

Using Molecular Beacons as Paste Series Computing Model of a Logic Gate

Jingjing Tang, Jingyi Shan

College of Science, Anhui University of Science and Technology, Huainan
Email: 595455103@qq.com

Received: Nov. 1st, 2013; revised: Dec. 10th, 2013; accepted: Dec. 22nd, 2013

Copyright © 2014 Jingjing Tang, Jingyi Shan. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. In accordance of the Creative Commons Attribution License all Copyrights © 2014 are reserved for Hans and the owner of the intellectual property Jingjing Tang, Jingyi Shan. All Copyright © 2014 are guarded by law and by Hans as a guardian.

Abstract: In recent years, it has become a research focus of DNA computer to realize the Boolean logic gates by using DNA sticker model and paste system, with the research of DNA computing and DNA computer going on. But making the molecular beacon as part of the paste series has not been seen, and this paper realizes the DNA computing model of logic gates with such a new method. This model has all the advantages that the previous general sticker model has. Also it is easier to operate and observe, and has higher reliability of results. With high specificity and sensitivity of molecular beacon itself, the logic gate has a more intuitive operation and observation.

Keywords: DNA Sticker Model; Molecular Beacon; Logic Gate; DNA Computing

用分子信标作为粘贴串的 DNA 计算模型的逻辑门

唐静静, 单静怡

安徽理工大学理学院, 淮南
Email: 595455103@qq.com

收稿日期: 2013 年 11 月 1 日; 修回日期: 2013 年 12 月 10 日; 录用日期: 2013 年 12 月 22 日

摘要: 近年来, 随着 DNA 计算及 DNA 计算机研究的发展, 运用粘贴模型和粘贴系统^[1]来实现布尔逻辑门已经成为 DNA 计算机研究的热点。而把分子信标作为粘贴串的一份子还未曾见到过, 此论文用这样一种新的方法来实现逻辑门的 DNA 计算模型。该模型同时拥有以往普通粘贴模型的优点。它更易于操作和观察, 使结果有更强的可靠性。由于分子信标本身所具有的高特异性和高灵敏度, 逻辑门有了更直观的操作和检测。

关键词: DNA 粘贴模型; 分子信标; 逻辑门; DNA 计算

1. 引言

1994 年, 美国加利福尼亚大学的 L. M. Adleman 博士用 DNA 计算的方法解决了有向 Hamilton 路问题 (Directed Hamiltonian Path Problem, 简记为 DHPP)^[2], 开创了 DNA 分子计算的新纪元。目前, 人们在研究 NP 完全问题的 DNA 计算模型的同时, 也在开始探索新的潜在的应用^[3]。构建基于 DNA 分子的逻辑电路就是其中一个研究方向。例如 2004 年刘文斌教授等人提出了一个基于吡啶二聚物的诱导发夹结构的逻辑“与非”门的 DNA 计算模型^[4]。

DNA 计算是一种以生物分子 DNA 或 RNA 作为媒介的计算, 运用生物分子之间的杂交反应来实现所求问题的解, DNA 计算机是以 DNA 计算作为基础的一种高科技, 而逻辑计算是计算机的基础, 有 DNA 分子逻辑和数字逻辑电路, 而数字逻辑电路是数字计算机体系结构的基础。计算机的硬件电路是由不同类型的逻辑门电路所构成的。与数字逻辑电路是计算机体系结构的产生基础类似, 基于 DNA 的分子逻辑电路是 DNA 计算机体系结构的产生基础和 DNA 计算机实现技术的硬件基础。基于此, 关于 DNA 逻辑计算的研究成为了 DNA 计算机研究的基础和热点, 现已有很多这方面的研究呈现出来且各有各的优缺点。此时我在他们的研究基础上发现了一个新的关于粘贴 DNA 计算模型的分子逻辑门, 即运用分子信标作为粘贴串的 DNA 分子作为计算模型的逻辑门来实现逻辑或门运算。

2. 粘贴模型的概念

粘贴模型^[1]最早由 Roweis 等人于 1996 年第二届 DIMACS 专题讨论会上提出的。是目前关于 DNA 计算理论研究中较为成功的限制型之一。

粘贴模型是一种基于分子操作和随机访问内存的一种 DNA 计算模型, 是一种通用的计算机系统。在一般的 DNA 计算中对于问题的编码要么采用单链, 要么采用双链, 而粘贴模型采用单链与双链的混合形式进行编码: 将一条长链划分为若干段, 其中有些是单链, 有些是双链, 单双链随机分布。若用单链表示数据 0, 用双链表示数据 1, 则一条这样的长链可用来表示一个二进制数据。由于单链和双链根据不同的生物操作可发生变化, 因而 DNA 链相当于一个随机数据存储单元。粘贴模型的优点是在运算过程中不需要 DNA 链的延伸, 也不需要酶的作用, 并且 DNA 链可以重复使用。文献[5]是粘贴模型建立的首篇开创性的文章。

关于编码问题, 在对二进制数的进行表述时, 粘贴模型(Sticker Model)使用两组基本的单链 DNA 分子。将长度为 n 个碱基的存储链(memory strand)分为 k 个区段, 每区段的长度为 m 个碱基(因此 $n = km$)。在计算的过程中, 每个区段实际定义为一个二进制位元(bit)(或等价于一个布尔变量)。同时设计 k 个不同的粘贴串(sticker), 每个粘贴串有 m 个碱基长, 并且与 k 个存储区中的一个且只有一个区具有 Watson-Crick 互补关系。如果一个粘贴串在给定的存储链上与所观察的位元发生退火反应, 那么该二进制所对应的位元在这条链中定义为开(on), 反之定义为关(off)。图 1 举例说明了这种表示方法。图中的从 5'到 3'的碱基序列定义为存储链, 长度 $n = 30$ bp, 共分为 2 个区段, 每区段 $m = 15$ bp, 3'到 5'的碱基序列定义为粘贴串, 长度为 15 bp。通常也可以将上述的 5'到 3'的序列用 3'到 5'的序列来定义, 这样, 其粘贴串就正好反过来, 从 5'到 3'; 两者的区别只是下位粘贴或上位粘贴。

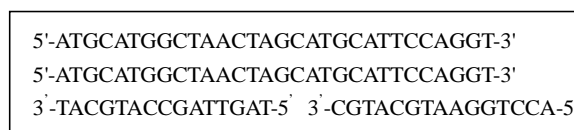


Figure 1. A store chain and the associated sticker string (collectively, the storage complex) represents a string
图 1. 一个存储链和所关联的粘贴串(统称存储合成物)表示一个串

图中上层代表一个存储链, 其两个位串全为关(off), 下层的存储合成物有两个退火的粘贴物。

要求在一个存储链中, 每个粘贴串恰好与 n 个子链中的一个互补。存储链的每个子链被视为一个位元的位置。如果一个粘贴串被退火于存储链的匹配子链上, 则这个特殊的子链为“关”(相当于数字逻辑电路中的低电平); 否则为“开”(相当于数字逻辑电路中的高电平)。因此, 存储合成物可用来表示一个二进制数, 其中一个子链为“开”表示该位元的数为“0”, 其中一个子链为“关”表示该位元的数为“1”。

3. 分子信标的概念

分子信标是一种由寡聚核苷酸形成的发夹型分子(图 2)。它包括一个环(Loop)和一个干(Stem)结构, 其中环

由与靶分子互补的核酸碱基序列组成, 一般有 15~20 个碱基; 干为两列互补的碱基序列, 一般是 5~7 个碱基对。

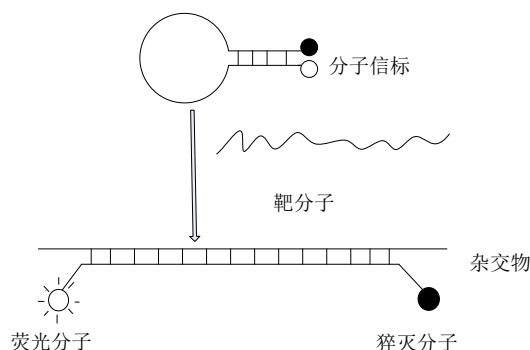


Figure 2. Molecular beacon combined with the target molecule
图 2. 分子信标与靶分子结合

它的 5'端与 3'端分别用荧光素和猝灭剂两种分子标记, 基本结构和反应原理如图 2。当分子信标自身互补杂交时, 它的两个末端靠近, 荧光分子和猝灭分子之间发生荧光共振能量转移。在这个过程中, 猝灭分子猝灭了荧光分子发出的荧光, 并以热的形式散发, 荧光消失。而当分子信标与靶分子互补结合时, 荧光分子和猝灭分子距离增大, 荧光恢复。

4. 分子信标作为粘贴串

粘贴串的设计:

构建分子信标作为粘贴串模型需要三个方面: 一是设计分子信标的环和干部分的 DNA 分子内部和链间互补配对编码问题; 二是设计存储链 DNA 的分子和长度问题; 三是整个系统 DNA 序列的设计, 着重考虑, 让分子信标的环和存储链互补配对, 让当分子信标 A 和 B 同时输入时, 分子信标 A 和 B 的干能很好的互补配对, 以避免干扰荧光的强度。另外由于运算方式和原理问题, 需要设计相应的分子信标, 然后再通过编码条件及限制条件, 筛选出合适的分子信标作为粘贴串, 设计过程中要重点注意避免碱基错配和 DNA 多聚体等问题。

分子信标的环是由与靶分子互补的核酸碱基序列组成, 现在把分子信标的环做成与存储链互补的碱基序列, 用它作为粘贴串, 当它与存储链上的相应位元接触时, 就会发生退火反应, 那么分子信标的发夹结构就会打开, 从而荧光得以恢复。设计过程及图如下:

- 1) 选择存储链作为已有的碱基序列: 5'-ATGCATGGCTAACTAGCATGCATTCCAGGT-3', 如图 3:

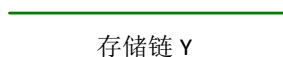


Figure 3. Storage chain Y
图 3. 存储链 Y

- 2) 制造与存储链互补的特殊分子信标的环 A: 3'-TACGTACCGATTGAT-5' 环 B: 3'-CGTACG-TAAGGTCCA-5, 如图 4:

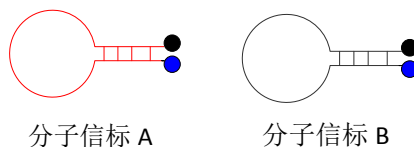


Figure 4. Molecular beacon A and Molecular beacon B
图 4. 分子信标 A 和分子信标 B

3) 让分子信标与存储链混合, 使分子信标与存储链的相应位元发生退火反应, 分子信标的发夹结构被破坏, 荧光出现。

4) 图形及反应过程如图 5:

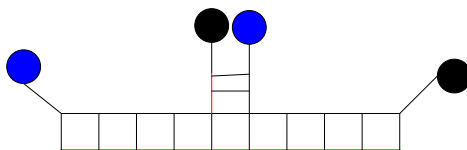


Figure 5. The reaction conditions of enter the molecular beacon A and B
图 5. 输入分子信标 A 和分子信标 B 时的反应状况

在图 5 中, 当同时输入分子信标 A 和 B 时, 分子信标 A 和 B 的环序列的碱基分子分别与存储链上相应位元发生碱基互补配对, 以致使分子信标的发夹结构打开, 荧光得以恢复。

为了使分子信标 A 和 B 相邻的荧光部分和猝灭部分能更好的靠近, 使其猝灭部分能很好的吸收荧光部分的光, 再此特意制作了两个分子信标的干部分的碱基也互补, 即设计了分子信标 A 和 B 干部分的碱基分别对应相同, 而环部碱基不同的两个分子信标。A 的淬灭分子和 B 的荧光分子就会靠的比较近, 于是猝灭的效果会更好, 淬灭了 B 的荧光恢复, 所以只有 A 的荧光得以恢复, 即荧光出现一次, 强度和单个分子信标和靶序列结合时的强度相当。

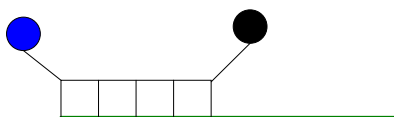


Figure 6. The output results when input molecular beacon A
图 6. 输入分子信标 A 时, 输出的结果

在图 4 中, 当只输入分子信标 A 时, A 的环序列碱基与存储链上的相应位元发生退火, 使发夹结构打开, 荧光恢复。即荧光出现一次, 强度和单个分子信标和靶序列结合时的强度相当。

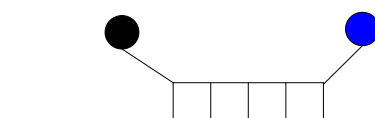


Figure 7. The output results when input molecular beacon B
图 7. 输入分子信标 B 时, 输出的结果

在图 7 中, 原理同 6 相似, 分子信标 B 环序列碱基与存储链上的相应位元发生退火反应, 从而分子信标的发夹结构被破坏而打开, 即荧光出现一次, 强度和单个分子信标和靶序列结合时的强度相当。

Figure 8. The output results when no input molecular beacons
图 8. 不输入分子信标时, 输出的结果

在图 8 中, 当没有分子信标输入时, 存储链仍旧一条单链, 则没有荧光出现。

在上述的 DNA 计算中, 我们采用要么有荧光出现, 要么无荧光出现, 而此模型采用有荧光出现的形式进行编码: 将一条长链划分为两段, 其中一段要么有荧光, 要么无荧光。若用无荧光表示数据 0, 用有荧光表示数据 1, 则一条这样的长链可用来表示一个二进制数据。由于有无荧光出现可以根据不同的生物操作可发

生变化, 因而 DNA 链相当于一个随机数据存储器。

分析以上反映图形, 图 5~8 都分别有分子信标输入, 结果都有荧光出现, 所以我们可以用此表示数据 1, 而图 8 无荧光出现, 表示数据 0。因此我们根据有无荧光实现了逻辑门的或门(见表 1)。

Table 1. The truth table of OR gate

表 1. 或门真值表

X	Y	AND
1	1	1
1	0	1
0	1	1
0	0	0

The truth table of the OR gate of DNA of molecular beacon as sticker string
分子信标作为粘贴串的 DNA 或门真值表

分子信标 A	分子信标 B	变化后的 Y
+	+	+
+	-	+
-	+	+
-	-	-

5. 结论

本文以粘贴模型和粘贴系统为基础展开的研究, 并结合分子信标的原理, 把分子信标当作“粘贴串”, 使之与存储链上的相应位元发生退火, 导致分子信标的夹夹结构破坏, 发出荧光, 从而荧光出现在存储合成物中。根据是否有荧光出现就可以判断真值, 实现了或门的真值计算。

本文操作简单, 灵敏度高, 且容易检测和观察。在这次的逻辑或门的实现过程中, 当同时输入分子信标 A 和分子信标 B 时, 需要对分子信标 A 和分子信标 B 的干给与相应的特别碱基。在实验过程中也可能会发生错误杂交, 需要在实验的基础上给与验证和改进。

自 DNA 计算诞生以来, 许多国家的著名研究机构和大学都相继开展了这一领域的研究工作, 很多研究小组及机构进行跨国界的相互协作, 并定期进行各种交流和讨论。从 1995 年召开第一届 DNA 计算国际学术会议以来, 每年都要举行一次 DNA 计算机国际学术会议。第一本关于 DNA 计算机的学术专著 DNA Computing-New Computing Paradigms 已于 1998 年由德国的 Springer 出版社出版^[15], 其中译本已于 2004 年 9 月由清华大学出版社出版。与此同时, 在诸如 DNA 计算应用于 NP 类问题解法, DNA 计算的运算问题, DNA 计算在生物实验方面的研究^[6-11], DNA 计算机的语言系统^[12-15], 联想记忆问题 DNA 计算在密码学上的应用等方面都取得了不少的进展。国际上最具权威的两个杂志“Nature”和“Science”每年都有关于 DNA 计算和 DNA 计算机方面的最新研究成果的报道。

项目基金

国家自然科学基金资助项目(61170172, 61073102)。

参考文献 (References)

- [1] 黄布毅, 许进 (2005) DNA 计算中若干理论问题的研究. 硕士学位论文, 华中科技大学, 武汉.

- [2] Adleman, L., et al. (1994) Molecular computations to combinatorial problems. *Science*, **266**, 1021-1024.
- [3] 刘文斌, 朱翔鸥, 王向红, 陈丽春 (2006) DNA 计算的研究进展. *电子学报*, **11**, 2053-2057.
- [4] Liu, W.B., et al. (2004) A new DNA computing model for the NAND gate based on induced hairpin formation. *BioSystems*, **77**, 87-92.
- [5] Hwang, K. (1991) *Advanced computer architecture*. McGraw Hill, New York.
- [6] 高琳, 马润年, 许进 (2002) 基于质粒 DNA 匹配问题的分子算法. *生物化学与生物物理进展*, **5**, 820-822.
- [7] 张连珍, 许进 (2004) 基于质粒的 DNA 计算模型研究. *计算机工程与应用*, **4**, 51-52.
- [8] Cukras, A., Faulhammer, D., Lipton, R., et al. (1998) Chess games: A model for RNA-based computation. *Biosystems*, **52**, 35-45..
- [9] Liu, Q.H., Guo, Z., Condon, A.E., et al. (1998) A surface-based approach to DNA computation. *Proceedings of the 2nd Annual Meeting on DNA Based Computers*, DIMACS: Series in Discrete Mathematics and Theoretical Computer Science. American Mathematical Society, Providence, *Journal of Computational Biology*, **5**, 255-267.
- [10] Liu, Q.H., Guo, Z., Condon, A.E., et al. (2000) DNA computing on surfaces. *Nature*, **403**, 175-179.
- [11] Liu, Q., Frutos, A., Wang, L., et al. (1999) Progress towards demonstration of a surface based DNA computation: A one word approach to solve a model satiability problem. *Biosystems*, **52**, 25-33.
- [12] Takahara, A. (2002) On the computational power of insertion-deletion systems. *Proceedings of 8th International Meeting on DNA Based Computers*, Sapporo, 10-13 June 2002, 139-150.
- [13] 许进, 董亚非 (2003) 粘贴 DNA 计算模型理论及应用. *科学通报*, **12**, 205-212, 299-307.
- [14] Liu, W.B., Wang, S.D. and Xu, J. (2003) Solving the 3-SAT Problem Based on DNA Computing. *Journal of Chemical Information and Computer Sciences*, **43**, 1872-1875.
- [15] Nishikawa, A., Hagiya, M. and Yamamura, M. (1999) Virtual DNA simulator and protocol design by GA. *Proceedings of the Genetic and Evolutionary Computation Conference*, Orlando, 13-17 July 1999, 1810-1816.