智能核酸计算系统的构建及生物医学应用

赵雨萌, 韩达*

上海交通大学医学院附属仁济医院分子医学研究院, 上海 200127

*通讯作者, E-mail: dahan@sjtu.edu.cn

收稿日期: 2019-02-15; 接受日期: 2019-04-25; 网络版发表日期: 2019-06-04

国家青年千人计划资助项目

摘要 智能分子计算系统可以精确操控纳米器件、原位监控生物化学反应和智能调控生物体内的基因表达, 因此在生物医学领域具有优越的应用前景. 核酸作为遗传信息的载体, 具有纳米级尺寸、易于化学合成、可预测的碱基互补识别能力以及重要的生物调控功能等特性, 是构建智能分子计算系统的理想材料. 目前, 研究人员已经设计出包括布尔逻辑门、分子电路、人工神经网络在内的大量基于核酸的智能分子计算系统, 可以用来解决各种生物医学领域的难题. 本文重点介绍了基于核酸的智能核酸分子计算系统的构建及其在生物医学领域的研究进展.

关键词 核酸, 核酸适体, 逻辑计算系统, 生物医学应用

1 引言

20世纪90年代以来, 基于核酸的分子计算技术开始兴起. 其主要的动力来自于基于硅基底的半导体技术从制造工艺上已经到达了无法进一步变小的瓶颈. 作为生物自然选择的遗传物质, 核酸有着优越的微型化信息储存能力, 其储存1 bit的信息仅需要50个原子^[1,2]. 同时, 快速程序化的化学合成以及低廉的合成成本也为核酸分子在逻辑计算中的大规模应用提供了优良的物质基础. 早在1994年, Adleman^[3]就利用核酸的计算能力解决了一个数学逻辑计算领域的难题: 哈密顿路径问题. 他将哈密顿路径问题的节点编码为不同的单链DNA序列, 利用核酸碱基之间的识别能力及聚合酶链式反应(PCR)扩增技术解码出正确的哈密顿路径. 这一研究开创了核酸计算领域的先河. 经过近20

年的发展, 人们已经证明了由核酸分子组成的逻辑系统可以执行各种模拟计算机中的逻辑计算指令, 如信号储存、放大、反馈和联级等. 但是, 核酸分子本身的局限性, 如信号传导速度慢和难以设计成为固定的模块系统等, 使核酸分子计算技术在计算速度与计算容量上远远无法接近传统半导体计算机技术.

随着生物技术的进一步发展, 人们发现核酸分子具备越来越多的重要生物学功能, 如调控基因表达、干预蛋白质功能等, 而核酸分子计算技术则可以把核酸分子的这些功能有机整合起来, 实现对生物功能分子更精细、智能的调控效果. 所以, 把核酸分子电路应用于生物医学中成为了其新的发展方向. 利用核酸分子的可编程性、纳米级尺寸、生物调控功能等优势, 并与纳米技术相结合, 研究者们构建出具有不同功能的智能核酸计算系统. 本文重点介绍了智能核酸计

引用格式: Zhao Y, Han D. Nucleic acid based intelligent systems and their biomedical applications. *Sci Sin Chim*, 2019, 49: 1149–1155, doi: 10.1360/SSC-2019-0018

算系统的构建及在生物医学应用方面的最新重要进展。

2 智能核酸计算系统的构建

对于核酸的结构与物理化学特性的深入了解, 包括双螺旋结构、碱基互补配对原则等理论知识, 为构建核酸计算系统提供了理论基础。一般来说, 智能核酸计算系统需包含三个模块: 信号输入、信号处理和信号输出^[4]。此外, 还需要采用特定的反应机制来实现信号在不同模块之间的传递。

2.1 基于链取代反应的核酸计算系统

核酸的链取代反应在智能核酸计算系统的构建中起了重要作用^[5]。典型的链取代反应从取代链上一段被称为“toehold”的区域开始(图1(a))。取代链上的toehold序列可与被取代的双链核酸中的互补区域特异性结合, 进而引发取代链与被取代链之间发生链迁移和链解离反应, 最终用取代链将被取代链从双链中解离出来, 使信号发生传递。在这个反应中, 取代链的碱基数目要高于被取代链, 反应由生成的具有更多互补碱基的双链所引发的吉布斯自由能的降低来驱动。同时, 链取代反应的速率可以通过改变toehold序列的碱基数来调节^[6]。利用链取代反应, Zhu等^[7]成功构建了一系列基于G四聚体的YES、NOT、OR、AND、INHIBIT布尔逻辑门。这些逻辑门可以智能分析致病的基因片

段, 并给出高准确率的综合检测结果。Li等^[8]巧妙地利用一个具有相同输出域不同识别域的环状核酸结构, 构建了一种基于链取代反应的三输入逻辑门。该逻辑门的关键特征是只要其中两个输入端为真值, 则其输出为真值(图2(a))。Seelig等^[9]利用链取代反应构建了可实现3~4层联级反应的逻辑门, 该逻辑门可智能分析输入端的不同microRNA种类并实现AND、OR、NOT等布尔逻辑门计算的功能。

2.2 基于链交换反应的核酸计算系统

尽管链取代反应为构建智能核酸计算系统提供了有效的信号传递手段, 但是此方法仅适用于构建规模较小的系统。主要原因是链取代反应要求取代链的长度要多于被取代链, 这就导致系统中负责信号传递的核酸的序列长度沿着信号传递的方向逐渐变短, 无法用于多步的信号传递。为了实现核酸信号的多步传递以提高核酸计算系统的规模, Zhang和Winfree^[10]开发了链交换反应(图1(b))。在链交换反应中, 被取代链的互补链中也包括一部分“移位”的“toehold”序列, 该“toehold”序列可以继续引发下一步反应。此设计所用的取代链和被取代链的长度一致, 所以反应的进行方向是可逆的。在体系中加入高浓度的“fuel”序列所产生的熵增益, 可以驱动反应向一个方向进行。利用这个反应机理, 人们构建了更大规模的核酸计算系统。Qian和Winfree^[11]设计了一个包含130条核酸链的核酸计算系统, 该系统可以通过预设的逻辑, 正确解答四比特位的

图 1 (a) DNA链取代反应示意图; (b) DNA链交换反应示意图^[2c](网络版彩图)

Figure 1 (a) Scheme of DNA toehold-mediated displacement reaction; (b) scheme of DNA toehold-exchange displacement reaction^[2c] (color online).

图 2 (a) 基于DNA链交换反应的逻辑门结构设计示意图^[8]; (b) 基于DNA脱氧核酶的计算系统设计示意图^[16] (网络版彩图)
Figure 2 (a) Architectural design of a three-input majority gate based on DNA strand displacement [8]; (b) general design of a computing unit module using libraries of DNAzyme subunits and substrates [16] (color online).

平方根计算问题。他们还创建了一个可模拟神经网络的核酸计算系统, 用四个由链交换反应连接的多层核酸网络模块组成的“人工神经元”实现“Hopfield”联想记忆。这些人工神经元可以在系统引入不完整的信息输入时, 进行逻辑分析, 找到与正确结果最相似的答案^[12]。Qian等^[13]的另一项开创性工作是利用核酸计算进行图像识别。基于核酸的神经网络计算系统可以通过“赢家全选”原则接受训练, 并且成功地将不同的测试图像进行分类。在训练过程中, 该核酸计算系统可将不同人手写的数字1~9正确识别并分类。

2.3 基于脱氧核酶的核酸计算系统

脱氧核酶(DNAzyme)作为一种催化分子, 可以特异性地切割底物RNA位点, 为核酸计算系统中输入端和输出端之间的信号传递提供了另一种有效的策略。Stojanovic及其同事^[14]提出利用脱氧核酶对底物核酸片段的催化切割能力, 实现了将被切割的核酸片段作为下级反应的信号输入的联级反应。随后应用这一理论, 首次通过设计不同的核酸信号输入序列来控制两种不同的脱氧核酶的活性, 建立了AND、OR和NOT等布尔逻辑门。在此基础上, 他们还进一步开发了多

层的基于脱氧核酶的智能计算系统, 该系统可以与人类进行“tic-tac-toe”的简单游戏, 并能保证在每次游戏中不输给人类玩家^[15]。Willner及其同事^[16]利用脱氧核酶及其底物构建了一种核酸计算系统, 可以实现半加法器、半减法器等不同逻辑门的运算(图2(b))。综上所述, 我们对可用于构建核酸计算系统的三种不同反应机理进行了归类比较, 具体内容详见表1。

3 智能核酸计算系统的生物学应用

基于核酸的计算系统可以用来精确调控分子功能以及化学生物反应中的时空动力学^[17], 因此在智能药物的开发、仿生系统的构建等生物学应用领域具有优秀的应用前景。本章将对智能核酸计算系统的生物学应用的重要进展做重点介绍。

3.1 核酸计算系统在疾病监测与分析中的应用

智能核酸系统在疾病监测与分析中展现了巨大的应用潜力。华盛顿大学的Seelig等^[18]构建了基于mRNA的计算系统, 将核酸碱基的识别能力与机器学习技术相结合, 成功地实现了宿主细胞内病毒性和细菌性呼

吸系统感染不同基因表达病例的鉴别. Hemphill和Deiters^[19]也设计了在细胞内同时监控microRNA的计算系统, 并指出逻辑门将来有望用于更准确地检测和治疗人类疾病, 尤其是癌症. 哈佛大学医学院的Yin等^[20]也设计并实现了可以在大肠杆菌细胞内精确执行多步“和门”、“或门”、“非门”等逻辑门的RNA逻辑系统, 并指出该系统可用于细胞内基因疾病的治疗. 本课题组^[21]在体外实现了使用核酸和酶构建可模拟获得性免疫系统基本功能的智能核酸系统(AIRS), 该系统可智能识别病原体DNA, 产生特异性的免疫反应和对病原体序列的记忆效应. 这项研究成果, 为深度研究获得

性免疫系统提供了分子学基础, 同时, 也为建立复杂人工模拟生物系统提供了方法(图3(a))^[21]. 此外, 我们还证实了该人工免疫系统可在原型细胞中发挥作用(图3(b))^[22]. 当入侵原型细胞的病原体DNA片段数目高于系统容忍值时, 该系统被迅速激活并执行一系列分子运算反应, 产生大量的人工抗体模拟物, 捕获病原体DNA片段. 被捕获的病原体DNA片段在限制性内切酶的作用下被选择性降解, 从而达到移除病原体DNA片段的目的. 将智能核酸计算系统作为原型细胞计算核心的策略, 对构建可用于未来生物学研究的人工细胞或纳米机器人具有很好的启发作用.

表 1 基于链取代、链交换、脱氧核酶的核酸计算系统特性对比

Table 1 Comparison of toehold-mediated, toehold-exchange and DNAzymes reactions

	反应速率	可联级层数	辅助条件要求	生物体系兼容性
链取代反应	根据toehold长度进行调节	3~4层	镁离子	高, 不需要酶等其他条件
链置换反应	根据toehold长度进行调节	无限制	镁离子	高, 不需要酶等其他条件
脱氧核酶反应	速度不可调	1~2层	金属离子	低, 稳定性低, 需要重金属离子作为辅助因子

图 3 (a) 模拟获得性免疫系统的智能核酸系统(AIRS)工作原理图^[21]; (b) 构建于原型细胞内的人工获得性免疫系统工作原理图^[22] (网络版彩图)

Figure 3 (a) Working principle of the artificial adaptive immune response simulator (AIRS) [21]; (b) schematic figure of protocells with built-in DNA-based adaptive immune response simulator [22] (color online).

上述核酸计算系统主要针对的是基因层面的检测与分析. 相比而言, 针对细胞膜上靶标蛋白进行逻辑分析的计算系统具有分析速度快、准确性高等优势, 在癌症等疾病的诊断方面同样具有巨大的潜力. Douglas 等^[23]开发了一种基于DNA折纸术的计算分析系统, 能够检测到细胞膜表面不同的抗原蛋白信息, 通过对抗原信息进行综合分析后智能选择靶细胞并进行药物输送. Stojanovic 等^[24]也设计了基于单克隆抗体的, 可同时监控细胞膜上不同CD抗原的计算分析系统, 来进一步提高细胞分型的精确度. 本课题组^[25]也设计了一种构建于细胞膜上的基于核酸适体的计算系统, 该系统能够对细胞膜上核酸适体的靶标物进行分析, 特异性选择满足系统识别条件的靶细胞, 激活光敏剂并产生光动力治疗效果. 通过核酸适体的特异性识别与结合, 使得该核酸计算系统能够将靶标癌细胞与健康细胞区分开来, 同时只在癌细胞周围产生高浓度的单线态氧, 从而减少对附近健康细胞的损伤. 此外, You和Tan 等^[26,27]也构建了一系列基于核酸适体的智能核酸计算系统, 这些系统通过在细胞膜上进行逻辑运算, 实现智能的药物递送功能(图4(a)). 这些智能核酸计算系统在疾病监测与分析中均展现了巨大的应用潜力.

3.2 核酸计算系统在蛋白功能调控方面的应用

智能核酸系统在蛋白功能的智能调控中同样展现出优秀的应用前景. 长期以来, 根据作为理论基础的核酸碱基互补配对原则, 智能核酸计算系统的调控对象主要为核酸分子本身, 无法在体外实现对其他目标物如蛋白质的调控. 而核酸适体作为一种能够特异性识别蛋白质的分子, 可以为构建具有蛋白质调控能力的智能核酸系统提供简单有效的手段. 因此, 使用核酸计算系统直接调控蛋白质的功能成为了智能核酸计算系统生物医学应用的一个重要切入点. Kolpashchikov 和Stojanovic^[28]首次将核酸适体应用到核酸计算电路中. 他们构建了基于核酸适体的布尔逻辑门, 可以对蛋白质或小分子等输入信号进行逻辑运算. 本课题组^[29]近来也发现, 通过在核酸计算系统中引入核酸适体, 可以实现凝血酶功能的精确调控. 我们构建了一个基于核酸和核酸适体精确调控凝血酶功能的药物智能识别体系. 该系统可以智能地感知体系中凝血酶的浓度. 当酶的浓度低于阈值时, 系统处于休眠状态; 而当酶的浓度高于设定阈值时, 系统将自动释放酶抑制剂(核酸适体)来抑制酶的活性, 防止其过度凝血. 这个体系可以实现智能调控酶的功能, 将为个性化药物的开发提

图 4 (a) 构建于细胞膜上的基于DNA适配体的智能核酸药物递送系统原理示意图^[27]; (b) 智能核酸计算系统监测细胞膜上的蛋白质分子相互作用原理示意图^[30](网络版彩图)

Figure 4 (a) Scheme of aptamer-based DNA logical circuit for smart drug delivery ^[27]; (b) schematic illustration of intelligent DNA system for studying dynamic protein encounters on a live cell membrane ^[30] (color online).

供有效的方法和模型。

本课题组^[30]也将智能核酸计算系统调控蛋白功能的应用范围从体外进一步扩展到细胞膜上。细胞及细胞膜上的蛋白质分子相互作用, 不仅对完成细胞信号传导及功能具有重要的影响, 也是反映细胞健康状态的重要信息。因此, 检测细胞膜上的蛋白质分子相互作用一直是生物医学领域的一个热点课题。荧光共振能量转移(FRET)和共聚焦显微镜等通用的检测方法, 无法对蛋白质分子瞬时碰撞产生的微弱信号进行检测。因此无法对这些蛋白质分子的碰撞动力学进行准确表征。我们利用核酸的链取代反应, 在细胞膜上建立了可进行联级反应的智能核酸计算系统。该系统可以将每一次蛋白质分子在细胞膜上的瞬时相互作用转化和放大为可以累加的荧光信号, 并固定于细胞膜表面(图4(b))。随着越来越多的蛋白质分子相互作用的发生, 累加的荧光信号逐渐达到一个可以被检测的量级, 进而可以通过流式细胞仪检测到。该研究建立的细胞及细胞膜上分子相互作用检测技术, 为进一步研究细胞膜和膜蛋白分子的运动提供了有效的手段。同时, 这项重要发现对促进通过分析膜蛋白分子运动揭示细胞健康状态的深度研究提供了重要条件, 成为

了智能核酸系统在生物医学领域的一个重要应用方向。

4 总结与展望

目前, 基于核酸的智能计算系统已经发展到在体外以可预测的方式合理地设计具备任何逻辑功能的阶段。此外, 使用计算机软件可对系统的序列设计、动力性和热力学行为进行预测和模拟, 大大提高了系统构建的成功率。因此, 在过去的10年中, 基于核酸的计算系统无论在质量还是在数量上, 都经历了急剧的增长。然而在将其大规模转化至生物医学领域的应用之前, 依旧存在一些难题等待克服, 如相对较慢的反应动力学和较高的错误率^[31]。此外, 生物稳定性的缺乏和在复杂环境下的协同性也给其在生物体内的应用设下了难题。但是, 随着各种新技术的层出不穷, 这些问题会得到逐步解决。可以确定的是, 研究人员将会设计出更多的智能核酸计算系统, 并进一步拓展其在生物医学中的应用, 为模拟生命的功能和创新重大疾病的诊疗提供变革性新技术, 为人类探索生命基本规律和干预疾病信号通路提供重要的理论与技术支撑。

参考文献

- 1 Ball P. *Nature*, 2000, 406: 118–120
- 2 (a) Goldman N, Bertone P, Chen S, Dessimoz C, LeProust EM, Sipos B, Birney E. *Nature*, 2013, 494: 77–80; (b) Church GM, Gao Y, Kosuri S. *Science*, 2012, 337: 1628; (c) Han D, Kang H, Zhang T, Wu C, Zhou C, You M, Zhuo C, Zhang X, Tan W. *Chem Eur J*, 2014, 20: 5866–5873
- 3 Adleman LM. *Science*, 1994, 266: 1021–1024
- 4 Benenson Y. *Mol Biosyst*, 2009, 5: 675–685
- 5 Zhang DY, Seelig G. *Nat Chem*, 2011, 3: 103–113
- 6 Turberfield AJ, Mitchell JC, Yurke B, Mills AP, Blakey MI, Simmel FC. *Phys Rev Lett*, 2003, 90: 118102
- 7 Zhu J, Zhang L, Li T, Dong S, Wang E. *Adv Mater*, 2013, 25: 2440–2444
- 8 Li W, Yang Y, Yan H, Liu Y. *Nano Lett*, 2013, 13: 2980–2988
- 9 Seelig G, Soloveichik D, Zhang DY, Winfree E. *Science*, 2006, 314: 1585–1588
- 10 Zhang DY, Winfree E. *J Am Chem Soc*, 2009, 131: 17303–17314
- 11 Qian L, Winfree E. *Science*, 2011, 332: 1196–1201
- 12 Qian L, Winfree E, Bruck J. *Nature*, 2011, 475: 368–372
- 13 Cherry KM, Qian L. *Nature*, 2018, 559: 370–376
- 14 Stojanovic MN, Mitchell TE, Stefanovic D. *J Am Chem Soc*, 2002, 124: 3555–3561
- 15 Pei R, Matamoros E, Liu M, Stefanovic D, Stojanovic MN. *Nat Nanotech*, 2010, 5: 773–777
- 16 Elbaz J, Lioubashevski O, Wang F, Remacle F, Levine RD, Willner I. *Nat Nanotech*, 2010, 5: 417–422
- 17 Franco E, Friedrichs E, Kim J, Jungmann R, Murray R, Winfree E, Simmel FC. *Proc Natl Acad Sci USA*, 2011, 108: E784–E793
- 18 Lopez R, Wang R, Seelig G. *Nat Chem*, 2018, 10: 746–754

- 19 Hemphill J, Deiters A. *J Am Chem Soc*, 2013, 135: 10512–10518
- 20 Green AA, Kim J, Ma D, Silver PA, Collins JJ, Yin P. *Nature*, 2017, 548: 117–121
- 21 Han D, Wu C, You M, Zhang T, Wan S, Chen T, Qiu L, Zheng Z, Liang H, Tan W. *Nat Chem*, 2015, 7: 835–841
- 22 Lyu Y, Wu C, Heinke C, Han D, Cai R, Teng IT, Liu Y, Liu H, Zhang X, Liu Q, Tan W. *J Am Chem Soc*, 2018, 140: 6912–6920
- 23 Douglas SM, Bachelet I, Church GM. *Science*, 2012, 335: 831–834
- 24 Rudchenko M, Taylor S, Pallavi P, Dechkovskaia A, Khan S, Butler VP, Rudchenko S, Stojanovic MN. *Nat Nanotechnol*, 2013, 8: 580–586
- 25 Han D, Zhu G, Wu C, Zhu Z, Chen T, Zhang X, Tan W. *ACS Nano*, 2013, 7: 2312–2319
- 26 You M, Peng L, Shao N, Zhang L, Qiu L, Cui C, Tan W. *J Am Chem Soc*, 2014, 136: 1256–1259
- 27 You M, Zhu G, Chen T, Donovan MJ, Tan W. *J Am Chem Soc*, 2015, 137: 667–674
- 28 Kolpashchikov DM, Stojanovic MN. *J Am Chem Soc*, 2005, 127: 11348–11351
- 29 Han D, Zhu Z, Wu C, Peng L, Zhou L, Gulbakan B, Zhu G, Williams KR, Tan W. *J Am Chem Soc*, 2012, 134: 20797–20804
- 30 You M, Lyu Y, Han D, Qiu L, Liu Q, Chen T, Sam Wu C, Peng L, Zhang L, Bao G, Tan W. *Nat Nanotech*, 2017, 12: 453–459
- 31 Chen X, Ellington AD. *Curr Opin Biotech*, 2010, 21: 392–400

Nucleic acid based intelligent systems and their biomedical applications

Yumeng Zhao, Da Han*

Institute of Molecular Medicine, Renji Hospital, School of Medicine, Shanghai Jiao Tong University, Shanghai 200127, China

*Corresponding author (email: dahan@sjtu.edu.cn)

Abstract: Artificial intelligent molecular systems are playing a significant role in the field of biomedical applications. These logical systems could be used as processors to precisely control nanodevices *in vitro*, monitor biochemical reactions *in situ* and intelligently regulate gene expression *in vivo*. Nucleic acids are promising materials for construction of artificial intelligent molecular systems due to their nanometer-size, easy chemical synthesis, predictable base-pairing behavior and significant regulation ability on biomolecules. A number of nucleic acid-based intelligent systems have been constructed in last couple decades, including Boolean logical gates, molecular circuit or artificial neural network. They have been utilized to handle various challenges in the biomedical applications. This review will focus on the recent progress of nucleic acid-based intelligent systems and their biomedical applications.

Keywords: nucleic acids, aptamers, intelligent systems, biomedical applications

doi: [10.1360/SSC-2019-0018](https://doi.org/10.1360/SSC-2019-0018)