



DNA 纳米技术与 生物编程*

文 / 李娜¹ 丁宝全^{1*} 颜颢^{2**}

1 中科院/教育部 国家纳米科学中心 北京 100190

2 美国亚利桑那州立大学 坦佩市 85287

【摘要】 DNA 纳米技术是设计和构建具有一定用途的人工核酸纳米结构。DNA 纳米技术中的一个重要挑战就是如何有效设计和构建具有明确功能的纳米尺度的结构和器件。计算机辅助工具是预测、设计、建模以及描绘 DNA 结构的最强有力的工具。计算机辅助设计 DNA 结构以及识别有效的自组装路径让 DNA 作为一种独特的材料在构建纳米结构领域中显露锋芒。文章总结了 DNA 纳米结构设计的原理,介绍了一些用于 DNA 结构设计的算法以及一些用户友好软件。同时,对 DNA 纳米技术中构建 DNA 逻辑门以及 DNA 计算方面的最新研究进展做了介绍。

【关键词】 DNA 纳米技术,人工核酸纳米结构,化学自组装,计算机辅助设计,生物编程

DOI 10.3969/j.issn.1000-3045.2014.01.008

DNA 分子是遗传信息的携带体,它可简单精准地传递遗传信息,其应用渗透于我们生活中的许多领域。从遗传咨询到信息取证,从基因组学到基因治疗,所有这些以及其他密切影响人类生活的应用,都依赖于 DNA 作为遗传物质的功能。DNA 对于生命体有着众所周知的重要性,但在其生物学功

能之外,着眼于 DNA 的结构特性时,就会了解到 DNA 在构建功能纳米图形及结构方面具有得天独厚的优势,由此发展出一个全新的领域——DNA 纳米技术。DNA 纳米技术是设计和构建具有一定用途的人工核酸纳米结构。DNA 纳米技术中的一个重要挑战就是如何有效设计和构建具有明确功能的纳米尺度的结构和器件。化学自组装是构建纳米结构和器件的有效方法之一。DNA 是适合自组装的良好材料,因为它们可以构建具有清晰轮廓的结构,而且可对预

* 基金项目:国家自然科学基金(21222311),中科院“百人计划”项目

** 通讯作者

修改稿收到日期:2014年1月5日



中国科学院

先的设计进行编程,并且在生物系统中具有良好的适配性^[1]。DNA在纳米尺度下的精确建模和可寻址能力在化学、分子计算以及纳米医药的新兴领域中都将是具有极为可观的应用价值^[2,3]。构建一个新的DNA纳米结构通常包括以下步骤^[4]:(1)建立模型;(2)设计序列;(3)DNA序列的合成、纯化和自组装过程;(4)DNA纳米结构的表征。近年来,计算机辅助设计DNA结构以及识别有效的自组装路径让DNA作为一种独特的材料在构建纳米结构领域中显露锋芒,并且可在某种程度上超越纳米技术中“自下而上”(bottom up)的合成方法,使“自上而下”(top-down)的合成方法进一步得到延伸。

计算机辅助工具是预测、设计、建模以及描绘核酸纳米结构的研究中不可分割的一部分^[5]。预测方法(如图1A所示)一般开始于从序列数据中预测二级结构,然后通过比对实验数据(如:X射线晶体学或核磁共振)来预测三级结构的相互作用以及三维构型。设计方法(如图1B所示)是从构建一个三维模型或一个简单的二维模型开始,然后是设计序列,使其能够折叠成理想结构。将设计好的序列合成出来,用来自组装成理想结构,然后将其与理论模型进行比较,并且通过生物物理等方法进行表征。预测方法的工作流程是用于直接鉴别和描绘由于进化而演变的生物结构,而设计模型则采取了一种循环路径,用于优化由科学家构建的理想结构。

计算机辅助设计DNA纳米结构的起始阶段是较小几何结构的理论计算,如DNA立方体^[6]。第二阶段是设计较大复杂结构的DNA瓦片(tile),该设计能被特异性识别并且在识别过程中可以执行算法计算^[7,8]。第三阶段是自组装有效路径的实现,该途径通过一个较大骨架链的高序折叠来实现^[9]。设计DNA结构的通用方法是用已有的三维模块(motif)来构建全新的三维结构,然后设计序列,使其自组装生成所设计的结构。

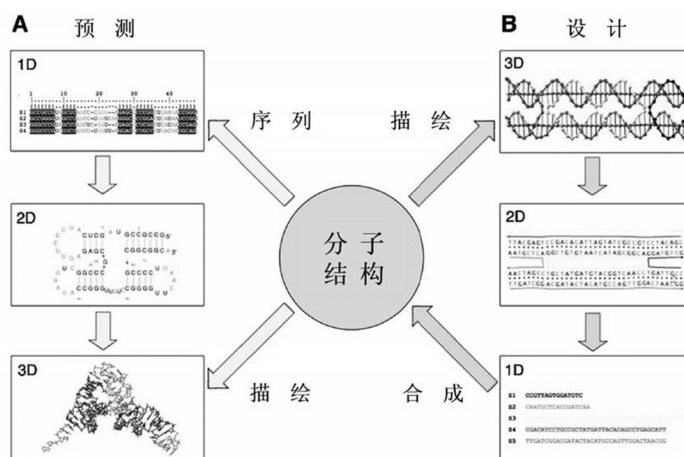


图1 计算机辅助预测及设计核酸纳米结构的流程图

1 从结构到序列

最早用于DNA结构设计的实用方法是采用序列对称性最小化(sequence symmetry minimization, SSM)原理来构建稳定的DNA螺旋四臂结(four-way junctions)^[10],如图2A所示。该程序名为SEQUIN^[11],由DNA纳米技术的奠基人,纽约大学的纳德里安·西曼(Nadrian Seeman)教授推出。SEQUIN通过减少重复序列以及减小在螺旋连接处可替换碱基对的可能性来筛选序列。

SSM模型已成功整合至图形用户界面的软件,如:Tiamat^[12],并且进一步扩展到用于较大分支结构的序列设计软件,如:Uniquimer-3D^[13]。另一种设计方法是利用自由能折叠算法来保证目标结构的自由能最小。然而,这种方法不能保证所选序列可高概率地按照目标折叠进行组装。一个可行的方案是在未确定它是总体能量最小值的情况下降低目标折叠的能量。这种解决方案对目标结构具有更高的亲和力,然而如果总体能量最小值存在的话,就有可能导致目标结构形成的概率更低。如果将两种设计方案结合起来,如:在序列选择步骤中结合SSM和自由能最小化^[14],就有可能得到更好的结果。

2 设计原理和自组装路径

为实现利用计算机模拟现实模型,最重要的

一点就是遵循设计原理,设计原理要与已知的自组装路径兼容。基于瓦片(tile)自组装模型的设计已非常成熟(如图2A所示),且

叠产物的自由能最小。

Rothemund在2006年提出一个全新的自组装方法,被称作DNA折纸术^[22]。该方法

通过将一条较长的DNA单链(7.2kb的单链M13噬菌体基因组DNA)折叠成预先设计的结构,然后与200—250条短的锚定链(又称为订书钉链, staple strands)进行退火杂交,就可自组装成特定的图形或结构,如图3A所示。该方法为直接自组装提供了迄今为止不可限定的可能性,因为250个短的锚定链从理论上来说,可从折纸结构表面延伸出来作为250个独立的可寻址像素,位于折纸结构表面,像素可达6 nm(如图3C所示)。自Rothemund提出DNA折纸结构以来,该方法被成功用来构建蛋白纳米阵列^[25]、金纳米粒子阵列等^[26],并为算法计算提供了相应的支撑^[27]。

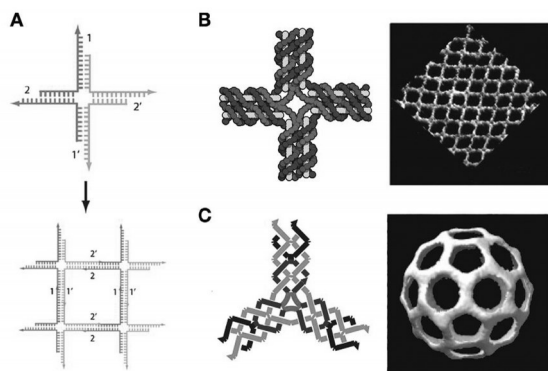


图2 基于tile自组装DNA结构。(A)上:四臂结^[18],箭头代表3'端,1,1'以及2,2'为粘性末端;下:四个四臂结通过1,1'以及2,2'彼此杂交所形成的结构;(B)左:十字tile^[9];右:十字tile组装的二维网格的AFM图^[9];(C)左:三角星tile^[19];右:三角星tile组装的DNA巴基球的冷冻电镜重建图^[16]

成功用于构建DNA的二维和三维网格图形或结构(如图2B,C所示)^[6,15-17]。早期的研究是利用手动步进式方式让多条链按照正确的顺序进行组装^[6],后来的方法演变成利用浓度和几何构型来控制产物结构的大小和形状^[16]。最近的研究表明,合理设计tile就可自组装成宏观可见的三维晶体^[17]。

然而,基于tile结构的建模方法并不是组装具有特定形状和大小结构的最佳方法。自然界中产生功能RNA结构的自组装路径是将RNA的一条单链进行折叠,使之成为最佳结构。采用与之相似的方法,可将一条含有1.7 kb的DNA单链与一些短的辅助单链(helper strands)进行自组装,从而可用来合理设计DNA的八面体结构^[20]。采用SSM原理进行序列设计,然后用Mfold软件(现已被UNAFold软件所取代)^[21]进行反复验证,最终可确定折

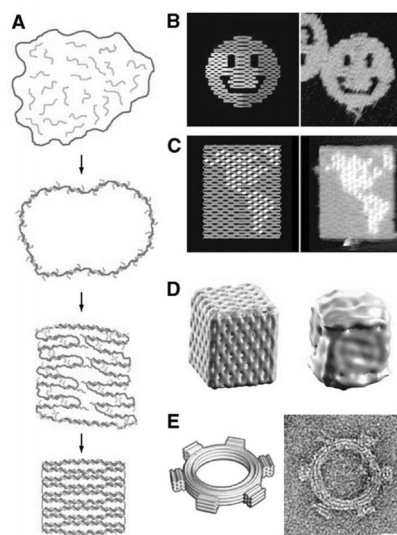


图3 DNA折纸结构的设计以及自组装。(A)DNA折纸结构自组装路径示意图,黑色代表环状DNA骨架链,灰色代表短锚定链;(B)利用DNA折纸术构建的笑脸结构^[22]; (C)利用DNA折纸术构建的长方形,上面用哑铃状的发卡结构作为像素修饰排布南、北美洲图形^[22]; (D)利用DNA折纸术构建的 $42 \times 36 \times 36 \text{ nm}^3$ 的盒子,左为理论模型,右为冷冻电镜得到的实验模型^[23]; (E)控制DNA折纸结构的扭曲度和弯曲度从而构建具有弯曲度的齿轮图形,左为理论模型,右为TEM图^[24]

每个锚定链沿着 DNA 的长链在两个或更多的位点进行杂交,从而可以构建预先设计的二维图形或结构。然而该方法面临着序列设计问题,即选用的一条较长的 DNA 单链,其序列是确定的(如 M13 噬菌体基因组 DNA),但是却需要最少的重复区域以及稳定的发卡结构。这种技术被证明可形成任意图形(如图 3B 所示),并且具有独特的表面寻址能力(如图 3C 所示)。基于计算机辅助设计的软件极大地促进了 DNA 结构设计,计算得到基于 M13 单链 DNA 骨架形成 DNA 折纸结构大约需要 250 个锚定短链,并且软件的最新改进可以发展 DNA 折纸术用于构建三维结构(如图 3D, E 所示)。

3 构建二维或三维结构的方法

在建立模型过程中,实际图形通常被用来辅助 DNA 结构的设计,建立模型要求设计者具有良好的 DNA 结构方面的基础知识,如螺距、直径、碱基堆积、磷脂键长等。模型建立好后可以给予设计者关于 DNA 结构的尺寸、形状和稳定性等直接信息。在设计 DNA 结构时最重要的因素是 DNA 结构的自由能最小化。目前,许多实验室成功开发了不同的程序如 Mfold^[21]、GIDEON^[28]、Tiamat^[12]和 NanoEngineer-1^[29]等用于辅助建模和自由能预测。另外,目前存在相当数量的可以直接利用的 DNA 纳米结构,可通过改编、拼接和剪裁已知的序列来自组装构建新的结构。

序列设计中的一个最重要的原则是序列对称性最小化。在分支结构中,序列对称性最小化能够有效避免参加结构形成的 DNA 序列之间的错配和交叉点的位移。需要指出的是,我们只需要在设计单条 DNA 链时避免序列的对称性,但同一个 tile 中多条链间的对称性是不需要完全回避的。DNA 序列都可以利用计算机程序进行自动化设计,如 SEQUIN^[11]、Tiamat^[12]、Uniquimer-3D^[13]等,在设计过程中,程序会自动避免序列对称性。这些程序提供了交互式用户界面,可以更好地促进复杂二维和/或三维结构的设计,并且一些程序

还包括一些完整的模块数据库以及三维模型和/或序列设计的精炼算法。用于 DNA 结构设计的各种软件的比较如表 1 所示。

为构建 DNA 模型更为复杂的结构,其应用程序需要精准诠释结构的几何设计。还需注意的是,所设计的 DNA 纳米结构是由随机序列构建生成的,要求序列的对称性最小,无二级结构产生。1985 年,Seeman 设计了一条基于公式翻译程序(Fortran program)的命令行,可以创建相对简单的 DNA 纳米结构图形^[32]。但这一程序缺少图形接口,导致直接观察三维 DNA 结构非常困难。最近, Birac 和 Seeman 设计了 GIDEON 程序^[28], GIDEON 是三维 DNA 渲染程序,且非常有用,它允许使用者设计复杂精准的 DNA 纳米结构。该程序可以让设计的结构从顶端、侧面和前面可见。GIDEON 具有预测构建 DNA 纳米结构中所发生的真实畸变的能力。总体来说, GIDEON 在设计 DNA 结构方面是一款非常不错的软件。然而, GIDEON 在有效构建较大结构设计上面有一些缺陷,如,其序列的生成与构建的模型不能整合。生成 DNA 序列占主导地位的软件是基于公式翻译程序的 SEQUIN^[11]。SEQUIN 可将多个 DNA 双螺旋连接在一起并且可以按照一次一个的方式产生一组序列。将多个 DNA 双螺旋连接在一起,要求使用者首先要画一个二维的结构线条图,然后将结构输入到程序里,且要标明臂的长度以及将多条臂连接起来。一旦结构在程序中执行,使用者可以得到该结构的序列,通过新序列与原结构中的现有序列的比对,从而避免序列的重复与可能的碱基错配。结构生成器不是随机的,使用者需要决定生成的序列是否是理想序列,或者决定让计算机重新生成一组序列,并且将较大的结构输入成一条单独的命令行是一项非常繁重的工作。

先前的设计程序的另一缺陷,就是数据结构的设计和可视化不够灵活,不足以设计较大的并且复杂的结构。对设计者来说即使是在一个简单

表 1 DNA 结构设计工具

工具	类型	输入	输出	年份/参考文献
SEQUIN	序列对称性最小	2D 结构 多条链	序列表	1982 ^[10]
DNA-origami	DNA 折纸术	2D 结构 手动	序列表	2006 ^[22]
GIDEON	编辑器, 3D 界面; 张力最小	3D 结构 手动	3D 结构	2006 ^[28]
Tiamat	编辑器, 3D 界面; 序列对称性最小	3D 结构 手动	序列表	2009 ^[12]
NanoEngineer-1	编辑器, 3D 界面; 能量最小	3D 结构 手动	3D 结构	2008 ^[29]
Uniquimer-3D	编辑器, 3D 界面; 对称和能量最小	3D 结构 手动	序列表	2009 ^[13]
SARSE	编辑器, 2D 界面; DNA 折纸术	2D 结构 自动	3D 结构 序列表	2008 ^[30]
caDNAno	编辑器, 2D 界面; DNA 折纸术	2D 结构 手动	序列表	2009 ^[31]

设计内的大量的分子结构也有可能造成比较大的困惑,并且设计者也会被图形硬件折磨不堪。Sean Williams 和 Hao Yan 等设计了一个名为 Tiamat^[12]的软件可以弥补上述软件的缺陷。该软件具有 3 个明显的优点:首先,设计界面在结构建模上非常灵活,结构允许使用通用的设计进行建模;其次,具有一个可以整合随机序列的生成器,该生成器可直接检查所提供模型的序列,从而避免形成二级结构;最后,该软件具有非常快速并且是动态层面的细节算法,可以使较大的结构进行精简显示,但却可以囊括重要的信息。Tiamat 可以同时给出设计结构的不同视图,包括前面、侧面、顶端以及三维视图。在结构设计中并不像 GIDEON 那样用 DNA 螺旋作为设计单元,而是采用碱基作为最小设计单元,从而更具有灵活性。Tiamat 结构 DNA 纳米技术中是一款非常实用的软件,

Yan 等利用 Tiamat 设计了 4×4 DNA tile, 并且将单一的 4×4 DNA tile 外延出较大的二维阵列图形^[9]。该二维阵列结构是通过十字结上每个臂中 5 个碱基的粘性末端彼此垂直连接而构建。并且,该二维阵列采用波浪结构的设计从而可以使 tile 能够形成较大的二维阵列,而不是自身折叠形成管状结构。如图 4A 所示为 Tiamat 输出的 4×4 DNA tile 的设计模型,图 4B 为用 4×4 DNA tile 外延较大二维阵列图形的设计模型以及其构建的二维晶体的 AFM 图。最近 Yan 等又利用 Tiamat 设计了基于平行螺旋结构的单向 DNA 折纸结构(parallel helix (PH) origami)^[33],如图 4C、D、E 所示。单向 DNA 折纸结构中的平行螺旋结构只需要设计独特的骨架折叠路径就可以实现,而摆脱了以往折纸结构中所必需的反平行的限制,因此所设计的结构就更为灵活。



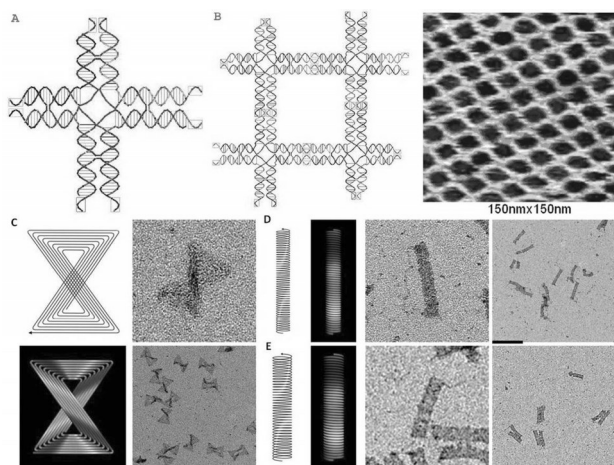


图4 (A) 4×4 DNA 瓦片的 Tiamat 模型^[9]; (B) 4个4×4 DNA 瓦片彼此垂直连接的模型, 以及其构建的二维晶体的 AFM 图^[9]; (C, D, E) 3种平行螺旋 DNA 折纸结构的单向折纸路径的简单示意图、立体灰度图以及 TEM 图^[33]; (C) “蝴蝶结”结构, 灰度图中灰色代表5'端, 黑色代表3'端; (D) 7.4 nm 直径的弹簧结构^[33]; (E) 9.0 nm 直径的弹簧结构^[33]; TEM 图范围为 200 nm×200 nm, 标尺为 200 nm

软件 NanoEngineer-1^[29]可以简化 DNA 双螺旋结构并且可以很容易地在 DNA 双螺旋结构间创建交叉结构。该程序有一个关联数据库可以发布一些模型和模块用于构建全新的结构。简化的 DNA 双螺旋结构允许计算机输出较大的结构, 并且可以在手动构建结构模式中比较快速地计算最小张力^[28]或能量^[13, 29]。使用完整的原子模型可以为计算能量最小化或分子动力学模拟方面提供更为精准的模型, 但是这种方法由于繁重的计算量, 仅适用于计算结构的一小部分, 代替的方法是最低限度的使用小模块作为构建结构的基础^[34, 35]。

Liu 和 Seeman 等利用 NanoEngineer-1 设计了两种正交折纸 tile^[35], 如图 5(a) 中所示的 tile A 和 tile B。正交折纸结构包含两个平面区域, 一个平面在另一个平面之上。由于小的交叉 DNA 模块具有一定的曲率^[9], 在设计时将相同的 tile 平面旋转 90° (实际上 tile A 的一个平面与 tile B 的另一个平面杂交即可, tile A 与 tile B 只是粘性末端不同), 由于每个正交折纸结构中两个 tile 平面具有相反的取向, 这样在构筑二维平面结构时, 一个 tile 中的上层平面会与另一个 tile 中的下层平面相连, 这样就可以有效避免曲率的产生。用所设计

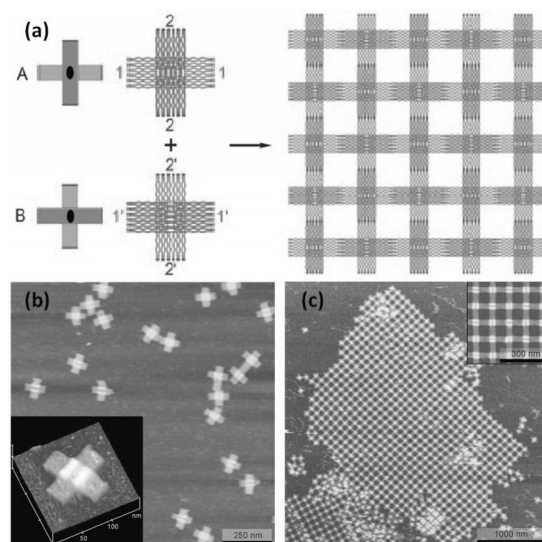


图5 正交折纸 tile 及二维阵列^[35]。(a) 用 NanoEngineer-1 绘制的正交折纸 tile 及二维阵列图; (b) 正交折纸 tile 的 AFM 图; (c) 二维阵列的 AFM 图

的 tile A 和 tile B 构筑折纸结构的二维晶格, 通过 AFM 可以观察到正交折纸 tile 可以组装成较大范围的二维晶格图形, 如图 5(c) 所示。

Andersen 开发了用于设计 DNA 折纸结构的软件, 名为 SARSE^[23, 27]。SARSE 是利用 DNA 折纸技术自动设计较大 DNA 结构的软件。它可以从图形文件中导入一个图形并可以自动将骨架链进行折叠, 而且允许使用者在二维序列编辑器中对自动生成的设计进行修改。该程序可以导出三维模型并可以与三维视图相结合, 如, 可以在构建 DNA 盒子结构中进一步设计三维结点 (如图 3D 所示)^[23]。Andersen 利用 SARSE 设计了可以灵活摆尾的海豚^[30]。该程序使用一个编辑器和一个总体视窗, 这样可以使较大 DNA 折纸结构的编辑和导航更为便捷 (如图 6A 所示)。程序中还有一个历史窗口允许使用者撤销不想要的更改和记录文件, 从而有助于复制复杂的设计。工具箱是用来激活程序, 这些程序可以分析和更改设计或者产生在三维可视窗口中用于检查的原子模型 (如图 6B、C 所示)。SARSE 程序允许在折纸结构设计中插入一条缝隙, 将位图分为两部分, 可以更简单地构建海豚的眼睛, 所设计的海豚位图在背部和尾

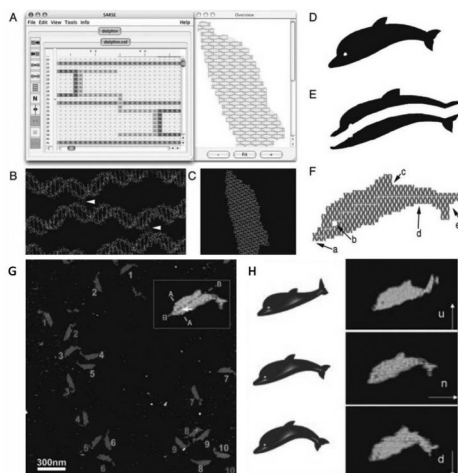


图6 DNA折纸海豚的设计过程^[30]。(A)SARSE编辑器窗口显示了DNA折纸设计的一小部分,骨架链用深灰色表示,锚定链用黑色和灰色表示;(B)在设计过程中产生的原子模型,可在三维结构显示软件,如PyMol中可见^[36],白色箭头指示交叉位置;(C)较低倍数下与B相同的原子模型;(D)输入的原始海豚形状;(E)输入的形状被放大1.125倍,沿形状的中心位置分割,中间插入一条缝隙,修改背鳍及尾鳍;(F)DNA折纸设计的原子模型,箭头和字母指示其特有属性,a:鼻子;b:眼睛;c:90°角的背鳍;d:狭窄的尾部;e:分支的尾鳍,中间的折线表示缝隙的位置;(G)海豚图形的AFM图;(H)海豚的不同尾部构象的示意图(左)及AFM图(右),箭头表示尾部的方向,u:上,n:正常,d,下

鳍上做了改动,这样更有利于简单折叠(如图6E、F所示)。SARSE编辑器可以更改设计并且可以在设计边缘添加T-loops和T-extensions来防止内部分子的螺旋堆积作用^[22]。通过SARSE完成设计之后,让DNA链进行自组装,用高分辨AFM进行观察,可以看到大量的海豚折纸结构(如图6G所示),并且通过在尾部设计交叉结构可以使尾部灵活摆动(如图6H所示)。

构建复杂的三维DNA纳米结构可显著扩充DNA纳米技术的应用范围,但增加了设计的复杂度。Douglas设计了一个快速构建三维DNA折纸结构的软件,名为CaDNAno^[31],可以利用6个螺旋束^[37]或方形格子图形^[38]手动构建多层折纸结构。该程序进一步被用来构建具有球形扭曲或弯曲的折纸结构,这样就可以设计出大量全新的图

形,如锯齿形(如图3E所示)^[24]。CaDNAno具有图形用户界面,可以更好地辅助设计者进行三维DNA纳米结构的设计。设计过程包括4个主要步骤:首先,目标形状被拆分成层状密堆积骨架;其次,指定分配与骨架链互补的锚定链;再次,锚定链被拆分成较短的片段,约含18—49个碱基长度;最后,骨架链用理想模板的DNA序列填充(如:7.2 kb长度的M13噬菌体基因组DNA),从而决定与之互补的锚定链序列。caDNAno将使用者从繁重的寡核苷酸序列设计的工作中解脱出来,极大地缩短了设计时间,从以前的几天或几周到现在的几个小时内就可以完成。caDNAno的程序界面包括切面图、路径图以及渲染窗口。切面图可以显示蜂巢晶格的横断面,在设计中可以添加螺旋。路径图提供了一个编辑界面,可以编辑铺展的二维骨架和锚定链。渲染窗口,提供了一个实

时的三维设计模型。Douglas利用CaDNAno设计了7种不同结构的蜂巢晶格^[31](如图7所示)。由于CaDNAno软件进行设计的基本单元是基于密堆积(蜂巢状或方形)的DNA螺旋结构,因此,所设计的三维结构都是具有层状堆积的结构特点。对于更为复杂的结构设计可能不如Tiamat灵活。

计算机辅助预测和设计三维DNA折纸结构的软件目前还比较稀少,但是这些软件在复杂结构,特别是涉及到弯曲和扭曲结构的设计上具有特殊的价值。Bathe等^[39]设计了一个用于DNA折纸结构的计算机辅助工程的软件CanDo(computer-aided engineering for DNA origami)。该软件能够基于caDNAno设计文件用有限元法计算三维DNA折纸图形,预测三维结构在溶液中的



中国科学院

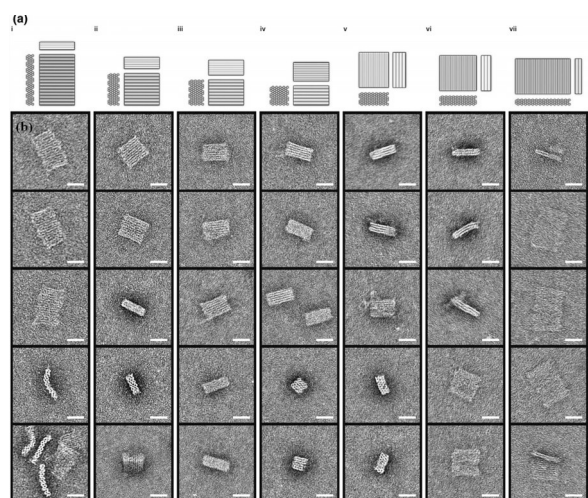


图7 不同结构的蜂巢晶格示意图和TEM图^[31]。(a)圆柱模型示意图,设计以 $m \times n$ 进行命名, m 指密堆积行, n 指密堆积列,(i) 15×4 模块;(ii) 10×6 模块;(iii) 8×8 模块;(iv) 6×10 模块;(v) 4×16 模块;(vi) 3×20 模块;(vii) 2×30 模块;(b)不同模块负染后的TEM图

状态以及稳定性(如图8所示)^[40]。为了计算三维结构,CanDo首先创建了一个初始结构,即在caDNAno源文件中的所有双螺旋在空间中都是线性排列的。这个初始结构与caDNAno软件中3个设计面板中的一个显示的结构相同。CanDo会进一步施加外力使附近的螺旋变形,从而可能使稳定的交叉结构被置于双螺旋结构之间。接下来释放

这些外力会使结构松弛,用非线性有限元法分析当交叉结构强加的连通性与B-DNA默认的几何构型不相容时就会导致变形和内应变。CanDo用市售的有限元分析软件ADINA(automatic、dynamic、incremental linear analysis; Adina R&D)进行数值分析。CanDo可以给出松弛结构的畸变图形并且可以显示热导波动的里氏震级的热点图,可以表明畸变结构的灵活性,由此可以检测目标DNA折纸结构的稳定性。所有的输出都可以用三维视图显示软件进行可视化显示。CanDo可以提供待评估结构的反馈,特别是对于含有大量的弯曲或者扭曲结构的设计来说具有重要意义。交互使用caDNAno和CanDo软件进行设计和分析,可以极大地降低复杂DNA折纸结构的设计风险。目前,CanDo只支持在线分析,可以提供具有热点图的结构图形,并且还会提供三维结构在溶液中震动的模拟视频。

4 DNA编程与生物计算机的构建

在过去的十几年里,人们证明了基于核酸分子的生物化学方法可以成功地用于解决计算机问题。DNA不仅可以用于逻辑操作,也可以用于解决复杂的计算机难题。核酸分子在分子计算上的潜力依赖于核酸分子本身可以高度并行性的碱基配对反应^[41]。核苷以0.35 nm的间隔连续排列,使得DNA具有巨大的信息储存密度($1 \text{ bit}/\text{nm}^3$)^[42],即 $7 \text{ Gbits}/\text{in}^2$ ($1 \text{ in}^2 = 6.4516 \times 10^{-4} \text{ m}^2$)。DNA的另一个重要性质是碱基互补组成的双链。碱基A和T、C和G可以结合起来形成碱基对。核酸逻辑门以及DNA计算机的设计都要最终依赖于这一基本性质。

核酸逻辑门经历了近20年的发展,取得了丰硕的成果。从理论计算到实验模拟,核酸逻辑门的快速发展显示出核酸分子在生物计算机上的巨大优势。构建分子逻辑门是设计分子计算机的先决条件^[43]。设计DNA逻辑门主要依赖于DNA分子的一些特性,如碱基互补、单链置换、DNAzyme切割、核酸适配体结

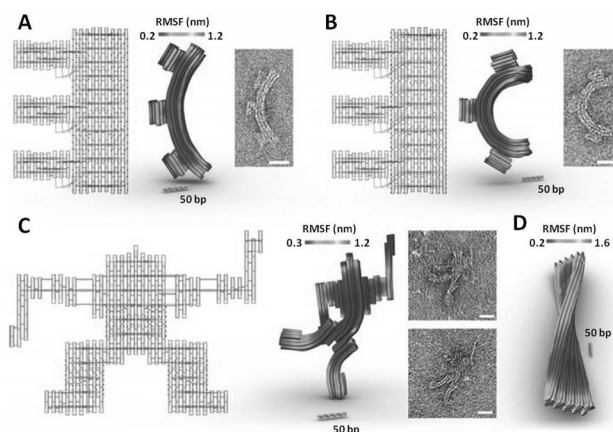


图8 (A-C)左:多层DNA折纸结构的caDNAno设计图;中:CanDo三维结构和结构的灵活性预测,热点图显示均方根波动现象(root-mean-square fluctuations, RMSFs);右:具有代表性的负染TEM图,标尺:20 nm;(D)CanDo三维结构以及caDNAno设计的60个螺旋束组成的四聚物的灵活性预测,结果显示该四聚物并不是以平面结构存在,而是具有一定的扭曲

合、i-motif 结构互变、DNA 镊子以及利用光化学核酸序列特异性切割^[44, 45]、PCR^[46]等。

Ghadiri 等设计了一种利用 DNA 单链置换的逻辑门^[47]。输入的寡核苷酸首先与双链中的短的单链支点区域互补结合,然后通过分支迁移机制不断地置换后面的互补碱基。支点区域允许输入 DNA 比输出 DNA 形成更长的双链。AND 逻辑门:如图 9A 所示,AND-输出链(黑色),荧光素标记的 28-mer 与中间的 AND-固定链互补杂交,AND-固定链是通过 5' 修饰的生物素 55-mer 与亲和素修饰的硅珠连接。固定链的两末端可作为两个输入信号的支点。AND-输入信号 1 与 AND-输入信号 2 各包含 24 个核苷酸。能够与固定链两末端特定的 10 个核苷酸结合。当 10 个核苷酸开始结合时导致随后的 14 个核苷酸逐渐取代 28 个核苷酸的输出链。当只有一个 AND 输入信号加入时,输出链由于只有一半被置换仍然会结合在固定链上。而当同时加入两种 AND 输入信号时,输出链则被完全置换出来,释放到溶液中。监测溶液中的荧光强度就可验证 AND 逻辑门。

OR 逻辑门:如图 9B 所示,OR-输出信号是荧光素标记的 24 个碱基寡核苷酸,与 30-mer 的 OR-固定链互补 14 个碱基。当无输入信号时,OR 逻辑门包含两种不同的单链,一个是输出链,一个是固定链。OR-输入信号 1 包含 24 个碱基,可以与固定链的 10 个碱基支点区结合,随后置换掉输出链的 14 个碱基互补区,相似地,OR-输入信号 2 则可以从固定链的另一端置换掉输出链,两者都可以将荧光标记的输出链释放到溶液中。因此,只要有一种输入信号存在,输出链就会被置换出来。这里溶液中的荧光强度便会增强。

XOR 逻辑门:如图 9C 所示,将上面使

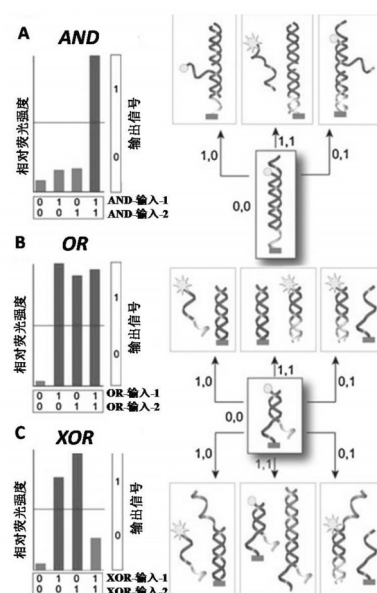


图 9 AND、OR、XOR 逻辑门的示意图。输入链将荧光标记的输出链释放到上清液中,监测的荧光强度增强。如果输出链仍然固定在固体介质中,上清液中的荧光强度则很弱^[47]

用的 OR 输入链的末端连接 20 个碱基得到 44 个碱基的寡核苷酸 XOR-输入信号 1 和 XOR-输入信号 2。这种附加的 20 个碱基在两 XOR 输入信号之间会形成 34 个碱基的互补区。像 OR 逻辑门一样,当加入一种输入信号时,荧光标记的输入信号便会被置换出来。而当同时加入两种输入信号时,两种输入信号由于存在 34 个碱基互补而更易形成互补双链,输出链不会被置换出来。荧光强度测定实验也证实了这一点。

设计逻辑门的主要用途是构建 DNA 计算机。DNA 计算机分为普通计算和减法计算。普通计算通常用杂交、切割和扩增将不正确的序列从体系中去掉。至今,用 DNA 计算方式主要用来解决一些经典的计算机难题^[48, 49],如 Hamiltonian 路径难题^[50]、可满足性问题^[51]、最大集团问题^[52];其他还有设计 DNA 计算^[53],如利用 DNA 自组装进行细胞内计算^[54]、在固体介质表面进行计算^[49]、设计布尔逻辑回路^[55]、Guarnieri 等设计一种



中国科学院

程序将两个非负整数相加^[56], Klein 等证明过逻辑可逆计算^[57]等。

1994 年 Adleman^[50]首次使用 DNA 解决了著名的 NP-complete 计算机难题——旅行商人难题的简化例子。这种问题是假设有 n 个城市(顶点)并有 m 条路(边)连接它们, 是否有这样一条路径: 它开始于一个给定的顶点 V_{in} , 穿过每一个顶点一次, 并最终在顶点 V_{out} 处结束, 这样的路径称为哈密顿路径。对于很小的 n 值, 现在的计算机可以很容易解决, 但随着 n 变得很大, 计算机用于搜索与验证每个可能的答案所用的时间将呈指数级增加, 这种计算对现有的计算机来说是不可行的。Adleman 解决的问题只包含 7 个顶点, 这对于现在的计算机来说不算什么, 但他使用 DNA 来解决这种问题标志着一个新学科的开始。他们的策略是一种简单的暴力破解算法: 产生随机路径, 去掉那些不在 V_{in} 开始并在 V_{out} 结束的路径, 去掉那些不通过 n 个顶点的路径, 再去掉那些顶点不止一次的路径 (如图 10 所示)^[58]。

Adleman 开始时先合成一些随机的 20 个碱基

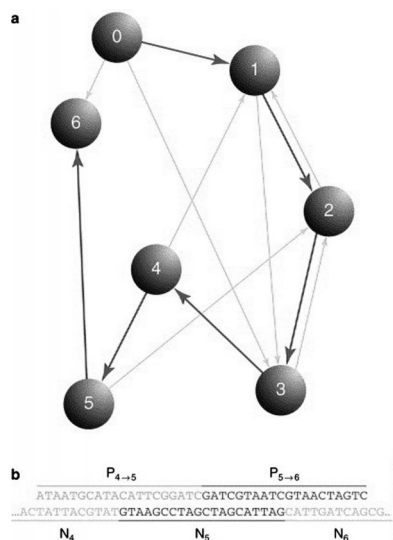


图 10 Hamiltonian 路径难题^[58]。A: 正确的路径用黑线标记, 错误的用灰线标记; B: 20 个碱基序列代表一个节点, 节点之间的路径用分别与前一个节点的后 10 个碱基序列和后一个节点的前 10 个碱基序列代表。这种序列可分别与前、后节点的 10 个碱基序列互补。如路径 $P_{4 \rightarrow 5}$ 可与节点 N_4 和 N_5 互补结合

对的 DNA 寡核苷酸 (20-mers) 来代表每个顶点, 另外再合成一系列 20-mers 来代表边。这些 20-mers 边的序列有特定的规律: 前 10 个碱基与一个顶点的后 10 个碱基互补, 后 10 个碱基与另一个顶点的前 10 个碱基互补。然后, 将混合的 DNA 变性、退火, 并用 T_4 DNA 连接酶连接后, 那些边 DNA 会与合适的顶点 DNA 互补结合, 而 DNA 连接酶则会把所有可能的组合连接起来。这种批量处理方式可以在分子水平大量地平行测试所有可能答案。

这种变性、杂交、连接步骤可以产生超过 10^{13} 个 DNA 片段。其中至少有 1 个哈密顿路径。下一步, 所有在 V_{in} 开始、在 V_{out} 结束的片段可以以相应的引物用 PCR 选择性扩增出来。那些未通过 7 个顶点的路径可以用琼脂糖凝胶在 140 碱基产物中剔除掉。最后, 为了剔除那些通过顶点超过一次的路径, 将上一次筛选的产物用亲和层析纯化将产物变性, 去除生物素标记的互补链。第一个亲和靶物是与第二个顶点互补的 20 个碱基单链, 因此那些包含第二个顶点的 DNA 序列被保留下来。反复重复这一过程直到完全纯化。最后出现的 DNA 条带说明有哈密顿路径存在。验证一条路径是否真实存在, 还需要进行梯度 PCR。

这次实验是 DNA 计算的首次实验验证。该计算需要 7 天的实验工作, 因此对于更大的 n , 这一过程变得不可行。但他首次证明了计算过程可在 DNA 水平完成, 这一开创性的工作引领了更多的后人研究 DNA 逻辑门和 DNA 计算。

受 Adleman 的启发, Lipton 提出 DNA 计算可以解决另一种 NP-complete 问题——可满足性问题 (SAT)^[51]。该类难题中, 需要对任一长度或复杂度的逻辑表达式进行判断, 看它能否满足问题要求。随着变量的线性增加, 需要解决此问题的时间呈指数增加。Lipton 提出用 DNA 计算和分子生物学方法杂交、连接、限制性内切核酸酶切割等操作解决含有两个变量的 SAT 问题。后来 Schmidt 解决了含 4 个变量的 SAT 问题^[59], Hagiya 解决了含 10 个变量的 SAT 问题^[60], Adleman 解决了含有 20

个变量的 SAT 问题^[61]。

不管核酸计算机最终是否可行,核酸逻辑门及计算机对学科交叉已经做出不少贡献,并让人们对于计算原理产生好奇,使计算机学科与生物学之间有了沟通的桥梁。

5 挑战与未来

用于 DNA 结构设计的程序发展很快,但是在核酸研究中仍有许多实际问题亟待改进,迫切需要新的设计方法让核酸成为更加稳定的构建纳米结构的材料。一个重要的进程就是基于热力学参数来进行严格的序列设计,构建诸如 DNA 折纸结构那样较大的分子结构。当然,这同时也需要较大的计算资源。基于热力学参数的序列设计也可能被用来控制自组装路径,以最优规则形成结构域,从而提高自组装的时间和产率。另一方向是在核酸结构中引入化学修饰基团或使用人造核酸,例如,锁核酸 (locked nucleic acids, LNA) 和肽核酸 (peptide nucleic acids, PNA),使组装的结构具有全新的功能。对于 DNA 的结构设计来说,设计方法的最佳标准测试具有最大的价值。基于此,需要制定相应的评估标准和代表数据集,综合比较所有方法,让使用者对于特定的问题可以确定最好的执行算法。

目前,人工设计的单链结构可以很容易地克隆进入细胞,我们已看到了几个可以体内复制的人造 DNA 纳米结构的实例^[62, 63]。为进一步实现该目标,发展一种可以选择序列,特别是适用于体内自组装和复制的设计方法是非常有前途的。人造结构甚至有可能被用来设计进行强制突变并且/或者进行最优进化^[64],这种构想可以通过人造结构的进化仿真来实现,并且有助于预测有害进化的风险。更深层次的进展,是希望核酸结构的设计可以为合成生物学做出相应的贡献。

用于构建 DNA 结构模型的编辑器有许

多不同的用户界面和特征,但是大多数编辑器的算法(包括设计算法)只在网络服务或者可执行文件中才可使用。网络服务和可执行文件的一个主要问题是具有很多种不同类型的输入和输出模式,这就对申请和比较不同的软件造成很多困难。最佳的解决方案是使所有的软件在相同的软件环境中均适用,不论是台式机还是网络服务。将设计工具整合在相同的软件环境中可以有利于分享相同的工具,注释标准以及使用相同的结构模块数据库。

在结构设计领域,生物物理学表征是一个主要的表征手段。随着核酸建模工具的不断发展和需要进行原子级分辨率的实验数据可能变得就不那么重要了。不需要用 X 射线晶体学来分析功能核酸结构的自由度和动力学,也不需要核磁共振来分析较大结构。低分辨率技术在决定核酸的结构和动力学方面有望扮演一个更加重要的角色。原子力显微镜已被证明是研究 DNA 折纸结构的主要手段^[15, 22],冷冻电子显微镜是用来研究三维 DNA 折纸结构的主要技术^[23, 38]。小角 X 射线散射已被用来直接在溶液中研究自组装结构,并且可以非常有效地提供与理论模型进行比较的数据^[23]。所设计的结构可以很容易地用荧光化学基团进行标记,采用荧光共振能量转移技术研究相应的动力学性质^[23]。随着计算机辅助工具和新的实验技术的发展,将会极大地推动我们对核酸结构的理解力和控制力,为 DNA 纳米技术的发展提供前所未有的推动力。

DNA 计算可能是未来的趋势,为化学、计算机科学以及生物学的科技进展铺平道路。但现阶段,它还有几个挑战:(1)使用的材料 DNA 是不可重复利用的。硅电子计算机可以用电作为唯一的输入信号,而 DNA 计算机则需要不断加入原料和清除废物;



中国科学院

(2)DNA 计算的分子组件比较专一。如果要解决 Knight 难题则需要大量不同的寡核苷酸;(3)如何降低错误率。在体内,当细胞内其他分子机器运转正常时,碱基错配的概率是每 10 亿次发生 1 次。但在体外,DNA 计算的反应条件并不如体内,因此错误率明显比体内高。为了超越普通计算机,需要想出办法降低错误率,达到可以接受的水平;(4)最明显的缺点是完成每次计算需要的时间,普通计算机解决旅行商人的 7 个城市问题需要不到 1 秒的时间,而 Adleman 用了 7 天。使用 DNA 芯片或其他方法可能最终自动化,这将节省大量时间,只有比电子计算机更具优势的 DNA 计算技术才能用于实际。目前已经看到 DNA 计算潜在的巨大优势:(1)大量的并行运算。暴力算法可同时在 1 000 万亿个分子中找到一个正确的答案;(2)微型化。当程序正常运行时,原材料花费得很少^[65]。当今计算机最快的运算速度是 10^{12} 次/s,而 Adleman 的分子计算过程运算超过了 10^{14} 次,其实,这一数字可扩大到 10^{20} ;(3)耗能少。每焦耳可支持 2×10^{19} 次运算,根据热力学第二定律理论计算,可达 34×10^{19} ;(4)极高的存储能力。1bit 的信息可以存储在 1 nm^3 的 DNA 中,比现有的存储介质高 10^{12} 倍。因此,分子计算机有潜力成为比现在的计算机更快、更有效的计算机。尽管十几年过去了,DNA 计算仍有待进一步发展,但研究者已设计了几种算法来解决经典的难题,其中一些已经验证且有效。尽管有这些进步,要解决更复杂的问题仍有很长的路要走^[66]。

参考文献

- Seeman N. An overview of structural DNA nanotechnology. *Mol Biotechnol*, 2007, 37(3): 246-257.
- Simmel F C. Towards biomedical applications for nucleic acid nanodevices. *Nanomedicine*, 2007, 2(6): 817-830.
- Gothelf K V, Labeau T H. DNA-programmed assembly of nanostructures. *Organic & biomolecular chemistry*, 2005, 3(22): 4023-4037.
- Liu H, He Y, Ribbe A E et al. Two-dimensional(2D) DNA crystals assembled from two DNA strands. *Biomacromolecules*, 2005, 6(6): 2943-2945.
- Andersen E S. Prediction and design of DNA and RNA structures. *New biotechnology*, 2010, 27(3): 184-193.
- Chen J, Seeman N C. Synthesis from DNA of a molecule with the connectivity of a cube. *Nature*, 1991, 350(6319): 631-633.
- Rothmund P W, Papadakis N, Winfree E. Algorithmic self-assembly of DNA Sierpinski triangles. *PLoS biology*, 2004, 2(12): e424.
- Yan H, Labeau T H, Feng L et al. Directed nucleation assembly of DNA tile complexes for barcode-patterned lattices. *Proceedings of the National Academy of Sciences*, 2003, 100(14): 8103-8108.
- Yan H, Park S H, Finkelstein G et al. DNA-Templated Self-Assembly of Protein Arrays and Highly Conductive Nanowires. *Science*, 2003, 301(5641): 1882-1884.
- Seeman N C. Nucleic acid junctions and lattices. *Journal of Theoretical Biology*, 1982, 99(2): 237-247.
- Seeman N C. De novo design of sequences for nucleic acid structural engineering. *Journal of biomolecular structure and dynamics*, 1990, 8(3): 573-581.
- Williams S, Lund K, Lin C et al. Tiamat: a three-dimensional editing tool for complex DNA structures. *DNA Computing*. Springer. 2009, 5347: 90-101.
- Zhu J, Wei B, Yuan Y et al. UNIQUIMER 3D, a software system for structural DNA nanotechnology design, analysis and evaluation. *Nucleic acids research*, 2009, 37(7): 2164-2175.
- Dirks R M, Lin M, Winfree E et al. Paradigms for computational nucleic acid design. *Nucleic Acids Research*, 2004, 32(4): 1392-1403.
- Winfree E, Liu F, Wenzler L A et al. Design and self-assembly of two-dimensional DNA crystals. *Nature*, 1998, 394(6693): 539-544.
- He Y, Ye T, Su M et al. Hierarchical self-assembly of DNA into symmetric supramolecular polyhedra. *Nature*, 2008, 452(7184): 198-201.
- Zheng J, Birktoft J J, Chen Y et al. From molecular to macroscopic via the rational design of a self-assembled 3D DNA crystal. *Nature*, 2014, 511(7511): 189-192.

- ture, 2009, 461(7260): 74-77.
- 18 Kallenbach N R, Ma R-I, Seeman N C. An immobile nucleic acid junction constructed from oligonucleotides. *Nature*, 1983, 305(5937): 829-831.
- 19 He Y, Chen Y, Liu H et al. Self-Assembly of Hexagonal DNA Two-Dimensional(2D) Arrays. *Journal of the American Chemical Society*, 2005, 127(35): 12202-12203.
- 20 Shih W M, Quispe J D, Joyce G F. A 1.7-kilobase single-stranded DNA that folds into a nanoscale octahedron. *Nature*, 2004, 427(6975): 618-621.
- 21 MFOLD. (<http://mfoldrna.banyed.edu/?q=mfold/download-mfold>)
- 22 Rothmund P W K. Folding DNA to create nanoscale shapes and patterns. *Nature*, 2006, 440(7082): 297-302.
- 23 Andersen E S, Dong M, Nielsen M M et al. Self-assembly of a nanoscale DNA box with a controllable lid. *Nature*, 2009, 459(7243): 73-76.
- 24 Dietz H, Douglas S M, Shih W M. Folding DNA into twisted and curved nanoscale shapes. *Science*, 2009, 325(5941): 725-730.
- 25 Williams B A, Lund K, Liu Y et al. Self-Assembled Peptide Nanoarrays: An Approach to Studying Protein-Protein Interactions. *Angewandte Chemie*, 2007, 119(17): 3111-3114.
- 26 H Gberg B, Olin H. DNA-scaffolded nanoparticle structures; proceedings of the *Journal of Physics: Conference Series*, F, 2007. IOP Publishing.
- 27 Fujibayashi K, Hariadi R, Park S H et al. Toward reliable algorithmic self-assembly of DNA tiles: A fixed-width cellular automaton pattern. *Nano Letters*, 2007, 8(7): 1791-1797.
- 28 Birac J J, Sherman W B, Kopatsch J et al. Architecture with Gideon, a program for design in structural DNA nanotechnology. *Journal of Molecular Graphics and Modelling*, 2006, 25(4): 470-480.
- 29 Sims M. NanoEngineer-1. (<http://www.nanoengineer-1.net/>).
- 30 Andersen E S, Dong M, Nielsen M M et al. DNA origami design of dolphin-shaped structures with flexible tails. *ACS nano*, 2008, 2(6): 1213-1218.
- 31 Douglas S M, Marblestone A H, Teerapittayanon S et al. Rapid prototyping of 3D DNA-origami shapes with caDNA. *Nucleic acids research*, 2009, 37(15): 5001-5006.
- 32 Seeman N C. Interactive design and manipulation of macro-molecular architecture utilizing nucleic acid junctions. *Journal of Molecular Graphics*, 1985, 3(2): 34-39.
- 33 Han D, Jiang S, Samanta A et al. Unidirectional Scaffold-Strand Arrangement in DNA Origami. *Angewandte Chemie International Edition*, 2013, 52(34): 9031-9034.
- 34 Sherman W B, Seeman N C. Design of minimally strained nucleic acid nanotubes. *Biophysical journal*, 2006, 90(12): 4546-4557.
- 35 Liu W, Zhong H, Wang R et al. Crystalline Two-Dimensional DNA-Origami Arrays. *Angewandte Chemie*, 2011, 123(1): 278-281.
- 36 Delano W L. The PyMOL Molecular Graphics System. <http://www.pymol.org/>. 2002.
- 37 Douglas S M, Dietz H, Liedl T et al. Self-assembly of DNA into nanoscale three-dimensional shapes. *Nature*, 2009, 459(7245): 414-418.
- 38 Ke Y, Douglas S M, Liu M et al. Multilayer DNA origami packed on a square lattice. *Journal of the American Chemical Society*, 2009, 131(43): 15903-15908.
- 39 Kim D-N, Kilchherr F, Dietz H et al. Quantitative prediction of 3D solution shape and flexibility of nucleic acid nanostructures. *Nucleic acids research*, 2012, 40(7): 2862-2868.
- 40 Castro C E, Kilchherr F, Kim D-N et al. A primer to scaffolded DNA origami. *Nature methods*, 2011, 8(3): 221-229.
- 41 Ezzi Z. DNA computing: applications and challenges. *Nanotechnology*, 2006, 17(2): R27.



- 42 Chen J, Wang Y, Deaton R. Large scale genomic monitoring or profiling using a DNA-based memory and microarrays. 24th Army Science Conf, 2004,.
- 43 Reif J H. Successes and Challenges. Science, 2002, 296(5567) : 478-479.
- 44 Ogasawara S, Kyoi Y, Fujimoto K. Nonenzymatic Parallel DNA Logic Circuits. ChemBioChem, 2007, 8(13) : 1520-1525.
- 45 Ogasawara S, Ami T, Fujimoto K. Autonomous DNA Computing Machine Based on Photochemical Gate Transition. Journal of the American Chemical Society, 2008, 130(31) : 10050-10051.
- 46 Nojima T, Yamamoto T, Kimura H et al. Polymerase chain reaction-based biochemical logic gate coupled with cell-free transcription-translation of green fluorescent protein as a report gate. Chemical Communications, 2008, (32) : 3771-3773.
- 47 Voelcker N H, Guckian K M, Saghatelian A et al. Sequence-Addressable DNA Logic. Small, 2008, 4(4) : 427-431.
- 48 Chen J, Wood D H. Computation with biomolecules. Proceedings of the National Academy of Sciences, 2000, 97(4) : 1328-1330.
- 49 Liu Q, Wang L, Frutos A G et al. DNA computing on surfaces. Nature, 2000, 403(6766) : 175-179.
- 50 Adleman L M. Molecular computation of solutions to combinatorial problems. Science, 1994, 266(5187) : 1021-1024.
- 51 Lipton R. DNA solution of hard computational problems. Science, 1995, 268(5210) : 542-545.
- 52 Ouyang Q, Kaplan P D, Liu S et al. DNA Solution of the Maximal Clique Problem. Science, 1997, 278(5337) : 446-449.
- 53 Beaver D. Computing with DNA. Journal of Computational Biology, 1995, 2(1) : 1-7.
- 54 Winfree E, Liu F, Wenzler L A et al. Design and self-assembly of two-dimensional DNA crystals. Nature, 1998, 394(6693) : 539-544.
- 55 Boneh D, Dunworth C, Lipton R J et al. On the computational power of DNA. Discrete Applied Mathematics, 1996, 71(1-3) : 79-94.
- 56 Guarnieri F, Fliss M, Bancroft C. Making DNA Add. Science, 1996, 273(5272) : 220-223.
- 57 Klein J P, Leete T H, Rubin H. A biomolecular implementation of logically reversible computation with minimal energy dissipation. Biosystems, 1999, 52(1-3) : 15-23.
- 58 Cox J C, Cohen D S, Ellington A D. The complexities of DNA computation. Trends in Biotechnology, 1999, 17(4) : 151-154.
- 59 Schmidt K A, Henkel C V, Rozenberg G et al. DNA computing using single-molecule hybridization detection. Nucleic Acids Research, 2004, 32(17) : 4962-4968.
- 60 Sakamoto K, Gouzu H, Komiya K et al. Molecular Computation by DNA Hairpin Formation. Science, 2000, 288(5469) : 1223-1226.
- 61 Braich R S, Chelyapov N, Johnson C et al. Solution of a 20-Variable 3-SAT Problem on a DNA Computer. Science, 2002, 296(5567) : 499-502.
- 62 Lin C, Rinker S, Wang X et al. In vivo cloning of artificial DNA nanostructures. Proceedings of the National Academy of Sciences, 2008, 105(46) : 17626-17631.
- 63 Liz, Wei B, Nangreave J et al. A replicable tetrahedral nanostructure self-assembled from a single DNA strand. Journal of the American Chemical Society, 2009, 131(36) : 13093-13098.
- 64 Wagner A. Robustness and evolvability: a paradox resolved. Proceedings of the Royal Society B: Biological Sciences, 2008, 275(1630) : 91-100.
- 65 Benenson Y, Adar R, Paz-Elizur T et al. DNA molecule provides a computing machine with both data and fuel. Proceedings of the National Academy of Sciences, 2003, 100(5) : 2191-2196.
- 66 Ruben A J, Landweber L F. The past, present and future of molecular computing. Nature Reviews Molecular Cell Biology, 2000, 1(1) : 69-72.

DNA Nanotechnology and Bio-Programming

Li Na¹ Ding Baoquan¹ Yan Hao²

(1 National Center for Nanoscience and Technology, Beijing 100190, China

2 Department of Chemistry and Biochemistry, Arizona State University, Tempe AZ 85287-1604, USA)

Abstract DNA nanotechnology is the design and manufacture of artificial nucleic acid structures for technological uses. A central challenge of DNA nanotechnology is to efficiently design and construct well-defined functional structures and devices at the nano scale. Computational tools are the most powerful means of predicting, designing, modeling, and characterizing nucleic acid structure. Computer-aided design of nucleic acid structure and the identification of efficient self-assembly pathways have revealed nucleic acids as unique material for the construction of nano-components. In this review, we have summarized the principles of the DNA nanostructures design, and introduced some algorithms and user-friendly software for nucleic acid design. Meanwhile, the recent progress in DNA logic gates and DNA computing is also included.

Keywords DNA nanotechnology, artificial nucleic acid structure, computer-aided design, bio-programming

李娜 中科院/教育部国家纳米科学中心助理研究员, 博士。主要研究方向包括: 生物纳米材料的设计与制备; 金属及半导体纳米结构的合成及自组装。E-mail: Lin@nanoctr.cn

丁宝全 中科院/教育部国家纳米科学中心研究员, 博导。2000年毕业于吉林大学化学系, 2006年获纽约大学化学系博士学位, 2006—2009年在美国劳伦斯伯克利国家实验室进行博士后研究, 2009—2010年在亚利桑那州立大学作研究助理教授。2010年入选“百人计划”。主要研究方向: 生物纳米材料主要是核酸纳米结构的合理设计与制备, 其与多种其他纳米材料的复合结构的性能研究及可能应用。E-mail: dingbq@nanoctr.cn

颜颢 (Yan H.) 美国亚利桑那州立大学终身教授, Milton D. Glick 杰出教授, 生物设计研究院分子设计与生物模拟研究中心主任。1993年毕业于山东大学化学系, 2001年获纽约大学化学系博士学位, 2001—2004年在杜克大学任教, 2004年至今在亚利桑那州立大学任教。主要研究方向: 生物纳米技术, 纳米电子学、纳米光子学, 大分子结构解析及生物分子成像。发表学术论文近百篇, 其中 *Nature* 和 *Science* 文章多篇。E-mail: Hao.Yan@asu.edu



中国科学院