

DNA 分子链的电子局域性质及电导的研究^{*}

刘小良[†] 徐 慧 马松山 邓超生 郭爱敏

(中南大学物理科学与技术学院,长沙 410083)

(2005 年 12 月 6 日收到,2006 年 2 月 27 日收到修改稿)

将一维随机二元固体模型应用于 DNA 分子链,利用传输矩阵方法来研究系统电子态的局域性质并进而讨论系统的导电性质.对一个链长为 50000 个碱基对的 DNA 序列,数值分析了局域长度和电导随碱基对的摩尔百分数、本征能量和无序度的变化关系.结果表明,系统的局域长度和电导强烈地依赖于能量,在能带中心部分局域长度大于边沿部分,无序度也在一定程度上影响着局域长度,双方成反向变化的关系.对有限长度的 DNA 分子链,局域长度体现出明显的对碱基对摩尔百分数的依赖关系,对正常成分比例的随机 DNA 序列,在所有能量范围内系统的态都是局域的,系统的电导很小,系统呈现绝缘体行为.仅当一种碱基对在序列中所占比例很小时,系统中可以发现与特定分立能量值相对应的“扩展态”存在,处在这些态下的系统有较大的电导,但这些扩展态是不稳定的,在热力学极限之下会消失.

关键词:DNA 分子链,电子局域,局域长度,电导

PACC:8715,7125V

1. 引 言

近年来,对 DNA 分子链电学性质的研究已成为物理学及生物学领域的研究热点.其独特的电子输运性能,为其在生物、材料和信息领域的应用带来了许多潜在的应用前景^[1,2],如 DNA 芯片、生物合成以及制备生物纳米材料等^[3,4].

自 20 世纪 60 年代,Eley 等人^[5]认为 DNA 分子链可以看成一维系统以来,对 DNA 电学特性的研究就一直激起了许多科学工作者的浓厚兴趣^[6,7].在理论上,通常可将 DNA 分子链视为一维无序体系,其电子波函数应当是局域化的^[8],其局域化特性必然会对导电特性产生重大的影响.在实验上,Barton 等人^[9]用荧光猝灭法得出电子在 DNA 中高速传递而成为分子导体,还有人直接采用伏安法测量了 DNA 分子链的电阻,发现其导电特性非常复杂,得到的结果差别极大,可能呈现绝缘体、半导体或导体行为,甚至还可能是超导体.最近,宋骏等人^[10]研究了 DNA 分子链的能带结构及电子态,高绪团等人^[11]研究了位置涨落对 DNA 分子链电子结构的影响,他们对 DNA 分子链的电学性质进行了有益的探讨.

Albuquerque 等人^[12]考虑了 DNA 分子链中中程关联对电子局域化特性及电子输运的影响,得出了有益的结论.特别是 Liu 等人^[13]在对电子局域化特性的研究中,进一步考虑了多电子情况,研究了具有库仑相互作用的布洛赫电子波函数的局域化行为.

但在绝大部分对 DNA 分子链电学性质的研究中,对其电子波函数局域化特性、影响其电子波函数局域化程度的因素以及 DNA 分子链中碱基对组分或者说碱基对的摩尔百分数对其电学特性的影响未曾有深入的讨论.为弄清 DNA 分子链的导电机理,研究 DNA 分子链中的电子局域态特性,探讨 DNA 分子链中电子波函数局域长度及其影响因素,对探讨其导电特性无疑是非常有意义的,对阐明具有不同结构 DNA 片段的电子输运和生物功能也是很重要的.

为简化问题的研究,本文将 DNA 分子链抽象成一维随机二元固体模型,采用单电子紧束缚近似^[14-19],只计及最近邻交换积分,将处理无序系统的一般方法,如传输矩阵方法^[20,21]、负本征值理论以及无限阶微扰理论^[22-24],应用于这种系统,具体计算了链长为 50000 个碱基对的 DNA 分子链中电子的本征能量、局域长度、电导等,并探讨了格点能量、

^{*} 湖南省自然科学基金(批准号 05JJ40135)和中南大学文理研究基金(批准号 0601052)资助的课题.

[†] E-mail: xlliucsu@sina.com

无序度和碱基对的摩尔百分数对电子波函数局域长度及电导的影响.

2. 模型的建立

在构成 DNA 双螺旋链状结构的 4 种核苷酸中, 胞嘧啶(C)和鸟嘌呤(G), 腺嘌呤(A)和胸腺嘧啶(T) 相互配对构成碱基对 C-G 和 A-T, 而碱基对之间的堆积相互作用势和其他相互作用使相邻碱基对平面之间形成旋转而呈现双螺旋结构. 因此, 在理论上, 可将 DNA 双螺旋结构链看成是由两种基本碱基对按一定的分布规律(如随机、重复或关联分布)所构成的一维二元固体系统^[25]. 本文采用一维随机二元固体模型, 将每个碱基对抽象为一个格点, 与两类碱基对 C-G 和 A-T 相对应的格点位能分别记为 ϵ_{CG} 和 ϵ_{AT} , 把碱基对之间的相互作用用格点间的交换积分 V 表示. 为突出碱基对组分对 DNA 分子链导电性能的影响, 只考虑格点之间的最近邻交换积分, 并且将其视为常数. 在格点轨道波函数基 $\{|n\rangle\}$ 之下, 考虑周期性边界条件的一维随机二元 DNA 分子链, 其单电子紧束缚模型的哈密顿量可表示为

$$H = \sum_{n=1}^N \epsilon_n |n\rangle \langle n| + V \sum_{n=1}^N (|n\rangle \langle n+1| + |n\rangle \langle n-1|), \quad (1)$$

其中 N 为系统的链长, 即 DNA 的碱基对总数; ϵ_n 为格点位能, 以几率 p 和 $1-p$ 随机取 ϵ_{CG} 和 ϵ_{AT} ; $W \equiv |\epsilon_{CG} - \epsilon_{AT}|$ 代表格点位能的变化范围, 即代表了系统的对角无序度; V 为两最近邻格点间电子波函数的交换积分. 对于一维无序固体模型, 通常只考虑两个参量, 最近邻格点之间的相互作用 V 和系统的无序度 W , 这样控制参数就是比率 $W/|V|$. 典型情况下, V 可固定, 这样固定了系统的能量标度, 改变 W 就可研究不同的无序程度所产生的结果. 为简化运算, 我们取 $V = -1$, 即以 $|V|$ 作为能量标度, 而对于格点的位能, 参照文献^[26]的处理方法, 并利用势的相对性, 取 $\epsilon_{CG} = 0.5$, $\epsilon_{AT} = 0$ 进行数值计算.

能量为 E 的电子波函数 ψ 可按格点轨道波函数基展开, 即

$$\psi = \sum_n a_n |n\rangle, \quad (2)$$

其中 $|a_n|^2$ 代表在格点 n 处发现电子的几率. 借助传输矩阵方法^[22, 27], 可将系统的薛定谔方程写成如

下的形式:

$$E a_n = \epsilon_n a_n + V(a_{n+1} + a_{n-1}), \quad (3)$$

其递推形式为

$$\begin{bmatrix} a_{n+1} \\ a_n \end{bmatrix} = \begin{bmatrix} (E - \epsilon_n)V & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} a_n \\ a_{n-1} \end{bmatrix}. \quad (4)$$

3. 局域长度

参照文献^[28, 29]的做法, 采用如下的表达式来定义局域长度 L :

$$\frac{1}{L} = \lim_{N \rightarrow \infty} \frac{1}{N} \ln \left| \frac{a_N}{a_0} \right|, \quad (5)$$

选择归一化的初始条件 $(a_0, a_1) = (\sqrt{2}/2, \sqrt{2}/2)$ 对系统某一给定的能量 E 和一组随机给定的格点能量 $\{\epsilon_n\}$, 由(4)式的迭代过程可得出 a_N , 再由(5)式便可得出一维随机 DNA 分子链中电子波函数的局域长度.

通过对多种随机格点能量分布求平均, 我们讨论了在碱基对的不同摩尔百分数下系统局域长度随系统本征能量(图 1)及无序度(图 2)的变化关系.

图 1 中计算了不同碱基对组分下系统局域长度随本征能量的变化情况. 系统的能量分布范围为 $[\epsilon_{CG} - 2|V|, \epsilon_{CG} + 2|V|] \cap [\epsilon_{AT} - 2|V|, \epsilon_{AT} + 2|V|]$, 在我们这种参数设置之下, 能量范围约为 $-1.5 < E < 2.0$, 且系统的能量中心点随 C-G 的成分的增加向高能方向略有移动. 结果显示, 在能带的中心部分, 局域长度总体比边沿区域的要大, 说明中央部分电子波函数的局域化程度要弱, 这一结论和莫特的关于迁移率边的理论是一致的, 在莫特的理论中, 系统中的扩展态总是率先出现在能带的中心, 而局域态则往往出现在带尾处.

在实际的 DNA 分子链中, 碱基对 C-G 的摩尔百分数约为 0.6, 但考虑到 DNA 某些片段中 C-G 含量可以很低, 我们对一些极端的摩尔百分数也做了讨论. 在图 1(a)中, $p = 0.005$, 可以发现在一些分立的能量值之下存在大于系统大小的局域长度(如图中虚线所示), 预示着系统中有扩展态存在; 当 p 逐渐增大时, 如图 1(b)和(c)所示, 系统的局域长度迅速减少, 而且当 p 大于某一临界值 p_c 时, 扩展态消失, 所有能量范围上的态都是局域的. 显然在 C-G 的成分较小时, 可将 C-G 视为系统的“杂质”, 而将 A-T 视为系统的“本征格点”, 当“杂质”浓度增大时, 系统的无序程度增加, 其局域长度自然减少.

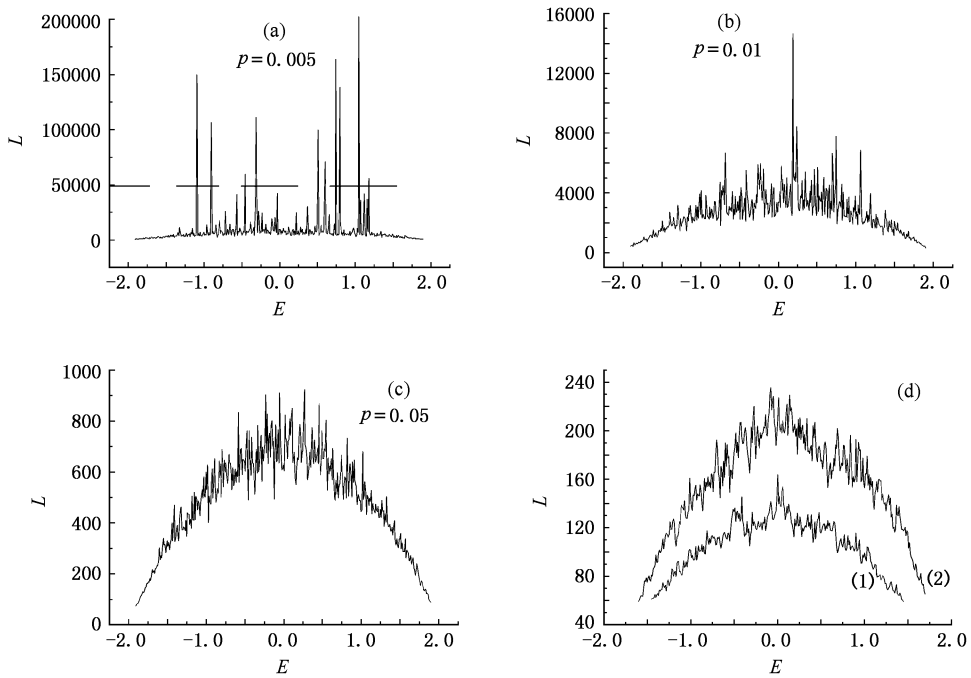


图 1 局域长度与本征能量的变化关系,其中 $N = 50000$, (a) $p = 0.005$, (b) $p = 0.01$, (c) $p = 0.05$, (d)(1) $p = 0.5$, (2) $p = 0.8$

图 1(d)中,C-G 的摩尔百分数更接近于实际 DNA 分子链的情况,可以看出,在正常的碱基对成分比范围内,系统中的电子态是局域的,电子的扩展范围最大只能达到 230 个碱基对的距离,根据最近的一些 DNA 分子链的导电性的实验测量和理论解释,随机的 DNA 序列应该是绝缘体,我们的结果支持了这一观点.当 p 由 0.5 再继续增大时,此时局域长度不再继续减少反而增大,因为当 $p > 0.5$ 时“本征格点”与“杂质格点”的地位互换,随着 p 值的增加“杂质”A-T 浓度减少,系统愈加接近由 C-G 作为本征格点的有序结构.

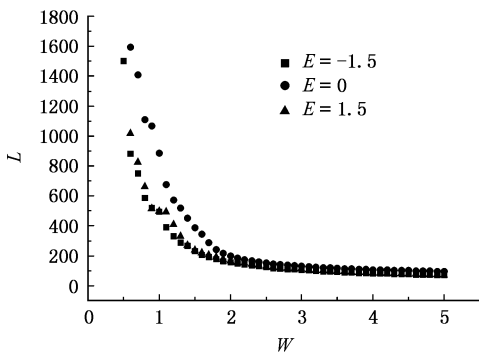


图 2 系统局域长度随无序度的变化关系,其中 $p = 0.01$, $N = 50000$

图 2 是能量本征值为 -1.5 和 1.5 时局域长度随无序度的变化情况,容易看出,局域长度随无序度的增加而减小,符合一般的无序系统电子输运理论,无序度的增加,使电子受到的散射作用加强,扩散受阻,局域化程度愈强.而当无序度趋向于零时,局域长度无限增大,对应于有序条件下的扩展态.比较三者的位置关系还可以发现,能量越接近能带中心,局域长度越大,这也印证了前文的结论.

由一般标度理论^[25,30],在一维无序系统中所有的电子态都是局域的,我们数值计算中所预言的“扩展态”的存在并不意味着对一般标度理论的违背,因为我们讨论的是有限大小的系统,对系统在热力学极限,即 $N \rightarrow \infty$ 时的行为,我们有必要进行标度分

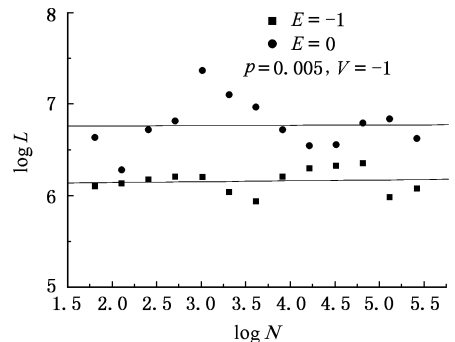


图 3 标度关系 $L \propto N^\xi$, $p = 0.005$

析.

图 3 中我们描述了不同能量值下系统的标度关系 $L \propto N^\xi$, 显示的结果分别对应于 $E = -1$ 和 $E = 0$ 的情况, 系统的大小范围为 $2^6 \sim 2^{18}$, 为避免数据的过度振荡, 此处的局域长度 L 为一个小能量区间 $\Delta E = 0.1$ 的平均值, 由拟合的实线可得出指数 $\xi = 0$, 即对所有的系统大小值, 局域长度保持恒定, 这样对一维随机 DNA 分子链, 在热力学极限下, 不会有新的扩展态的生成, 即使在有限尺寸下有扩展态的出现, 在此极限之下, 系统必将变成局域态.

4. 电导和电导率

利用兰多尔公式^[31], 对一维随机 DNA 分子链, 其电导可以表示为

$$g = (e^{2N/L} - 1)^{-1}, \quad (6)$$

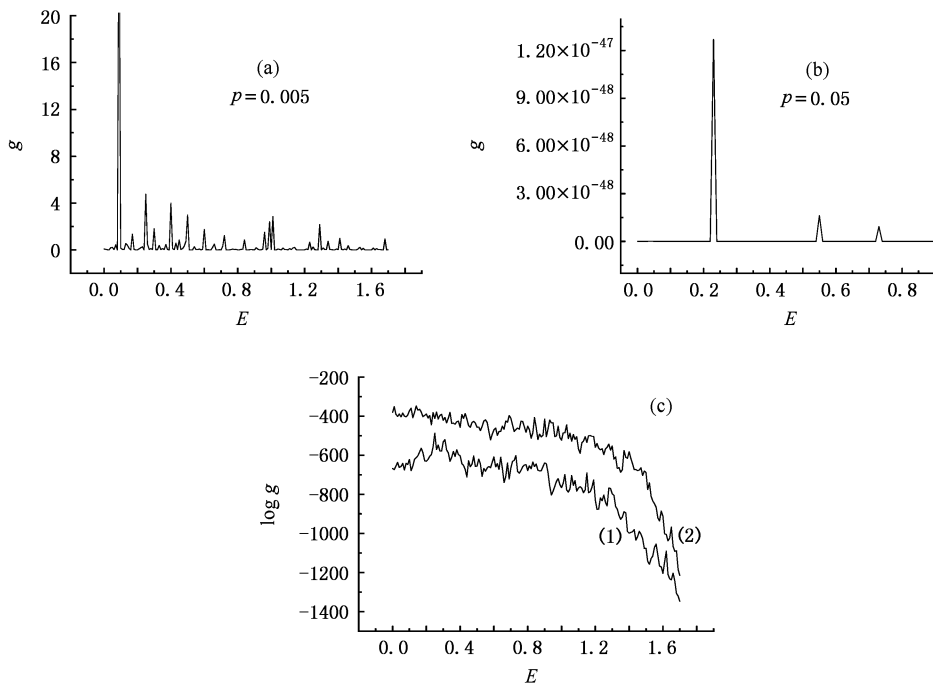


图 4 电导随能量的变化 (a) $p=0.005$, (b) $p=0.05$, (c) (1) $p=0.5$, (2) $p=0.8$

在最近的一次 DNA 电导率的实验测定中, Danny Porath 等人^[6, 32]在适当的条件下, 针对大约包含 1800 个碱基对的 DNA 片段测出其电导率约为 $1.0 \times 10^5 \Omega^{-1} \cdot \text{m}^{-1}$. 若令 DNA 双螺旋的直径为 2 nm, 碱基对间距离约为 0.34 nm, 50000 个碱基对的长度约为 1.7 mm, 对上述各图的峰值电导可得出相应的电导率, 图 4(a) 约为 $8.0 \times 10^5 \Omega^{-1} \cdot \text{m}^{-1}$, 图 4(b) 为 $10^{-42} \Omega^{-1} \cdot \text{m}^{-1}$ 的数量级, 而图 4(c) 对应的电导率更

其中 g 是以 $e^2 / (\pi \hbar)$ 为单位, 量纲为 1 的电导, N, L 分别为系统的碱基对数目和局域长度.

图 4 中, 我们给出了 DNA 分子链在不同碱基对组分下, 其电导对能量的依赖关系. 总体上可看出, 处于能带中心区域的电子态有较大的输运能力, 其电导要大于边沿区域能量值下的电导. 当 p 很小时, 如图 4(a), 在一些分立的能量处系统有较大的电导, 为前文中观察到的扩展态的存在提供了证据. 当 $p < 0.5$ 时, 系统的电导随 p 值的增大迅速减少, 但当 $p > 0.5$ 时, 系统的电导随 p 值的增大而增大. 在图 4(b) 中, 只在能量的中心部分存在几个峰值, 第一个峰值出现在 $E = 0.2$ 附近, 是因为系统的能带中心发生了移动所致. 图 4(c) 中, 我们算出的电导更小, 为避免数值计算的困难, 对电导取了自然对数. 对一维随机 DNA 分子链来说, 电导很小, 其通过电子或空穴跳跃来实现的导电能力是非常微弱的.

小, 数量级约为 $10^{-190} \Omega^{-1} \cdot \text{m}^{-1}$. 我们从理论上计算的电导率偏小, 说明我们的模型还有待进一步改进. 在我们的模型中, 格点的能量取为独立的无规变量, 没有考虑格点间能量的关联效应, 而现实中完全无序的系统是不存在的, 作为更贴近于现实的理论模型, 已经有些文献^[28]尝试将关联纳入模型之中, 期待得出更符合实际的关于 DNA 导电性的理论解释. 但一维随机 DNA 分子链模型在定性解释 DNA

的导电性方面是比较成功的,很多结论已经为实验所证实.

5. 结论和讨论

对一维随机 DNA 分子链,采用单电子近似下的安德逊模型,并且只考虑最近邻格点间交换积分,进行数值处理的结果表明,系统的局域长度和电导强烈地依赖于能量,在能带的中心部分局域长度总体要大于边沿部分,即在能带中心处,电子的局域化程度弱,系统具有较大的电导,而在带尾处,电子的局域化程度高,相应地系统的电导较小.

无序度也在一定程度上影响着局域长度,双方成反向变化的关系,当无序度趋向于零时,系统会向扩展态转化,对应于有序固体时的情况.

对有限长度的系统,局域长度对碱基对摩尔百分数有强烈的依赖关系,当一种组分的摩尔百分数 p 低于某一临界值时,一维系统中可以发现扩展态的存在,而当 p 大于此临界值时,对所有的能量范围系统的态都是局域的.

对于未考虑长程关联(包括格点能量之间的长程关联和格点交换积分的长程关联)的一维随机 DNA 分子链,在正常的摩尔百分数之下,系统的电子态是局域的,电导很小,可以认为是很好的绝缘体.而且从有限系统出发所得到的局域态是稳定的,通过对其标度关系的研究发现,系统的局域长度不会因为系统大小的增加而增加,在连续极限之下,不会有新的扩展态的生成.但当考虑系统中的长程关联时^[33,34],这个结论就不再成立,我们将在以后的工作中进一步讨论这个问题.

- [1] Murphy C J , Arkin M R , Jenkins Y *et al* 1993 *Science* **262** 1015
- [2] Arkin M R , Stemp E D A , Holmlin R E *et al* 1996 *Science* **273** 475
- [3] Dondliker P J , Holmlin R E , Barton J K 1997 *Science* **257** 1465
- [4] Lewis F D , Wu T , Zhang Y *et al* 1997 *Science* **277** 673
- [5] Eley D D , Spivory D I 1962 *Trans. Faraday Soc.* **58** 411
- [6] Porath D , Bezryadin A , Vries S D *et al* 2000 *Nature*(London) **403** 635
- [7] Kasumov A Y , Kociak M , Gueron S *et al* 2001 *Science* **291** 280
- [8] Pedro C , Pedro B G , Plamen C I *et al* 2002 *Nature*(London) **418** 955
- [9] Arhin M R , Stemp E D A , Barton J K *et al* 1996 *Science* **273** 475
- [10] Song J , Chen L , Liu D S *et al* 2004 *Acta Phys. Sin.* **53** 2792(in Chinese)[宋 骏、陈 雷、刘德胜等 2004 物理学报 **53** 2792]
- [11] Gao X T , Fu X , Song J *et al* 2006 *Acta Phys. Sin.* **55** 952(in Chinese)[高绪团、傅 雪、宋 骏等 2006 物理学报 **55** 952]
- [12] Albuquerque E L , Vasconcelos M S , Lyra M L *et al* 2005 *Phys. Rev.* **71** 021910
- [13] Liu W M , Liang J Q , Chui S T 2001 *Phys. Rev. B* **65** 033102
- [14] Liu X L , Xu H , Ma S S *et al* 2006 *Acta Phys. Sin.* **55** 2949(in Chinese) [刘小良、徐 慧、马松山等 2006 物理学报 **55** 2949]
- [15] Ma S S , Xu H , Liu X L *et al* 2006 *Acta Phys. Sin.* **55** 3170(in Chinese) [马松山、徐 慧、刘小良等 2006 物理学报 **55** 3170]
- [16] Liu X L , Xu H , Ma S S *et al* 2006 *Physica B* **383** 226
- [17] Liu X L , Xu H , Ma S S *et al* 2006 *Acta Phys. Sin.* **55** 2492(in Chinese) [刘小良、徐 慧、马松山等 2006 物理学报 **55** 2492]
- [18] Ma S S , Xu H , Liu X L *et al* 2006 *Chin. Phys.* **15** 190
- [19] Anderson P W 1958 *Phys. Rev.* **109** 1492
- [20] Davids P S 1995 *Phys. Rev. B* **52** 4146.
- [21] Dean P 1972 *Rev. Mod. Phys.* **44** 127
- [22] Kramer B , MacKinnon A 1993 *Rep. Prog. Phys.* **56** 1469
- [23] Xu H , Zeng H T 1992 *Acta Phys. Sin.* **41** 1666(in Chinese) [徐 慧、曾红涛 1992 物理学报 **41** 1666]
- [24] Xu H , Song Y P , Li X M 2002 *Acta Phys. Sin.* **51** 143(in Chinese) [徐 慧、宋 璞、李新梅 2002 物理学报 **51** 143]
- [25] Yu Z G , Song X Y 2001 *Phys. Rev. Lett.* **86** 6018
- [26] Roche S , Bicout D , Macia E *et al* 2003 *Phys. Rev. Lett.* **91** 228101
- [27] Francisco A B F M , Marcelo L L 1998 *Phys. Rev. Lett.* **81** 3735
- [28] Carpena P , Bernaola-Galvan P , Ch Ivanov P *et al* 2002 *Nature* **418** 955
- [29] Shima H , Nomura T , Nakayama T 2004 *Phys. Rev. B* **70** 075116
- [30] Landauer R 1970 *Phil. Mag.* **21** 863
- [31] Anderson P W , Thouless D J , Abraham E *et al* 1980 *Phys. Rev. B* **22** 3519
- [32] Fink H W , Christian S 1999 *Nature* **398** 407
- [33] Xiong S J , Zhang G P 2003 *Phys. Rev. B* **68** 17420
- [34] Rodriguez A , Malyshev V A , Sierra G *et al* 2003 *Phys. Rev. Lett.* **90** 027404

The localized properties of electronic states and conductivity of DNA sequence^{*}

Liu Xiao-Liang[†] Xu Hui Ma Song-Shan Deng Chao-Sheng Guo Ai-Min

(College of Physical Science and Technology , Central South University , Changsha 410083 , China)

(Received 6 December 2005 ; revised manuscript received 27 February 2006)

Abstract

According to the tight-binding model of the single electron , we establish a one-dimensional random binary solid model for DNA molecules . Using the transfer-matrix approach , we calculate numerically the localization length and the conductivity of a DNA sequence with 50000 nucleotides base pairs . The results show that the random DNA sequence is an insulator with very small conductivity , that the localization length and conductivity strongly depend on energies , and there exist greater localization lengths at the band center than at the band edge , which implies there is higher conductivity at the band center . The results also show the localization lengths are affected by disorder degrees to some extent and decrease with the increasing of the degree of disorder for all energies . Especially with finite-size DNA sequence , the localization length and conductivity have obvious effect on nucleotide base pairs ' mol percentage .

Keywords : DNA sequence , electronic localization , localization length , conductivity

PACC : 8715 , 7125V

^{*} Project supported by the Natural Science Foundation of Hunan Province of China (Grant No.05JJ40135) and the Arts and Science Fund of the Central South University of China (Grant No. 0601052).

[†] E-mail : xlliucsu@sina.com