PHYSICAL REVIEW E VOLUME 50, NUMBER 6 DECEMBER 1994

Lattice ribbons: A model of double-stranded polymers

E.J. Janse van Rensburg,[1] E. Orlandini,[2] D.W Sumners,[3] M.C. Tesi,[2] and S.G. Whittington[2]
[1]*Department of Mathematics, York University, Downsview, Ontario, Canada M3J 1P3*
[2]*Department of Chemistry, University of Toronto, Toronto, Ontario, Canada M5S 1A1*
[3]*Department of Mathematics, Florida State University, Tallahassee, Florida 32306-3027*
(Received 24 August 1994)

We introduce a discrete ribbon model for double-stranded polymers (such as duplex DNA) where the ribbon is constrained to lie on the simple cubic lattice Z^3. The ribbon is made up of a sequence of plaquettes and can either be open or closed. We investigate the growth of the number of ribbons as a function of the number of plaquettes and use Monte Carlo methods to estimate the dimensions of the ribbon, the writhe of the backbone and, in the case of orientable closed ribbons, the linking number of the two boundary curves.

PACS number(s): 05.50.+q, 87.15.−v, 02.70.Lq

The traditional model of the conformational properties of linear polymers in dilute solution is the self-avoiding walk [1,2], i.e., a one-dimensional piecewise linear curve on a lattice, subjected to particular geometrical constraints (self-avoidance). In spite of its simplicity, this model captures several of the essential features which determine the large scale properties of such molecules and has been adapted to include attractive forces (to model collapse in polymers), closed to form a ring (to examine topological features such as knotting), and extended to a variety of related models relevant to branched polymers. However, some biologically important polymers such as duplex DNA and RNA exist as double-stranded molecules, where the two strands of complementary nucleotides are wound as right-handed helices around each other and around a common axis [3]. In addition, the double helix can wind in space to form a new helix of higher order, in which case the molecule is said to be supercoiled. Supercoiling in DNA can result from the binding of DNA to proteins (histones) or, in the case of a closed circular molecule, from topological and energetic constraints. Indeed if the molecule is closed to form a ring the two backbone strands of the double helix form circles which can be linked, in the sense that the strands cannot be separated without breaking one or the other. The linking of closed circular DNA depends only on the topological state of the strands and is maintained throughout all conformational changes that occur in the absence of strand breakage: indeed, the distinctive conformational properties of closed circular DNA are a consequence of this invariance. Moreover, the linking has important biological consequences. For instance, in the replication of closed circular duplex DNA, the unwinding of the DNA strands and the separation of the daughter chromosomes are obstructed by linking of DNA strands. In the biological situation, these topological obstructions are removed by enzymes (topoisomerases) which pass DNA strands through each other by means of enzyme-bridged transient breaks in the DNA backbone strands.

It is clear that the introduction of simple models, which take account of the double helix structure and the linking of the two strands, would represent an important step towards understanding the conformational properties of DNA-like molecules and their topological and geometrical properties such as linking and supercoiling. Unfortunately these features cannot be captured by models derived from self-avoiding walks, because of their one-dimensional nature; a higher level of complexity must be introduced.

A model which has proved useful in modeling the double helix structure and the topology of DNA is the ribbon model [4,5], in which the hydrogen bonds between the nucleotides of the two strands form an orientable ribbon surface whose boundary is precisely the two backbone strands. A particularly interesting result for this model is a conservation theorem relating the twist and writhe of orientable ribbons to the linking number (giving a quantitative description of the linking) of the two boundary curves [6,7].

In this paper we introduce a lattice version of the ribbon model. This has several important advantages. First it allows us to use techniques from self-avoiding walk theory to prove theorems about the model, which are very helpful in understanding the asymptotic behavior. In addition, the discrete nature of the model is particularly suitable for Monte Carlo methods, which can be used to supply additional results.

We shall be concerned with the simple cubic lattice Z^3 and we define a *plaquette* as a unit square, with vertices having integer coordinates. We define an *open ribbon* as an ordered sequence of plaquettes labeled $i=1,2,\ldots,n$ such that the following are true.

(i) Every two adjacent plaquettes ($|i-j|=1$) in the sequence have a common edge.

(ii) Two plaquettes (i and j) cannot be incident on a common edge unless $|i-j|=1$.

(iii) Two nonadjacent plaquettes cannot be incident on a common vertex unless they are also incident on a common plaquette.

(iv) Not more than three plaquettes can be incident on a common vertex.

We call the the number of edges which a plaquette has in common with other plaquettes the *degree* of the plaquette. In an open ribbon the first and last plaquettes have degree 1, and all other plaquettes have degree 2. We write w_n for the number of open ribbons with n plaquettes, where two ribbons are considered distinct if they cannot be superimposed by translation.

In a *closed ribbon* every plaquette has degree 2 and, also in this case, two closed ribbons are considered distinct if they cannot be superimposed by translation. Closed ribbons can be orientable (i.e., having two boundary curves) or nonorientable (having only one boundary curve). We write r_n for the number of closed ribbons with n plaquettes, and respectively r_n^o and r_n^u for the numbers of closed ribbons with n plaquettes which are orientable and nonorientable: clearly $r_n = r_n^o + r_n^u$.

It is very useful to have upper and lower bounds on w_n and r_n, since such bounds are essential in proving rigorous results about the asymptotic behavior of the model. Moreover, they represent a crucial comparison for the numerical results. For w_n, it is not difficult to show that $3(2\times 2)^{n-1} \leq w_n \leq 36(3\times 3)^{n-2}$. The upper bound can be obtained by considering objects defined as follows. The first plaquette is in any of the three coordinate planes, and the second is incident on one of the four edges of the first plaquette, but is not superimposed with the first plaquette. The kth plaquette is added so that it has an edge incident on one of the edges of the $(k-1)$th plaquette, other than the edge on which the $(k-1)$th and $(k-2)$th plaquettes are both incident. In addition, the kth and $(k-1)$th plaquettes are not superimposed. Clearly $w_n \leq 36(3\times 3)^{n-2}$ since the first plaquette can be embedded in the lattice in three ways, the second can be added to this in 4×3 ways (choose an edge in four ways, and an orientation in three ways), and subsequent ones in at most 3×3 ways (choose an edge in three ways, since one has already been used, and an orientation in three ways).

To obtain the lower bound consider the number of open ribbons with n plaquettes which have the property that the barycenter of the ith plaquette has at least one coordinate larger, and no coordinate smaller, than the corresponding coordinate of the barycenter of the $(i-1)$th plaquette. Such objects are certainly examples of open ribbons and give the inequality $w_n \geq 3(2\times 2)^{n-1}$ since the first plaquette can be embedded in the lattice in three ways and, when adding subsequent plaquettes, the edge to which the next plaquette is to be added can be chosen in two ways, and the orientation in two ways.

The growth of the number of open ribbons is controlled by the growth constant ρ, defined as $\log\rho = \lim_{n\to\infty} n^{-1}\log w_n$. One can use concatenation arguments to prove that this limit exists. Moreover, from the two bounds obtained above it is clear that $4 \leq \rho \leq 9$. Using the same techniques it is also possible to prove that the growth constant exists for closed ribbons, for closed orientable ribbons, and for closed nonorientable ribbons, and that in all three cases it is equal to the growth constant ρ of open ribbons.

We have used a grand canonical Monte Carlo algorithm to generate a sample of closed ribbons (orientable and nonorientable). The algorithm is based on sampling along a realization of a Markov chain, defined on the set of all closed ribbons. The proposed moves are of two types. One type is closely related to the pivot moves for polygons [8], which make large changes in the conformation of the ribbon but do not change the number of plaquettes. The second type is related to the local moves of the Berg–Foester–Aragao de Carvalho–Caraccioli–Frölich (BFACF) algorithm [9–11], which can change the number of plaquettes in the ribbon. The details of these moves, and a proof of ergodicity, will be given elsewhere [12]. Writing K for the fugacity associated with the number of plaquettes, we define the generating function

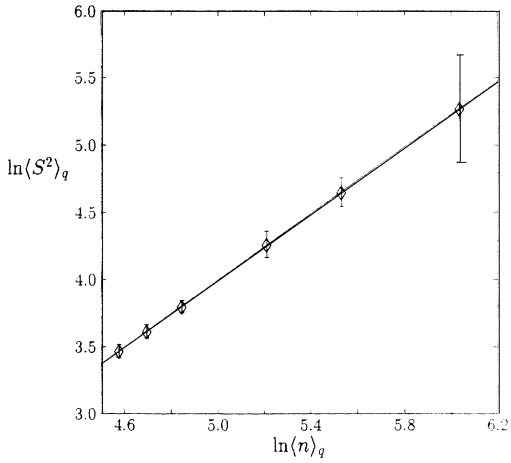

FIG. 1. The n dependence of the mean-square radius of gyration for closed ribbons.

$$G_q(K) = \sum_n r_n K^n n^q \tag{1}$$

which is expected to be singular at $K = 1/\rho$ and to behave like

$$G_q(K) \sim A/(1-\rho K)^{\alpha-2+q} \tag{2}$$

for K smaller than but close to $1/\rho$. The average number, $\langle n\rangle_q$, of plaquettes in the ribbon should then behave as

$$\langle n\rangle_q \sim \frac{K(\alpha-2+q)\rho}{1-\rho K} \tag{3}$$

so that the value of ρ can be obtained by plotting $1/\langle n\rangle_q$ against $1/K$. We have carried out runs at a number of values of K, corresponding to values of $\langle n\rangle_3$ of up to about 400. Analyzing these data in this way gives an estimate of ρ of $\rho = 4.33 \pm 0.09$, so that the lower bound derived above ($\rho \geq 4$) is reasonably close to the numerical estimate.

At the same time we have calculated the mean-square radius of gyration $\langle S^2\rangle_q$ of the ribbon in this ensemble. In this case we have considered orientable and nonorientable ribbons separately, in order to compare the two cases. In each case we expect the behavior to be given by

$$\langle S^2\rangle_q \sim \langle n\rangle_q^{2\nu}(1 + B\langle n\rangle_q^{-\Delta}) \tag{4}$$

and we show the dependence of $\langle S^2\rangle_3$ on $\langle n\rangle_3$ in Fig. 1. The results for the orientable and nonorientable cases are essentially superimposed and we give only the average values. Fitting the data to the form given by (4) with Δ taken to be 1/2 [13] gives an estimate of ν of 0.591 ± 0.016. [The error estimate takes account only of statistical errors, and not any systematic errors induced by terms omitted from (4).] As

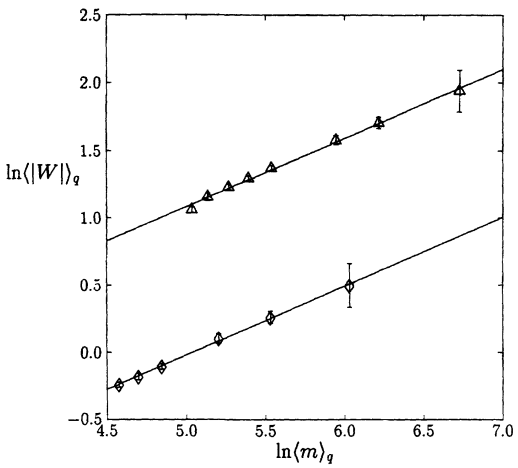

FIG. 2. The mean of the absolute value of the writhe of the boundary curve as a function of the average length of the boundary curve. The upper curve is for nonorientable ribbons and the lower curve is for orientable ribbons. The results plotted are for $q=3$.

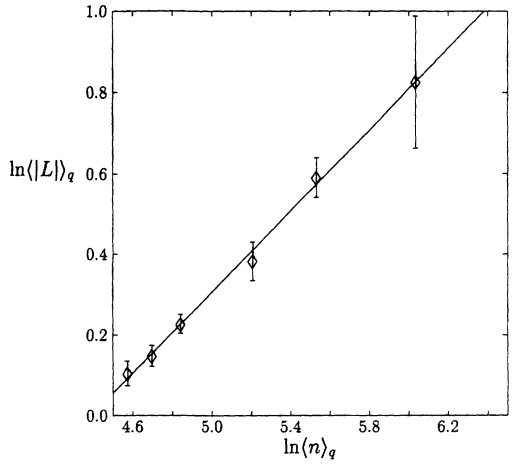

FIG. 3. The mean of the absolute value of the linking number of the two boundary curves of an orientable ribbon, as a function of the average length of the ribbon. The results are plotted for $q=3$.

expected, this is equal, within the error bars, to the value for a self-avoiding walk in $d=3$ [14].

An interesting measure of the entanglement complexity of an object such as a polygon is the *writhe* [6,7]. In a similar way we can use the writhe of the boundary curve (or curves) of the ribbon as a measure of the entanglement complexity of the ribbon. In order to define writhe, consider any simple closed curve in R^3, and project it onto R^2 in some chosen direction. In general, the projection will have crossings and, for almost all projection directions, these crossings will be transverse, so that we can associate a sign $+1$ or -1 with each crossing. For this projection we form the sum of these signed crossing numbers and average over all projection directions. This average quantity is the writhe W of the curve [7]. Writhe is a geometrical quantity (since it is not invariant under ambient isotopy) and is a real number which measures the extent to which the curve is supercoiled. We have calculated the average writhe of the two boundary curves of orientable ribbons, and of the single boundary curve of nonorientable ribbons, as a function of the length (m) of the boundary curve. (This is conveniently accomplished by making use of a theorem of Lacher and Sumners [15] which relates the writhe to the average linking number of the curve with its pushoffs in certain directions.) For polygons, the mean of the absolute value of the writhe behaves as [16]

$$\langle|W_m|\rangle \sim m^{\zeta}, \tag{5}$$

and we expect similar power law behavior in the case of ribbons. In Fig. 2 we show a log-log plot for the dependence of $\langle|W|\rangle_q$ on $\langle m\rangle_q$, the average of the length of the boundary. The lower curve is for orientable ribbons and the upper curve is for nonorientable ribbons. In each case the behavior is quite linear and our estimates of the exponent are $\zeta=0.51\pm0.02$ (nonorientable) and $\zeta=0.51\pm0.01$ (orientable). These values are very close to the corresponding value for a polygon, where it is known that the value is bounded below by 1/2 [16].

In the case of the orientable ribbons we can ask about the behavior of the linking number (L) of the two boundary curves. [To define this, orient the two boundary curves of the ribbon in parallel, and project the oriented boundary curves onto R^2. The sum of the signed crossing numbers for the *overcrossings* of one boundary curve (A, say) with the other (B, say) is the linking number $L(A,B)$. This quantity is invariant under ambient isotopy and under interchange of A and B.] In Fig. 3 we show a log-log plot of the mean of the absolute value of the linking number against the average length of the boundary curves. The behavior is quite linear with an estimated slope of 0.50 ± 0.01, so that $\langle|L|\rangle_q$ appears to scale as $\langle m\rangle_q^{\delta}$ with δ being about 1/2. Hence it appears probable that $\delta=\zeta$, so that these two measures of entanglement complexity grow at the same rate. (Experimental work suggests that changes in the linking number translate into changes in twist and writhe in the ratio 0.28:0.72 in DNA [17].)

In summary, we have introduced a discrete ribbon model on the cubic lattice to investigate the behavior of double-stranded molecules such as DNA and RNA. We have shown that, in spite of its higher complexity, the ribbon model shares many of the geometrical properties of its simpler random walk counterparts. In particular, as far as the mean-square radius of gyration and the writhe are concerned, we found that ribbons and self-avoiding walks seem to belong to the same universality class. On the other hand, ribbon models are richer, in that topological features such as the linking of the two boundary curves can be investigated. We found that the average linking number increases with n following a power law of the form $\langle n\rangle_q^{1/2}$, exactly as in the case of the writhe. This interesting result suggests that the writhe and the linking number, in spite of their different nature, capture a universal feature of the entanglement complexity of the ribbon.

There are many challenging questions which remain about the properties of the ribbon model. It will be interesting to see if the universal behavior of the writhe and of the linking number can be extended to other measures of the entanglement complexity of the ribbon. In addition, by add-

ing suitable fugacities in the model it should be possible to analyze, by Monte Carlo simulations, the occurrence of particular structures (plectonemic and solenoidal forms, branched conformations, etc.) in supercoiled DNA and to study how their occurrence depends on the ionic strength of the solvent in which the molecule is dissolved. Some work has already been carried out in this direction, both for a cylinder model (see [18] for a review) and by using a walk-like model with a Yukawa potential [19], and it will be interesting to compare these results with those obtained for the ribbon model.

We are pleased to acknowledge financial support from NSERC and from NSF.

[1] W. Kuhn and H. Kuhn, Helv. Chim. Acta **26**, 1394 (1943).
[2] N. Madras and G. Slade, *The Self-Avoiding Walk* (Birkhäuser, Boston, 1993).
[3] J. D. Watson and F. H. C. Crick, Nature (London) **171**, 737 (1953).
[4] W. R. Bauer, F. H. C. Crick, and J. H. White, Sci. Am. **243**, 118 (1980).
[5] J. H. White and W. R. Bauer, J. Mol. Biol. **189**, 329 (1986).
[6] J. H. White, Am. J. Math. **91**, 693 (1969).
[7] F. B. Fuller, Proc. Natl. Acad. Sci. USA **68**, 815 (1971).
[8] N. Madras, A. Orlitsky, and L. Shepp, J. Stat. Phys. **58**, 159 (1990).
[9] B. Berg and D. Foester, Phys. Lett. **106B**, 323 (1981).
[10] C. Aragao de Carvalho and S. Caracciolo, J. Phys. (Paris) **44**, 323 (1983).
[11] C. Aragao de Carvalho, S. Caracciolo, and J. Frölich, Nucl. Phys. B **251**, 209 (1983).
[12] E. Orlandini, E. J. Janse van Rensburg, and S. G. Whittington (unpublished).
[13] J. C. Le Guillou and J. Zinn-Justin, Phys. Rev. B **21**, 3976 (1980); J. Phys. (Paris) **50**, 1365 (1989).
[14] B. Li, N. Madras, and A. Sokal (unpublished).
[15] R. C. Lacher and D. W. Sumners, in *Computer Simulations of Polymers*, edited by R. J. Roe (Prentice-Hall, Englewood Cliffs, NJ, 1991), p. 365.
[16] E. J. Janse van Rensburg, E. Orlandini, D. W. Sumners, M. C. Tesi, and S. G. Whittington, J. Phys. A: Math. Gen. **26**, L981 (1993).
[17] N. R. Cozzarelli, Proc. Symp. Appl. Math. **45**, 1 (1992).
[18] A. V. Vologodskii and N. Cozzarelli, Annu. Rev. Biophys. Biomol. Struct. **23**, 609 (1994).
[19] M. C. Tesi, E. J. Janse van Rensburg, E. Orlandini, D. W. Sumners, and S. G. Whittington, Phys. Rev. E **49**, 868 (1994).