

DNA 单分子弹性理论*

欧阳钟灿[†]

(中国科学院理论物理研究所 北京 100080)

摘要 随着单分子操纵技术的发明与发展,人们已经可以对单个生物大分子施以力或力矩,并测量它们的物理性质. DNA 单分子的力学实验表明,在分子尺度上理解生物大分子的生化过程,力与能量是同等重要的结构与功能参数. 一个梯子模型被用来描述双链 DNA 的外力拉伸曲线,在这个模型中, DNA 是由许多碱基对(梯子的横杆,横杆之间存在吸引势)连接两条聚核苷酸虫链(梯子的两侧)形成的高分子. 利用路径积分法得出的理论曲线与实验曲线吻合得很好. 对于单链 DNA,用分立的杂化高分子链统计理论的母函数方法来计算其弹性行为,得出与实验相符合的外力引起的解链相变结果. 此外,对于抑瘤蛋白 p53 识别序列 DNA 微环弹性进行分析,发现其弹性模量只是通常随机序列的三分之一.

关键词 生物大分子,单分子,DNA 力学

Elastic theory of single DNA molecules

OU-YANG Zhong-Can[†]

(Institute of Theoretical Physics, Chinese Academy of Sciences, Beijing 100080, China)

Abstract With the development of single molecule manipulation, the properties of single biological macromolecules can be tested by applying a force or torque to it. Mechanical experiments suggest that both force and energy are equally important structural and functional elements for the understanding of bio-chemical processes in biological macro-molecules. A ladder model where the double strand DNA consists of a number of base pairs (parallel rungs of the ladder with attractions between them) linking to two worm-like nucleotide chains (two long sides of the ladder) is presented to describe the force-extension curve of DNA. The theoretical curve obtained by the path integral method agrees well with experimental results. The generated function method in the hybridized statistics theory of polymer chains is used to calculate the elastic behavior of single strand DNA, and the theoretical results are consistent with the force-induced unzipping phase transition observed in experiments. Moreover, it is found that the elastic modulus for DNA sequences determined by p53 (the tumor-suppressor protein) is about one-third of the value for random-sequence DNA.

Key words bio-polymer, single-molecule, DNA mechanics

1 引言

20 世纪 90 年代以前,对细胞内生物大分子的物理、化学性质的研究,都是根据集群(ensemble)测量得出的,这只是反映大量(bulk)分子的平均行为. 最近十年,这种“试管生物物理”有了革命性的进展,由 20 世纪 80 年代物理学家发明的单分子操纵技术,如由隧道扫描显微镜(STM)开发出来的原子力显微镜(atomic force microscope, AFM)、光镊(optical tweezers)、流场拖曳(hydrodynamic drag)、磁钳(magnetic tweezers)等,人们已经可以对单个生物大分子施以力或力矩(如磁钳),并测量它们的物理性质(如 DNA 弹性、蛋白质的力学变性等)及力学生化反应(如分子马达). 分子马达是那些能够利用化学能来做功的生物大分子的统称,它们包括肌球蛋白

tical tweezers)、流场拖曳(hydrodynamic drag)、磁钳(magnetic tweezers)等,人们已经可以对单个生物大分子施以力或力矩(如磁钳),并测量它们的物理性质(如 DNA 弹性、蛋白质的力学变性等)及力学生化反应(如分子马达). 分子马达是那些能够利用化学能来做功的生物大分子的统称,它们包括肌球蛋白

* 2003-09-08 收到

[†] E-mail: oy@itp.ac.cn

白(myosin)、驱动蛋白(kinesin)、RNA 聚合酶(RNA - polymerase)和拓扑异构酶(topoisomerases)等. 1993—1994 年期间,三个小组(Block S, Spudich J 和 Yanagida T)首次测量肌球与驱动蛋白的单分子马达力信号. 1997 年, Kinoshita K 领导的小组看到了单个 ATP 酶($F_1 - ATPase$)分子中的 γ - 亚单元在“生物燃料分子”三磷酸腺苷(ATP)供应下的“步进电机”式转动,即 360° 是由三次 120° 完成,而 120° 转动是由 90° 加 30° 两个档级进行的. 这些单分子实验克服以往“试管试验”的两个根本性弱点,静态不整(static disorder)与动态不整(dynamic disorder),而给出分子实时行为与性质的分布,避免了对集群测量苛刻的同步(synchronization)要求. 利用单分子技术研究酶的催化反应已可以把底物结合、水解或其他催化步骤区分开来. 酶促变量(如催化速率等)作为时间的分布函数的积分同时也给出了集群的性质. 虽然大多数集群性质可由其单分子数据平均得到,但反之不成立. 如 DNA 的解链(unzipping)、蛋白质的折叠(folding)是不可能由集群性质得到. 除了单分子操纵,单分子实时视见(visualizing in real time)也是近几年单分子生物学的发展热点. 在细胞活体(in vivo)研究中,荧光蛋白(GFP 和 DsRed)的发现开辟了荧光单分子检测(fluorescence single - molecule detection)的新领域,并涉及物理光学的最新技术,如近场扫描光学显微镜(near - field scanning optical microscope)和广野显微镜(wide - field microscope)等.

2 皮牛顿力学

生物大分子(如分子马达)运动或转动主要靠 ATP 水解获得能量. 一个 ATP 分子水解能约 $20k_B T$, 这里 k_B 是玻尔兹曼常数, T 是温度, $k_B T = 4 \times 10^{-12} \text{ J}$, 分子马达移动尺度约 10 nm ($1 \text{ nm} = 10^{-9} \text{ m}$), 因此一个 ATP 驱动的力约为 5 pN (皮牛顿, $1 \text{ pN} = 10^{-12} \text{ N}$). 在 DNA 解链实验中,需要打开互补碱基对的氢键(hydrogen bonding),在蛋白质变性(去折叠)实验中,涉及氨基酸残基间“非共价”的键,以及疏水(hydrophobic)作用;它们所需的力约在 100 pN 范围. 蛋白质分解、DNA 和 RNA 酶切涉及断开共价键,所需的力是最强的单分子力,约 1000 pN . 但这对于无生命的固态物理实验,只相当于在 0.1 nm 晶格做 1 eV 功所需的力. 因此说,单分子生物物理所需的力是非常微弱的力,其检测实现要晚于固态物理

是可以想见的.

在对 DNA、RNA 拉伸时,这些生物高分子的一端固定在固体表面,另一端联结在一种力敏感单元(force sensor),它们可由 AFM 或光镊直接施力操纵,它们可能是微米大小的颗粒或悬臂(cantilever),测量其位移便可确定生物大分子所受的力. AFM 的悬臂可测量到 1 毫秒时间内 0.1 nm 的位移及大于 10 pN 的力. 一般的 AFM 玻璃微丝在体积上大于悬臂,因此时空分辨率不如后者,但其测力精度可达到皮牛顿或 0.1 pN . 光镊可测到亚秒时间内的 pN 力与 nm 的位移. 由巴黎高等师范学院统计物理实验室发明的磁镊技术的力敏单元是一种磁性颗粒,因此垂直移动磁场或旋转磁场可同时对生物大分子施加力与力矩. DNA 单分子的力学实验表明,在分子尺度上理解生物大分子的生化过程,力与能量是同等重要的结构与功能参数^[1]. 例如,在基因调控与表达时,酶对 DNA 要施加皮牛顿的力才能完成其生物功能^[2]. 因此理解 DNA 单分子在外力作用下的弹性应变成成为理论生物物理的一个挑战性的课题.

3 DNA 单分子弹性理论

对双链 DNA(dsDNA)、单链 DNA(ssDNA)以及 RNA 二级结构的单分子力学实验的理论解释已带来高分子物理的一场革命,生物高分子不同于传统高分子,具体表现在:序列的非均匀性、序列之间相互作用的特异性、以及序列高级结构引起的序列之间的“远程作用”. 这里的“远程”不是空间意义上的远距离作用,而是由于高分子链随机弯曲而使沿链相隔甚远的序列在空间变成近邻的相互作用. 这种复杂相互作用,使得生物大分子的单分子力学(无论是平衡还是远离平衡)都带来一些基本的统计物理新问题. 但不管怎么说,一条高分子链的自由能($F = U - TS$)是由内能 U 与熵能($-TS$)组成的,这里 T 是温度, $S = k_B \ln \omega$ 是构形熵, ω 是构形数. 如果单分子力还不能引起共价键的破坏,则其影响的仅是熵能的改变. 显然,如果链的首尾端距很小,这时的链有无数的构形(S 大,其自由能低)如果首尾端距增大,则构形数必定减少(在极端的直链情况下构形只有一种),即自由能升高. 因此用单分子力拉伸生物高分子的过程会像拉弹簧一样增加链的自由能,因此 $-TS$ 也叫熵弹性. 由于生物大分子的序列有关的复杂性,这种熵弹性已经不是诺贝尔奖得主 Flory 发展起来的传统高分子理论,如高斯链

(Gaussian chain), 自由连接链(free jointed chain)及虫链(worm - like chain)等模型所能描述. 例如, ds-DNA 单分子拉伸实验曲线在 70pN 左右出现一个上述理论无法解释的阶跃伸长平台, 其伸长量达到 1.7 倍 DNA B 形式的长度, 被称之为超伸展形式, 即 S 形式^[3]. DNA 的 B 形式即沃森与克里克于 1953 年发现的 DNA 双螺旋标准形式. 软物质倡导者、诺贝尔物理学奖得主 de Gennes 在对朱棣文的 DNA 单分子流场拉曳实验结果评论时, 特别以“分子的个人主义(molecular individualism)”刻画了这种不可捉摸的弹性行为^[4].

为描述双链 DNA 外力作用下的 B→S 结构相变, 我们于 1999 年提出 dsDNA 的梯子模型^[5]. DNA 是由许多碱基对(梯子的横杆)连接两条聚核苷酸虫链(梯子的两侧)形成的高分子; DNA 的双螺旋结构是由于碱基对的堆积作用(吸引势)形成的. 由于单个 DNA 分子从宽度上看是纳米尺度的微观客体, 但从其长度(从微米到米)看是宏观客体(如人体细胞核的染色体在 B 形式约 2m 长, 而百合花染色体则达到 20m), 因此一个分子就是一个统计系综. 其力/拉伸曲线就是对 DNA 无数构形的统计平均. 如果把 DNA 每个构形当成运动粒子的空间轨迹, 把高分子的弧长与时间对应起来, 则 DNA 高分子在外力拉伸过程中可以用粒子在力场作用下量子力学的路径积分加以描述, 其格林函数满足一个有趣的薛定谔方程(时间变量换为弧长变量, 普朗克常数与 $k_B T$ 对应), 求出方程的基态波函数, 并对“伸长算符”进行波函数积分, 我们算出了符合实验的拉伸曲线(图 1), 得到国外同行的重视^[6]. DNA 的结构相变可以由“量子力学”波函数来描述是一个有深远意义的启示: 许多不可思议的基因复制过程或许可以对应着“量子激发态”的跃迁, 例如右旋 DNA 到左旋(z -DNA)的跃迁^[7].

在双链模型进展的基础上, 我们从 2000 年起转入更困难的力作用下的 DNA 单链 - 发夹结构相变研究, 这更涉及到 DNA 在基因复制与调控的结构相变过程. 这时的链是非均匀的(有单链和双链混合, 见图 2), 即 DNA 的每对碱基都可以处在未分离(发夹态)或打开(单链)两种状态之一, 即 DNA 可视为 2 种单元组成的杂化高分子链, 因此不能用处理双链的连续的路径积分格林函数方法描述, 我们转而是用描述分立的杂化高分子链统计理论的母函数方法^[8]来计算非均匀结构的单链 DNA 弹性行为, 得出与实验符合的外力引起的解链相变结果(图 3)^[9].

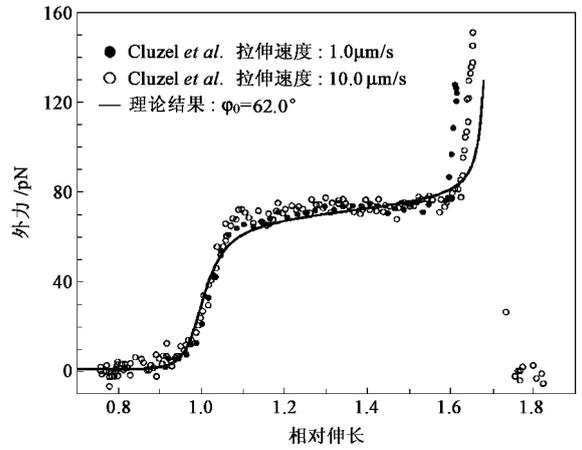


图 1 外力 - 拉伸曲线

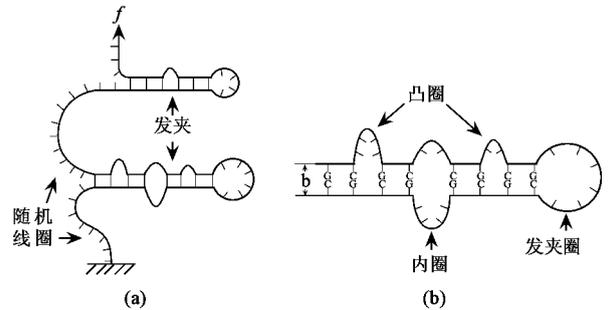


图 2 (a) 单链拉伸示意图 (b) 发夹结构

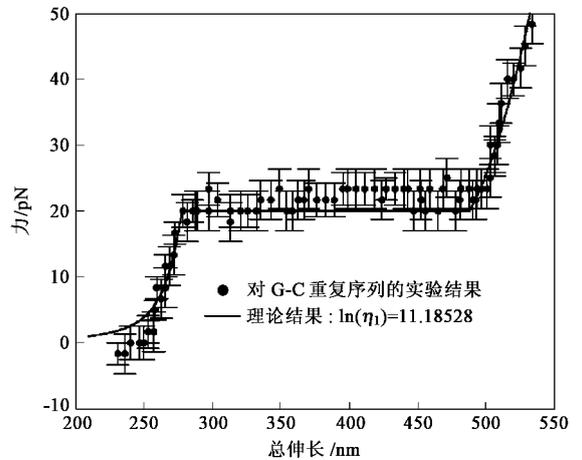


图 3 单链力 - 拉伸曲线

为更接近于生物实验, 我们进一步把溶液离子浓度德拜 - 休克尔(Debye - Hückel)静电势引入 DNA 弹性自由能, 利用蒙特卡罗方法模拟获得 DNA 分子的熔解(双链分离)相变依赖离子浓度的定量模型^[10]. 这个模型已为法、美实验工作者应用于分析他们的实验^[11].

4 肿瘤抑制蛋白 p53 识别序列 DNA 微环弹性

DNA 单分子弹性力学研究不仅仅是单纯描述这类系统的复杂熵弹性行为,而且要深入探讨其在生物功能中的作用,主要是 DNA 与蛋白质的相互作用。一个典型的例子是 Science 1993 年的年度分子抑癌蛋白质 p53 对细胞生命周期的调控。当生物体的细胞受到可破坏其 DNA 分子的病毒入侵时, p53 蛋白就被激活,被激活的 p53 蛋白能识别超过几十种不同 DNA 反应单元并捆扎到它们上面,使这些反应单元出现局域的大形变,导致受损染色体的再分裂过程终止,以确保病毒改变了的信息不会被传递和放大。实验表明, p53 和 DNA 结合的亲和度,特别是其结合专一性主要由相应 DNA 序列的可形变性而不是序列本身决定,把 p53 捆扎到 DNA 序列上去要在这些反应单元上产生很大的弯曲形变,所以这些特定的反应单元必须是非常柔软的。深入了解这种识别与捆扎的物理化学机理,一直是一个具有挑战性的问题。美国亚里桑那州立大学实验工作者把这些有关序列用 T4 DNA 连接酶催化链接,反应形成形状各异、长度为几百个碱基对(bp)的微环,然后用原子力显微镜(AFM)观测其构形。1997 年他们发现有一类序列(A - trac)的微环呈平面多边形(kink)变化[Nature , 1997 , 386 , 563], 本文作者的研究组在国际上首次用单分子弹性解析这种现象^[12], 引起对方重视, 随后对方通过 Internet 寄来 125 个另一类序列(Waf1)构成 168bp 长的微环的构形实验测量数据包, 要求我们作理论分析。这些构形呈现复杂的非平面弯翘变化。我们的理论工作分为两部分: 第一, 算出这 125 微环在几何上可分为两类, 其平均扭曲数分别为 0.109 ± 0.01 和 -0.098 ± 0.011 ; 第二, 从单分子两项弹性能(弯曲与缠绕)出发计算出平衡翘曲构形的平均扭曲数与上述观察

数据符合很好, 而且发现 Waf1 的弹性模量只是通常随机序列的 1/3, 这种由序列变异引起的柔软性是 p53 可以识别并绑上变异位点而阻止致癌染色体分裂的物理机制, 结果得到美方的高度赞许, 合作论文以中国科学院理论物理研究所作为第一单位, 对方物理系(Lindsay S M 小组)与生物系(Harrington R E 小组)为第二、三单位在重要的 JMB 杂志发表^[13]。

参 考 文 献

- [1] Allemand J F *et al.* Proc. Natl. Acad. Sci. USA , 1998 , 95 : 1415 ; Beyer M *et al.* Science , 1999 , 283 : 1727
- [2] Hegner M , Smith S B , Bustamante C. Proc. Natl. Acad. Sci. USA , 1999 , 96 : 10109 ; Léger J F *et al.* Proc. Natl. Acad. Sci. USA , 1998 , 95 : 12295 ; Wang M D *et al.* Science , 1998 , 282 : 902 ; Yin H *et al.* Science , 1995 , 270 : 1653
- [3] Cluzel P *et al.* Science , 1996 , 271 : 792 ; Smith S B , Cui Y , Bustamante C. Science , 1996 , 271 : 796
- [4] De Gennes P G. Science , 1997 , 276 : 1999
- [5] Zhou H J , Zhang Y , Ou - Yang Z C. Phys. Rev. Lett. , 1999 , 82 : 4560 ; Phys. Rev. E , 2000 , 62 : 1045 ; Zhang Y , Zhou H J , Ou - Yang Z C. Biophys J. , 2000 , 78 : 1979
- [6] Clausen - Schaumann H *et al.* Curr. Opin. Chem. Biol. , 2000 , 4 : 524 ; Krautbauer R , Clausen - Schaumann H , Gaub E. Angew. Chem. Int. Ed. , 2000 , 39 : 3912
- [7] Zhou H , Ou - Yang Z C. Mod. Phys. Lett. B , 1999 , 13 : 999
- [8] Lifson S. J. Chem. Phys. , 1964 , 40 : 3705
- [9] Zhou H , Zhang Y , Ou - Yang Z C. Phys. Rev. Lett. , 2001 , 86 : 356 ; Zhou H , Zhang Y. J. Chem. Phys. , 2001 , 114 : 8694
- [10] Zhang Y , Zhou H , Ou - Yang Z C. Biophys J. , 2001 , 81 : 1133
- [11] Dssinges M N *et al.* Phys. Rev. Lett. , 2002 , 89 : 248102
- [12] Zhao W , Zhou H , Ou - Yang Z C. Phys. Rev. E , 1998 , 58 : 8040
- [13] Zhou H *et al.* J. Mol. Biol. , 2001 , 306 : 227

封 面 说 明

定量生物学年会是 1933 年美国冷泉港实验室主任 R. Harris 发起的、后来成为分子生物学领域最有影响的学术会议, 许多诺贝尔奖项的论文, 如 1953 年 Watson 发现 DNA 双螺旋、1956 年 McClintock 提出转座基因、1966 年发现全套遗传密码, 最初都是在冷泉港年会上宣读的。封面照片是树立在冷泉港实验室的雕塑, 名为“盘旋的时代”, 是艺术家 Charles Jencks 2000 年的作品。

左上角的小图是当年 Watson 和 Crick 搭建 DNA 双螺旋模型时用的金属片和支架, 现存伦敦科学博物馆。两幅图皆摘自今年 Nature 杂志纪念 DNA 双螺旋 50 周年的专刊。

(北京大学物理学院 赵凯华 供稿)