

嫁接于平行板的受限紧密高分子链末端距分布函数的研究*

苏加叶¹ 章林溪^{2**}

(¹ 浙江大学物理系 杭州 310027) (² 温州大学物理系 温州 325027)

摘 要 采用剪除增加 Rosenbluth 方法 (Pruned-enriched-Rosenbluth method, PERM) 算法计算了嫁接于平行板的受限紧密高分子链的末端距分布函数. 由于受限紧密高分子链具有各向异性, 重点研究了平行板方向 x 轴上的分布函数 $P(x)$, 发现 $P(x)$ 可以表示为 $\ln[P(x)/P_m(x)]/ND^{-5/3} = a_0 + a_1 u + a_2 u^2 + a_3 u^3$ (其中 $u = x/ND^{-2/3}$). 这里 N 为链长, $P_m(x)$ 为分布函数 $P(x)$ 的最大值, 两平行板的间距为 $D+1$. 通过计算 $P(x)$ 的 Shannon 熵发现末端距分布函数 $P(x)$ 的 Shannon 熵可以用来描述高分子链受限的程度, Shannon 熵对平行板间距的变化非常敏感, 对于同一链长 N , $P(x)$ 的 Shannon 熵会随着 D 的增大而迅速减小, 超过临界值 D_c 会趋向一个定值, 即当 $D \geq D_c$ 时 Shannon 熵将趋于稳定, 也说明了此时受限条件对紧密高分子链影响非常小. 同时临界值 D_c 与链长 N 有关, $D_c \sim N^\lambda$, 其中 $\lambda = 0.543$, 并进行了一定的理论分析.

关键词 受限紧密高分子链, 末端距分布函数, Shannon 熵, PERM 算法

高分子链在受限情况下会表现出一些截然不同的性质, 这是因为受限高分子链的很多构象无法实现. 近年来, 受限高分子链的研究引起了许多研究者的兴趣, 如受限圆孔的 Monte Carlo 模拟^[1], 受限平行板的二维链的 PERM (pruned-enriched-Rosenbluth method) 模拟^[2] 及三维链的 Monte Carlo 模拟^[3], 受限圆形管道的高分子单链的 Monte Carlo 模拟^[4, 5]. 我们也做了一些工作, 如紧密高分子链通过一个小孔拉伸进行序列的预测^[6]、紧密高分子链在半径不同的圆桶中的拉伸^[7], 以及高分子在通过微孔时引起的自由能势垒等^[8]. 同时, 高分子链嫁接于不同的表面以及不同形状的容器中也会表现出不同的性质, 如嫁接于球形表面高分子多链的 Monte Carlo 模拟^[9, 10], 以及嫁接于三维圆形管道底部的高分子单链的 Monte Carlo 模拟^[1, 2]. 自由空间下无规行走链的末端距分布为高斯分布, 但在受限情况下其末端距分布将会发生改变, 显然这种分布与链所受到的限制大小有关, 即末端距分布函数里必定包含了限制空间的尺寸 D . 在三维限制中 D 可以是圆球的直径或是立方体的边长, 对于二维限制中 D 可以是无限长圆管或方管的直径. 紧密高分子是一类重要的高分子, 因为它是球形蛋白质的主要结

构^[11-13]. 紧密高分子与一般的高分子链的主要区别是, 紧密高分子链内部的原子之间存在较强的吸引相互作用, 因此它通常会形成比较紧密的结构. 这里采用两块无限大平行板对紧密高分子链的限制, 研究紧密高分子链末端距的横向分布, 并将传统信息理论中的 Shannon 熵运用到高分子链末端距分布函数的计算中, 找出末端距分布函数随受限大小 D 与链长 N 的关系.

1 计算方法

本文采用的模型如图 1 所示, 以两块无限大的平行板作为紧密高分子链的受限边界, 平行板的下表面为 $z = -1$, 平行板上表面为 $z = D$ (平行板的间距为 $D+1$), 即紧密高分子链的各个原子只能在 $z = 0, 1, 2, \dots, D-1$ 平面上行走, 两平行板以外的地方不能有原子存在, 紧密高分子链一端固定于坐标原点, 另一端为自由端, 图 1 中的 x 代表了自由端原子的 X 轴坐标值. 采用紧密高分子链的三维简立方格点模型, 平行板在这里只起到了纯粹的边界作用, 与紧密高分子链之间没有相互作用, 因此紧密高分子链体系的能量为:

$$E = \sum_{i < j} \epsilon_{ij} \Delta(r_i - r_j) \quad (1)$$

* 2006-04-10 收稿, 2006-06-21 修稿. 国家自然科学基金(基金号 20274040, 20574052)、教育部“新世纪优秀人才支持计划”(项目号 NCET-05-0538)、浙江省自然科学基金青年人才项目(基金号 R404047)和浙江省自然科学基金(基金号 Y405011)资助项目; ** 通讯联系人, E-mail: xzhang@hzncnc.com

即体系的能量全部来源于链内部的紧密接触相互作用能,这里 ϵ_{ij} 为第 i 和第 j 个原子之间的紧密接触能,取 $\epsilon_{ij} = -1$ (以 $k_B T$ 为单位)^[6-8],如果第 i 和第 j 个原子在空间上相邻但在高分子链上为非相邻原子则 $\Delta(r_i - r_j) = 1$,其它情况 $\Delta(r_i - r_j) = 0$. 这时紧密高分子链体系的配分函数可以表示为:

$$Z = \sum_i \exp(-E_i/k_B T) \tag{2}$$

式中 $\sum_i$ 为对紧密高分子体系的所有构象求和. 由于紧密高分子链的构象随着链长的增加呈指数递增,因此在本文中采用 PERM 算法^[14]. 在 高分子链生成过程中 PERM 算法会根据能量对高分子链行走进行控制,能量越低的方向选择的概率就越大. Grassberger^[2,3,14] 用 PERM 算法模拟了高分子链,并且他的结果可以证明 PERM 算法是模拟三维简立方格点模型最有效的算法.

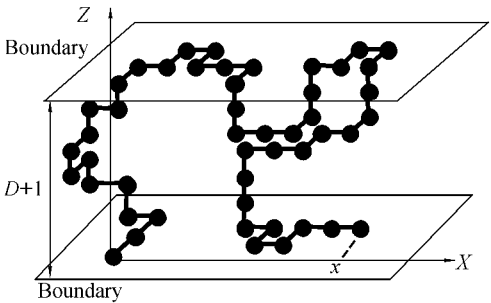


Fig. 1 A schematics of the grafted compact polymer chain confined between the two parallels
The distance of two parallels is $D + 1$, and no monomer can be adsorbed on the parallel

通过计算链的配分函数 Z 和链自由端落在 x 轴上各个点所对应的配分函数 Z_x ,可以得到链自由端落在 x 轴上各点的几率即:

$$P(x) = \frac{Z_x}{Z} \tag{3}$$

需要指出的是,在计算的时候把链自由端落在 x 负半轴上各点的概率都累加到 x 正半轴相应的点上去了,这样 $P(x)$ 满足归一化条件 $\sum P(x) = 1$. 在本文中,由于平行板平行于 $x - y$ 平面,且平行板是无限大的,由于对称性,紧密高分子链自由端落在 x 轴和 y 轴上各点是等效的,即 $P(x)$ 和 $P(y)$ 有着相同的行为,而 z 方向上由于平行板的限制,自由端只能落在 $z = 0$ 到 $z = D - 1$ 上,由于在 z 方向上的分布函数 $P(z)$ 极依赖于 D ,同时其

分布很集中,因此人们更喜欢研究 x 方向的末端距分布函数 $P(x)$ ^[1,2].

2 结果与讨论

2.1 紧密高分子链末端距分布函数 $P(x)$

计算了不同链长以及不同板间距时末端距的分布函数 $P(x)$. 首先讨论一种极限情况即 $D = 1$ 时. 这时平行板的位置在 $z = -1$ 和 $z = 1$,紧密高分子链只能存在于 $z = 0$ 的平面上,即紧密高分子链只在二维空间上行走,是三维链受限的极端情况. 图 2 给出了 $D = 1$ 时紧密高分子链末端距的分布函数 $P(x)$. 从图 2 中可以看出,各条曲线过了相应的极大值后 $P(x)$ 都均匀地减小,随着链长从 30 增大到 100, $P(x)$ 的峰值 $P_m(x)$ 从 0.184 减小到 0.112,而且峰值往右移动. $N = 30, 50, 70$ 时的峰值都出现在 $x = 1$ 的位置,而 $N = 100$ 时峰值出现在 $x = 2$. 还注意到随着链长增大 $P(x)$ 各点的值会更均匀化,它们的离散程度会减低,事实上这也是由于链长增大以后链受限变得更明显引起的. 图 3 分别给出了在同一个链长下不同板间距所对应的分布和在同一个板间距下不同链长所对应的分布. 如图 3(a) 所示,当链长 $N = 100$ 时, $P(x)$ 的峰值随着 D 的增大而增大, D 从 2 增大到 6 时, $P(x)$ 的峰值 $P(x)$ 也从 0.189 增大到 0.315. 同样随着 D 的减小 $P(x)$ 各点的值也会变地更加均匀,只是这种变化不象图 2 中那么明显,特别是 D 从 6 减小到 4 的时候几乎只有峰值在变,其它的点都变化不明显,只有 $D = 2$ 的时候才有较大的变化. 图 3(b) 的变化类似于图 2,随着 N 从 30 增大到 100,相应的峰值由 0.352 减小到

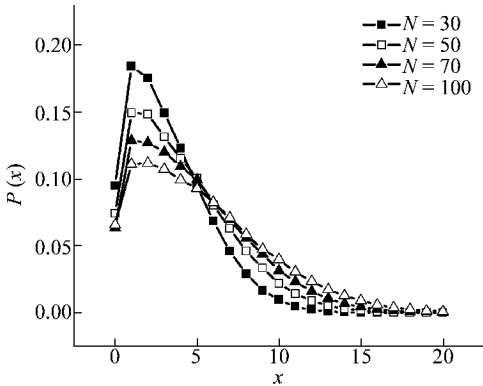


Fig. 2 The distribution function of end-to-end distance $P(x)$ as a function of x for the grafted compact polymer chain with $N = 30, 50, 70$ and 100, respectively
Here $D = 1$

0.245.与图 2 不同的是,图 3(b) 曲线的形状变尖了,并且在峰值过后 $P(x)$ 的值迅速减小, $P(x)$ 的值主要分布在 $x=0$ 到 $x=5$ 这样一个小范围内,也就是链的自由端大部分都落在了 0 到 5 这么小的区间里.

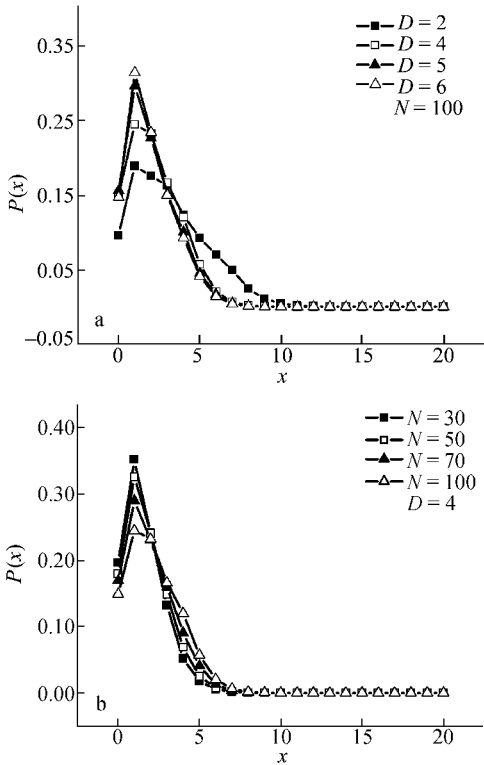


Fig. 3 The distribution function of $P(x)$ for various values of N and D
a) $N=100, D=2, 4, 5, 6$; b) $D=4, N=30, 50, 70, 100$

为了找到 $P(x)$ 分布与链长 N 及 D 之间的关系,将坐标经过了一系列变换,以 $u = x/ND^{-2/3}$ 作为新的横坐标,以 $F(u) = \ln(P(x)/P_m(x))/ND^{-5/3}$ 作为新的分布函数,其中 $P_m(x)$ 为 $P(x)$ 的最大值,图 4 给出了在新坐标下 $F(u)$ 随不同 N 和 D 的分布情况,发现 $F(u)$ 满足表达式:

$$\ln[P(x)/P_m(x)]/D^{-5/3} = a_0 + a_1 u + a_2 u^2 + a_3 u^3 \tag{4}$$

式中 a_0, a_1, a_2, a_3 都是常数,对于不同的 N 和 D 总可以通过调节这些系数达到最佳的拟合.比如对 $N=70, D=5$,这些系数可以拟合为:

$$F(u) = -0.052 + 0.927u - 15.69u^2 + 6.219u^3 \tag{5}$$

得到的相关系数为 0.997,拟合得非常好,如图 4 中的黑体曲线所示.因此表达式(4)包含了链长 N

和平行板间距 D ,它体现了高分子链末端距分布随链长和平行板间距的变化规律.

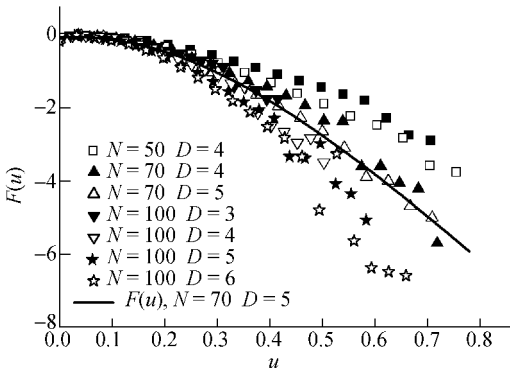


Fig. 4 The relative distribution $F(u) = \ln(P(x)/P_m(x))/ND^{-5/3}$ as a function of $u = x/ND^{-2/3}$ for various values of N and D

2.2 末端距分布函数的 Shannon 熵

在传统的信息理论中 Shannon 熵(H)可以表示为:

$$H = - \sum_{i=1}^n P_i \log P_i \tag{6}$$

其中 P_i 表示在由 n 个相互独立的事件组成的集合中第 i 个事件出现的概率, P_i 满足归一化条件 $\sum_i P_i = 1$. Shannon 熵平均地描述了整个事件集合中每个事件的不确定程度,事实上如果在事件集合中某一件事出现的概率为 1 其余为 0,则熵达到最小值 $H=0$,即此时体系是最确定的,如果 n 件事都以等概率出现,则熵达到最大值 $\lg n$,即此时的体系是最不确定的,所以 $0 \leq H \leq \lg n$. 在本文中,如果把紧密高分子链的自由端落在 x 轴上各个点看作是一个个相互独立的事件,则可以计算出这整个事件集合的 Shannon 熵,并且前面已经提到分布函数 $P(x)$ 满足归一化条件 $\sum_{x=0} P(x) = 1$, 所以得到末端距分布 $P(x)$ 的 Shannon 熵:

$$H(N, D) = - \sum_x P(x) \lg P(x) \tag{7}$$

注意到在不同的 N 和 D 下所得到的分布函数 $P(x)$ 是不同的,所以 Shannon 熵是 N 和 D 的函数.在前面分析中知道受限尺寸大小对分布函数 $P(x)$ 有影响,但不知道其影响的整体程度如何.引入 Shannon 熵的目的就是为了了解分布函数 $P(x)$ 的整体效果.这与传统意义上的热力学熵不一样,热力学熵是物理系统无序状态的描述.图 5(a)给出了在不同 D 时 H 与链长 N 的关系,从图

中可以看出, H 与 N 有很好的线性关系,同时直线的斜率与 D 有关,随着 D 的增大,斜率减小,最后斜率会趋向一个定值,所以不难得出以下关系:

$$H(N,D) = K(D)N + b \tag{8}$$

对于 $D = 2, K(2) = 5.5 \times 10^{-3}, b = 1.61$ 对于 $D = 4, K(4) = 3.5 \times 10^{-3}, b = 1.46$ 对于 $D = 6, K(6) = 2.6 \times 10^{-3}, b = 1.47$ 对于 $D = 10, 11, 12$ 三组数据所拟合出来的直线基本上都叠到了一起,对应于图中那条粗的黑线,此时 $K(10) = K(11) = K(12) = 2.1 \times 10^{-3}, b = 1.50$. 此外,还注意到了图 5(a) 中对于某一 N, H 的值会随着 D 的增大而减小,最后也将趋向于一个定值. 事实上这些结果也是很好理解的,从 2.1 节的讨论中已经知道了对于同一链长 N , 分布函数 $P(x)$ 的值会随着 D 的减小而趋于均匀化,即离散程度会降低,这必然导致了 Shannon 熵的增大. 而对于某一固定的板间距 $D, P(x)$ 的分布会随着 N 的增大而趋于均匀化,这正是导致 Shannon 熵线性增加的原因. 将方程(7)代入方程(8)会得到 $P(x)$ 分布所遵循的另一表达式:

$$-\sum P(x)\ln P(x) = K(D)N + b \tag{9}$$

这样便通过 Shannon 熵这个中间量再一次找出了末端距分布随着链长和平行板间距变化的关系. 虽然表达式(4)和表达式(9)的形式有所不同,但都描述了末端距分布规律,从数学上来说表达式(4)的形式更为解析,因为它体现了变量 x 与函数 $P(x)$ 之间的一一对应关系,但两个表达式都很好地体现了末端距分布 $P(x)$ 与 N 和 D 的一一对应关系.

从上面的讨论已经知道了随着 D 的增大,斜率 $K(D)$ 会趋于一个定值,而对于同一个链长 N, D 增大后 Shannon 熵也会减少并趋于一个定值,事实上这是由于 D 增大后平行板对紧密高分子链的限制作用逐渐减弱引起的,因为 D 增大后紧密高分子链上能够接触到平行板上附近的原子数会急剧减少,平行板上表面对高分子链行走的限制作用已经很弱,此时分布函数 $P(x)$ 会趋于分散,不会再随着 D 的增大而改变了,因此 Shannon 熵也就趋向了一个定值. 在理论上对于普通无规行走链,当限制空间尺寸 $D < R_F \sim N^\nu$ (R_F 为 Flory 半径,在三维中 $\nu \approx 0.587$) 时,高分子链才会受到边界的限制作用,但在本文中将会发现 D 的数值

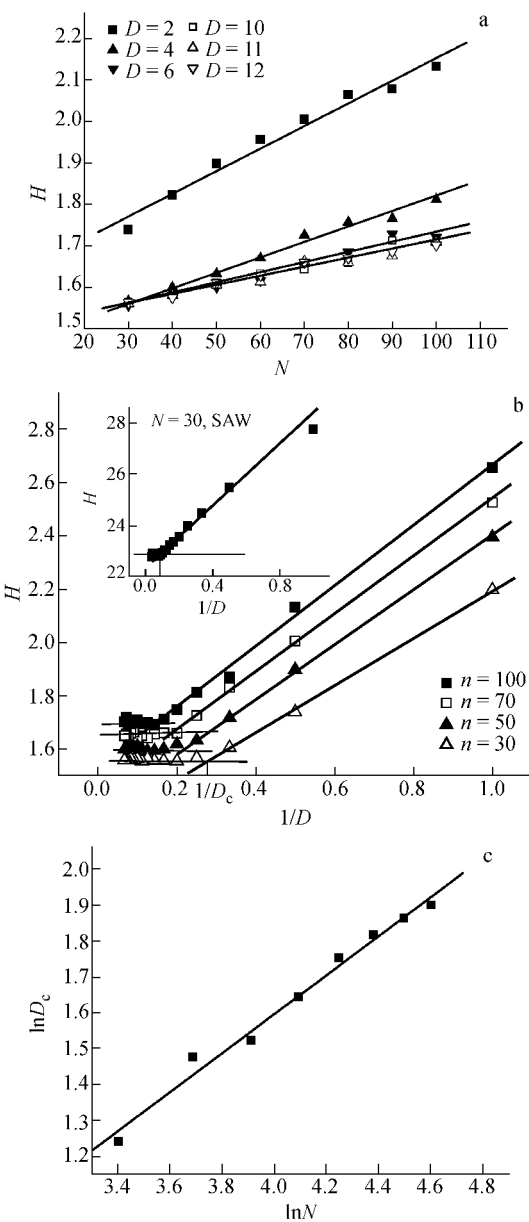


Fig. 5 The Shannon entropy of end-to-end distribution $P(x)$ as a function of (a) N for $D = 2, 4, 6, 10, 11, 12$, (b) D for $N = 30, 50, 70, 100$ and (c) the lg-lg plot of the transition D_c against N (c)

应当取得更小一些紧密高分子链才会较为明显地受到平行板的限制作用,这就是本文中引入 Shannon 熵这个物理量的另一目的. 凭借 Shannon 熵对 D 变化所具有的敏感性,当 D 增大到某一数值即 D_c 后 Shannon 熵的值基本上不变了,因此,这里认为 D_c 就是紧密高分子链受到明显限制作用的临界点,即 $D \leq D_c$ 的时候高分子链才会受到明显的限制作用. 图 5(b) 给出了 H 和 D 的关系,这里为了更准确地找出 D_c ,将横坐标轴定为 $1/D$. 从图 5(b) 中可以看出,随着 D 的增大即 $1/D$

的减小, Shannon 熵会迅速地减小并趋于一个定值, 比如对于 $N = 30$, 图中拟合出了 H 迅速减小和 H 趋于一个定值两条直线, 两直线的交点便确定了 D_c 值. $N = 30, 50, 70, 100$ 时 D_c 分别为 3.5, 4.6, 5.8, 6.7, 出现小数并不奇怪, 因为 H 随着 D 的增大而减小总是一个渐进的过程, 本文所确定的 D_c 只是这个过程中最为明显的点, 可以认为当 D 小于或等于 6.69 的时候 $N = 100$ 的链就会受到较明显的限制作用. 为了比较链段之间相互作用对 Shannon 熵的影响, 在图 5(b) 的小图给出了自避行走链 (self-avoiding walk chain, SAW, $\epsilon_{ij} = 0$) 的 Shannon 熵与 D 的变化关系, 可以看出相互作用对 Shannon 熵的影响也是很大的. 在相同的 D 下, 自避行走链的 Shannon 熵的数值明显要大一些, 这可以反映出在无相互作用下, 高分子链的行走显得更加离散. 另外, 小图中的值为 12.0, 明显要大于相同链长紧密高分子链所对应的数值 3.5, 这说明确实对自避行走链, 平行板对其行走的限制区域更广, 这也是很合理的, 因为紧密高分子链会形成比较紧密的结构, 所以它的空间尺寸必然要比自避行走链要小的多, 因此只有 D 取得较小的时候才会受到限制作用.

通过上面讨论可以确定 D_c 这样一个临界点, 使得本文能够定量判断限制空间该取多大才能对紧密高分子链产生较为明显的限制, 这一点今后研究紧密高分子链受限的问题很有帮助. 同时, 本文由于计算机的运算能力有限, 仅计算了链长 100 以内的紧密高分子链, 但又非常希望能够通过短的链计算对长的紧密高分子链进行预测, 为此, 本文进一步找出了链长 N 和临界点 D_c 之间的关系, 图 5(c) 给出了它们之间的双对数关系, 从图中得到了它们的近似关系:

$$D_c = \mu N^\lambda \tag{10}$$

其中 $\mu = 0.563, \lambda = 0.543$, 注意到 $\lambda < \nu = 0.587$, 这个结果非常有意思, 因为前面已经提到了, 在理论上限制空间尺寸 $D < R_F \sim N^\nu$ 的时候高分子链才会受到边界的限制作用, 而本文所得到的临界值 D_c 略小于自避行走链的理论值 (因为本文的研究对象是紧密高分子链, 紧密高分子链本身在结构上比自避行走链要紧密). 最后计算了不同 D 下不同链长的均方末端距, 结果见图 6. 不同链

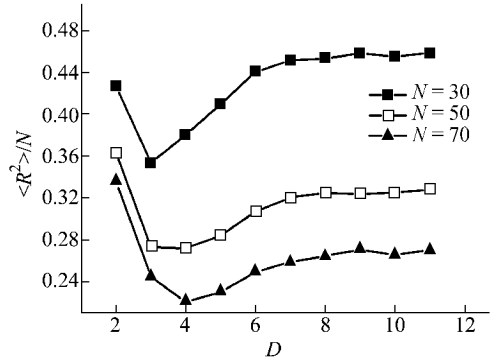


Fig. 6 The mean-square of end-to-end distance per bond $\langle R^2 \rangle / N$ as a function of D for $N = 30, 50$, and 70

长的 $\langle R^2 \rangle / N$ 随 D 变化的共同趋势是, 随着 D 的增加 $\langle R^2 \rangle / N$ 值都会先较为迅速地减小到一个最小值然后再缓慢地增加最后趋于一个平稳的值. 并且发现不同链长 $\langle R^2 \rangle / N$ 的最小值都出现在相应 D_c 的附近而略小于 D_c , 对于链长 30, 70, $\langle R^2 \rangle / N$ 的最小值所对应的 D 分别是 3, 4, 而链长 50 的时候, $\langle R^2 \rangle / N$ 的最小值出现在 D 介于 3 和 4 之间, 不妨取为 3.5, 本文认为 $\langle R^2 \rangle / N$ 的最小值都出现在 D_c 的附近并非偶然的, 这与 D_c 必定存在某种关系, 将在今后的工作中进一步去研究.

REFERENCES

- 1 Chen B C. J Chem Phys, 2005, 123 074702-1 ~ 7
- 2 Hsu H P, Grassberger P. Eur Phys J B, 2003, 36 209 ~ 214
- 3 Hsu H P, Grassberger P. J Chem Phys, 2004, 120 2034 ~ 2041
- 4 Sotta P. J Chem Phys, 2000, 113 6966 ~ 6973
- 5 Sotta P. J Chem Phys, 2000, 112 1565 ~ 1573
- 6 Zhang L X, Su J Y. Polymer, 2006, 47 735 ~ 741
- 7 Zhang L X, Chen J. Polymer, 2006, 47 1732 ~ 1740
- 8 Sun T T, Zhang L X, Su J Y. J Chem Phys, 2006, 125 (3) 034702-8
- 9 Toral R, Chakrabarti A. Phys Rev E, 1993, 47 4240 ~ 4246
- 10 Lindberg E, Elvingsson C. J Chem Phys, 2001, 114 6343 ~ 6352
- 11 Chan H S, Dill K A. Macromolecules, 1989, 22 4559 ~ 4573
- 12 Zhang Linxi (章林溪), Jin Jinsheng (金进生), Wang Xianghong (王向红), Ye Gaoxiang (叶高翔), Zhao Delu (赵得禄). Acta Polymerica Sinica (高

分子学报),2002,(4) 497 ~ 501

13 Chen Jin(陈进),Li Jing(李靖),Zhang Linxi(章林溪),Jiang Zhouting(姜舟婷),Xia Agen(夏阿根),Ye Gaoxiang(叶高翔),Zhao Delu(赵得禄).Acta Polymerica Sinica(高分子学报),2003,(2) 171 ~ 174

14 Grassberger P.Phys Rev E,1997,56 3682 ~ 3641

DISTRIBUTION OF END-TO-END DISTANCE OF A GRAFTED COMPACT
POLYMER CHAIN CONFINED IN TWO PARALLELS

SU Jiaye¹, ZHANG Linxi²

(¹ Department of Physics, Zhejiang University, Hangzhou 310027)
(² Department of Physics, Wenzhou University, Wenzhou 325027)

Abstract The end-to-end distance distribution of a grafted compact polymer chain confined between two parallels is studied by using the PERM (pruned-enriched-Rosenbluth method) algorithm. Because of the difference in different directions of confined polymer chains, in this paper we just consider the distribution function $P(x)$ in the x -axis direction, and it can be expressed as $\ln[P(x)/P_m(x)]/ND^{-2/3} = a_0 + a_1 u + a_2 u^2 + a_3 u^3$, here $u = x/ND^{-2/3}$, N is the chain length, $P_m(x)$ is the maximum value of $P(x)$, $D+1$ is the distance between the two parallels. On the other hand, we calculate the Shannon entropy according to $P(x)$, which is very sensitive to the variation of the distance of two parallels, and it can be used to describe the restriction of the polymer chain. With a given chain length N , the Shannon entropy of the distribution $P(x)$ decreases rapidly first and then tends to a fixed value with increasing D , by which we can find a transition point D_c . When $D > D_c$ the Shannon entropy tends to a constant, and at the same time the restriction of polymer chain performed by the two parallels can be ignored. We also find that the relationship between D_c and N as $D_c \sim N^\lambda$, here $\lambda = 0.543$.

Key words Confined compact chain, Distribution function, Shannon entropy, PERM