembly of canonical nucleosome and the assembly of H2A.Z nucleosome, INO80 can assist the reverse process [102] (color online).

这一动态组装过程的催组剂(图7(b)).

除了常见的五种组蛋白之外, 还存在着多种组蛋白变体^[107,108]. 组蛋白变体与常规组蛋白序列相似, 但在结构上存在一定的差别, 因此起到了不同的生物学功能. 核小体中组蛋白种类的不同会影响染色质的性质, 故对组蛋白进行替换是对染色质进行调控的一种重要手段. 组蛋白的替换需要原有核小体的解组装, 进而再重新组装成新的核小体. 组蛋白的替换是在多种组蛋白分子伴侣的协助下进行的^[100]. 2004年, Wu等^[109]发现了一类称为SWR1的组蛋白伴侣, 它可以特异性地促进染色质中的H2A-H2B二聚体解组装离开核小体, 并且可以促进H2A.Z-H2B与核小体的组装, 也就是说将H2A替换为H2A.Z, 从而激活特定位点的基因. 此外, 还存在另外一种组蛋白伴侣称为INO80, 则可以特异性地将H2A.Z替换为H2A^[102,110]. 这两个过程都是典型的催组过程(图7(c)). 相较于帮助折叠的分子伴侣, 目前大多数帮助分子间组装的分子伴侣的机制还不清楚, 有待于进一步的深入研究.

上述分子伴侣是调控这些组装和解组装动态过程的催组剂. 有时单个催组剂可以在一个复杂组装过程的不同阶段起到不同的作用, 例如上文提到的FACT

在转录不同阶段中起到不同的作用; 有时则需要多种催组剂参与到一个复杂的组装过程的不同阶段中, 例如上文提到的RuBisCO的组装过程, 单体的折叠过程需要GroEL-GroES的协助, 而组装过程则需要RbcX₂的协助.

这些分子伴侣的主要功能不再局限于帮助分子折叠和组装, 而是通过改变组装体的结构和状态, 进而调控其功能. 将这些调控组装体结构和状态的分子称为“分子伴侣”也许不太能准确描述其功能. 另外, 高级生命组装体都是经由多级组装, 对其精准把握不仅需要分层次的清晰脉络, 而且需要系统地考虑全面图像和通用语言. 对于各个不同层级的过程的描述, 若仍然使用“分子伴侣介导的组装过程”一词似乎也不如用“催组过程”的适用面更广. 因此, 使用“催组”和“催组剂”也许更有利于阐释上述过程的科学内涵, 并为上述体系的研究提供新的视角.

4.3 小分子催组剂

在生命体中, 催组剂的范畴远不止于分子伴侣. 除了大分子的分子伴侣之外, 生命体中也存在对组装过程起调控作用、但不被称为分子伴侣或组装伴侣的众

多小分子, 其中很多都起着催组剂的作用。

在一些病毒颗粒的形成过程中, 很重要的步骤是其前体的组装。例如, HIV-1病毒前体的基本单元是Gag蛋白的六聚体, 该六聚体的核心部分是由6个Gag单体中的SP1片段组成的六聚螺旋, 其形成是整个组装过程中最重要的步骤。SP1片段中包括两个赖氨酸残基, 带有较强的正电荷, 使得六个单体之间相互排斥, 难以形成六聚体。Dick等报道6-磷酸肌醇可以促进该六聚体的形成^[111~113]。他们通过结构生物学的手段, 发现该带负电的6-磷酸肌醇分子出现在六聚体的中心, 很好地平衡了赖氨酸的正电荷, 从而促进了六聚体的形成。他们进一步观察到, 在组装前体熟化为病毒颗粒时, SP1片段会被切割, 6-磷酸肌醇也随之离开组装体(图8)。在整个形成病毒的组装过程中, 6-磷酸肌醇扮演了催组剂的角色。

人们在对酶的研究中, 发现了各种别构因子的结合可以对酶的活性进行调控。例如ATP结合到天冬氨酸转氨甲酰酶(ATCase)的调节亚基之后, ATCase就会被激活, 对底物的结合常数大大上升, 酶的活性也得到了显著提升^[114~116]。别构因子使得酶发生变构, 促进了酶与底物的识别和组装。有些别构因子会自发地扩散离去, 或者接受下游产物的反馈离去, 导致酶又转入低活性状态。这些能够自发离去的别构因子实际上起着一定催组剂的作用。

参与细胞信号转导的各种激素分子往往也扮演着催组剂的角色。例如, 当激素结合到膜上的G蛋白偶联受体(GPCR)之后, 可以促进GPCR的变构, 引起G蛋白磷酸化和解组装成为 α 和 $\beta\gamma$ 两个蛋白, 进而引起了下游的信号传导^[117~119]。激素分子在完成使命后, 也会扩散离开GPCR。这些激素分子也在一定程度上起着催组

剂的作用。

鉴于细胞体内存在着数量众多、性质各异的小分子, 随着研究的广泛和深入开展, 基于小分子的催组剂将会得到更多的发掘。

4.4 搬运工分子是否也扮演着催组剂的角色?

组装基元只有到达了特定的空间位置, 组装过程才能顺利开始并进行。众所周知, 细胞内的环境十分复杂, 很多组装基元被限域在特定的细胞器内; 而在胞浆内, 粘稠的溶液犹如密集的人群, 限制了特定基元传质过程的顺利进行。与人工组装中常用的稀溶液中的传质相比, 生命体中的组装基元传质过程必然有其自身的特点, 这是复杂分子组装过程中很重要但尚未受到重视的关键步骤。因此, 细胞内的传质是否高效和成功显得尤为重要, 常常需要各种“搬运工”分子参与其中。

细胞内的运输可分为两大类, 一类是受限空间内的运输, 另一类是跨膜运输。迄今一般认为, 在受限空间内, 物质的运输主要是通过无规则热运动来实现^[120]。通过不停的热运动, 组装基元就可以找到并识别另一个组装基元, 并在组装的基础上进行后续的反应。但是, 无规则热运动在拥挤环境下存在着效率低下的问题, 显然还可能存在着更有效的机制^[120]。例如, 在细胞基质内存在肌动蛋白在微管上的囊泡运输等主动运输方式, 从而实现定向运输^[36,37]。

大多数的局域化都是由磷脂双分子层分隔的, 由于膜系统存在着物理阻隔, 因此需要各种“搬运工”分子协助“货物”实现跨膜运输。跨膜运输包括被动运输和主动运输两种。被动运输是顺着浓度梯度进行的跨膜运输, 有些不需要“搬运工”的参与, 例如 O_2 等; 有些则需要搬运工的参与, 例如 Ca^{2+} 通道蛋白, 这些蛋白一般被称为通道蛋白^[121]。主动运输需要消耗能量, 此能量可来自ATP, 也可利用浓度梯度的偶联, 在少量细菌中还可以通过光提供。这些搬运工一般被称为载体蛋白^[122]。

显而易见, 没有上述“搬运工”分子的存在, 要维持生命体这一复杂的组装体的存在并且行使如此复杂的功能是不可能的。“搬运工”分子将特定的组装基元搬运到特定地点, 促进了组装的进行。因此, 我们认为, 搬运工应归为一类特殊的催组剂。迄今对于细胞内尤其是胞浆中搬运工参与的传质过程的物理图像、搬运

图8 6-磷酸肌醇分子可以促进HIV-1病毒前体的组装, 是一种小分子催组剂^[111] (网络版彩图)

Figure 8 6-IP act as catassembler for the assembly of HIV-1 ^[111] (color online).

工作机理远未清楚, 鉴于它们是生命体组装的重要环节, 有必要重视和加强这方面的系统研究.

5 如何向生命学习, 发展催组装

通过分析生命体系中分子组装特别是催组装体系的特点, 我们更加认识到, 现有的人工分子组装与生命体内的分子组装在组装的各种层面上还相距甚远. 分子组装研究的主要趋势应该是向生命学习, 不止局限于自组装, 而是走向更加可控的复杂组装, 进而缩小两者之间的鸿沟. 这其中, 催组装将很有可能成为一个重要的研究方向. 若此, 我们该如何推进催组装研究呢?

5.1 如何开展复杂体系的分子组装研究

越为复杂和多级次的组装越需要催组装, 催组装研究需要建立在复杂组装的发展之上. 包括催组装在内的复杂分子组装研究有两大策略: 一是直接研究生命体系的相关组装现象和行为, 二是自下而上地人工构筑复杂组装体系, 进而研究其中的催组装行为. 前者是为了全面认知复杂组装体系乃至各级生命过程的机制, 后者是在学习自然的基础上, 模仿并探索超越自然的各种途径以及包括仿生材料与器件在内的广泛应用. 后者可进一步分为以下四种可能途径.

第一, 需要理性设计具有更多作用位点和更具可控变构性的基元. 相较于生命体系, 目前分子组装研究使用的绝大多数组装基元都过于简单, 很难在催组剂、别构因子等的作用下发生变构, 很难实现更复杂的组装. 设计更具可控变构性的组装基元将给合成化学带来巨大的挑战. 可能的途径包括对天然的蛋白质、核酸进行化学修饰和改造, 例如, 将有研究基础的主客体分子等修饰到蛋白质上来构筑新的组装基元^[23,123]; 也可仿生合成类似于生物大分子的组装基元, 以此构筑具更高可控响应性的复杂组装基元.

第二, 需要设计并使用更多种类的组装基元. 它们并不需要都是最终组装体的组成成分, 它们可以参与到组装的不同阶段和不同级次中. 这些组装基元之间的相互关系也可以不再局限于协作关系, 可以是竞争或者拮抗关系, 以便形成与生命体系更相似的复杂组装体系和环境. 也可扩展具有分子识别的组装基元的范畴, 组装基元不仅可以是分子, 也可以是组装体等,

以形成面向生命体的多级次的组装体. 还可构建包含正反馈和负反馈机制的复杂分子组装网络, 由此发现更为有趣或有用的组装体系^[124-126].

第三, 需要在关注组装基元的基础上, 进一步关注组装过程的研究. 目前的分子组装研究, 其组装过程往往是一个黑箱, 很有必要研究这其中的各种中间体, 进而揭示其组装机理. 这离不开各种实时原位表征手段的发展. 对组装过程的详细研究, 才能发现哪些过程是组装过程的关键步骤, 便于设计催组剂或者其他助组剂, 进而研究它们是如何改变组装过程的. 对组装过程的研究是进入复杂组装和多级次组装乃至生命体分子组装的关键.

第四, 需要研究复杂组装的纠错机制. 对于复杂组装, 导致其发生错误组装的原因既可能来自组装基元的缺陷, 也可能由于组装过程复杂而陷入某种中间的错误组装状态^[54]. 因此我们不仅要研究成功的组装过程, 也应重视那些组装不成功的、会出错的组装过程, 搞清楚是组装基元还是组装过程的原因. 进而需要对出错的组装过程, 设计特定的助组剂或者改变组装环境, 发展纠错机制. 重视组装过程动力学, 对于提高组装设计的成功率、发展有关分子机制来纠正错误组装以及治疗相关疾病都具有重要意义.

此外, 还可人为构筑复杂的包含多个子体系的大组装体系, 研究不同子体系之间的相互干扰和协同; 针对局域化的组装过程, 人为构筑一些囊泡和微液滴以包裹分子; 仿生设计搬运工分子, 研究其作用机制^[127,128].

总之, 需要用以上各种方式, 将目前的组装研究从组装基元、组装过程、组装方法等多个角度推向复杂组装. 越复杂的分子组装越需要协助和调控, 只有在复杂组装的研究框架下, 催组装这一高效、高选择性的组装新策略才能充分发挥其重要作用.

5.2 如何设计催组装体系

在复杂组装的研究基础之上, 催组装研究需要从全体系和全过程的大框架下设计和构筑更多新型催组装体系. 对于一个典型的、单一级次的催组装过程可能包括以下四个步骤: 催组剂识别组装基元、催组剂活化组装基元、活化后的组装基元之间发生组装、催组剂的离去.

催组装所针对的组装基元往往结构复杂, 组装基

元上存在着一定的空间位阻, 或者存在局域电荷的排斥, 从而导致组装基元之间难以相互识别组装. 如图9所示, 组装基元A与B之间的组装相互作用位点并没有暴露出来而导致无法自发组装. 其他可能导致组装无法自发完成的原因还包括组装的路径复杂而停滞在一定的中间态、组装基元无法通过简单传质到达特定组装位点等. 对于这些复杂组装过程, 需要催组剂的帮助才可以顺利进行. 下面我们将对这四个步骤的设计加以具体分析.

如图9所示, 首先催组剂要识别组装基元, 这要求两者之间存在一定的特异相互作用. 由于催组剂最终要离开组装体, 因此催组剂和组装基元之间的作用力最好应是非共价相互作用, 这些多位点的弱相互作用可以通过减少作用位点或降低作用强度以方便催组剂的离去.

其二, 催组剂识别组装基元之后, 催组剂应该能够活化组装基元使其发生变构, 暴露出相应组装作用位点. 这是生命体的催组装给予我们的启示之一, 与目前的自组装研究体系很不相同. 组装基元和催组剂需要较强的可控变构性, 对催组装的各个节点和步骤具有响应性. 这些组装基元应具有多个作用位点, 既有和催组剂作用的位点, 又有和其他组装基元作用的位点. 值得注意的是, 这种活化是通过多位点的非共价作用方式进行的, 故要求很高的系统设计能力.

其三, 被催组剂活化后的组装基元将和其他组装基元相互作用, 形成组装体. 组装基元之间总的相互作用力应该强于组装基元与催组剂的作用力, 这样才能保证催组剂的离去和组装体的结构稳定.

最后, 催组剂需要离开组装体, 这是催组装区别于助组装、共组装的最重要特征. 在没有外部能量输入的情况下, 催组剂离开的方式可能有以下三种: (1) 组装体形成后, 与催组剂作用的组装基元再次发生变构, 组装基元和催组剂之间的相互作用消失或者变弱而离开; (2) 组装体与催组剂发生分相, 组装体进入新相, 催组剂则留在未组装的基元所在的相中; (3) 催组剂与其他组装基元共有相同的作用位点, 被其竞争下去. 例如, 在DNA催组装体系中, 组装体完成后催组剂链会被其他链替换进而脱落^[129,130]. 而对于有能量输入的催组装过程, 催组剂的离去方式则不局限于以上三种, 它可以消耗能量主动离开组装体.

从能量角度来分析催组装的驱动力, 可以大致分

图9 一个典型的催组装过程包括识别、活化、组装和离去四个步骤(网络版彩图)

Figure 9 A typical catalytic assembly process has four steps: recognition, activation, assembly and leaving (color online).

为两大类情况. 对于无外部能量输入的催组装过程, 其驱动力主要来自组装过程的自由能变, 催组剂的作用是帮助组装基元降低组装能垒和改变组装路径. 而对于有能量输入的非平衡态和具有耗散性质的催组装过程, 催组剂需要与来自外界的化学能、光能、电能作用, 驱动催组装过程. 因此, 可以利用外场、外力等各种方法发展电催组装、光催组装、光电催组装. 这不仅是催组装研究的几个重要的新分支, 也是模拟对外部环境做出响应的生命现象和过程的一个重要途径.

总之, 催组装体系的设计跟现有的自组装体系的设计既有关联更有很大的差异. 催组装体系的成功与否有两个难点: 一是在组装初始阶段, 催组剂通过非共价方式活化组装基元; 二是在组装完成时, 组装基元的顺利离去. 这是催组装区别于其他助组装的显著特征.

值得注意的是, 上述讨论针对的是一个简单的单级次催组过程. 对于更为复杂的多级次分子组装, 下一级组装的启动与前一级的催组剂离去很可能有密切协同关系, 上述活化和离去可以与相邻级次组装过程的偶联和信息反馈相关联. 这些特征广泛存在于生命体中, 可能是自然进化的生命体具有复杂精巧且功能强大的奥秘之一. 因此, 我们向生命和自然学习, 便是要探索和建立通盘考虑多层级协同递进的全系统组装的思维框架和研究范式. 在组装基元的设计、催组剂的设计、组装过程动力学的研究上都需要提出新思路并发展组装、表征和理论的新方法.

5.3 如何发展适用于催组装的表征方法

分子组装是生命科学和软物质科学的重要基础科学问题, 发展针对复杂组装动力学和催组装动力学的

表征方法将极大地推进这两大领域的进展。但遗憾的是,目前人工分子组装研究采用的大多数表征技术尚不能完全满足分子组装对表征技术的需求。分子组装主要涉及多种弱键、多位点的协同作用,具有多尺度、高度动态的特点。催组装又进一步涉及单个或多个催组剂与不同组装基元的更复杂作用(包括对组装路径和动力学的主动调控)。如何在分子水平上揭示这一复杂、动态过程的机制是发展催组装面临的重大挑战之一。

催组装过程中,组装基元和催组剂的结构相对复杂,基元间通过多种弱键进行多位点的协同作用和连接,且复杂组装过程存在多个路径和可能的多种中间体,不同路径和中间体间还存在动态的协同与反馈,有必要分别从高的能量、时间和空间分辨率和高灵敏度给予全面分析,发展原位(*in-situ*)或在体(*in-vivo*)表征技术显得尤为重要^[131,132]。

首先,研究弱键相互作用的表征方法需要有高的能量分辨(如谱学分辨率),可解析不同弱键相互作用的连接和解离(其键能一般为10–50 kJ/mol,对应于远红外和微波区间^[133]),以提供高分辨的结构信息和重要的相互作用细节。表征技术需要高能量分辨,才能给出催组装过程的不同组装路径、催组剂作用机理、识别反馈调节机制等关键组装信息。其次,虽然组装过程的时间尺度相比化学反应更长(一般为微秒至秒区间,个别组装过程可达小时级或者更久^[134]),相应的组装速率也比化学反应更慢,但研究其过程仍需较高的时间分辨率,以分辨从组装基元到组装产物过程中不同组装中间体的结构、变化规律及其动力学信息,这对诠释组装机理至关重要。再次,催组装涉及大分子乃至多级次组装过程,各基元或分子基团间的空间排列对组装体结构和性质影响很大,若表征方法有高的空间分辨率,便可提供组装基元间的空间信息。此外,以上分辨率都需要足够高的检测灵敏度才能实现,确保表征能够在合理的时间尺度和浓度范围内完成。

目前常规表征方法大部分通过单一方式,如电磁波、微观粒子或探针等与被测物进行作用并采集相应的信息,因此使用单一表征方法很难兼顾两种以上的分辨率需求,很难全面表征分子组装过程。我们认为非常有必要针对分子组装的特点,结合表征方法原理,对有关仪器进行改进或与其他仪器联用,使表征技术同时满足两种或以上高分辨率。需要对常规表征仪器

进行改进,甚至从原理创新来实现同时具备两种以上分辨率的表征方法。

显微成像技术是最常见且直观的表征技术,能够直接给出组装体的形貌。根据光源的不同可大致分为光学显微镜、电子显微镜等技术,此外还有根据力的大小成像的原子力显微镜等。显微镜领域的研究人员已不再满足于空间分辨率的提高,开始向其他分辨率的提高发起挑战。例如李栋等^[135]发展了具有纳米级空间分辨和毫秒级时间分辨的超分辨荧光显微镜,并成功利用此技术研究了内质网与微管间的相互作用,发现了新的内质网重塑机制。

对于相对稳定的组装体或中间体,表征方法可与分离方法连用,分别研究分离后的组分从而给出更多组装信息。目前随着各种高通量、高精度分析测试手段的发展,研究复杂的混合物已经变得可能。高通量色谱、质谱技术等代谢组学中广泛使用的技术也可以在催组装中得到应用。利用这些技术分析溶液组成,进行时间序列的分析可得到催组装过程的动力学图景。

谱学表征技术是最常用的表征方法,根据使用的电磁波波长的不同,可分为微波/太赫兹频段、红外/可见光/紫外频段、X射线/ γ 射线频段等。对于研究弱键相互作用最有优势的技术当属微波段的核磁共振表征技术,它具有非常高的能量分辨,能够给出原子级结构信息和体系动态信息。但是,没有任何技术是完美的,NMR的低时间分辨和低灵敏度的劣势阻碍了其在分子组装过程研究中的应用。我们于近期发展了微流控-核磁联用技术,首次将其应用到组装过程研究。通过在核磁探头中引入微流控芯片,组装基元在芯片管路中原位组装,控制流速调控采样区的组装进度,已获得组装开始1 s时的核磁谱,且一维谱和二维谱的时间分辨率均被提升到1 s,在一种表征技术中同时实现了高能量分辨和较高的时间分辨^[136]。

此外,单分子力谱表征方法对研究纳米尺度的催组装具有优势,例如利用单分子光镊、磁镊等单分子力谱表征生物大分子的催组装过程^[101,104]。这些手段具有很高的空间分辨率,可以实时原位观察和研究组装体局部的组装和解组装过程,而且可以通过控制外力的大小和作用点,将组装体调制到不同的阶段上。由于组装过程具有一定的随机性,使用高通量的单分子技术,结合统计学等数据处理方法,可给出组装过

程的统计学图像。

针对催组装发展相应的表征技术既重要又存在较大的挑战性, 我们认为突破现有表征技术的瓶颈和研发范式, 需要多学科交叉、协同研究, 引进仪器研发和信号处理领域的前沿研究成果, 对涉及大量数据的表征方法引入人工智能方法等, 都将推动表征技术向前发展。上述各种表征技术获得的信息可进一步为理论计算提供关键的实验证据。

5.4 如何发展适用于催组装的理论方法

催组装的理论基础是理解各类人工或者生命体催组装过程、精准设计催组装体系及其催组剂的关键。催组装过程具有多物种、多尺度、多路径、短程与长程作用力耦合等特点, 尤其是生命体中的催组装体系更为庞大、复杂且时间跨度长, 这对催组装乃至整个分子组装理论的发展提出了前所未有的挑战^[137]。基于以上特点, 催组装理论方法需从物理化学基础和计算方法两方面入手, 注重催组装过程的活性中间体、催组剂作用的动态过程及不同种类催组剂的作用机制等环节, 获得催组装过程的物理化学图像以及揭示催组机理。

从物理化学角度, 首先针对催组装过程的多位点弱键的协同反馈机制, 需关注催组装过程中的关键物理化学参数的非线性变化, 特别要考虑熵驱动和熵焓互补的物理化学特征^[138], 全面系统地从热力学和动力学上建立催组装的物理化学基础。其次, 传质传能过程将直接影响催组装体系尤其是生命体的催组装的组装路径, 其组装过程可能与传质过程处于同一时间和空间尺度, 因此需升级现有的反应-扩散方程, 并关注各种传统与新型传质方式^[139](如主动运输、湍流等)与催组装过程的耦合机制。特别应当指出的是, 生命体系中的催组装体系往往处于非平衡状态和复杂拥挤涨落的环境中, 亟待发展有效的非平衡统计理论^[140,141]。

从计算方法角度, 针对催组装过程的多种弱键相互作用, 需升级现有描述弱相互作用的泛函^[142], 校对生物大分子力场(如AMBER^[143,144]、CHARMM^[145,146]等力场)的参数, 配合混合精度的组合方法详细分析弱键的协同方式^[147], 并结合分子动力学模拟和增强采样等技术来研究催组剂的动态作用过程。针对催组装体系的多级次特点, 将量子力学方法、全原子模型分子力场和粗粒化模型分子力场整合到统一框架下, 发挥

各种方法和模型的优势, 构建兼顾精度和效率的多尺度模拟方法, 并将该方法与动力学蒙特卡洛模拟(KMC)^[148]或巨正则蒙特卡洛模拟(GCMC)^[149]相衔接, 发展基于机器学习的高效可靠的图像识别和构造方法, 搜寻和实现在不同的精度和空间分辨率下描述催组装活性中间体的结构, 以期建立多层级、多尺度的理论模型。

应当指出, 目前催组装的理论方法尚处于初期探索阶段, 可用的物理化学模型及模拟方法仍十分匮乏, 需吸收最新的理论模拟方法(如机器学习辅助统计抽样等), 以期建立普适的理论模型, 获得催组装过程的清晰物理化学图像, 建立从微观到介观乃至宏观层次的理论与实验观测之间的桥梁。探索催组装的理论基础也需借鉴生命、物理和数学研究中的理论方法, 融合系统论、控制论及信息论中的概念^[150-152], 揭示各类催组剂的作用方式及其协同反馈机制, 不断革新现有组装及催组装理论。我们将在该系列文章的后续中试图全面阐述催组机理, 并讨论详细的理论基础。

6 总结与展望

我们在撰写本文的过程中, 愈发感到对生命体组装现象和过程理解的浅薄, 此文仅仅是抛砖引玉。迄今尚未建立(可能更准确地说是尚未初步建立)可全面解释生命体组装过程的科学图像, 有关理论框架和研究方法更是初步和碎片化。该科学问题若要有大的突破可能需要科学认识论的提升。在此方面, 2020年3月逝世的著名物理学家P. W. Anderson^[153]给我们指明了方向。他于1972年在*Science*上发表了里程碑式的文章, 题为“More is different: Broken symmetry and the nature of the hierarchical structure of science”。人们将“more is different”翻译为多种中文版本, 例如“多即异”或“多者异也”, 其内涵与人们较熟悉的马克思的名言“量变引起质变”类似, 不仅是影响科学的哲学基础, 更是当今面对日益复杂体系的科学认识论的基础。

Anderson^[153]在文中深刻地指出: “大型和复杂的基本粒子集合体的行为, 并不能按照少数基本粒子性质的简单外推来理解。事实上, 在复杂性的每一个层级, 都会有崭新的性质出现; 在每一个层级上, 新的定律、概念和原理都是必不可少的, 其所需要的想象力与创造力丝毫不亚于前一个层级。心理学不是应用生

物, 生物学也不是应用化学.....我相信, 在每一个层级上, 我们都会遇到迷人的、非常基本的问题, 即: 将不那么复杂的部分组合为一个更为复杂的系统, 并理解由此而来的本质上新型的行为.....我们必须摆脱一些分子生物学家的傲慢, 那些分子生物学家力图将人体组织或机能完全还原为化学, 从普通的感冒和各种精神疾病一直到宗教本能.....”。

众所周知, 大自然是以多个层级的方式组装和组织起来的, 关于生命整体不是所有部分的总和的观点并不是Anderson的独创。但难得的是, 约半世纪前他采用令人信服的科学术语来表述, 进而导致了“演生”(emergence)的概念和方法论的提出^[154]。在客观世界的形成和发展中, 任何事物和知识都是从更原始的层级上逐级演生而来的(如从物理、化学到细胞生物学、脑科学、人体生理学)。虽然更高层级遵循更低层级的定律, 但其可能表现出的行为仅凭其低层级的定律无法理解, 其主要原理和核心定律也不是来自低层级的结果。例如, 生命科学不违反化学科学的基本定律, 但具有不是从化学定律直接推导的思想和原理。

Anderson^[153]还指出, “当我说尺度变化引起根本性的变化时, 即新尺度上的现象可能服从根本不同的基本定律, 比如, 宇宙学尺度上需要用广义相对论, 原子尺度上则要用量子力学。我想还应该承认, 所有普通物质都服从电动力学和量子力学”。这表明, 各层级之间并不是隔离断开而自行其是, 特别是低层级对高层级(如化学对分子生物学和细胞生物学)的支撑是必不可少的。若抛开各类弱相互作用在生命体组装中的协同(如氢键、 $\pi\pi$ 堆积、亲疏水作用等), 要深入系统认知生命体内的组装机制则是难以想象。故化学工作者与生命工作者的协同合作似乎是十分必要的。

总之, Anderson所阐明的科学论点对于本文所探讨的生命体系组装和催组装的定位和方法论极为重要, 应当成为该方向的指南。我们也希望本文可以成为追思和敬献给这位科学大师的一朵小花。

本文首先从组装基元、组装过程、组装环境和组装体功能四个角度对比, 辨析生命体分子组装与人工分子组装的主要异同, 指出前者是原子和分子数目的增多到一定层次而导致的复杂体系。生命体是基于多组分、多路径、多级次的可控复杂组装, 形成了从分子到细胞器、细胞到器官、组织, 乃至个体等多个层级的组装体。在较高层级的组装中, 同时伴随着远离

平衡态的多方式(物质、能量与信息)的多种交流。由于这些“多”造成了显著的“异”, 即人工分子组装体系迄今可望而不可及的复杂且多样的高级功能, 造就了丰富多彩的生命现象。

人工分子组装的研究或许长期以来受到自组装概念中“自”的束缚, 更多地依赖和关注依靠自发过程构筑组装体, 由此所能获得的组装体系和功能都较为简单。与之截然不同, 生命体内众多的精巧复杂的分子组装之所以能够实现, 绝不是只单纯依靠组装基元之间的热力学自发过程, 而是体系内很多其他分子的参与、协助与调控。这种“多即异”的具体体现之一便是催组装, 它作为一种高效、高选择性的调控手段, 广泛存在于生命组装体系中。例如, 协助蛋白质折叠的分子伴侣, 协助染色质、RuBisCO等组装的催组剂, 参与病毒组装、细胞信号传导等过程的小分子催组剂, 以及通过主动运输实现高效传质的“搬运工”分子等。当然, 催组装并不是仅仅是生命体组装所独有的, 可以预见, 随着人工分子组装体系向生命学习的过程中逐渐趋于复杂和性能强大, 将愈发需要催组装这样高效的组装调控手段。

在Anderson的“多即异”思想的指导下, 催组装这一组装调控新策略可能会更好地发挥作用。但是应当注意到, 生命组装体包含多个层级, 同一层级的组装体还有不同类别。故首先需要分层分类的清晰思路。本文主要探讨的组装体系是生命体的最基本层级(DNA、蛋白质等)和次低层级(染色质、病毒等), 需要从简到繁, 首先理清以这两个层级为基础的“多即异”的具体内涵, 方可进一步提出乃至建立相应层级的新的原理和定律。

我们认为, 研究催组装可以从两个方向入手。一是直接研究生命体系的催组装现象, 二是自下而上地人工构筑催组装复杂体系。构筑并研究催组装体系, 需要从复杂组装体系着手, 需要提升组装基元结构、种类及相互作用的复杂性, 从组装基元、组装过程、组装体、催组剂等多角度, 选择具有特殊功能的体系, 全方位地推动复杂组装体系的研究。

我们提出了单级催组装体系的四个可能步骤, 包括催组剂识别组装基元、活化组装基元、活化后的组装基元之间发生组装、以及催组剂的离去, 指出了其中的关键点是活化和离去。而对于更为复杂的多组分多级次催组装, 其组分之间的相互作用与协同将变得

更为复杂, 其催组装过程甚至可以通过相邻级次的偶联和信息反馈等来实现. 通过向自然和生命学习, 我们感到, 催组装的思路很可能覆盖于生命体系的多个组装层级, 也有利于探索和建立多层次协同递进的全系统组装的思维框架和研究范式.

催组装理论方法尚处于非常初期的探索阶段, 需借鉴数学、物理、生命科学、信息科学的理论方法, 结合催组装体系的实验结果, 在“多即异”的思想指导下, 大胆突破现有的组装理论框架, 积极引进可研究复杂体系的新方法, 例如机器学习和人工智能, 逐步建立相应的理论模型, 进而为催组装体系的设计及功

能应用的开发提供理论指导. 有关催组装理论和实验研究皆离不开表征方法的发展, 故亟需有针对性地发展高能量、时间和空间分辨以及高灵敏度的、原位或在体的(联用)表征手段.

总之, 催组装是广泛存在于生命体分子组装体系中的高效、高选择性组装调控策略, 针对催组装开展深入、全面的研究, 可能成为人工分子组装体系从简单迈向复杂, 从难以控制到实现高度可控的有效途径, 也可能为深入揭示生命体系的机制和奥秘做出贡献, 并有望促进软物质科学研究从基础理论到功能应用的进一步提升与拓展.

致谢 衷心感谢自然科学基金委的“催组装研究方法 with 理论基础”重大项目(21991130)、“可控分子自组装体系及其功能化”重大研究计划(91427304)的支持. 衷心感谢吴家睿、张希、刘鸣华、樊春海、李国红、梁好均、徐昕、高毅勤、杨海波、张文彬、昌增益、王梅祥、陈玲玲、李伟、陈萍、程俊等的有益讨论和建议.

参考文献

- 1 Lehn JM. *Science*, 2002, 295: 2400–2403
- 2 Lehn JM. *Angew Chem Int Ed*, 2013, 52: 2836–2850
- 3 Shen JC, Zhang X. *Sci & Tech Rev*, 1994, 12: 6–8 (in Chinese) [沈家骢, 张希. 科技导报, 1994, 12: 6–8]
- 4 Shen JC, Zhang Xi. *Chin Sci Bull*, 2003, 48: 1477–1478 (in Chinese) [沈家骢, 张希. 科学通报, 2003, 48: 1477–1478]
- 5 Lehn JM. *Pure Appl Chem*, 1978, 50: 871–892
- 6 Lehn JM. *Supramolecular Chemistry: Concepts and Perspectives*. New York: Wiley, 1996
- 7 Zhou Y, Yan D. *Chem Commun*, 2009, 1172–1188
- 8 Kim HC, Park SM, Hinsberg WD. *Chem Rev*, 2010, 110: 146–177
- 9 Mai Y, Eisenberg A. *Chem Soc Rev*, 2012, 41: 5969–5985
- 10 Ulman A. *Chem Rev*, 1996, 96: 1533–1554
- 11 Love JC, Estroff LA, Kriebel JK, Nuzzo RG, Whitesides GM. *Chem Rev*, 2005, 105: 1103–1170
- 12 Barth JV. *Annu Rev Phys Chem*, 2007, 58: 375–407
- 13 Whitesides GM, Grzybowski B. *Science*, 2002, 295: 2418–2421
- 14 Halley JD, Winkler DA. *Complexity*, 2008, 14: 10–17
- 15 Del Valle EMM. *Process Biochem*, 2004, 39: 1033–1046
- 16 Lagona J, Mukhopadhyay P, Chakrabarti S, Isaacs L. *Angew Chem Int Ed*, 2005, 44: 4844–4870
- 17 Caltagirone C, Gale PA. *Chem Soc Rev*, 2009, 38: 520–563
- 18 Hasenknopf B, Lehn JM, Kneisel BO, Baum G, Fenske D. *Angew Chem Int Ed Engl*, 1996, 35: 1838–1840
- 19 Jones MR, Seeman NC, Mirkin CA. *Science*, 2015, 347: 1260901
- 20 Aida T, Meijer EW, Stupp SI. *Science*, 2012, 335: 813–817
- 21 Brunsveld L, Folmer BJB, Meijer EW, Sijbesma RP. *Chem Rev*, 2001, 101: 4071–4098
- 22 Hartgerink JD, Beniash E, Stupp SI. *Science*, 2001, 294: 1684–1688
- 23 De Greef TFA, Smulders MMJ, Wolffs M, Schenning APHJ, Sijbesma RP, Meijer EW. *Chem Rev*, 2009, 5687–5754
- 24 Tian J, Zhou TY, Zhang SC, Aloni S, Altoe MV, Xie SH, Wang H, Zhang DW, Zhao X, Liu Y, Li ZT. *Nat Commun*, 2014, 5: 5574
- 25 Tian J, Wang H, Zhang DW, Liu Y, Li ZT. *Natl Sci Rev*, 2017, 4: 426–436
- 26 Yao W, Ma C, Ma H, Fu L, Lu S, Lv A, Cai S, Hang X, Singh M, Shi H, An Z, Huang W. *Giant*, 2020, 1: 100007

- 27 Zheng J, Birktoft JJ, Chen Y, Wang T, Sha R, Constantinou PE, Ginell SL, Mao C, Seeman NC. *Nature*, 2009, 461: 74–77
- 28 Harata K. *BCSJ*, 1984, 57: 2596–2599
- 29 Wang Y, Lin HX, Ding SY, Liu DY, Chen L, Lei ZC, Fan FR, Tian ZQ. *Sci Sin-Chim*, 2012, 42: 525–547 (in Chinese) [王宇, 林海昕, 丁松园, 刘德宇, 陈亮, 雷志超, 范凤茹, 田中群. 中国科学·化学, 2012, 42: 525–547]
- 30 Wang Y, Lin HX, Chen L, Ding SY, Lei ZC, Liu DY, Cao XY, Liang HJ, Jiang YB, Tian ZQ. *Chem Soc Rev*, 2014, 43: 399–411
- 31 Vantomme G, Meijer EW. *Science*, 2019, 363: 1396–1397
- 32 Tame JRH, Vallone B. *Acta Crystlogr D Biol Crystlogr*, 2000, 56: 805–811
- 33 Schmidli C, Albiez S, Rima L, Righetto R, Mohammed I, Oliva P, Kovacic L, Stahlberg H, Braun T. *Proc Natl Acad Sci USA*, 2019, 116: 15007–15012
- 34 Song F, Chen P, Sun D, Wang M, Dong L, Liang D, Xu RM, Zhu P, Li G. *Science*, 2014, 344: 376–380
- 35 Perlmutter JD, Hagan MF. *Annu Rev Phys Chem*, 2015, 66: 217–239
- 36 Hirokawa N, Noda Y, Tanaka Y, Niwa S. *Nat Rev Mol Cell Biol*, 2009, 10: 682–696
- 37 Verhey KJ, Kaul N, Soppina V. *Annu Rev Biophys*, 2011, 40: 267–288
- 38 Winkler WC, Breaker RR. *Annu Rev Microbiol*, 2005, 59: 487–517
- 39 Serganov A, Nudler E. *Cell*, 2013, 152: 17–24
- 40 Rothmund PWK. *Nature*, 2006, 440: 297–302
- 41 Pinheiro AV, Han D, Shih WM, Yan H. *Nat Nanotech*, 2011, 6: 763–772
- 42 King NP, Sheffler W, Sawaya MR, Vollmar BS, Sumida JP, André I, Gonen T, Yeates TO, Baker D. *Science*, 2012, 336: 1171–1174
- 43 Butterfield GL, Lajoie MJ, Gustafson HH, Sellers DL, Nattermann U, Ellis D, Bale JB, Ke S, Lenz GH, Yehdego A, Ravichandran R, Pun SH, King NP, Baker D. *Nature*, 2017, 552: 415–420
- 44 Yeates TO. *Annu Rev Biophys*, 2017, 46: 23–42
- 45 Anfinsen CB. *Biochem J*, 1972, 128: 737–749
- 46 Dobson CM. *Nature*, 2003, 426: 884–890
- 47 Chiti F, Dobson CM. *Annu Rev Biochem*, 2006, 75: 333–366
- 48 Gregersen N, Bross P, Vang S, Christensen JH. *Annu Rev Genom Hum Genet*, 2006, 7: 103–124
- 49 Hartl FU, Bracher A, Hayer-Hartl M. *Nature*, 2011, 475: 324–332
- 50 Dill KA, MacCallum JL. *Science*, 2012, 338: 1042–1046
- 51 Sweeney P, Park H, Baumann M, Dunlop J, Frydman J, Kopito R, McCampbell A, Leblanc G, Venkateswaran A, Nurmi A, Hodgson R. *Transl Neurodegener*, 2017, 6: 6
- 52 Hamley IW. *Chem Rev*, 2012, 112: 5147–5192
- 53 Herschlag D. *J Biol Chem*, 1995, 270: 20871–20874
- 54 Lei ZC, Yin X, Wang X, Ke G, Cao X, Fan C, Yang CJ, Liang H, Tian ZQ. *Giant*, 2020, 1: 100008
- 55 Bowie JU. *Nature*, 2005, 438: 581–589
- 56 Skach WR. *Nat Struct Mol Biol*, 2009, 16: 606–612
- 57 Gao YQ, Yang W, Karplus M. *Cell*, 2005, 123: 195–205
- 58 Conde C, Cáceres A. *Nat Rev Neurosci*, 2009, 10: 319–332
- 59 Harano Y, Kinoshita M. *BioPhys J*, 2005, 89: 2701–2710
- 60 Imai T, Harano Y, Kinoshita M, Kovalenko A, Hirata F. *J Chem Phys*, 2006, 125: 024911
- 61 Ball P. *Chem Rev*, 2008, 108: 74–108
- 62 Cui D, Ou S, Patel S. *Proteins*, 2014, 82: 3312–3326
- 63 Milo R. *BioEssays*, 2013, 35: 1050–1055
- 64 Zimmerman SB, Minton AP. *Annu Rev Biophys Biomol Struct*, 1993, 22: 27–65
- 65 Ellis RJ. *Trends Biochem Sci*, 2001, 26: 597–604
- 66 Zhou HX, Rivas G, Minton AP. *Annu Rev Biophys*, 2008, 37: 375–397
- 67 Ando T, Skolnick J. *Proc Natl Acad Sci USA*, 2010, 107: 18457–18462
- 68 Dhar A, Samiotakis A, Ebbinghaus S, Nienhaus L, Homouz D, Gruebele M, Cheung MS. *Proc Natl Acad Sci USA*, 2010, 107: 17586–17591

- 69 Hyman AA, Weber CA, Jülicher F. *Annu Rev Cell Dev Biol*, 2014, 30: 39–58
- 70 Bergeron-Sandoval LP, Safaee N, Michnick SW. *Cell*, 2016, 165: 1067–1079
- 71 Alberti S. *Curr Biol*, 2017, 27: R1097–R1102
- 72 Banani SF, Lee HO, Hyman AA, Rosen MK. *Nat Rev Mol Cell Biol*, 2017, 18: 285–298
- 73 Boeynaems S, Alberti S, Fawzi NL, Mittag T, Polymenidou M, Rousseau F, Schymkowitz J, Shorter J, Wolozin B, Van Den Bosch L, Tompa P, Fuxreiter M. *Trends Cell Biol*, 2018, 28: 420–435
- 74 Williams RJ, Smith AM, Collins R, Hodson N, Das AK, Ulijn RV. *Nat Nanotech*, 2009, 4: 19–24
- 75 Soetan N, Erwin WR, Tonigan AM, Walker DG, Bardhan R. *J Phys Chem C*, 2017, 121: 18186–18194
- 76 Baccile N, Selmane M, Le Griel P, Prévost S, Perez J, Stevens CV, Delbeke E, Zibek S, Guenther M, Soetaert W, Van Bogaert INA, Roelants S. *Langmuir*, 2016, 32: 6343–6359
- 77 Prins LJ, Timmerman P, Reinhoudt DN. *J Am Chem Soc*, 2001, 123: 10153–10163
- 78 Leandri V, Yang W, Gardner JM, Boschloo G, Ott S. *ACS Appl Energy Mater*, 2018, 1: 202–210
- 79 Zhang KQ, Liu XY. *Nature*, 2004, 429: 739–743
- 80 Liu H, Xu Y, Li F, Yang Y, Wang W, Song Y, Liu D. *Angew Chem Int Ed*, 2007, 46: 2515–2517
- 81 Kim BG, Kim MS, Kim J. *ACS Nano*, 2010, 4: 2160–2166
- 82 Ellis J. *Nature*, 1987, 328: 378–379
- 83 Cheng MY, Hartl FU, Martin J, Pollock RA, Kalousek F, Neupert W, Hallberg EM, Hallberg RL, Horwich AL. *Nature*, 1989, 337: 620–625
- 84 Horwich AL, Fenton WA. *Quart Rev Biophys*, 2020, 53: e4
- 85 Ritossa F. *Experientia*, 1962, 18: 571–573
- 86 Tissières A, Mitchell HK, Tracy UM. *J Mol Biol*, 1974, 84: 389–398
- 87 Bardwell JCA, Craig EA. *Proc Natl Acad Sci USA*, 1984, 81: 848–852
- 88 Schuh S, Yonemoto W, Brugge J, Bauer VJ, Riehl RM, Sullivan WP, Toft DO. *J Biol Chem*, 1985, 260: 14292–14296
- 89 Langer T, Pfeifer G, Martin J, Baumeister W, Hartl FU. *EMBO J*, 1992, 11: 4757–4765
- 90 Braig K, Otwinowski Z, Hegde R, Boisvert DC, Joachimiak A, Horwich AL, Sigler PB. *Nature*, 1994, 371: 578–586
- 91 Bukau B, Horwich AL. *Cell*, 1998, 92: 351–366
- 92 Bukau B, Weissman J, Horwich A. *Cell*, 2006, 125: 443–451
- 93 Laskey RA, Honda BM, Mills AD, Finch JT. *Nature*, 1978, 275: 416–420
- 94 Ellis RJ. *Trends Biochem Sci*, 2006, 31: 395–401
- 95 Ellis RJ. *Phil Trans R Soc B*, 2013, 368: 20110398
- 96 Liu C, Young AL, Starling-Windhof A, Bracher A, Saschenbrecker S, Rao BV, Rao KV, Berninghausen O, Mielke T, Hartl FU, Beckmann R, Hayer-Hartl M. *Nature*, 2010, 463: 197–202
- 97 Hauser T, Bhat JY, Miličić G, Wendler P, Hartl FU, Bracher A, Hayer-Hartl M. *Nat Struct Mol Biol*, 2015, 22: 720–728
- 98 Wilson RH, Hayer-Hartl M. *Biochemistry*, 2018, 57: 3210–3216
- 99 Luger K, Mäder AW, Richmond RK, Sargent DF, Richmond TJ. *Nature*, 1997, 389: 251–260
- 100 Hammond CM, Strømme CB, Huang H, Patel DJ, Groth A. *Nat Rev Mol Cell Biol*, 2017, 18: 141–158
- 101 Chen P, Dong L, Hu M, Wang YZ, Xiao X, Zhao Z, Yan J, Wang PY, Reinberg D, Li M, Li W, Li G. *Mol Cell*, 2018, 71: 284–293.e4
- 102 Wang J, Gao S, Peng X, Wu K, Yang S. *IJMS*, 2019, 20: 4591
- 103 Orphanides G, LeRoy G, Chang CH, Luse DS, Reinberg D. *Cell*, 1998, 92: 105–116
- 104 Li W, Chen P, Yu J, Dong L, Liang D, Feng J, Yan J, Wang PY, Li Q, Zhang Z, Li M, Li G. *Mol Cell*, 2016, 64: 120–133
- 105 Jeronimo C, Poitras C, Robert F. *Cell Rep*, 2019, 28: 1206–1218.e8
- 106 Liu Y, Zhou K, Zhang N, Wei H, Tan YZ, Zhang Z, Carragher B, Potter CS, D'Arcy S, Luger K. *Nature*, 2020, 577: 426–431
- 107 Talbert PB, Henikoff S. *Nat Rev Mol Cell Biol*, 2016, 18: 115–126
- 108 Buschbeck M, Hake SB. *Nat Rev Mol Cell Biol*, 2017, 18: 299–314
- 109 Mizuguchi G, Shen X, Landry J, Wu WH, Sen S, Wu C. *Science*, 2004, 303: 343–348
- 110 Eustermann S, Schall K, Kostrewa D, Lakomek K, Strauss M, Moldt M, Hopfner KP. *Nature*, 2018, 556: 386–390
- 111 Dick RA, Zadrozny KK, Xu C, Schur FKM, Lyddon TD, Ricana CL, Wagner JM, Perilla JR, Ganser-Pornillos BK, Johnson MC, Pornillos O,

- Vogt VM. *Nature*, 2018, 560: 509–512
- 112 Dick RA, Mallery DL, Vogt VM, James LC. *Viruses*, 2018, 10: 640–651
- 113 Mallery DL, Faysal KMR, Kleinpeter A, Wilson MSC, Vaysburd M, Fletcher AJ, Novikova M, Böcking T, Freed EO, Saiardi A, James LC. *Cell Rep*, 2019, 29: 3983–3996.e4
- 114 Honzatko RB, Crawford JL, Monaco HL, Ladner JE, Edwards BFP, Evans DR, Warren SG, Wiley DC, Ladner RC, Lipscomb WN. *J Mol Biol*, 1982, 160: 219–263
- 115 Kantrowitz ER. *Archives Biochem Biophys*, 2012, 519: 81–90
- 116 Lipscomb WN, Kantrowitz ER. *Acc Chem Res*, 2012, 45: 444–453
- 117 Dorsam RT, Gutkind JS. *Nat Rev Cancer*, 2007, 7: 79–94
- 118 Rosenbaum DM, Rasmussen SGF, Kobilka BK. *Nature*, 2009, 459: 356–363
- 119 Wootten D, Christopoulos A, Marti-Solano M, Babu MM, Sexton PM. *Nat Rev Mol Cell Biol*, 2018, 19: 638–653
- 120 Bressloff PC, Newby JM. *Rev Mod Phys*, 2013, 85: 135–196
- 121 Catterall WA. *Cold Spring Harbor Perspectives Biol*, 2011, 3: a003947
- 122 Alberts B. *Molecular Biology of the Cell*. New York: W. W. Norton & Company, 2014
- 123 Sakai F, Yang G, Weiss MS, Liu Y, Chen G, Jiang M. *Nat Commun*, 2014, 5: 4634
- 124 Ashkenasy G, Hermans TM, Otto S, Taylor AF. *Chem Soc Rev*, 2017, 46: 2543–2554
- 125 Ludlow RF, Otto S. *Chem Soc Rev*, 2008, 37: 101–108
- 126 Mattia E, Otto S. *Nat Nanotech*, 2015, 10: 111–119
- 127 Gale PA, Davis JT, Quesada R. *Chem Soc Rev*, 2017, 46: 2497–2519
- 128 Wu X, Howe ENW, Gale PA. *Acc Chem Res*, 2018, 51: 1870–1879
- 129 Turberfield AJ, Mitchell JC, Yurke B, Mills AP, Blakey MI, Simmel FC. *Phys Rev Lett*, 2003, 90: 118102
- 130 Yin P, Choi HMT, Calvert CR, Pierce NA. *Nature*, 2008, 451: 318–322
- 131 Peterlinz KA, Georgiadis R. *Langmuir*, 1996, 12: 4731–4740
- 132 Bai W, Yager KG, Ross CA. *Macromolecules*, 2015, 48: 8574–8584
- 133 Müller-Dethlefs K, Hobza P. *Chem Rev*, 2000, 100: 143–168
- 134 Engel J. *Biophysics*. Berlin: Springer, 1983
- 135 Guo Y, Li D, Zhang S, Yang Y, Liu JJ, Wang X, Liu C, Milkie DE, Moore RP, Tulu US, Kiehart DP, Hu J, Lippincott-Schwartz J, Betzig E, Li D. *Cell*, 2018, 175: 1430–1442.e17
- 136 Fang H, Sun Y, Wang X, Sharma M, Chen Z, Cao X, Utz M, Tian Z. *Sci China Chem*, 2018, 61: 1460–1464
- 137 Klein ML, Shinoda W. *Science*, 2008, 321: 798–800
- 138 Schwierz N, Frost CV, Geissler PL, Zacharias M. *J Am Chem Soc*, 2016, 138: 527–539
- 139 Sevim S, Sorrenti A, Franco C, Furukawa S, Pané S, deMello AJ, Puigmartí-Luis J. *Chem Soc Rev*, 2018, 47: 3788–3803
- 140 Ragazzon G, Prins LJ. *Nat Nanotech*, 2018, 13: 882–889
- 141 Hess H, Ross JL. *Chem Soc Rev*, 2017, 46: 5570–5587
- 142 Su NQ, Xu X. *Annu Rev Phys Chem*, 2017, 68: 155–182
- 143 Duan LL, Feng GQ, Zhang QG. *Sci Rep*, 2016, 6: 31488
- 144 Salomon-Ferrer R, Case DA, Walker RC. *WIREs Comput Mol Sci*, 2013, 3: 198–210
- 145 Brooks BR, Brucoleri RE, Olafson BD, States DJ, Swaminathan S, Karplus M. *J Comput Chem*, 1983, 4: 187–217
- 146 Jo S, Cheng X, Lee J, Kim S, Park SJ, Patel DS, Beaven AH, Lee KI, Rui H, Park S, Lee HS, Roux B, MacKerell Jr AD, Klauda JB, Qi Y, Im W. *J Comput Chem*, 2017, 38: 1114–1124
- 147 Guo W, Wu A, Zhang IY, Xu X. *J Comput Chem*, 2012, 33: 2142–2160
- 148 Kaiser J, Castellano M, Gnanndt D, Koslowski T. *J Comput Chem*, 2020, 41: 1105–1115
- 149 Kevrekidis IG, Samaey G. *Annu Rev Phys Chem*, 2009, 60: 321–344
- 150 Han J, Li M, Guo L. *J Sys Sci Compl*, 2006, 19: 54–62
- 151 Guo L. *J Sys Sci & Math Sci*, 2011, 9: 1014–1018 (in Chinese) [郭雷. 系统科学与数学, 2011, 9: 1014–1018]
- 152 Li M, Lü L, Deng Y, Hu MB, Wang H, Medo M, Stanley HE. *Natl Sci Rev*, 2020, 7: 1296–1305

153 Anderson PW. *Science*, 1972, 177: 393–396

154 Horgan J. *Sci Am*, 1994, 271: 34–35

Some thoughts about controllable assembly (II): catassembly in living organism

Zhi-Chao Lei², Xinchang Wang¹, Hang Qu¹, Cen Zhou³, Zhihao Li¹, Liulin Yang¹, Xiaoyu Cao¹, Zhong-Qun Tian^{1*}

¹ State Key Laboratory of Physical Chemistry of Solid Surfaces, Department of Chemistry, College of Chemistry and Chemical Engineering, and iChEM, Xiamen University, Xiamen 361005, China;

² Key Laboratory of Polymer Chemistry and Physics of Ministry of Education, Center for Soft Matter Science and Engineering, College of Chemistry and Molecular Engineering, Peking University, Beijing 100871, China;

³ Ocean College, Minjiang University, Fuzhou 350108, China

*Corresponding author (email: zqtian@xmu.edu.cn)

Abstract: Molecular assembly plays a critical role in constructing the physical world beyond molecules. Living organism shows the highest level of molecular assembly, inside which complex subunits assemble into functional assemblies through complex intermolecular interactions. Self-assembly is the dominant paradigm in the research of artificial molecular assembly, whereas most molecular assemblies in living organism are multi-pathway, multi-step, and hierarchical, thus requiring the assistance and guidance of other molecules. Among these assisting methods, catassembly uses catassemblers to modulate the kinetics and pathways of the assembly processes, exhibiting high efficiency and selectivity. In living organisms, catassemblers such as molecular chaperones, histone chaperones and allosteric factors are critical to protein folding, chromatin remodeling, and regulation of enzyme activities, respectively. In this review, we analyzed and then summarized various catassembly cases in living organisms, and envisioned that catassembly will be more important in more complex molecular assemblies. These investigations may provide a new perspective for the development of catassembly in artificial and living systems, and further improve the complexity, controllability and function of molecular assembly to a higher level.

Keywords: molecular assembly, controllable assembly, catassembly, catassembler, molecular chaperone

doi: [10.1360/SSC-2020-0127](https://doi.org/10.1360/SSC-2020-0127)