

DNA 及基于 DNA 链替换反应的分子计算

肖石燕 梁好均

DNA and DNA computation based on toehold-mediated strand-displacement reactions

Xiao Shi-Yan Liang Hao-Jun

引用信息 Citation: *Acta Physica Sinica*, 65, 178106 (2016) DOI: 10.7498/aps.65.178106

在线阅读 View online: <http://dx.doi.org/10.7498/aps.65.178106>

当期内容 View table of contents: <http://wulixb.iphy.ac.cn/CN/Y2016/V65/I17>

您可能感兴趣的其他文章

Articles you may be interested in

基于虚拟开车环境的自闭症儿童脑电样本熵

Sample entropy of electroencephalogram for children with autism based on virtual driving game

物理学报.2016, 65(10): 108701 <http://dx.doi.org/10.7498/aps.65.108701>

基于频率切片小波变换和支持向量机的癫痫脑电信号自动检测

Automatic seizure detection of electroencephalogram signals based on frequency slice wavelet transform and SVM

物理学报.2016, 65(3): 038703 <http://dx.doi.org/10.7498/aps.65.038703>

基于 AdaBoost 算法的癫痫脑电信号识别

Recognition of epilepsy electroencephalography based on AdaBoost algorithm

物理学报.2015, 64(12): 128701 <http://dx.doi.org/10.7498/aps.64.128701>

基于自适应模板法的脑电信号转移熵分析

Transfer entropy analysis of electroencephalogram based on adaptive template method

物理学报.2015, 64(8): 088701 <http://dx.doi.org/10.7498/aps.64.088701>

改进的相对转移熵的癫痫脑电分析

Analysis on relative transfer of entropy based on improved epileptic EEG

物理学报.2014, 63(21): 218701 <http://dx.doi.org/10.7498/aps.63.218701>

专题: 软物质研究进展

DNA 及基于 DNA 链替换反应的分子计算*

肖石燕¹⁾ 梁好均^{1)2)†}

1) (中国科学院软物质化学重点实验室, 能源材料化学协同创新中心, 中国科学技术大学高分子科学与工程系, 合肥 230026)

2) (合肥微尺度物质科学国家实验室, 合肥 230026)

(2016年5月26日收到; 2016年6月15日收到修改稿)

DNA 是生物遗传信息的载体, 同时也是一种理想的生物相容材料. 单链黏性末端 (toehold) 协助的链替换反应是 DNA 常温纳米技术的基础. 利用 DNA 的碱基互补特性和碱基序列的可编程性, 人们可以基于 DNA 链替换反应构建分子机器并对其运转进行精确调控, 实现各种复杂分子计算. 本文回顾了近年来 DNA 结构和力学性质方面的研究进展, 探讨 DNA 链替换反应的微观理解, 介绍 DNA 分子计算领域的最新成果, 以及它在 DNA 恒温自组装等方面的应用.

关键词: DNA 力学性质, 链替换反应, 分子计算, 自组装

PACS: 81.16.Fg, 87.85.-d, 87.10.Tf, 87.14.gk

DOI: 10.7498/aps.65.178106

1 引言

脱氧核糖核酸 (deoxyribonucleic acid, DNA) 是生物遗传信息的载体, 它是由腺嘌呤 (adenine, A)、胸腺嘧啶 (thymine, T)、胞嘧啶 (cytosine, C)、鸟嘌呤 (guanine, G) 四种碱基通过氢键互补配对, 以脱氧核糖 (五碳糖) 为骨架形成的双螺旋生物大分子 (图 1). 这些碱基沿着 DNA 长链所排列而成的序列, 可组成遗传密码, 指导蛋白质的合成. DNA 的骨架由脱氧核糖和磷酸基团通过磷酸二酯键相连构成并排列在外侧, 配对的碱基则通过 $\pi-\pi$ 作用堆叠在分子内部.

目前, 人们对 DNA 结构的认识主要基于 X 射线晶体衍射数据. B-DNA 通常被认为是生理环境中存在的最普遍的 DNA 构型. 标准 B-DNA 的横截面宽度为 2 nm, 一个螺旋周期为 3.4 nm, 包含 10.5 个碱基对; 两条反平行单链形成两个沟槽: 大沟宽度约 2.2 nm, 小沟宽度约 1.2 nm. 与离子等其他分子作用时, 大小沟槽的结合位点和结合强度

存在显著的差异. 与 DNA 作用的特异性蛋白, 可以通过大小沟槽与特定序列的 DNA 片段结合, 从而完成相应的生理功能. 除 B-DNA 外, A-DNA 和 Z-DNA 是另外两种较为常见的构型.

显然, DNA 具有精细的内部结构. 同时, 作为一种生物大分子, 长链 DNA 又是一种高分子. 当 DNA 链长超过 1000 个碱基对 (bp) 时, 人们采用蠕虫链模型 (worm-like chain, WLC) 来描述 DNA^[1]. 正常生理条件下, DNA 持久长度的测量值约为 50 nm, 即 150 bp 的长度. 当 DNA 足够长, 我们可以忽略碱基序列, 而视其为结构均一的大分子; 不过在真实的生理条件中, DNA 生物功能的完成往往与其短链尺度上的结构和柔顺性相关. 这对 DNA 物理性质的描述提出了挑战, 因此, 我们需要谨慎考虑 WLC 模型在这个尺度上是否适用. 研究 DNA 的物理性质, 不仅可以加深我们对 DNA 认识, 同时也为其在生物化学领域中的应用打下基础.

DNA 的碱基互补特性, 使其功能不只局限于作为遗传物质, 同时也是一种理想的生物相容材料. 在 DNA 双链上保留一定长度的单链结合位

* 国家自然科学基金 (批准号: 91427304, 21434007, 21574122, 51573175, 21404098) 和国家重点基础研究发展计划 (批准号: 2012CB821500) 资助资助课题.

† 通信作者. E-mail: hjliang@ustc.edu.cn

点(toehold端,即黏性末端),即可以与具有互补端的DNA双链相结合,从而启动分支迁移反应.基于DNA碱基互补配对特性,Seeman于1982年提出将经特殊设计的DNA序列通过自组装构建纳米器件^[2],为复杂DNA纳米器件的设计和构造提供了思路.经过数十年的努力,人们已经可以构建各种复杂的纳米结构,为DNA纳米器件的进一

步应用打下了基础.与此同时,Mirkin等将具有黏性末端的DNA接枝到金纳米粒子表面^[3,4],并通过连接链(linker)将两种不同的金纳米粒子连接在一起,从而实现DNA修饰的金纳米粒子的聚集反应.最近几年,人们基于这种方法构造了金纳米粒子的三维有序结构,取得可控有序组装的突破^[5].

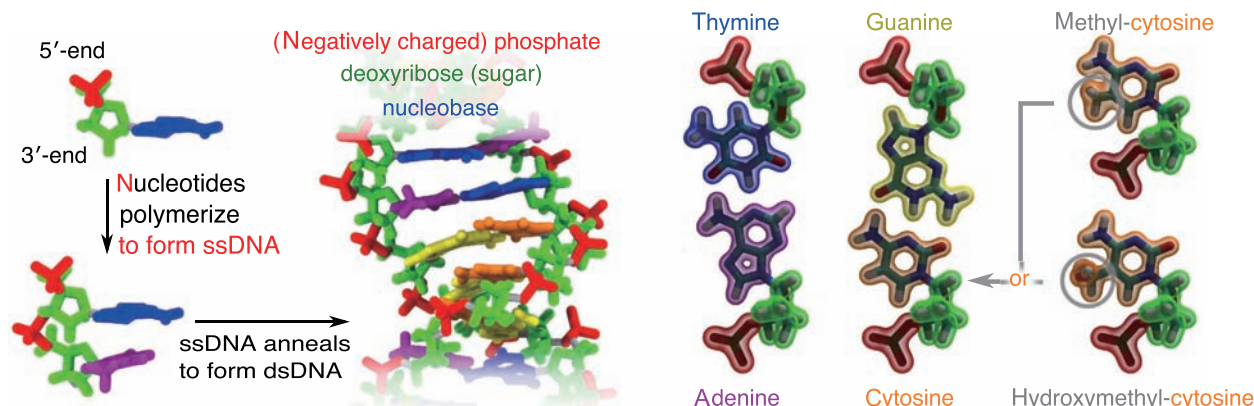


图1 碱基、核苷酸、单链和双链DNA结构图^[6]

Fig. 1. Chemical model of DNA^[6].

早期的DNA纳米器件构造是将长、短链DNA严格退火实现的.在细胞生理环境中,生物大分子是在恒温条件下组装形成纳米器件,完成生物功能,并能针对细胞或者外部环境调整其组装过程和组装结构.鉴于恒温DNA纳米技术在生物医学和生物纳米材料领域中的重要意义,近年来,DNA在特定环境中的恒温自组装受到了广泛的关注.DNA恒温自组装的基础是toehold调控的链替换反应.Toehold在反应中的作用是为DNA链分支迁移反应提供立足点.Toehold协助的链替换反应使人们能够在恒温条件下构建DNA分子机器和复杂DNA反应网络,从而实现特定的计算功能.

在本文中,我们回顾了DNA物理性质、DNA链替换反应以及DNA分子计算等方面的发展历程,介绍了本领域的重要研究方向和主要问题,并探讨未来的可能发展趋势.

2 DNA的结构和力学性质

2.1 DNA的结构和力学性质

长度只有十几至几十个碱基对的DNA,其自身构象的柔顺性和构象转变,以及DNA和蛋白质相互作用是核小体内DNA-组蛋白作用的关键.因

此,DNA自身结构和构象转变受到了科学家们的广泛关注^[6].近年来Olson和Zhurkin^[7]和Orozco等^[8]采用全原子分子动力学(MD)方法对溶液环境中DNA弯曲和构象转变做了大量的研究工作.尤其需要提到的是,北京大学孙育杰、苏晓东和哈佛大学谢晓亮等通过单分子全内反射荧光显微镜实验表明两个不同的DNA结合蛋白可以影响各自对DNA的结合能力,而且其变化随着两个蛋白之间DNA链的长度同时增强或变弱,呈现周期性变化,这个周期大约是10个碱基对,正好是DNA的一个双螺旋周期,并且这种效应的大小会随着两个蛋白之间的距离增加而衰减.他们证明这种效应是由蛋白质结合到DNA上导致的DNA双螺旋的构象变化即DNA的别构效应引起的^[9].Yin等^[10]通过实验结合理论研究了双链DNA错配碱基打开闭合的过程,并发现其打开的时间寿命为10 ms.这些研究工作大大加深了我们对DNA结构和构象转变的认识.

描述DNA长链特征最重要的参量是持久长度.细胞环境中带正电的金属离子,比如 Na^+ , Mg^{2+} , Ca^{2+} 等,会在DNA分子周围形成致密的反离子层.鉴于离子环境对DNA刚性的重要影响,不同离子环境和离子浓度中DNA的持久长度是人

们关注的重要问题. 通过光散射^[11-13]、流动双折射^[14,15]、光镊^[16]等手段, 人们发现当NaCl浓度小于生理浓度(约0.1 M)时, 随着浓度的降低(到约1.0 mM), DNA持久长度从50 nm左右增加到100 nm左右, 各种研究方法测量得到的持久长度和离子强度的关联性都比较一致. 当离子浓度较高, 尤其是当离子浓度高于0.1 M NaCl时, 部分实验数据表明离子浓度的增加对DNA持久长度影响非常微弱^[17-19]; 而另外一些结果^[11-15,17]则表明在大约1.0 M NaCl浓度中, DNA持久长度降低到30 nm左右. 前面的实验结果与Odijk, Skolnick和Fixman(OSF)的理论^[20,21]一致, 即当一价离子浓度高于0.05 M左右时, DNA持久长度与离子浓度关联极为微弱. 而后面的实验数据很好地符合了Manning的理论预测^[22]. 不论在实验和理论上, DNA持久长度在高离子浓度下的分歧暴露出人们对DNA物理性质理解的不足: 静电排斥作用和碱基对之间的 $\pi-\pi$ 堆叠作用是如何决定DNA刚性. 在OSF理论中, 认为DNA骨架的静电排斥作用会因为环境中的离子屏蔽而被减弱, 因此, DNA刚性主要由碱基对间的 $\pi-\pi$ 作用决定, 静电作用在其中的贡献很小. Manning的理论则正好相反, 通过计算中性DNA“null isomer”模型的持久长度, 他提出在生理环境下, $\pi-\pi$ 堆叠作用仅仅贡献DNA整体持久长度的约14%. 人们也尝试从实验上去回答这个问题, 但是很遗憾, 实验结果进一步加大了人们在这个问题上的分歧.

此外, 人们尝试采用多尺度分子动力学模拟(见图2)研究了DNA刚性的来源. Manning的DNA“null isomer”模型为人们理解DNA刚性提供了很好的研究思路, 即通过DNA中性化模型分

离静电作用, 从而有针对性地研究静电排斥作用对DNA结构和力学性质的影响. 2002年, Olson和Manning^[23]通过12个碱基对长度的全原子分子动力学模拟发现, 如果将双链DNA一侧的六个磷酸基团电荷减少, 会导致DNA自发弯向中性一侧(弯曲角度达到 $7^\circ-11^\circ$). 同时发现, 甲基磷酸基团将会显著地影响DNA的大小沟结构. 2008年美国密苏里大学Tan和Chen采用“tightly bound ion”理论研究了离子-DNA作用对DNA弯曲的影响^[24]. 他们发现DNA弯曲能和弯曲取向与金属离子的价态、离子半径和浓度密切相关. 2014年, 中国科学技术大学梁好均课题组^[25]采用全原子动力学模拟, 发现DNA中性化对DNA构象柔顺性的影响具有序列依赖性. 武汉大学谭志杰课题组^[26-28]研究了一价和多价金属离子对单链核酸结构和持久长度的影响, 发现离子价态对链塌缩的显著影响; 并通过全原子分子动力学和Monte Carlo方法研究发现6个碱基长度的短链DNA有显著的弯曲和拉伸柔顺性. 2011年, 来自美国北卡大学教堂山分校的Savelyev和马里兰大学Papouian采用分子重整化群粗粒化方法(MRG-CG)建立了DNA的粗粒化模型^[29], 通过分子动力学模拟计算正常DNA和中性DNA分子的持久长度, 他们声称正常DNA和中性DNA的持久长度相差约17 nm, 即静电作用对持久长度的贡献约为35%^[30]. 这个研究结论与OSF和Manning的理论预测都存在较大的差异. 鉴于粗粒化DNA模型对DNA分子结构细节的忽略, 尤其是对DNA序列差异性描述不够充分^[31], 采用全原子方法在更长的时间尺度和更大的空间尺度上对这个问题进行研究依然非常必要.

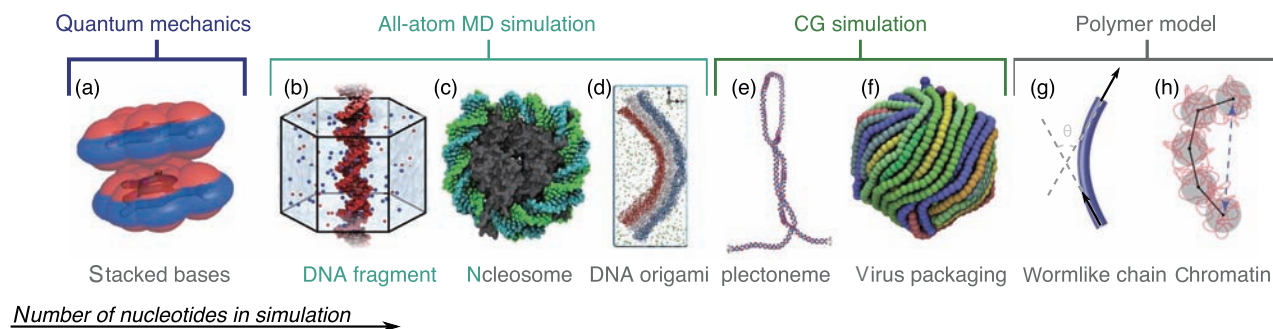


图2 多尺度DNA分子模拟方法^[6]

Fig. 2. Nucleic acid systems that represent the range of scales amenable to various computational and theoretical methods^[6].

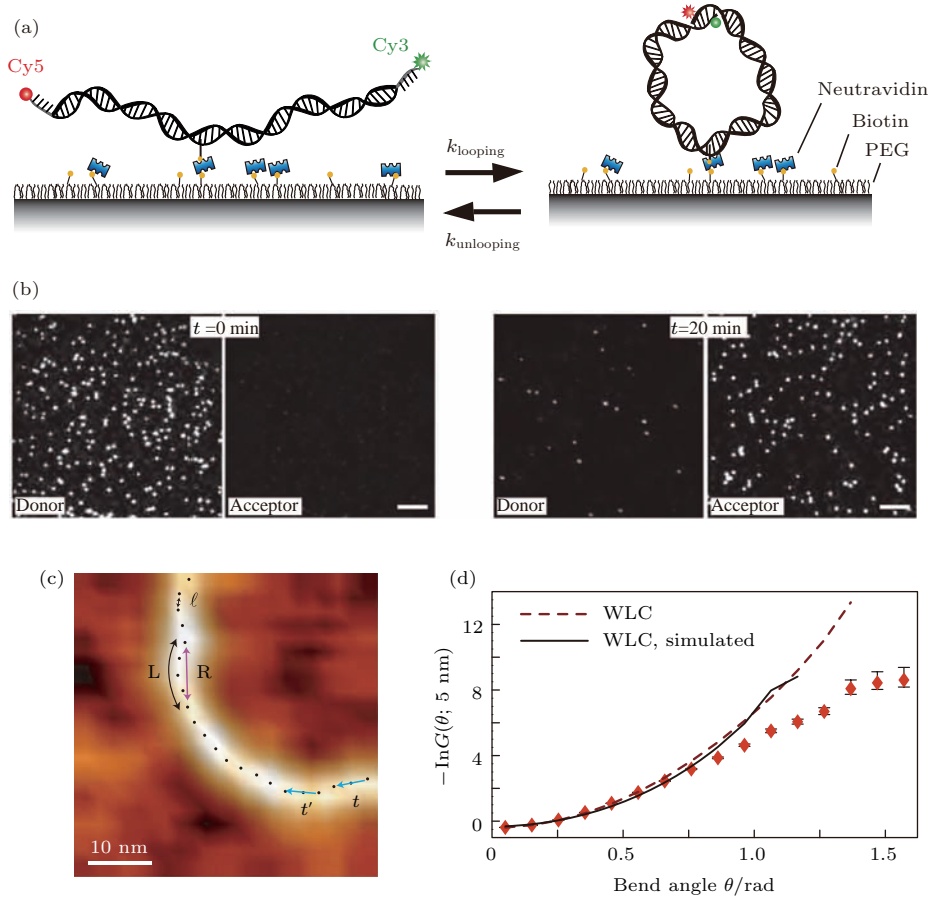


图3 (a) 使用生物素将被能量共振转移供体 (Cy3) 和受体 (Cy5) 分子修饰的 DNA 分子连接在表面上; 实验初始时, 处于缓冲液中的系统没有金属离子存在以保证 DNA 处于打开状态, 即 (b) 左图; 然后加入浓度为 1 M 的 NaCl 溶液, (b) 右图为 20 min 时的供体和受体荧光图; 比例尺为 5 nm; (c) DNA 的原子力显微镜 (AFM) 高清图; (d) DNA 弯曲角度与其概率分布对数取负做图, 红点为实验结果, 虚线为 WLC 理论结果, 黑色实线为 WLC 模型对应的模拟原子力显微镜 (AFM) 结果; (a) 和 (b) 为 Vafabakhsh 和 Ha 的实验结果 [46], (c) 和 (d) 为 Wiggins 等的实验结果 [44]

Fig. 3. (a) DNA molecules labeled using Donor (Cy3) and acceptor (Cy5) were immobilized on the surface via biotin-neutravidin interaction; (b) fluorescence images representing donor and acceptor channels are shown before (left panels) and 20 min after adding high salt (1 M NaCl) buffer (right panels), scale bar, 5 nm; (c) high-resolution AFM images of DNA chain; (d) negative logarithm of the observed probability distribution function. (a) and (b) are the experimental results taken from the studies of Vafabakhs and Ha [46], (c) and (d) are the experimental studies performed by Wiggins *et al.* [44].

2.2 短链 DNA 刚性及其理论模型

从高分子物理的角度, 对由 N 个长度为 l_0 的链段组成的大分子, Kratky-Prorod (KP) 模型 [32] 认为其能量可以用

$$\begin{aligned} E_{\text{KP}} &= \frac{B}{l_0} \sum_{i=1}^{N-1} (1 - \mathbf{t}_i \cdot \mathbf{t}_{i+1}) \\ &= \frac{B}{l_0} \sum_{i=1}^{N-1} (1 - \cos \theta_i) \end{aligned}$$

表示, 其中 $B = l_p k_B T$ 为弹性模量, \mathbf{t}_i 为第 i 个链段的取向, l_p 为持久长度, k_B 为玻尔兹曼常数, T 为

温度, θ_i 为相邻链段夹角. 如果 $l_0 \rightarrow 0$, $N \rightarrow \infty$, 且 DNA 线长度为 $L = Nl_0$, 则有

$$E_{\text{WLC}} = \frac{B}{2} \int_0^L (d\mathbf{t}/ds)^2 ds.$$

WLC 模型虽然没有考虑 DNA 的复杂结构, 却很好地描述了长链 DNA 分子的性质. 但是, WLC 模型能否适用于长度小于 100 bp 的 DNA 链段, 最近引起了人们的关注和争议 [33,34]. 2005 年前后, Cloutier 和 Widom [35,36] 在长约 100 bp 的双链 DNA 片段成环反应实验研究中发现, 其首尾相连成环的速率是 WLC 模型预测值的 3—5 倍. 为了解释这个出乎意料的现象, 人们开始寻找可能的理论

和实验证据. 人们首先研究DNA 纽结(kink) 结构在其中的可能作用. 早在1975年, Crick和Klug^[37]就提出, 短链DNA成环时, DNA的弯曲并不是均匀的, 而可能发生kink从而大大降低成环能; 而后在核小体的晶体结构中也确认了kink结构的存在^[38-40]. 因此人们尝试在理论上引入kink从而解释短链DNA的柔顺性. 他们建议, 如果在DNA链段碱基堆叠强度较弱的部分, 碱基对在保持配对的情况下以一定的可能性发生kink^[41], 局部单链(bubble)^[42]或者碱基对发生破缺而打开(opening)^[43], 从而导致短链DNA表现出更强的柔性. 在WLC模型中, DNA链段之间的弹性能是基于弹簧势, 即

$$E_{\text{WLC}} = \frac{1}{2}k_{\text{B}}T(l_{\text{P}}/l)\theta_i^2.$$

Wiggins等^[44]提出更柔的势能函数, 即 $E_{\text{LSEC}}(\theta) = \alpha|\theta|k_{\text{B}}T$, 来描述DNA链段之间的弯曲, 并建立了LSEC模型. 其中 α 为调控参数, 如果链段长度设为 $l = 2.5 \text{ nm}$, 则 $\alpha = 6.8$. 2014年, Cao等^[45]将WLC模型推广至C-WLC模型以期望描述DNA在短链尺度上的柔性.

WLC模型是否适用于短链DNA的争论依然在继续. Wiggins等^[44]采用AFM研究短链DNA弯曲的实验数据表明, DNA自发大角度弯曲远超WLC的预测(图3(b)). 2012年, Vafabakhsh和Ha^[46]将长度小于100 bp的DNA短链中部固定在芯片表面上, 然后使用单分子荧光共振能量转移(smFRET)研究其成环的速率. 研究结果再一次证明了DNA在短链尺度上不同寻常的柔性^[46](图3(a)). 同时, 也有学者尝试通过实验证明WLC模型在描述短链DNA柔顺性的可行性. Du等^[47]尝试从技术角度否认Cloutier和Widom的短链DNA成环实验结论. Mastroianni等^[48]通过大量的X-射线小角散射实验以期望证明WLC模型在短链尺度上同样适用, 并宣称其结果证明WLC模型可以描述链长为 $1/3l_{\text{p}}$ 到 $2/3l_{\text{p}}$ 的DNA的构象涨落. 2007年, Mazur^[49]采用MD模拟计算了短链DNA自发弯曲角度分布, 宣称短链DNA的模拟数据是完全支持WLC模型的, 而不支持Wiggins等关于短链DNA的AFM实验数据^[44]以及为解释AFM数据提出的SEC模型^[41]. 但作者同时也表示, DNA分子力场有可能导致数据不够准确^[49]. 2014年, Mazur^[50]采用AFM测量了长

度短至1个螺旋周期的DNA弯曲, AFM数据表明WLC模型完全适用于长度大于3个螺旋周期(即10.5 nm)的DNA.

3 DNA链替换反应

3.1 DNA链替换反应的基本原理

单链DNA通过碱基互补配对形成双链DNA分子. 利用DNA链间这种特殊的相互作用, DNA在生物体内能够有效进行遗传信息的储存和复制^[51]. 同时作为一种生物相容材料, 由于其精确的碱基互补配对、可预测和热可逆碱基相互作用、相对简单的化学合成和修饰以及序列可编程等特性, DNA分子被广泛地用于DNA纳米结构和纳米器件的构建、生物传感、药物输送等.

基于DNA碱基互补配对的特异性, Seeman^[2]于1982年提出将经特定设计的DNA序列通过自组装构建纳米器件. 基于这种想法, 人们通过合理设计, 将长短链DNA严格退火, 从而构建一维、二维^[52,53]和三维复杂纳米结构和纳米器件^[54-58].

在细胞生理环境中, 生物大分子是在恒温条件下组装形成纳米器件, 并能针对细胞或者外部环境调整其组装过程和组装结构. 鉴于恒温DNA纳米技术在生物医学和生物纳米材料领域中的重要意义, 近年来, DNA在特定环境中的恒温自组装受到了广泛的关注^[59,60].

DNA恒温自组装的基础是toehold调控的链替换反应^[61]. toehold调控的链替换反应的具体过程如图4.

图4(a)是双链DNA图示方法; 图4(b)是toehold协助的DNA分支迁移和链替换过程图. Toehold协助的链替换反应类似于化学反应中的置换反应, 反应物为具有toehold黏性末端 t^* 的 t^*-x^* 链与链 x 组成的双链复合物complex-1和单链 $t-x$. 具体反应过程为: 单链 $t-x$ 作为侵入链(invading strand), 其toehold黏性末端 t 与complex-1的 t^* 结合; 然后, 启动侵入链的 x 链段分支迁移反应逐步替换complex-1的结合链 x ; 最终的生成物为 $t-x$ 与 t^*-x^* 结合形成的新复合结构complex-2以及单链 x . Toehold在反应中的作用是为DNA链分支迁移反应提供立足点从而促进反应顺利进行. 因此, toehold端的长度在一定程度上决定了DNA链替换反应速率. 随着toehold长度增

3.2 DNA链替换反应的微观理解和速率调控

DNA链替换反应是DNA常温组装技术的基础,并已经被大量应用于DNA纳米器件的构建^[61].但是,目前人们对DNA链替换反应的微观物理机理缺乏了解,并难以精确调控其反应速率^[63].DNA链替换反应速率的精确调控是复杂反应网络设计和广泛应用的基础和关键^[64].DNA的结构和功能与溶液环境密切相关.在考虑温度效应下,溶剂-DNA相互作用对DNA分子结构和杂交的影响,以及溶剂分子在DNA结合和链替换过程中的参与机理尚未得到深入研究.通常,人们认为DNA链替换反应的分支迁移过程是一个随机的无规行走.研究表明,DNA链替换反应的分支迁移过程的每一步行走都要跨越一定的能垒 ΔG_s ,并具有序列相关性^[63];而不是通常人们所设想的随机无规行走.但到目前为止,在理论和实验上都没有获得精确的 ΔG_s 数值.

在反应速率调控方面,牛津大学Turberfield等^[65]在toehold和分支迁移链段之间引入空白链段(spacer)以及在分支迁移链段引入错配^[66]的方式来调节DNA链替换反应速率.加拿大滑铁卢大学Liu等^[67,68]通过在溶液中加入醇类等有机小分子来调节DNA beacon的杂交速率.他们发现,有机小分子使DNA熔融温度 T_m 下降,同时加快DNA杂交过程.华中科技大学夏帆等发现醇类可以促进DNA的链替换反应^[69].当把DNA接枝到芯片表面时,可以通过控制芯片表面DNA的接枝密度、表面曲率、调节DNA自身链长度、溶液pH值和离子浓度等方式调控DNA链替换反应速率^[70].另一方面,科学家们尝试通过理论和模拟的手段研究DNA链替换反应的微观机理.Zhang和Winfree^[62]以及Yurke和Mills^[71]对DNA链替换反应过程速率的早期讨论并未涉及反应的微观机理,Srinivas和Winfree于2013年提出了IEL模型(intuitive energy landscape model),以期望从最基本的DNA结合能数据计算得到DNA链替换反应速率^[63].DNA链替换反应速率的理论计算依赖于人们对DNA结合能的准确计算.目前,人们最常用的DNA结构预测和结合能计算软件是NUPACK^[72].除此之外,SantaLucia等在2000年和2004年针对不同相邻序列结合能、发卡结构、错配以及环

状(loop)DNA结构等情况提出了完整的热力学参数^[73,74].Pyshyi和Ivanova^[75]以及Protozanova等^[76]分别对碱基堆叠作用能和双链DNA中间单链断裂位点(nick site)提出了自己的解决方案.

DNA粗粒化模型的发展^[77]也为人们研究DNA杂化过程的微观机理提供了可能.2009年,de Pablo等采用3SPN模型^[78,79]研究DNA杂化过程发现,双链的杂化过程与DNA碱基序列相关^[80].对重复序列DNA链,只要有连续四个碱基发生配对,就可以通过“slithering”机理完成杂化;对随机序列DNA链,杂化过程遵循更为严格的机理,两条互补单链要么完全配对,要么以单链存在,即两条单链DNA某片段碰撞结合从而使整个DNA链段迅速完成杂化.采用相同的模型,天津大学何学浩课题组研究了受限条件下双链DNA的螺旋解开和恢复的过程,他们发现受限越强,熔融温度越高^[81].2013年,Ouldrige等采用oxDNA^[82,83]讨论了DNA杂化动力学过程^[84].他们发现,两条单链DNA的部分互补碱基随机结合配对之后,通过“拉链”(zippering)方式完成其他部分杂化.杂化反应的启动需要2—3个碱基完成互补配对以抵抗热力学扰动.模拟结果揭示,如果初始结合位点不在双链完全结合时的对应位置,DNA将通过“pseudoknot”和“inchworm” internal displacement两种方式最终实现双链完全互补.模拟结果同时解释了DNA杂化速率的序列差异.基于oxDNA模型,人们研究了DNA发卡的形成^[85],DNA镊子^[83],四臂DNA的形成^[86]以及其他DNA分子机器的运行机理^[87].

4 DNA分子计算和DNA恒温自组装

4.1 DNA计算和复杂反应网络的设计和实现

DNA具有绝大多数生物分子所不具备的碱基互补特性和序列可设计性,使其可以被用来设计各种复杂的分子机器和反应网络以实现特定的功能,比如信号放大和DNA计算等,从而为“智能”药物设计提供一种可能的解决方案.DNA计算最早由美国南加州大学Adleman教授提出,他于1994年利用DNA计算方法解决了著名的数学难题——“七顶点哈密尔顿路径”^[88].借助于限制性内切和连接酶^[89,90],或者DNA发卡结构^[91-93]等,

人们成功构建了类似于电子计算机的逻辑运算元件. 2000年, Yurke等^[61]提出了toehold引导的DNA链替换反应. 他们采用三条链, 构建了DNA镊子, 利用链替换反应并在燃料链的驱动下, 使镊子“闭合”和“打开”, 成功实现了DNA分子机器的运行. 加州理工学院的Zhang等^[94]基于toehold exchange反应建立了第一个DNA催化反应体系. 2006年, Seelig等^[95]基于链替换反应构建了OR, AND和NOT逻辑门, 并实现了信号放大和反馈, Seelig等的工作表明了采用DNA构建大型逻辑运算模块的可能性. 2007年, Zhang等^[94]基于链替换反应构建了分子机器反应网络. 2011年, Qian和Winfree^[96]采用可逆链替换反应设计了一种新的逻辑门, 可以压制信号从而过滤噪音. 在此基础上, 他们采用130条DNA链构建了极为复杂的逻辑电

路并实现了四位二进制数的开方运算. 同年, 他们利用DNA实现了神经网络计算^[97]. 该研究组还利用DNA反应设计了模拟生物化学体系的转录振荡系统^[98]; 并在理论上论证了DNA模拟复杂化学反应网络的可能性^[64]. 2015年, 湖南大学和美国佛罗里达大学谭蔚泓课题组借助于酶和DNA分子机器模拟了生物体适应性免疫反应过程^[99].

DNA序列的可编程性使人们可以利用DNA链替换反应设计基元反应, 并作为程序语句用于构建复杂反应网络(类似于“程序”)实现特定的功能. 2010年, Soloveichik等^[64]从理论上论证了利用DNA构建化学基元反应的可能性(图6(b)). 2013年, Chen等^[100]基于DNA链替换反应, 设计了 $A + B \rightarrow C$, $A + B \rightarrow C + B$, $A + B \rightarrow C + 2B$ 等基元反应, 并在此基础上设计并实现了由三个基

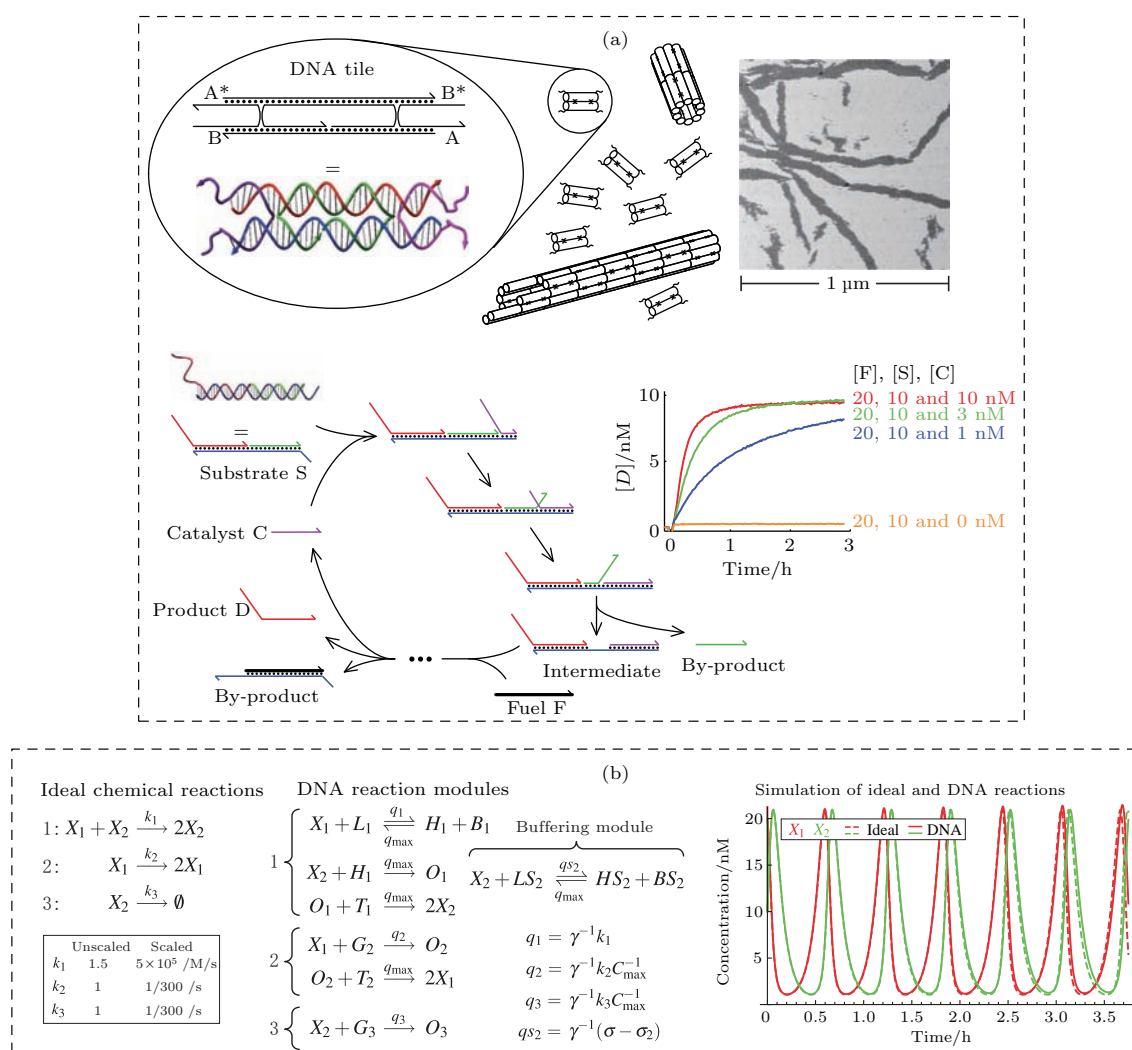


图6 (a) DNA tile组装原理图^[60], 在这个体系中, 通过循环放大反应释放大产物链D, 参与体系组装; (b) 采用DNA实现 Lotka-Volterra 化学震荡的理论模型^[64]

Fig. 6. (a) Structural and dynamic DNA nanotechnology^[60]; (b) Lotka-Volterra chemical oscillator designed using DNA strand displacement reactions^[64].

元反应构成的反应串联. 目前, 虽然 Winfree 等^[64]从理论上论证了采用 DNA 链替换反应构建 Lotka-Volterra 等化学振荡反应的可行性, 但是最终在实验上实现依然存在极大的挑战. 另一方面, 人们也尝试采用计算机模拟的方法来推动和协助 DNA 反应网络的设计. 比如 Phillips 发布了 DSD 程序包, 可以通过求解 ODE 方程的方法和蒙特卡罗方法模拟基于 DNA 设计的电路逻辑门, 以及其他多种 DNA 复杂反应^[101,102]. DNA 反应网络的设计和成功构建依赖于 DNA 链替换反应速率的设置. 但是, 人们目前无法直接从碱基对的配对和堆叠作用等最基本热力学参数推导得到 DNA 链替换反应速率. 因此, DNA 反应网络的理论设计一直是受到人们的关注并亟需得到解决的重要问题.

基于 DNA 链替换反应的一个重要问题是泄露 (Leak). 由于 DNA 纯度和常温下的热力学扰动等原因, DNA 复合双链可能在没有催化链存在的条件下即与环境中的互补链直接发生替换反应, 取代初始配对链从而引发体系反应. 尤其在多级催化反应体系中, Leak 反应会对体系的稳定性和系统的运行产生较大影响^[103]. 2014 年, 美国德克萨斯大学奥斯汀分校的 Jiang 等^[104]通过在 DNA 发卡结构的 Loop 部分引入错配的方案以抑制 Leak 反应, 在两级发卡催化反应体系中取得了相比于未引入错配的同体系高达 100 倍的信噪比. 这些工作为在多级催化体系中抑制 Leak 反应做了较好的尝试. 但是, Leak 反应依然是多级 DNA 反应网络中的巨大困扰.

4.2 DNA 恒温可控自组装

随着人们对 DNA 杂交热力学性质了解的加深^[73], 精确设计和预测 DNA 器件已经逐渐成为可能. DNA 合成和提纯技术的成熟, 也促进了 DNA 作为一种生物相容材料在纳米器件构建等方面的应用^[105]. 粗略地看, 基于 DNA 互补配对方法的 DNA 纳米技术可以分为 DNA 折纸和 DNA 恒温组装技术. 相比于通过 DNA 折纸获得稳定的平衡态结构, 基于 DNA 链替换反应的恒温可控组装技术则更侧重于系统的非平衡态性质.

2000 年, 英国牛津大学 Yurke 等^[61]采用 DNA 构建了分子镊子, 并第一次采用 toehold 协助的 DNA 链替换反应控制镊子的闭合和打开. 不同于以往人们主要通过改变环境变量 (比如盐浓度、

pH 值、温度等) 来改变和控制分子和分子器件的结构形态, Yurke 等的工作表明 DNA 链替换反应是控制分子器件运行的另外一种可能的有效途径. 2001 年, Simmel 和 Yurke 构建了类似的“纳米制动器”(nanoactuator)^[106]. 2002 年, Simmel 和 Yurke 构建了具有三个不同构象的纳米器件, 通过一次加入两对燃料链使纳米器件^[107]在三个构象之间依次转换. 2004 年, 美国普渡大学 Tian 和 Mao^[108]设计了由两个圆环组成的 DNA 齿轮, 通过加入 DNA 来驱动齿轮转动, 并能控制齿轮的转动方向. 当然, 链替换反应也可以用来调整甚至重构大尺寸的 DNA 纳米结构. 这方面的第一个例子是 Yan 等^[109]在 2002 年完成的工作. 他们构造了一个由四个小模块连接组成并可以用 AFM 直接观察的大尺寸纳米器件, 通过链替换反应, 可以使其中的模块旋转 180°, 从而在 PX 和 JX₂ 两个形态之间转变. 基于类似的设计思想, Chakraborty 等^[110], Zhong 和 Seeman^[111], Ding 和 Seeman^[112], Feng 等^[113], Lubrich 等^[114]和 Goodman 等^[115]用链替换反应实现了更为复杂的纳米器件结构调整甚至组装. Dirks 等^[116]基于 toehold 协助的链替换反应实现了 DNA 单链引发发卡 DNA 链的交替聚合形成数千个碱基长度的 DNA 双链, 随后 Yin 等^[93]更将其扩展到 DNA 组装路径和结构形态的控制. 清华大学刘冬生课题组^[117,118]利用 DNA 亚稳态构造隐式 toehold 控制分子反应活性. Ricci 等^[119,120]设计了基于 pH 值变化的 DNA 链替换反应调控方法. 梁好均课题组^[121]基于 pH 响应的 DNA 三链结构设计了 pH 控制的分子机器驱动纳米粒子聚集方案. 尤其需要提到的是, 清华大学刘冬生课题组在 pH 响应的 DNA 组装方向做出了大量重要的工作^[122,123]. 另一个有趣的应用是构建 DNA walker, 即利用链替换反应使 DNA 模块在已设计好的路线行走. 2004 年, 美国纽约大学 Seeman 课题组^[124]以及加州理工学院 Pierce 课题组^[125]构建了一维 DNA walker. 2010 年, Gu 等^[126]设计了一种精妙的 DNA walker, 它可以沿着设定的路线行走, 并可以抓起、运送并放下金纳米颗粒. 2012 年, 湖南大学和美国佛罗里达大学 Tan 等^[127]设计了一种光响应驱动的 DNA 行走系统, 并通过 AFM 直接观察到了 DNA 纳米器件上的行走过程^[128].

DNA 链替换反应也可以用来控制 DNA 自组

装路径, 从而提供一种相比于退火方法更加自由、精确的组装手段. 2004年, Dirks和Pierce^[116]引入了DNA杂交链增长反应(HCR)的概念, 通过加入引发链把DNA发卡打开, 形成DNA双链复合结构, 同时暴露出下一级发卡的黏性末端, 由此引发下一级链替换反应, 从而将两个发卡连接起来, 组成一个新的双链复合结构. 2007年, Pierce等^[129]把一系列链发卡通过HCR反应连接起来形成直链高分子. 2008年, Turberfield等^[130]通过链替换反应将双链DNA单体连接构造了直线链式结构. Yin等^[93]更进一步利用HCR方法构建了树枝型的分子. 香港科技大学Xuan和Hsing^[131]通过非线性HCR反应, 设计了一种树枝型DNA纳米结构. Zhang等Winfree实现了DNA催化体系的动态链构象调整^[132]并在恒温条件下利用DNA tile组装形成纳米管^[60] (图6(a)). 中国科学技术大学梁好均课题组将DNA toehold exchange链替换反应引入DNA-金纳米粒子聚集反应, 在一定程度上^[131]实现了金纳米粒子的恒温可控自组装, 有力支持了“催组装”理论^[133]; 并基于此方法实现了对DNA短链单碱基突变的检测^[134,135]. 青岛大学Bi等^[136]基于HCR实现了超支化DNA纳米结构, 构建了串联逻辑门, 实现了对三种DNA目标链的有效监测.

在DNA生物传感器的构建及其生物医学应用上, 近年来也有不少研究工作. 2013年, 中国科学技术大学梁好均课题组和美国佛罗里达大学谭蔚泓课题组合作, 针对难以获得高纯度的具有toehold的双链DNA结构的问题, 提出了一种新型光触发可控toehold形成的方法. 该方法利用可光切断的2-硝基苄基将两条互补的DNA链连接起来, 该长链DNA通过碱基互补形成发卡前驱体, toehold则被隐藏在发卡结构的Loop部分. 经紫外光照射将Loop释放形成具有活性toehold的纯的1:1的双链DNA结构^[137]. 该方法已经被进一步应用于多色及可擦写DNA光刻研究^[138]和构建功能化的DNA聚合物分子刷图案^[139]. 他们利用紫外光并结合光刻掩模选择性地切断表面DNA发卡结构并产生两个黏性末端, 该DNA黏性末端能够引发后续的DNA杂交链增长反应, 从而在表面获得图案化的DNA聚合物刷. 更重要的是, 利用该方法可以很容易地在图案化的DNA聚合物刷中引入功能单元并获得功能化的DNA聚合物刷

表面^[139].

4.3 细胞环境下的DNA纳米器件设计和反应网络设计及其调控

DNA链替换反应的精确可控, 使DNA纳米技术应用于细胞水平上的“智能”诊断和治疗以及“精准”药物释放等方面提供了可能性^[140-142]. DNA纳米器件体积小, 可控性强, 尤其DNA的高度生物相容性, 大大降低甚至避免了一般药物载体的高毒性. 同时, 核苷酸特有的互补配对结构可以使人工设计的纳米器件和核酸序列与细胞生理过程(比如与mRNA作用)相结合, 甚至直接介入细胞内的反应网络, 从而使人们可以重新设计和调控细胞生理过程. 当然, 细胞内生理过程是一个异常复杂的系统, 对其反应网络的介入和调控是一个巨大的挑战. 要实现这个目标, 不仅需要人们在实验技术和方法上的进步, 同时也需要在细胞调控理论上的密切配合.

在过去的十多年中, 科学家们为将DNA纳米技术导入细胞做了许多努力, 从最初的细胞外(cell-free)体系, 到细胞表面, 甚至将DNA纳米器件直接导入细胞进行相关的应用. Meldrum等^[143]研究了DNA折纸产物在细胞提取液中的稳定性. 他们观察到, DNA放入细胞提取液12h之后还能够被重新提取, 而单链和双链则很快被降解. 这可能是因为DNA纳米结构是将DNA在包含离子的缓冲液中通过折纸形成了更为稳定的形态结构. Conway等^[144]的实验研究也得到了相似的结果. 当然, DNA纳米器件在细胞环境中的稳定性也取决于其自身设计的差异^[145]以及细胞环境^[146]. 鉴于不同种类细胞和细胞器之间的巨大差异, DNA纳米器件在细胞环境中的稳定性需要更为系统和详尽的研究. 利用细胞表面的特定蛋白, 人们设计携带有DNA双链复合物的适配体与之结合. 通过加入特定DNA链段引发链替换反应, 就可以在细胞表面进行“计算”, 从而对细胞类型或者细胞状态进行鉴定^[147,148]. 近年来, Tan等^[149,150], Gartner和Bertozzi^[151]以及Francis等^[152]在将DNA链替换反应应用于细胞表面计算的方向上进行了大量的研究.

同时, 科学家也尝试将DNA纳米技术应用于细胞内环境监测、细胞内物质检测和细胞内分子计算. 2009年, Modi等^[153]利用富胞嘧啶DNA链

在酸性和碱性溶液中的构象差异, 将其导入细胞, 成功实现了对活性细胞内 pH 的监测. 2013 年, 同一课题组成功在一个细胞内运行两个不同的 DNA 纳米机器, 对细胞两种内吞路径的 pH 变化进行了实时监测等^[154]. Tsourkas 等^[155], Tan 等^[156] 利用分子信标 (molecular beacon)^[157] 对细胞内 mRNA 进行监测. 通常, DNA 链通过互补配对结合使两端的荧光基团猝灭并在中间形成 Loop. 在细胞环境中, 细胞内 mRNA 会与 Loop 或者预设的 toehold 段结合启动分支迁移从而发出荧光. 2015 年, Tan 课题组^[158] 通过设计 DNA 发卡, 并利用 mRNA 作为催化链, 在细胞内实现了 mRNA 检测信号的放大. Mirkin 等^[159] 将经过荧光基团修饰的 DNA 链与金纳米粒子表面的单链 DNA 结合, 在将其导入细胞后, 细胞内的 mRNA 通过结合预设的 toehold 端, 启动链置换反应并释放荧光信号, 从而实现了对细胞内 mRNA 的监测. 经改进, Mirkin 的方法已经被商业化^[160,161]. 在细胞内分子计算方面, 人们也进行了一些有意义的探索^[162,163].

5 结论和展望

过去数十年里, 人们对 DNA 自身的理化性质以及在生物纳米材料领域的应用做了大量的探索. 基于前面的讨论, 我们可以认识到, 人类对 DNA 的认识依然需要进一步深入. 我们认为 DNA 分子计算相关研究领域的未来发展趋势将体现在以下几个方面.

首先, 计算机模拟作为一种新兴的研究手段, 将在 DNA 相关研究中扮演更加重要的角色. 鉴于经典分子力场在描述键、键角和二面角作用采用的是弹簧势或其他类似作用势, 因此, 我们在使用经典分子力场描述 DNA 构象转变、RNA 结构等方面时需要非常注意. 针对这一问题, Šponer 等^[164,165] 进行了许多讨论. 另外, 极化力场的发展刚起步, 尚未对极化力场和非极化力场能否很好地描述 DNA 结构及其转变这个问题进行系统性的研究. 在这方面, 我们还有许多工作需要去完善. 但是随着 DNA 全原子力场和粗粒化力场的不断发展和完善, DNA 的计算机模拟将更加准确, 模拟尺度范围更大, 可研究的问题范围更广, 为实验之外研究 DNA 的结构、功能及其动力学性质提供了一条可行的途径. 随着计算机性能的提升和大规模并行计算的发

展, 分子动力学模拟方法将在 DNA 生物功能探究和 DNA 材料研究等领域发挥更大的作用.

其次, 光镊等实验方法的进一步成熟, 将会被更多地应用于 DNA 的结构性质、DNA 包装等课题的研究. 人们将可以对 DNA 的力学特性、包装动力学过程进行更直接的观测.

再次, DNA 分子计算将更多地应用于 DNA 纳米器件的组装、细胞鉴定和识别以及细胞检测. 基于 DNA 分子计算的 DNA 常温纳米技术将进一步成熟并被应用于药物载体设计、药物释放、基因治疗等领域. 在此基础上, 将 DNA 分子机器导入细胞并对细胞环境下的 DNA 反应网络进行调控, 从而实现在细胞层次上的疾病诊断及治疗.

参考文献

- [1] Bustamante C, Bryant Z, Smith S B 2003 *Nature* **421** 423
- [2] Seeman N C 1982 *J. Theor. Biol.* **99** 237
- [3] Alivisatos A P, Johnsson K P, Peng X, Wilson T E, Loweth C J, Bruchez M P, Schultz P G 1996 *Nature* **382** 609
- [4] Mirkin C A, Letsinger R L, Mucic R C, Storhoff J J 1996 *Nature* **382** 607
- [5] Macfarlane R J, Lee B, Jones M R, Harris N, Schatz G C, Mirkin C A 2011 *Science* **334** 204
- [6] Maffeo C, Yoo J, Comer J, Wells D B, Luan B, Aksimentiev A 2014 *J. Phys.: Condens. Matter* **26** 413101
- [7] Olson W K, Zhurkin V B 2011 *Curr. Opin. Struct. Biol.* **21** 348
- [8] Orozco M, Noy A, Pérez A 2008 *Curr. Opin. Struct. Biol.* **18** 185
- [9] Kim S, Brostromer E, Xing D, Jin J, Chong S, Ge H, Wang S, Gu C, Yang L, Gao Y Q, Su X D, Sun Y, Xie X S 2013 *Science* **339** 816
- [10] Yin Y, Yang L, Zheng G, Gu C, Yi C, He C, Gao Y Q, Zhao X S 2014 *Proc. Natl. Acad. Sci. U. S. A.* **111** 8043
- [11] Sobel E S, Harpst J A 1991 *Biopolymers* **31** 1559
- [12] Borochoy N, Eisenberg H, Kam Z 1981 *Biopolymers* **20** 231
- [13] Kam Z, Borochoy N, Eisenberg H 1981 *Biopolymers* **20** 2671
- [14] Harrington R E 1978 *Biopolymers* **17** 919
- [15] Cairney K L, Harrington R E 1982 *Biopolymers* **21** 923
- [16] Baumann C G, Smith S B, Bloomfield V A, Bustamante C 1997 *Proc. Natl. Acad. Sci. U. S. A.* **94** 6185
- [17] Rizzo V, Schellman J 1981 *Biopolymers* **20** 2143
- [18] Maret G, Weill G 1983 *Biopolymers* **22** 2727
- [19] Baumann C G, Smith S B, Bloomfield V A, Bustamante C 1997 *Proc. Natl. Acad. Sci. U. S. A.* **94** 6185
- [20] Fixman M 2010 *J. Phys. Chem. B* **114** 3185
- [21] Odijk T 1977 *J. Polym. Sci. Polym. Phys. Ed.* **15** 477

- [22] Manning G S 2006 *Biophys. J.* **91** 3607
- [23] Kosikov K M, Gorin A A, Lu X, Olson W K, Manning G S 2002 *J. Am. Chem. Soc.* **124** 4838
- [24] Tan Z J, Chen S J 2008 *Biophys. J.* **94** 3137
- [25] Xiao S, Zhu H, Wang L, Liang H 2014 *Soft Matter.* **10** 1045
- [26] Wang F H, Wu Y Y, Tan Z J 2013 *Biopolymers* **99** 370
- [27] Wu Y, Bao L, Zhang X, Tan Z, Wu Y, Bao L, Zhang X, Tan Z 2015 *J. Chem. Phys.* **142** 125103
- [28] Bao L, Zhang X, Jin L, Tan Z 2016 *Chin. Phys. B* **25** 018703
- [29] Savelyev A, Papoian G A 2010 *Proc. Natl. Acad. Sci. U. S. A.* **107** 20340
- [30] Savelyev A, Materese C K, Papoian G A 2011 *J. Am. Chem. Soc.* **133** 19290
- [31] Ortiz V, de Pablo J J 2011 *Phys. Rev. Lett.* **106** 238107
- [32] Kratky O, Porod G 1949 *Recl. des Trav. Chim. des Pays-Bas* **68** 1106
- [33] Peters J P, Maher L J 2010 *Q. Rev. Biophys.* **43** 23
- [34] Yuan C, Chen H, Lou X W, Archer L A 2008 *Phys. Rev. Lett.* **100** 018102
- [35] Cloutier T E, Widom J 2005 *Proc. Natl. Acad. Sci. U. S. A.* **102** 3645
- [36] Cloutier T E, Widom J 2004 *Mol. Cell* **14** 355
- [37] Crick F H C, Klug A 1975 *Nature* **255** 530
- [38] Richmond T J, Luger K, Mäder A W, Richmond R K, Sargent D F 1997 *Nature* **389** 251
- [39] Hogan M E, Rooney T F, Austin R H 1987 *Nature* **328** 554
- [40] Richmond T J, Davey C A 2003 *Nature* **423** 145
- [41] Wiggins P, Phillips R, Nelson P 2005 *Phys. Rev. E* **71** 021909
- [42] Yan J, Marko J 2004 *Phys. Rev. Lett.* **93** 108108
- [43] Theodorakopoulos N, Peyrard M 2012 *Phys. Rev. Lett.* **108** 078104
- [44] Wiggins P A, van der Heijden T, Moreno-Herrero F, Spakowitz A, Phillips R, Widom J, Dekker C, Nelson P C 2006 *Nat. Nanotechnol.* **1** 137
- [45] Xu X, Thio B J R, Cao J 2014 *J. Phys. Chem. Lett.* **5** 2868
- [46] Vafabakhsh R, Ha T 2012 *Science* **337** 1097
- [47] Du Q, Smith C, Shiffeldrim N, Vologodskaya M, Vologodskii A 2005 *Proc. Natl. Acad. Sci. U. S. A.* **102** 5397
- [48] Mastroianni A J, Sivak D A, Geissler P L, Alivisatos A P 2009 *Biophys. J.* **97** 1408
- [49] Mazur A K 2007 *Phys. Rev. Lett.* **98** 218102
- [50] Mazur A K, Maaloum M 2014 *Phys. Rev. Lett.* **112** 068104
- [51] Neidle S 2008 *Principles of Nucleic Acid Structure* (Boston: Academic Press)
- [52] Yan H 2003 *Science* **301** 1882
- [53] Malo J, Mitchell J C, Vénien-Bryan C, Harris J R, Wille H, Sherratt D J, Turberfield A J 2005 *Angew. Chem. Int. Ed.* **44** 3057
- [54] Rothmund P W K 2006 *Nature* **440** 297
- [55] Zheng J, Birktoft J J, Chen Y, Wang T, Sha R, Constantinou P E, Ginell S L, Mao C, Seeman N C 2009 *Nature* **461** 74
- [56] Klafter J, Sokolov I M 2005 *Phys. World* **18** 29
- [57] Andersen E S, Dong M, Nielsen M M, Jahn K, Subramani R, Mamdouh W, Golas M M, Sander B, Stark H, Oliveira C L P, Pedersen J S, Birkedal V, Besenbacher F, Gothelf K V, Kjems J 2009 *Nature* **459** 73
- [58] Han D, Pal S, Nangreave J, Deng Z, Liu Y, Yan H 2011 *Science* **332** 342
- [59] Zhang D Y, Seelig G 2011 *Nat. Chem.* **3** 103
- [60] Zhang D Y, Hariadi R F, Choi H M T, Winfree E 2013 *Nat. Commun.* **4** 1965
- [61] Turberfield A J, Yurke B, Mills A P, Simmel F C, Neumann J L 2000 *Nature* **406** 605
- [62] Zhang D Y, Winfree E 2009 *J. Am. Chem. Soc.* **131** 17303
- [63] Srinivas N, Ouldrige T E, Sulc P, Schaeffer J M, Yurke B, Louis A A, Doye J P K, Winfree E 2013 *Nucleic Acids Res.* **41** 10641
- [64] Soloveichik D, Seelig G, Winfree E 2010 *Proc. Natl. Acad. Sci. U. S. A.* **107** 5393
- [65] Genot A J, Zhang D Y, Bath J, Turberfield A J 2011 *J. Am. Chem. Soc.* **133** 2177
- [66] Machinek R R F, Ouldrige T E, Haley N E C, Bath J, Turberfield A J 2014 *Nat. Commun.* **5** 5324
- [67] Smith B D, Liu J 2010 *J. Am. Chem. Soc.* **132** 6300
- [68] Dave N, Liu J 2010 *J. Phys. Chem. B* **114** 15694
- [69] Kang D, Duan R, Tan Y, Hong F, Wang B, Chen Z, Xu S, Lou X, Wei W, Yurke B, Xia F 2014 *Nanoscale* **6** 14153
- [70] Ravan H, Kashanian S, Sanadgol N, Badoei-Dalfard A, Karami Z 2014 *Anal. Biochem.* **444** 41
- [71] Yurke B, Mills Jr A P 2003 *Genet. Program. Evolvable Mach.* **4** 111
- [72] Zadeh J N, Steenberg C D, Bois J S, Wolfe B R, Pierce M B, Khan A R, Dirks R M, Pierce N A 2011 *J. Comput. Chem.* **32** 170
- [73] SantaLucia J, Hicks D 2004 *Annu. Rev. Biophys. Biomol. Struct.* **33** 415
- [74] Bommarito S, Peyret N, SantaLucia Jr J S 2000 *Nucleic Acids Res.* **28** 1929
- [75] Pyshnyi D, Ivanova E 2002 *Russ. Chem. Bull.* **51** 1145
- [76] Protozanova E, Yakovchuk P, Frank-Kamenetskii M D 2004 *J. Mol. Biol.* **342** 775
- [77] Potoyan D A, Savelyev A, Papoian G A 2013 *WIREs Comput. Mol. Sci.* **3** 69
- [78] Sambriski E J, Schwartz D C, de Pablo J J 2009 *Biophys. J.* **96** 1675
- [79] Knotts T A, Rathore N, Schwartz D C, de Pablo J J 2007 *J. Chem. Phys.* **126** 084901
- [80] Sambriski E J, Schwartz D C, de Pablo J J 2009 *Proc. Natl. Acad. Sci. U. S. A.* **106** 18125
- [81] Li H, Wang Z, Li N, He X, Liang H 2014 *J. Chem. Phys.* **141** 044911
- [82] Ouldrige T E, Louis A A, Doye J P K 2011 *J. Chem. Phys.* **134** 085101
- [83] Ouldrige T E, Louis A A, Doye J P K 2010 *Phys. Rev. Lett.* **104** 178101
- [84] Ouldrige T E, Sulc P, Romano F, Doye J P K, Louis A A 2013 *Nucleic Acids Res.* **41** 8886

- [85] Mosayebi M, Romano F, Ouldrige T E, Louis A A, Doye J P K 2014 *J. Phys. Chem. B* **118** 14326
- [86] Matek C, Ouldrige T E, Levy A, Doye J P K, Louis A A 2012 *J. Phys. Chem. B* **116** 11616
- [87] Ouldrige T E 2014 *Mol. Phys.* **113** 1
- [88] Adleman L 1994 *Science* **266** 1021
- [89] Benenson Y, Paz-Elizur T, Adar R, Keinan E, Livneh Z, Shapiro E 2001 *Nature* **414** 430
- [90] Benenson Y, Adar R, Paz-Elizur T, Livneh Z, Shapiro E 2003 *Proc. Natl. Acad. Sci. U. S. A.* **100** 2191
- [91] Sakamoto K 2000 *Science* **288** 1223
- [92] Benenson Y, Gil B, Ben-Dor U, Adar R, Shapiro E 2004 *Nature* **429** 423
- [93] Yin P, Choi H M T, Calvert C R, Pierce N A 2008 *Nature* **451** 318
- [94] Zhang D Y, Turberfield A J, Yurke B, Winfree E 2007 *Science* **318** 1121
- [95] Seelig G, Soloveichik D, Zhang D Y, Winfree E 2006 *Science* **314** 1585
- [96] Qian L, Winfree E 2011 *Science* **332** 1196
- [97] Qian L, Winfree E, Bruck J 2011 *Nature* **475** 368
- [98] Kim J, Winfree E 2014 *Mol. Syst. Biol.* **7** 465
- [99] Han D, Wu C, You M, Zhang T, Wan S, Chen T, Qiu L, Zheng Z, Liang H, Tan W 2015 *Nat. Chem.* **7** 835
- [100] Chen Y, Dalchau N, Srinivas N, Phillips A, Cardelli L, Soloveichik D, Seelig G 2013 *Nat. Nanotechnol.* **8** 755
- [101] Lakin M R, Youssef S, Cardelli L, Phillips A 2012 *J. R. Soc. Interface* **9** 470
- [102] Phillips A, Cardelli L 2009 *J. R. Soc. Interface* **6** S419
- [103] Chen X, Briggs N, McLain J R, Ellington A D 2013 *Proc. Natl. Acad. Sci. U. S. A.* **110** 5386
- [104] Jiang Y S, Bhadra S, Li B, Ellington A D 2014 *Angew. Chem. Int. Ed.* **53** 1845
- [105] Carlson R 2009 *Nat. Biotechnol.* **27** 1091
- [106] Simmel F C, Yurke B 2001 *Phys. Rev. E* **63** 041913
- [107] Simmel F C, Yurke B 2002 *Appl. Phys. Lett.* **80** 883
- [108] Tian Y, Mao C 2004 *J. Am. Chem. Soc.* **126** 11410
- [109] Yan H, Zhang X, Shen Z, Seeman N C 2002 *Nature* **415** 62
- [110] Chakraborty B, Sha R, Seeman N C 2008 *Proc. Natl. Acad. Sci. U. S. A.* **105** 17245
- [111] Zhong H, Seeman N C 2006 *Nano Lett.* **6** 2899
- [112] Ding B, Seeman N C 2006 *Science* **314** 1583
- [113] Feng L, Park S H, Reif J H, Yan H 2003 *Angew. Chem.* **115** 4478
- [114] Lubrich D, Lin J, Yan J 2008 *Angew. Chem. Int. Ed.* **47** 7026
- [115] Goodman R P, Heilemann M, Doose S, Erben C M, Kapanidis A N, Turberfield A J 2008 *Nat. Nanotechnol.* **3** 93
- [116] Dirks R M, Pierce N A 2004 *Proc. Natl. Acad. Sci. U.S.A.* **101** 15275
- [117] Xing Y, Cheng E, Yang Y, Chen P, Zhang T, Sun Y, Yang Z, Liu D 2011 *Adv. Mater.* **23** 1117
- [118] Xing Y, Yang Z, Liu D 2011 *Angew. Chem. Int. Ed.* **50** 11934
- [119] Idili A, Porchetta A, Amodio A, Vallée-Bélisle A, Ricci F 2015 *Nano Lett.* **15** 5539
- [120] Idili A, Vallée-Bélisle A, Ricci F 2014 *J. Am. Chem. Soc.* **136** 5836
- [121] Yao D, Li H, Guo Y, Zhou X, Xiao S, Liang H 2016 *Chem. Commun.* **52** 7556
- [122] Liu D, Cheng E, Yang Z 2011 *NPG Asia Mater.* **3** 109
- [123] Dong Y, Yang Z, Liu D 2014 *Acc. Chem. Res.* **47** 1853
- [124] Sherman W B, Seeman N C 2004 *Nano Lett.* **4** 1203
- [125] Shin J S, Pierce N A 2004 *J. Am. Chem. Soc.* **126** 10834
- [126] Gu H, Chao J, Xiao S J, Seeman N C 2010 *Nature* **465** 202
- [127] You M, Chen Y, Zhang X, Liu H, Wang R, Wang K, Williams K R, Tan W 2012 *Angew. Chem. Int. Ed.* **51** 2457
- [128] Yang Y, Goetzfried M A, Hidaka K, You M, Tan W, Sugiyama H, Endo M 2015 *Nano Lett.* **15** 6672
- [129] Venkataraman S, Dirks R M, Rothmund P W K, Winfree E, Pierce N A 2007 *Nat. Nanotechnol.* **2** 490
- [130] Lubrich D, Green S J, Turberfield A J 2009 *J. Am. Chem. Soc.* **131** 2422
- [131] Xuan F, Hsing I M 2014 *J. Am. Chem. Soc.* **136** 9810
- [132] Zhang D Y, Winfree E 2008 *J. Am. Chem. Soc.* **130** 13921
- [133] Wang Y, Lin H X, Chen L, Ding S Y, Lei Z C, Liu D Y, Cao X Y, Liang H J, Jiang Y B, Tian Z Q 2014 *Chem. Soc. Rev.* **43** 399
- [134] Song T, Xiao S, Yao D, Huang F, Hu M, Liang H 2014 *Adv. Mater.* **26** 6181
- [135] Song T, Liang H 2012 *J. Am. Chem. Soc.* **134** 10803
- [136] Bi S, Chen M, Jia X, Dong Y, Wang Z 2015 *Angew. Chem. Int. Ed.* **54** 8144
- [137] Huang F, You M, Han D, Xiong X, Liang H, Tan W 2013 *J. Am. Chem. Soc.* **135** 7967
- [138] Huang F, Xu H, Tan W, Liang H 2014 *ACS Nano* **8** 6849
- [139] Huang F, Zhou X, Yao D, Xiao S, Liang H 2015 *Small* **11** 5800
- [140] Yurke B, Mills A P, Lai Cheng S 1999 *Biosystems* **52** 165
- [141] Dittmer W U, Reuter A, Simmel F C 2004 *Angew. Chem. Int. Ed.* **43** 3550
- [142] Ko S H, Liu H, Chen Y, Mao C 2008 *Biomacromolecules* **9** 3039
- [143] Mei Q, Wei X, Su F, Liu Y, Youngbull C, Johnson R, Lindsay S, Yan H, Meldrum D 2011 *Nano Lett.* **11** 1477
- [144] Conway J W, McLaughlin C K, Castor K J, Sleiman H 2013 *Chem. Commun.* **49** 1172
- [145] Hahn J, Wickham S F J, Shih W M, Perrault S D 2014 *ACS Nano* **8** 8765
- [146] Castro C E, Kilchherr F, Kim D N, Shiao E L, Wauer T, Wortmann P, Bathe M, Dietz H 2011 *Nat. Methods* **8** 221
- [147] Douglas S M, Bachelet I, Church G M 2012 *Science* **335** 831
- [148] Rudchenko M, Taylor S, Pallavi P, Dechkovskaia A, Khan S, Butler V P, Rudchenko S, Stojanovic M N 2013 *Nat. Nanotechnol.* **8** 580
- [149] You M, Zhu G, Chen T, Donovan M J, Tan W 2015 *J. Am. Chem. Soc.* **137** 667

- [150] You M, Peng L, Shao N, Zhang L, Qiu L, Cui C, Tan W 2014 *J. Am. Chem. Soc.* **136** 1256
- [151] Gartner Z J, Bertozzi C R 2009 *Proc. Natl. Acad. Sci. U. S. A.* **106** 4606
- [152] Chandra R A, Douglas E S, Mathies R A, Bertozzi C R, Francis M B 2006 *Angew. Chem. Int. Ed.* **45** 896
- [153] Modi S, M. G. S, Goswami D, Gupta G D, Mayor S, Krishnan Y 2009 *Nat. Nanotechnol.* **4** 325
- [154] Modi S, Nizak C, Surana S, Halder S, Krishnan Y 2013 *Nat. Nanotechnol.* **8** 459
- [155] Chen A K, Davydenko O, Behlke M A, Tsourkas A 2010 *Nucleic Acids Res.* **38** e148
- [156] Qiu L, Wu C, You M, Han D, Chen T, Zhu G, Jiang J, Yu R, Tan W 2013 *J. Am. Chem. Soc.* **135** 12952
- [157] Tyagi S, Kramer F R 1996 *Nat. Biotechnol.* **14** 303
- [158] Wu C, Cansiz S, Zhang L, Teng I T, Qiu L, Li J, Liu Y, Zhou C, Hu R, Zhang T, Cui C, Cui L, Tan W 2015 *J. Am. Chem. Soc.* **137** 4900
- [159] Seferos D S, Giljohann D A, Hill H D, Prigodich A E, Mirkin C A 2007 *J. Am. Chem. Soc.* **129** 15477
- [160] Prigodich A E, Seferos D S, Massich M D, Giljohann D A, Lane B C, Mirkin C A 2009 *ACS Nano* **3** 2147
- [161] Massich M D, Giljohann D A, Schmucker A L, Patel P C, Mirkin C A 2010 *ACS Nano* **4** 5641
- [162] Amir Y, Ben-Ishay E, Levner D, Ittah S, Abu-Horowitz A, Bachelet I 2014 *Nat. Nanotechnol.* **9** 353
- [163] Hemphill J, Deiters A 2013 *J. Am. Chem. Soc.* **135** 10512
- [164] Sponer J, Cang X, Cheatham T E 2012 *Methods* **57** 25
- [165] Šponer J, Banáš P, Jurečka P, Zgarbová M, Kührová P, Havrila M, Krepl M, Stadlbauer P, Otyepka M 2014 *J. Phys. Chem. Lett.* **5** 1771

SPECIAL TOPIC — Progress in Soft Matter Research

DNA and DNA computation based on toehold-mediated strand-displacement reactions*

Xiao Shi-Yan¹⁾ Liang Hao-Jun^{1)2)†}

1) (CAS Key Laboratory of Soft Matter Chemistry, Collaborative Innovation Center of Chemistry for Energy Materials, Department of Polymer Science and Engineering, University of Science and Technology of China, Hefei 230026, China)

2) (Hefei National Laboratory for Physical Sciences at the Microscale, University of Science and Technology of China, Hefei 230026, China)

(Received 26 May 2016; revised manuscript received 15 June 2016)

Abstract

Other than being a carrier of the code of life, DNA can also be used as a kind of ideal biomaterial with good biocompatibility. Considering the critical role of DNA less than 150 base pairs (bp) in cellular processes such as regulated gene expression, quantifying the intrinsic bend ability of DNA on a sub-persistence length scale is essential to understanding its molecular functions and the DNA-protein interaction. From the classical point of view, double-stranded DNA is assumed to be stiff and can be treated by semi-flexible chain, but recent studies have yielded contradictory results. A lot of studies tried to prove that the worm-like chain model can be used to fully describe DNA chain. However, recent theoretical and experimental studies indicated that DNA exhibits high flexibility on a short length scale, which cannot be described by the worm-like chain model. Further studies are needed to address the extreme flexibility of DNA on a short length scale. On the basis of the predictability of the double helical structure and the Watson-Crick binding thermodynamics for DNA, a class of DNA reactions can be defined, called toehold-mediated strand-displacement reaction, in which one complementary single-stranded DNA sequence first binds to the dangling toehold domain of the substrate in a pre-hybridized double-stranded DNA, then triggers the strand-displacement reaction, and finally results in the dissociation of the third strand previously bound to the substrate with partial complementarity. In dynamic DNA nanotechnology, isothermal toehold-mediated DNA strand-displacement reaction has been used to design complex nanostructure and nanodevice for molecular computation. The kinetics of the strand-displacement can be modulated using the toehold length. In order to weaken the coupling between the kinetics of strand-displacement and the thermodynamics of the reaction, the concept of toehold exchange was introduced by Winfree et al. to improve the control of strand-displacement kinetics. More importantly, the biomolecular reaction (BM) rate constant of toehold exchange can be analytically derived using the three-step model. Through utilizing strand-displacement reactions and taking advantage of its programmable sequences and precise recognition properties, DNA can be used to build complex circuits which can proceed robustly at constant temperature, achieving specific functions. DNA strand-displacement reaction can be employed to fabricate logic gates, and large and complex circuits for DNA computing, to mimic the naturally occurring occurrence of biological systems. Based on that, DNA circuit can then be used to direct the assembly of nanodevice following the designed pathway, and modulate the chemical reaction networks on the surface of living cell or in cellular systems for biosensing, even program the cellular machinery in the future for genetic diagnostic or gene therapy. In the present paper, we reviewed the proceedings in the fields of DNA structure and conformational changes, and DNA flexibility, discussed the mechanism of DNA strand-displacement reaction at the molecular level, and introduced the recent studies in DNA computation as well as the dynamic DNA nanotechnology, such as self-assembly.

Keywords: DNA flexibility, strand-displacement reaction, DNA computation, self-assembly

PACS: 81.16.Fg, 87.85.-d, 87.10.Tf, 87.14.gk

DOI: 10.7498/aps.65.178106

* Project supported by the National Natural Science Foundation of China (Grant Nos. 91427304, 21434007, 21574122, 51573175, 21404098) and the National Basic Research Program of China (Grant No. 2012CB821500).

† Corresponding author. E-mail: hjliang@ustc.edu.cn