

用模块化分子元件调控和编程生物系统

刘海燕*

(中国科学技术大学生命科学学院, 合肥 230026)

摘要 合成生物学的目标包括“通过合成来理解生命”以及用现代工程学方法设计合成复杂生物系统。其工程学目标的实现依赖于可集成、可调控、可重用、功能多样的蛋白质、RNA、DNA 等基本分子元件。以分子机制为基础, 合理设计与实验室进化相结合, 改造和创建生物分子的相互作用特异性、调控方式、定量活性等, 是实现生物系统人工调控与编程的重要策略, 同时为自下而上设计合成日益复杂的人工生物系统奠定基础。

关键词 合成生物学, 工程化需求, 人工合成生物系统, 基因组学, 功能集成

学科分类号 Q6

DOI: 10.3724/SP.J.1206.2011.00582

1 合成生物学的兴起及其工程学目标

据哈佛大学历史学教授 Keller 的考证, “合成生物学”一词的出现已有约 100 年的历史。然而, 合成生物学作为一个新兴科学前沿的起点, 是约 10 年前研究者报道人工设计构建的基因“阻遏振荡器”^[1]和基因“双态开关”^[2]。这两项研究试图证明, “通过在最基本原理指导下进行设计、构建和验证来真正理解研究对象”这一现代科学的基本理念, 也适用于生命对象。对生物系统的人工设计和编程很快从基因转录线路扩展到细胞内信号通路的重连接^[3]、胞内的 RNA 逻辑线路^[4]、细胞间通信的多细胞体系^[5]、基因线路计数器^[6]等。

这些研究为生命科学带来了新的视角。例如, 分子水平的随机性和噪音在生物网络结构和功能上的重要性^[7]; 当生物系统复杂性达到一定程度后, 仅凭“直觉”很难对其功能特性做出正确判断, 数学模型变得不可或缺^[8], 等等。可以说, 合成生物学作为生命科学的一种基础研究策略和一个新兴科学前沿已得到广泛承认。

随着研究内容的拓展, 人们很快认识到, 人工生物系统的设计合成还可以与现代工程学的基本理念自然地结合, 即在定量模型指导下, 用“模块化、标准化”的组成单元来设计构建复杂系统^[9](图 1)。工程化设计构建复杂生物系统的能力可使

生物技术的发展模式、速度发生革命性变化。原有生物技术的发展模式主要是从环境中筛选具有所需物质能量转化能力的天然生物。而新的生物技术模式可以是在细胞中构建由天然或非天然功能元件组成、自然界不存在的人工代谢途径, 形成有生物制造能力的细胞工厂。

为了方便在 DNA 水平上用模块化、标准化的“部件”(parts)拼装“机器”(machines), 美国麻省理工学院的研究者创立了“标准生物元件登记库”

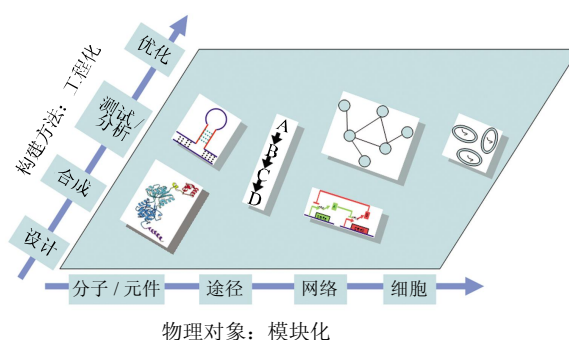


Fig.1 Modularity and engineering in synthetic biology

图 1 合成生物学中的模块化与工程化

* 通讯联系人。

Tel: 0551-36074510, E-mail: hylu@ustc.edu.cn

收稿日期: 2011-12-06, 接受日期: 2012-02-07

(<http://partsregistry.org/>), 用于登记能够编码特定生物学功能、在 DNA 水平上可以方便地组装的功能元件, 称为“BioBricks”. 加州大学伯克利分校的 Keasling 等利用酵母中的甲羟戊酸途径, 把 *amorphaadiene* 合成酶、细胞色素 P450 单加氧酶等元件组装到酵母中, 形成人工的青蒿酸生化合成通路, 青蒿素的生产成本可望显著降低^[10-11].

2010 年, Venter 等^[12]首次成功实现用全人工合成的基因组替换微生物的天然基因组. 首次实验证实可用全人造基因组来编程细胞的生命活动, 从而引起媒体和公众对人工生命潜在能力和对社会可能影响的广泛关注. 这一工作被《科学》(*Science*)杂志评为 2010 年十大科学突破第二位. Venter 等的工作还展示, 当前化学与分子生物学实验技术能力强, 已经可以在基因组规模“再现性”地制造复杂生命系统.

从设计生命的角度来看, Venter 等合成的基因组序列几乎百分之百来自于单种天然细菌, 显示目前人工设计生命的能力和合成基因组的能力之间存在巨大差距. 能否创立和发展有效的工程化策略和工具, 逐一突破设计生命的瓶颈, 是合成生物学发展的关键.

2 符合工程化需求的生物功能元件是合成生物学的基础和难点

之所以可能用工程化理念设计构建生物系统, 是因为生物系统从分子、途径、基因组到细胞等各个层次广泛存在的模块性. 在这里, 模块性是指一个复杂网络或系统可以分解成相对独立又互相耦合的组成单元. 合成生物学的工程学目标, 依赖于大量可集成、可调控、可重用、功能多样的模块化单元. 通过对模块化单元的功能进行数学抽象, 可以形成建立抽象数学模型的通用方法和计算程序, 使用系统分析、计算仿真等“干实验”手段设计、调试系统. 在此基础上, 才能做到通过设计、构建、调试优化的工程学循环来高效合成生物系统.

生物系统的最基本单元是分子元件. 如果把细胞看成化学工厂, DNA 序列就是其设计蓝图和控制程序. 由 DNA 编码的蛋白质、RNA 等是执行结构、转运、调控、催化等功能的基本元件. 这些基本分子单元在生命系统中表现出一定程度的通用性、模块性, 为基于它们设计构建新的人工生物系统提供了可能.

然而, 天然分子元件是自然进化的产物, 在模

块化、可组合、功能集成等方面远不符合通常的工程学需求. 例如, 生物系统与非生命工程系统的重要区别之一在于生命系统的组织和组装是自发的, 其更高层次的结构、动力学和功能由原子层次的化学结构和物理化学基本原理决定, 不需借助外力操作. 实现生物分子元件和模块可控的、功能化的自组织和自组装, 让不同单元之间、单元与途径和网络之间能够自发整合, 是工程化设计创建生物系统首先要面临的问题. 相应地, 大量已知功能的生物元件其自身性质、相互组装能力等还没有得到很好的定义和表征, 它们被整合到一起后不能协调一致地工作, 无法实现设计功能, 或实现效率太低, 是阻碍合成生物学实现其工程学目标的首要难点. 《自然》(*Nature*)2010 年发表的述评^[13]总结了现阶段合成生物学面临的五大困难, 它们都与这一难点密切相关.

3 现阶段合成生物学元件的数据化和标准化

在合成生物学发展的初期阶段, 元件标准化的主要思路一是通过设计接口处的核苷酸序列, 用标准化的实验组装流程把编码或非编码元件拼接在同一条载体 DNA 链上^[14], 二是提供一个对元件分类登记、查询、获取的统一平台. 美国麻省理工学院 2003 年建立了“标准生物元件登记库”(Registry of Standard Biological Parts)(<http://partsregistry.org/>), 首次把生物系统中的各类功能单元视为工程学的元件对象进行收集、分类描述以及分发. 从建立以来, 该元件库中的元件数量主要通过“国际遗传工程机器大赛”(iGEM)等途径快速增长, 已包括启动子、核糖体结合位点、蛋白质结构域、蛋白质编码序列等 10 余类共 30 000 多种元件. 对每一种元件有详略不一的功能特性描述. 很多 DNA “元件”已构建好标准的序列拼接接口, 通过简单分子生物学流程的反复循环能方便地拼装到一起. 这一元件库和 iGEM 竞赛对合成生物学发展产生了巨大推动作用, 其重要意义不容低估.

然而, 该元件库的设计思想也反映了早期对构建合成生物系统过于乐观. 仅仅在物理上把不同元件连接到同一条 DNA 链上, 远不能保证它们能够各自发挥预期作用、相互配合形成有完整功能的系统^[15]. 类似电子工程中物理设备与逻辑设备概念, 我们也可以把生物元器件组装分为基因组装与功能集成两个层次. 对单个宿主细胞中的生物线路来

说, 基因组装主要是将与元件对应的编码和非编码 DNA 导入宿主, 并保证其中编码蛋白质和 RNA 的部分能够正确表达. 功能集成则是元器件之间按预定方式相互作用和调控. 与电子线路中不同器件可用导线任意连接不同, 生物元器件的功能集成不能借助外力操作, 而要靠分子的扩散、定位与相互作用等来驱动. 如果器件之间和器件与环境之间不存在必须的相互作用、或者上下游器件的活性或强度不匹配, 都可能导致人工合成系统没有预期功能. 这些困难正是合成生物系统中元器件工程化的主要瓶颈. 完全依赖天然生物分子作为元器件不能克服这一瓶颈.

4 通过对天然分子的功能拓展创建符合需求的分子元件

4.1 通过功能拓展实现分子元件的按需连接和组装

在天然生物系统的进化过程中, 新的分子间组装和调控方式往往不是从头产生的, 而是原有蛋白质结构域、DNA 操纵子等功能单元的重新组合, 同时通过序列局部变化导致相互作用特异性改变等. 这同样是近年来对生物元件功能进行人工拓展的重要途径.

Skerker 等^[15]用细菌双组分系统, 展示了在生物信息学分析结果指导下, 仅仅通过两个界面残基的突变, 就能够改变上游组氨酸激酶和下游效应分子的连接特异性, 从而改变信号的传导方向. 新的相互作用特异性还可以通过在上下游分子中分别引入“插座”(adaptor)结构域和能够被 adaptor 特异性识别的配体结构域实现. Lim 等^[16]用这一策略改造了 MAPK 激酶信号传导通路中骨架(scaffold)蛋白的 adaptor 结构域, 实现了真核细胞中信号通路的重连接. 基于蛋白质相互作用结构域的脚手架蛋白还可以用于代谢途径上酶的顺序组装. Keasling 研究组^[17]把同一目标代谢途径上的酶与相互作用多肽(配体)融合, 同时用相应多肽的受体结构域(adaptor)融合构成合成骨架蛋白, 能够显著提高目标代谢途径的通量.

通过功能拓展实现元件自发按需连接组装, 除了创造必需的相互作用之外, 还可以消除不必要的相互作用. 例如, 通过改造核糖体, 创造与宿主正交的蛋白质翻译系统, 可以减少合成生物系统受到宿主的干扰^[18]. 在代谢路径改造中, 对关键酶进行改造, 可消除对目标产物产生有负面效果的产物抑

制效应^[19].

4.2 用分子元件人工调控生物系统

对生物器件进行功能拓展的另一个目的是对生物功能进行控制, 包括用环境信号控制合成生物系统, 或者借助人造元件来精确控制天然系统(图 2). 以信号响应蛋白质元件为例, 天然功能蛋白质分子往往包括信号感受、信号传递、效应器等功能单元. 这些单元大多与蛋白质结构域家族对应. 典型的信号感受单元如 PAS 结构域, 该家族的不同成员能够感受光、氧、代谢小分子等信号, 能引发构象变化^[20]. 这些结构域可称为“分子开关”. 分子开关结构域的构象变化可引起效应单元的功能发生改变, 从而调控下游系统. 在人工功能元件中, 响应同种信号的分子开关结构域可以与不同效应单元的整合. 例如, 感受光信号的 LOV 结构域(PAS 结构域的一个亚家族)被成功整合到组氨酸激酶、转录因子等功能蛋白中, 实现了多种人工光控器件^[21]. 这些器件不但能够用于用外部信号控制合成生物系统(例如细菌照相机), 还可用于对天然系统的高时空分辨控制(例如用光信号控制细胞形变和迁移). 用人工功能器件控制天然系统的另一例子是包含锌指结构域的人工核酸酶、人工转录因子^[22]. 它们所含的锌指结构域的 DNA 序列结合特异性在一定程度上是可设计的, 从而可以靶向控制天然系统中的基因表达或敲除. Kornmann 等^[23]把人工构建的、包含线粒体膜结合结构域和内质网膜结合结构域的人工蛋白导入酵母细胞, 用遗传互补方法发现了内源线粒体-内质网系连复合物.

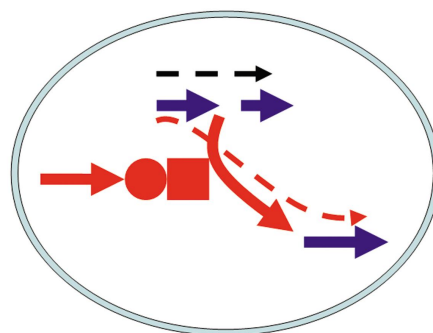


Fig.2 Regulate biological systems with synthetic molecular parts

图 2 用人工分子元件调控生物系统

蓝色箭头和黑色虚线代表细胞内源信息传导或调控路径. 红色几何图形代表人工元件, 红色箭头代表人工信息传导途径, 红色虚线代表新的信息传递途径.

4.3 元件定量活性的多样化

对合成生物系统而言, 元器件功能拓展的另一目的还可以是定量改变器件的活性, 以消除上下游单元之间活性的不匹配, 实现元件之间的协同工作, 最终让系统作为一个整体能高效完成设计目标或任务. Ellis 等^[24]通过构建启动子文库, 获得一系列有不同定量特性(转录活性)的启动子, 事先对其定量测试. 在后续设计的调控网络中, 依据数学模型, 直接选择使用具有所要求的定量特性的启动子元件. 这一工作很好地展示了首先通过序列变化产生有多样化定量特性的同类功能元件, 然后在数学模型指导下设计构建复杂生物系统的研究策略.

4.4 用基于 RNA 的分子元件实现生物系统的人工调控与编程

RNA 生物学和核酸工程的快速发展极大地推进了程序化地使用人工 RNA 调控生物系统的研究^[25-26]. 对合成生物系统来说, 人工 RNA 调控器件的重大优点是其特异性连接可以部分或全部依赖核苷酸序列的互补性, 从而使元件的集成具有很强的可编程性和模块性, 符合工程化的需求, 例如, 感受同种分子的 RNA 调控元件与不同的靶向序列组合, 就可获得调控不同下游目标的功能器件.

RNA 可通过与 mRNA 的反义结合、与小分子或蛋白质结合的构象变化、核酶的催化作用等机制发挥调控功能. 位于 5' 非翻译区(UTR)的 riboswitch 可作为小分子感受器, 用小分子信号调控蛋白质翻译^[25]. Riboswitch 区域还可以和核酶相连, 通过小分子控制核酶的催化功能^[26]. Riboswitch 还被用于设计能感受特定蛋白质表达等细胞状态变化的控制器. 利用 riboswitch 可设计感受单种小分子信号的“反义开关”^[4], 在特定小分子存在时实现(或解除)对特定 mRNA 的翻译抑制, 还可以实现用多种小分子信号的组合输入来调控蛋白质表达水平. Rinaudo 等设计验证了小干扰 RNA (siRNA) 的组合逻辑调控能力: 同一阻遏蛋白由两种具有不同 UTR 区的 mRNA 编码, 只有当两种相应的 siRNA 同时存在时, 其下游基因的抑制才能被解除. 利用类似原理可设计包括 3 个或更多输入的复杂逻辑表达式. 在高等真核生物中, 内含子也可以是 mRNA 层次调节的重要途径. 内含子可能的调控机制之一, 是增加 mRNA 成熟时间, 延缓蛋白质表达. Swinburne 等^[27]用一个含内含子和自抑制回路的人工基因器件验证了该器件能够产生脉冲式的基因表达, 并且脉冲的频率依赖于内含子

长度.

5 多学科交叉推动生物分子元件的按需设计与制造

5.1 计算生物学的发展提供了不同层次的合理设计工具

近 10 年来, 基于空间结构的蛋白质序列设计发展迅速, 完成多项有里程碑意义的成果. 1997 年, Mayo 等^[28]首次报道从头计算设计的多肽序列在不存在锌离子的情况下稳定折叠成类似锌指的空间结构. 2003 年, Kulhman 等^[29]报道从头设计的序列 TOP7 形成自然界尚未发现过的折叠类型. 2008 年和 2010 年, Baker 研究组先后报道了 3 种以天然蛋白为骨架、活性中心经重新设计的人工酶, 它们能够催化目前尚未发现存在天然催化酶的有机反应^[30-31]. 序列设计还被应用于蛋白质相互作用界面设计与优化^[32]. 随着计算能力的指数增长, 对蛋白质原子分辨率构象动力学进行计算机模拟的时空尺度快速提高. 2010 年, Shaw 等^[33]报道对蛋白质毫秒时间尺度动力学的计算机模拟, 使用同一物理模型成功再现了两种有不同结构类型的蛋白质从随机、无规卷曲的初始结构折叠成天然结构(与实验结构原子位置均方根偏差达 1 Å 左右)的过程. 这一工作被《科学》(Science)杂志评为 2010 年十大科学突破的第八位. 对生物分子构象动力学的长时间准确模拟是正确预测和设计生物分子相互作用特异性的基础. 对功能相关动力学行为的模拟是结构与动力学关系分析的依据, 其结果可指导通过结构域插入实现新别构调节功能等分子设计.

建立在生物学海量实验数据基础上的生物信息学技术稳步发展. 相关算法的通量、灵敏度、选择性等也有大幅提高, 使得同源序列家族及进化关系、蛋白质折叠家族、蛋白质相互作用网络等数据信息在快速增长的同时, 准确性显著提高. 借助生物信息学工具, 自然界生物系统化学结构、空间结构和功能的多样性和模块性可为发展合成生物学功能器件特别是信号感受器件和化学催化器件等提供大量原型分子, 同时也为发展模块化、工程化调控元件提供设计模板和指导规律. 例如, 分析不同蛋白质序列或序列上不同位置间的共进化关系, 可为相互作用设计、别构功能设计等提供指导. 通过系统生物学研究, 可基于实验数据建立生物网络的静态结构模型、稳态和瞬态动力学模型等数学模型. 通过对模型进行分析, 可发现网络中需改造(如解

除反馈抑制)或引入外源调控机制的关键节点。

5.2 实验室进化与合理设计的跨层次结合与相互推动

到目前为止,对生物功能器件或合成生物系统性能的充分优化,离不开实验室进化^[34-35]。实验室进化包括具有恰当多样性的突变文库的产生、文库的转化与功能表达、利用特征标记物或标记信号进行表型筛选等步骤。合理设计对实验室进化的重要作用可体现在以下方面。首先,通过合理设计可获得具有初步功能的人工系统(例如让一种分子开关结构域模块具有识别全新配基的初步活性),为实验室进化提供起点。在这一点上,目标功能越新颖,合理设计的作用就越大。其次,合理设计可大幅度(多个数量级)地缩减实验需要探索的多样性空间。以蛋白质氨基酸序列多样性为例,如果通过合理设计能够从上百个可能位点中确定突变位置,仅仅可能的双突变体数目就能减少数千倍。当有更多位点突变,或者合理设计能够进一步限制氨基酸残基的变化范围时,缩减幅度更大。最后也是最重要的,在人工生物系统和功能器件研究中,合理设计与实验室进化可以在不同层次进行(即分别考虑不同层次的对象),最大程度利用既有知识,提高效率和成功率。以代谢途径优化为例,可以利用网络层次的合理设计找出瓶颈节点和相互作用,再用实验室进化对相关分子进行突变筛选。在人工调控分子设计中,可基于对蛋白质分子整体层次的结构和动力学的分析,合理设计结构域的相对排列、热点残基位置,再对构建细节(个别位点残基类型、插入或连接片段的小范围变化)等进行实验筛选。在已有用感光分子开关结构域成功构建人工调控蛋白的例子中,通过先合理设计,实验筛选鉴定的变体数目都只有10余种^[36]。

需要指出的是,合理设计与实验室进化是互相补充而不是排斥的手段。在可以达到很高筛选通量的情况,实验室进化可以发现目前合理设计方法不能预测的多样性“盲区”,如对酪氨酸合成途径的合理设计确定了产物反馈抑制以及一种中间前体的合成是代谢路径优化的关键节点。通过实验改造这些节点确实成倍提高了目标产物产率。对同一体系采用完全随机的实验室进化,发现RNA聚合酶系统的突变也能大幅提高目标产物产率。后面这个靶标是用现有合理设计策略无法发现的^[37]。

5.3 化学方法实现生物分子可控自组装与功能化

通过生物分子自组装形成有可控结构和功能的

分子器件,不但可在有机体内进行,也可脱离有机体,完全用化学方法来实现。最有代表性的是利用DNA序列互补性组装构造各种功能模块与器件。

美国纽约大学化学系的Nadrian C. Seeman教授开创了结构DNA纳米技术这一研究领域,并在近年来得到迅猛的发展^[37]。这种基于核酸分子的构造方法具有三大优势:a. 优异的结构可设计性;b. 组装过程的精确性和高度并行性;c. 符合“自下而上”的生物体系构造原则。

目前,人工构筑的DNA结构已从简单的一维结构向二维甚至三维结构迈进,并由简单的周期性重复结构向非周期性结构拓展^[38]。另一方面,基于DNA分子构筑的核酸功能器件也引发了研究人员的强烈兴趣。Mao和Seeman等于1999年报道了一种基于B-Z构象变换的核酸机器^[39]。美国Bell实验室的科学家Yurke等^[40]随后在《自然》(*Nature*)杂志上发表论文向人们演示了一种可控“开”“合”的DNA分子镊。中国科学技术大学研究小组基于pH敏感的Hoogsteen氢键和DNA链置换反应,成功构造出可对特定DNA目标物进行准确识别与抓取的新型DNA分子镊^[41]。这些工作引发了人们对功能核酸分子和纳米器件的强烈兴趣。文献中利用DNA已经能够组装出模仿某些生物学功能的结构模块,如:可按“指令”合成氨基酸寡聚物的DNA机器^[42],以及对微管束(microtubule)上蛋白质机器运动的DNA步行器等^[43]。

6 国内发展展望

在生物大分子的人工合成领域,我国的前辈科学工作者在20世纪曾经创造过人工合成胰岛素和合成酵母丙氨酸转移核糖核酸等重大科学突破。改革开放以来,我国在生命科学基础研究领域实现跨越式发展,近10年间,在基因组学、系统生物学、生物信息学、计算生物学等方面都有了蓬勃的发展。与此同时,大力推进基因工程、蛋白质工程和代谢工程等生命科学工程领域的发展。以上为发展合成生物学研究奠定了很好的基础。近年来,国内科学界就我国合成生物学的发展方向展开了广泛深入研讨,先后召开“香山科学会议”、“东方论坛”、“中德圆桌会议”等进行专题讨论。从2007年开始,北京大学、中国科学技术大学、清华大学、天津大学等高校积极参加麻省理工学院组织的“国际遗传工程机器”(iGEM)比赛等,并取得了优异成绩。该比赛被公认对合成生物学的发展有重要

推动作用. 在 973 计划中, 2011 年正式启动了“人工合成细胞工厂”、“光合作用与‘人工叶片’”重大科学问题导向项目, 重点关注底盘细胞基因组合成、重大化工原料的高效生物制造和光能转化的合成生物学, 开始了国家层面对合成生物学研究的战略布局.

总体而言, 国内对人工合成生物系统的基础研究方面已经积累了较强的方法基础和团队基础. 以合成生物系统中的生物功能元件研究为例, 在从分子到网络的计算分析设计、实验室进化、化学可控自组装等, 国内都有具有一定实力的研究力量和团队, 并在不同局部取得了重要成果. 目前亟待解决的是, 要以合成生命体系的整体框架为指导, 实现不同学科、不同研究手段优势团队的深入交叉合作, 目标明确地对基础原理、普适方法和关键技术瓶颈等问题进行系统、深入、持续的研究, 为合成生物学工程化设计创造生命的发展目标提供支撑.

致谢 本文部分内容得益于与中国科学院生物物理研究所蒋太交研究员、杭海英研究员、中国科学技术大学邓兆祥教授、单革教授、天津工业生物技术研究所王钦宏研究员等的讨论及他们提供的参考资料, 特此感谢.

参 考 文 献

- [1] Elowitz M B, Leibler S. A synthetic oscillatory network of transcriptional regulators. *Nature*, 2000, **403**: 335–338
- [2] Gardner T S, Cantor C R, Collins J J. Construction of a genetic toggle switch in *Escherichia coli*. *Nature*, 2000, **403**: 339–342
- [3] Dueber J E, Yeh B J, Chak K, *et al.* Reprogramming control of an allosteric signaling switch through modular recombination. *Science*, 2003, **301**: 1904–1908
- [4] Bayer T S, Smolke C D. Programmable ligand-controlled riboregulators of eukaryotic gene expression. *Nat Biotechnol*, 2005, **23**: 337–343
- [5] Basu S, Gerchman Y, Collins C H, *et al.* A synthetic multicellular system for programmed pattern formation. *Nature*, 2005, **434**: 1130–1134
- [6] Friedland A E, Lu T K, Wang X, *et al.* Synthetic gene networks that count. *Science*, 2009, **324**: 1199–1202
- [7] Elowitz M B, Levine A J, Siggia E D, *et al.* Stochastic gene expression in a single cell. *Science*, 2002, **297**: 1183–1186
- [8] Kaern M, Blake W J, Collins J J. The engineering of gene regulatory networks. *Annu Rev Biomed Eng*, 2003, **5**: 179–206
- [9] Endy D. Foundations for engineering biology. *Nature*, 2005, **438**: 449–453
- [10] Martin V J, Pitera D J, Withers S T, *et al.* Engineering a mevalonate pathway in *Escherichia coli* for production of terpenoids. *Nat Biotechnol*, 2003, **21**: 796–802
- [11] Ro D-K, Paradise E M, Ouellet M, *et al.* Production of the anti-malarial drug precursor artemisinic acid in engineered yeast. *Nature*, 2006, **440**: 940–943
- [12] Gibson D G, Glass J I, Lartigue C, *et al.* Creation of a bacterial cell controlled by a chemically synthesized genome. *Science*, 2010, **329**: 52–56
- [13] Roberts K. Five hard truths for synthetic biology. *Nature*, 2010, **463**: 288–290
- [14] Shetty R P, Endy D, Jr Knight T F. Engineering BioBrick vectors from BioBrick parts. *J Biol Eng*, 2008, **2**: article 5
- [15] Skerker J M, Perchuk B S, Siryaporn A, *et al.* Rewiring the specificity of two-component signal transduction systems. *Cell*, 2008, **133**: 1043–1054
- [16] Bhattacharyya R P, Remenyi A, Yeh B J, *et al.* Domains, motifs and scaffolds: the role of modular interactions in the evolution and wiring of cell signaling circuits. *Annu Rev Biochem*, 2006, **75**: 655–680
- [17] Dueber J E, Wu G C, Malmirchegini G R, *et al.* Synthetic protein scaffolds provide modular control over metabolic flux. *Nat Biotechnol*, 2009, **27**: 753–759
- [18] Rackham O, Chin J W. A network of orthogonal ribosome x mRNA pairs. *Nat Chem Biol*, 2005, **1**: 159–166
- [19] Luetke-Eversloh T, Stephanopoulos G. L-tyrosine production by deregulated strains of *Escherichia coli*. *Appl Microbiol Biotechnol*, 2007, **75**: 103–110
- [20] Moeglich A, Ayers R A, Moffat K. Structure and signaling mechanism of Per-ARNT-Sim domains. *Structure*, 2009, **17**: 1282–1294
- [21] Moeglich A, Ayers R A, Moffat K. Design and signaling mechanism of light-regulated histidine kinases. *J Mol Biol*, 2009, **385**: 1433–1453
- [22] Beerli R R, Barbas III C F. Engineering polydactyl zinc-finger transcription factors. *Nat Biotechnol*, 2002, **20**: 135–143
- [23] Kornmann B, Currie E, Collins S R, *et al.* An ER-mitochondria tethering complex revealed by a synthetic biology screen. *Science*, 2009, **225**: 477–481
- [24] Ellis T, Wang X, Collins J J. Diversity-based, model-guided construction of synthetic gene networks with predicted functions. *Nature Biotechnol*, 2009, **27**: 465–471
- [25] Isaacs F J, Dwyer D J, Collins J J. RNA synthetic biology. *Nat Biotechnol*, 2006, **24**: 545–554
- [26] Win M N, Liang J C, Smolke C D. Frameworks for programming biological function through RNA parts and devices. *Chem Biol*, 2009, **16**: 298–310
- [27] Swinburne I A, Miguez D G, Landgraf D, *et al.* Intron length increases oscillatory periods of gene expression in animal cells. *Genes Dev*, 2008, **22**: 2342–2346
- [28] Dahiyat B I, Mayo S L. De novo protein design: Fully automated sequence selection. *Science*, 1997, **278**: 82–87
- [29] Kuhlman B, Dantas G, *et al.* Design of a novel globular protein fold with atomic-level accuracy. *Science*, 2003, **302**: 1364–1368

- [30] Siegel D J. Computational design of an enzyme catalyst for a stereoselective bimolecular diels-alder reaction. *Science*, 2010, **329**:309–313
- [31] Jiang L, Althoff E A, *et al.* De novo computational design of retro-aldol enzymes. *Science*, 2008, **139**: 1387–1391
- [32] Mandell D J, Kortemme T. Computer-aided design of functional protein interactions. *Nat Chem Biol*, 2009, **5**: 797–805
- [33] Shaw D E, Maragakis P, *et al.* Atomic level characterization of the structural dynamics of proteins. *Science*, 2010, **330**: 341–347
- [34] Schmidt-Dannert C, Arnold F H. Directed evolution of industrial enzymes. *Trends Biotechnol*, 1999, **17**: 135–136
- [35] Dietrich J A, McKee A E, Keasling J D. High-throughput metabolic engineering: Advances in small-molecule screening and selection. *Annu Rev Biochem*, 2010, **79**: 563–590
- [36] Toettcher J E, Voigt C A, *et al.* The promise of optogenetics in cell biology: interrogating molecular circuits in space and time. *Nat Methods*, 2011, **8**: 35–39
- [37] Zheng J P, Birktoft J, Chen Y, *et al.* From molecular to macroscopic *via* the rational design of a self-assembled 3D DNA crystal. *Nature*, 2009, **461**: 74–77
- [38] Douglas S M, Dietz H, Liedl T, *et al.* Self-assembly of DNA into nanoscale three-dimensional shapes. *Nature*, 2009, **459**: 414–418
- [39] Mao C D, Sun W Q, Shen Z Y, *et al.* A nanomechanical device based on the B-Z transition of DNA. *Nature*, 1999, **397**: 144–146
- [40] Yurke B, Turberfield A J, Mills A P, *et al.* A DNA-fuelled molecular machine made of DNA. *Nature*, 2000, **406**: 605–608
- [41] Han X G, Zhou Z H, Yang F, *et al.* Catch and release: DNA tweezers that can capture, hold and release an object under control. *J Am Chem Soc*, 2008, **130**: 14414–14415
- [42] He Y, Liu D R. Autonomous multistep organic synthesis in a single isothermal solution mediated by a DNA walker. *Nat Nanotechnol*, 2010, **5**: 778–782
- [43] Tian Y, He Y, Chen Y, *et al.* A DNzyme that walks processively and autonomously along a one-dimensional track. *Angew Chem Int Ed*, 2005, **44**: 4355–4358

Regulating and Programming Biological Systems With Modular Molecular Parts

LIU Hai-Yan*

(School of Life Sciences, University of Science and Technology of China, Hefei 230026, China)

Abstract The objectives of synthetic biology include to understand life through synthesis and to engineer complex biological systems based on general principles of modern engineering sciences. The engineering goal relies on the availability of molecular parts with diverse functions that can be predictably integrated, regulated and reused. These parts may include proteins, RNA, DNA and their complexes. Based on the understanding of molecular mechanisms and judicious combinations of rational design and laboratory evolution, specific interactions, regulatory functions and activities of biomolecules can be artificially changed or created. This comprises an important strategy to artificially regulate and reprogram existing biological systems. It also lays the foundation for the bottom-up design and construction of artificial biological systems of increasing complexity.

Key words synthetic biology, engineering approaches, artificial biological systems, gene assembly, functional integration

DOI: 10.3724/SP.J.1206.2011.00582

*Corresponding author.

Tel: 86-552-36074510, E-mail: hylu@ustc.edu.cn

Received: December 6, 2011 Accepted: February 7, 2012