PHYSICAL REVIEW E VOLUME 59, NUMBER 4 APRIL 1999

Quantitative measure of folding in two-dimensional polymers

Gustavo A. Arteca[1,2,*] and Shuangxi Zhang[2]
[1]*Fysikalisk-Kemiska Institutionen, Uppsala Universitet, Box 532, Uppsala S-751 21, Sweden*
[2]*Département de Chimie et Biochimie, Laurentian University, Ramsey Lake Road, Sudbury, Ontario, Canada P3E 2C6*

(Received 18 August 1998; revised manuscript received 23 October 1998)

The degree of folding of a single three-dimensional (3D) polymer configuration is a general concept associated with the pattern of interpenetrations between chain loops. In the present context, this notion applies to the state of a rigid chain, regardless of the polymer being permanently or only temporarily entangled. Folding features represent an important aspect of macromolecular shape, one whose characterization must take into consideration both the 3D geometry and the bond connectivity of the polymer. In this work, we present a measure of folding complexity for planar objects. These systems include self-avoiding walks on planar lattices used for modeling 2D (adsorbed) polymers. In 3D chains, folding patterns are usually compared in terms of the number of bond-bond *projected* crossings, averaged over all rigid projections (the so-called *mean overcrossing number*). The characterization of molecular shape in 2D systems must be based on different notions since bond-bond crossings may not occur. Here, we generalize the concept of ''overcrossings'' as a descriptor of folding complexity in 2D structures. We show that the resulting molecular shape descriptor exhibits a power-law scaling with the number of monomers, both in *regular* conformers and in a continuum of *random* configurations. The method can be applied to study the adsorption of polymers with various topologies, as well as the complexity of random structures, such as those in crack patterns, soap froths, and other cellular decompositions of the plane. [S1063-651X(99)11303-5]

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

I. INTRODUCTION

Under weak interactions, molecules can conserve certain *shape features* in spite of changes in nuclear positions. This situation is found in the electrostatic recognition between a ligand and its receptor, as well as in the response of polymers to flow, grafting, and confinement in nanopores. A quantitative measure of *molecular shape* is relevant to computer-aided design of new materials, including pharmaceutical drugs, plastics, and lubricants. Therefore, a great deal of effort has been devoted to building descriptors of polymer shape.

In this work, we address a particular aspect of this issue: the characterization of ''folding features'' for macromolecules on two-dimensional (2D) surfaces. These models are commonly used to study adsorbed polymers, as well as phase transitions in grafted, confined, or compressed polymers [1]. Two-dimensional random walks are also used to model disordered systems, including random cellular structures such as those found in crack patterns produced by thermal shock [2] and soap froths [3]. In addition, 2D chains are found in phenomena modeled by two-dimensional percolation theory [4], including liquid diffusion, crystal growth on a surface, formation of gels, distribution of oil and gas inside porous rocks, spreading of forest fires, and dielectric breakdown. In this work, we restrict ourselves to the analysis of 2D *linear* polymers. Nevertheless, the approach can be extended to study 2D structures with different topologies.

Whereas ''structure'' is expressed in terms of chemical composition and nuclear geometry, ''shape'' can be quantified by studying electron density surfaces [5,6] or molecular space curves [6,7]. In the latter case, one can make use of topological or geometrical properties of the curves. In order to discriminate between polymer configurations and assess their shape stability, geometrical descriptors are preferable [7]. While molecular size [8,9] and anisometry [10] descriptors rely on the nuclear positions, folding features take into account the bonding pattern (or ''chain connectivity''). These properties characterize how chain loops distribute in space and interpenetrate each other. Here, we deal with the characterization of folding complexity of a *single* macromolecule. In the case of a 3D polymer, these features are commonly referred to as ''*self-entanglements*.'' This broad term encompasses the polymer chains *permanently* (i.e., topologically) self-entangled, as well as those behaving as self-entangled in a kinetic sense (i.e., temporary or ''geometrical'' entanglements). In 2D polymers, this terminology may not be appropriate since these chains are not entangled in the usual sense. However, the basic issue remains: 2D polymers exhibit folding features stemming from the distribution of loops on the plane. Below, we present a definition of ''folding complexity'' for 2D chains that generalizes that for 3D polymers.

In 3D polymers, chain folding can be described in terms of the *bond-bond crossings* (or ''overcrossings'') in 2D projections of polymer backbones [11–13]. The mean overcrossing number, denoted by $\overline{N}$, is the number of bond-bond crossings in a 2D (rigid) projection of the backbone, averaged over all possible projections in three-space. Recent studies indicate that the value of $\overline{N}$ averaged over all accessible configurations (denoted by $\langle\overline{N}\rangle$) is a ''weak topological descriptor'' of knot complexity, correlating with the gel diffusion velocity of knotted DNA [14]. Its main analytical property is the occurrence of scaling with respect to the num-

---

[*] Author to whom correspondence should be addressed. Present address: University of Uppsala, Uppsala, Sweden. Electronic address: Gustavo@nickel.laurentian.ca

ber of monomers, n. Numerical work on lattice polymers [15], off-lattice polymers [16], and protein native states [16,17] shows that

$$\langle \overline{N} \rangle \approx n^{\beta}, \quad n \gg 1, \tag{1}$$

with an exponent β between 1.1 and 1.4. These estimates are consistent with results from a path-integral representation of $\overline{N}$, which gives a *rigorous* upper bound in knots ($\beta \leqslant \frac{4}{3}$) [18] and a *conjectured* upper bound for polymers with excluded volume ($\beta < 1.4$) [19].

Scaling is an important property in a shape descriptor. It allows one to classify polymers in *universality classes*, whereby molecules with different chemical compositions share similar shape features. A well-known example is the scaling in the configurationally averaged mean radius of gyration, $\langle R_G^2 \rangle^{1/2}$, which follows the law $\langle R_G^2 \rangle^{1/2} \sim n^{\nu}\{1 + O(n^{-\Delta})\}$ [20], where ν is the *size exponent* and Δ is the first correction-to-scaling exponent. The size exponent depends on the polymer-solvent interaction and the *dimensionality* of space, D. In dilute solutions of 3D linear polymers $(D=3)$, the behavior is as follows: (i) $\nu = \frac{1}{2}$ in an ideal solvent (or at the θ temperature) [8,9], (ii) $\nu = \frac{1}{3}$ in a poor solvent (when the polymer collapses to a spheroid), and (iii) $\nu = 0.588 \pm 0.002$ in a good solvent (when the polymer resembles a self-avoiding walk) [21]. In 2D polymers $(D=2)$, the scaling behavior is still in debate. Whereas we have exactly $\nu = \frac{1}{2}$ in a poor solvent, the values in ideal conditions and in a good solvent are believed to be $\nu = \frac{4}{7}$ [22] and $\nu = \frac{3}{4}$ [23], respectively.

In summary, while the properties of 3D polymers in terms of molecular size and folding are now known, the analysis of 2D polymers has been restricted so far to molecular size [21–23] and anisometry [10]. In this work, we complete their description by introducing a measure of folding complexity for 2D polymers. We make two main contributions. First, we present how the notion of ‘‘overcrossing’’ can be extended to planar structures. The resulting $\overline{N}$ function is the natural generalization to $D=2$ of the mean number of overcrossings defined in three-space. Second, we establish the scaling behavior of $\overline{N}$ in regular and random polymer conformations.

II. FOLDING COMPLEXITY IN TWO-DIMENSIONAL POLYMER BACKBONES

In three-space, ‘‘overcrossings’’ are the bond-bond intersections observed in a projection along a (one-dimensional) line of sight. Although 2D self-avoiding walks *do not* exhibit actual self-intersections, we can still define a similar shape descriptor by computing the *number of bonds intersected when a 2D polymer is observed along a one-dimensional line of sight on the molecular plane*. The resulting ‘‘line-intersection descriptor’’ is the rigorous extension of the notion of ‘‘overcrossings’’ from three to two dimensions.

In two dimensions, an n-node chain is represented parametrically as a curve $\mathbf{W}(t)$, with node coordinates $\{\mathbf{W}_i = (X_i, Y_i)\}$. When taking the chain's centroid as the origin $\mathbf{O}$, the node coordinates will be denoted as $\{\mathbf{w}'_j = (X'_j, Y'_j)\}$. Line-of-sight directions can then be taken as the *diameter lines* of the smallest circle, centered at $\mathbf{O}$, that completely encloses the polymer. Let $\mathbf{p}_1$ be the first of these directions. The number of intersections M_{int} between $\mathbf{p}_1$ and the molecular backbone can be used to define a *formal* ‘‘two-dimensional’’ number of overcrossings for a given ‘‘projection’’ line. To produce a description consistent with the one for 3D polymers, we define that a 1D rod has *no* overcrossings. Since a diameter line that is *not* parallel to the rod *must intersect the latter once*, then we choose $N = M_{\mathrm{int}} - 1$ as the formal ‘‘overcrossing number’’ for that diameter line. Note that the case where a line $\mathbf{p}_1$ coincides with a bond has zero measure with respect to the *mean* N value.

The algorithm to compute N is simple. Let $\mathbf{A} = (X_A^{(1)}, Y_A^{(1)})$ and $\mathbf{B} = (X_B^{(1)}, Y_B^{(1)})$ be the coordinates defining the diameter line $\mathbf{p}_1$. Now, we establish whether the $\mathbf{p}_1$ line intersects a generic bond between consecutive nodes $\mathbf{w}'_{i-1}$ and $\mathbf{w}'_i$. An intersection exists if, and only if, the equations

$$X = t_1(X'_i - X'_{i-1}) + X'_{i-1} = t_2(X_B^{(1)} - X_A^{(1)}) + X_A^{(1)}, \tag{2a}$$

$$Y = t_1(Y'_i - Y'_{i-1}) + Y'_{i-1} = t_2(Y_B^{(1)} - Y_A^{(1)}) + Y_A^{(1)}, \tag{2b}$$

have solutions for $t_1 \in [0,1]$ and $t_2 \in [0,1]$, simultaneously. The pair (X,Y) corresponds to the intersection point. By solving Eqs. (2) for $i = 2,3,\ldots$, we can compute the number of overcrossings $N_{(1)}$ for the diameter line $\mathbf{p}_1$. Finally, by repeating this procedure for m different diameters lines $(m \gg 1)$, the mean number of overcrossings $\overline{N}$ becomes

$$\overline{N} = \frac{1}{m} \sum_{j=1}^{m} N_{(j)}, \tag{3}$$

where $N_{(j)}$ is the number of overcrossings for the jth diameter line. In addition, we can define an overcrossing probability for a 2D chain, as it is done in 3D chains [12]. If m_N is the number of diameter lines producing N overcrossings, the probability of observing N ‘‘overcrossings’’ is $A_N \approx m_N/m$ (a strict equality for $m \to \infty$). With this definition, we have for a *linear* chain

$$\overline{N} = \sum_{N=0}^{n-2} N A_N, \tag{4}$$

where $\max N = n-2$ is the maximum possible overcrossing number in a 2D linear chain. (Since the chain has $n-1$ bonds, any diameter line cannot produce more than $n-1$ intersections.) In random walks, this maximum is usually not reached, and many of the A_N are zero. As commented before, the value of $\overline{N}$ in Eq. (4) is simply a geometrical property of planar curves, and it must *not* be interpreted in terms of the notion of overcrossings in three-space. However, the definition of $\overline{N}$ given here constitutes the proper extension of the concept of overcrossings to the case of 2D figures.

The properties of the overcrossing probability distribution are simple: (a) An ‘‘*unfolded*’’ 2D chain (e.g., a 1D rod) has an $\{A_N\}$ distribution with a dominant peak at $N=0$, and a small mean overcrossing number; (b) a ‘‘*folded*’’ 2D chain (e.g., a compact 2D coil) has many $N \gg 1$ values contributing to the overcrossing probabilities, and thus a large value for $\overline{N}$. As a result, $\overline{N}$ can be used to monitor globule⇄coil or coil⇄rod *shape transitions* in 2D chains, as it is done for 3D chains [24]. Accordingly, we consider $\overline{N}$ to be a *geometrical*

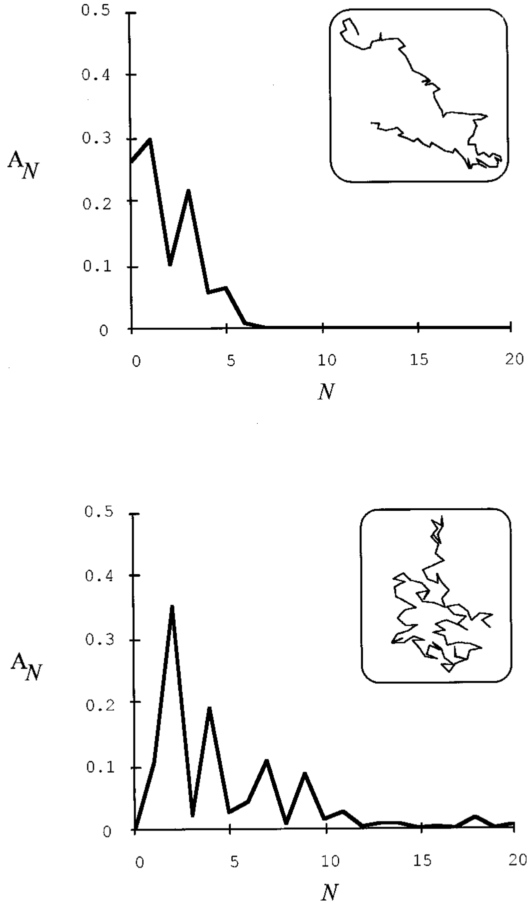

FIG. 1. Distribution of overcrossing probabilities (''overcrossing spectra'') for two random configurations of a 2D polymer with $y=\frac{1}{3000}$ and $n=90$. [The top diagram corresponds to an ''open'' conformer (inset), characterized by high A_N values for low N. The bottom diagram corresponds to a ''closed'' configuration of the same polymer (inset). This conformer exhibits larger overcrossing numbers.]

measure of ''2D folding complexity'' (or, loosely speaking, ''*2D entanglements*''). Using an analogy with single 3D polymer chains, we could say that a compact 2D self-avoiding random walk is more ''entangled'' (in the sense above) than the linear rod.

Figure 1 illustrates these ideas with two configurations of a model polymer, where the chain is represented as a 2D off-lattice self-avoiding walk with excluded volume interaction. The polymer is built as a sequence of bonds, linking n nodes with constant bond length b and random bond angles. These configurations are subject to two constraints: (i) there are no bond-bond intersections during the construction of the walk (i.e., self-avoidance); (ii) nodes not linked by bonds cannot be at a distance smaller than r_{ex} (i.e., a radius of excluded volume). In this scheme, all random walks approach a *single configuration* (the 1D rod) if $r_{\mathrm{ex}} \rightarrow 2b$. Note also that the folding properties depend on the dimensionless radius of the excluded volume $y=r_{\mathrm{ex}}/2b$, $y \in [0,1]$, and not on the individual values of b and r_{ex}. Low y values correspond to chains embedded in ''poor'' solvents, whereas large y values correspond to ''good'' solvents.

Figure 1 shows two chain configurations with $n=90$ and $y=\frac{1}{3000}$. The top diagram gives the distribution $\{A_N\}$ for an ''open'' configuration. As expected, low overcrossing values dominate (here, $N=0$ and 1). The bottom diagram presents the results for a compact conformation. In this case, all ''projections'' produce overcrossings (i.e., $A_0=0$). While the most frequent number of overcrossings is $N=2$ (35%), values up to $N \approx 20$ are found.

We can compute the average folding features in 2D polymers by evaluating the configurationally averaged mean overcrossing number, denoted by $\langle \overline{N} \rangle$. Accurate estimations of $\langle \overline{N} \rangle$ can be derived by sampling 2D walks with a ''naive'' Monte Carlo approach [20]. In our case, partial chains that fail the conditions of self-avoidance or excluding volume are rejected and not continued. The resulting set contains uncorrelated conformers. The results for $\overline{N}$ in regular and random walks are discussed in the following sections.

We note here another property. If $n \gg 1$, one can replace $\overline{N}$ by an integral [cf. Eq. (4)]. From the mean value theorem for integrals and the normalization of $\{A_N\}$, we have

$$\overline{N} \cong \int_0^{n-2} N A_N \, dN \leq \max\{N\} \int_0^{n-2} A_N \, dN = n-2. \tag{5}$$

Equation (5) implies a bound on $\overline{N}$ (and thus $\langle \overline{N} \rangle$). As a result, if a scaling relation such as Eq. (1) exists in 2D linear polymers, we deduce that $\beta \leq 1$. The following sections test this property in exact results for regular conformers and in numerical simulations for random conformers.

III. SCALING BEHAVIOR OF SELF-ENTANGLEMENTS IN REGULAR 2D POLYMER CONFIGURATIONS

We consider first the behavior of shape descriptors in *regular*, as opposed to *random*, conformations. In regular conformations, bond and dihedral angles take periodically repeated values. In 3D proteins, these conformations include elements of secondary structure, e.g., α helices, 3_{10} helices, and β strands [25]. The mean shape properties of these conformers are known in terms of the number of monomers n [26]. In α helices and strands, $\langle \overline{N} \rangle$ exhibits *linear* scaling [i.e., $\beta=1$ in Eq. (1)]. It is unclear whether a similar scaling is found in two dimensions. Here, we address this question by using a family of simple regular conformations.

The all-*trans* linear chain (a ''zigzag-like'' polymer) serves as a 2D model of an α helix. This polymer is defined by a constant bond length, b, and a constant bond angle, Φ. If the first node is the origin, $\mathbf{W}_1=(0,0)$, the coordinates $\mathbf{W}_j=(X_j,Y_j)$ of the remaining nodes are

$$X_j=(j-1)b\sin\frac{\Phi}{2}, \quad j=1,2,3,...;$$

$$Y_j=\begin{cases} b\cos\dfrac{\Phi}{2}, & j=2,4,... \\ 0, & j=1,3,.... \end{cases} \tag{6}$$

Figure 2 shows an example, with $n=10$ monomers. Using these coordinates, we can compute the mean overcrossing number as discussed in Sec. II. In this particular case, the chain model is simple enough to allow the *exact* computation of overcrossing probabilities and the determination of the scaling behavior for $\overline{N}$.

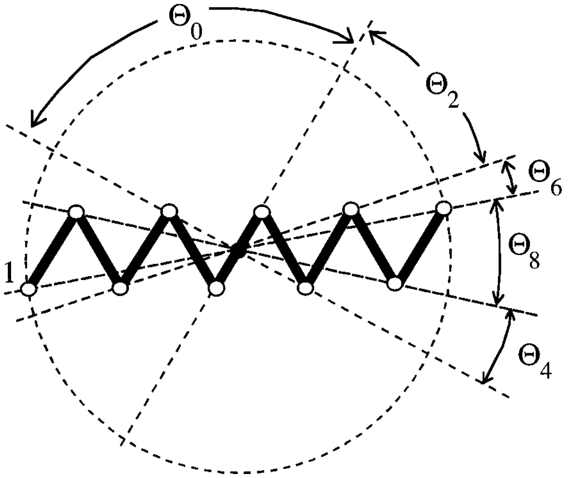

FIG. 2. Regular (all-*trans*) conformation of a linear polymer (a model of a 2D helix). [The regions indicated as Θ_N define the areas where the diameter lines intersect the chain $N+1$ times. The overcrossing probabilities are simply $A_N = \Theta_N/\pi$.]

The exact evaluation of A_N is as follows. Consider the chain in Fig. 2 with $n=10$. Let Θ_N be the angle (in radians) for the circular sector comprising the diameter lines leading to N overcrossings. The overcrossing probabilities are thus fractional areas of the circle:

$$A_N = \Theta_N/\pi. \tag{7}$$

It is clear that, in regular all-*trans* chains, N can only be even when n is even, whereas N is odd if n is odd. In Fig. 2, only $N=0,2,4,6,8$ are possible. Note that the Θ_N values, for $N<n-2$, are unaffected when the chain is lengthened by two bonds. The last angle, Θ_{n-2}, is given as: $\Theta_{n-2} = \pi - (\Theta_0 + \Theta_2 + \cdots + \Theta_{n-4})$, even n.

We analyze now the case of Θ_N for even-n chains. (The analysis of odd n is similar, the only difference being a shift in the origin **O**.) Using center-of-mass coordinates,

$$X'_j = X_j - (n-1)\frac{b}{2}\sin\frac{\Phi}{2}, \quad Y'_j = Y_j - \frac{b}{2}\cos\frac{\Phi}{2}, \tag{8}$$

we can determine Θ_0 from the position of the $(n/2)$ node and the $[(n/2)+2]$ node (cf., Fig. 2). From the scalar product $\cos\Theta_0 = \mathbf{w}'_{n/2}\cdot\mathbf{w}'_{n/2+2}/\|\mathbf{w}'_{n/2}\|\|\mathbf{w}'_{n/2+2}\|$, we obtain

$$\Theta_0 = \arccos\left\{ \frac{\cos^2\frac{\Phi}{2} - 3\sin^2\frac{\Phi}{2}}{\left(\cos^2\frac{\Phi}{2} + 9\sin^2\frac{\Phi}{2}\right)^{1/2}} \right\}. \tag{9}$$

A similar analysis can be used for all other Θ_N (and thus A_N) values. After some algebra, the general form for the overcrossing probabilities is found to be

$$A_N = \frac{1}{\pi}\arccos\left\{ \frac{\cos^2\frac{\Phi}{2} + (N-1)(N+3)\sin^2\frac{\Phi}{2}}{\left[\cos^2\frac{\Phi}{2} + (N-1)^2\sin^2\frac{\Phi}{2}\right]^{1/2}\left[\cos^2\frac{\Phi}{2} + (N+3)^2\sin^2\frac{\Phi}{2}\right]^{1/2}} \right\}, \tag{10}$$

where $N=0,2,4,\ldots,n-4$. For the largest number of overcrossings ($N=n-2$), we have

$$A_{n-2} = 1 - \sum_{N=0}^{n-4} A_N. \tag{11}$$

Finally, the *exact* $\overline{N}$ value is computed with Eqs. (4), (10), and (11). The $\overline{N}$ values derived with these formulas are indistinguishable, within the error bars, from those obtained numerically with $m=10\,007$ projections.

Using this formulation, we have computed $\overline{N}$ for $n \leqslant 15\,000$, in order to detect scaling behavior. From our results, compiled in Fig. 3, it is clear that $\overline{N}$ follows a *logarithmic scaling law*, $\overline{N} \propto \ln n$, a result that is consistent with Eq. (5). A linear regression using $\Phi = 60°$ and $n = 500,1000,1500,\ldots,15000$ gives

$$\overline{N} \approx (1.1031 \pm 0.0001)\ln n \quad \text{for } n \gg 4, \tag{12}$$

with 95% confidence. The prelogarithmic constant in Eq. (12) depends on the bond angle Φ. As the all-*trans* chain approaches the linear rod, the constant decreases. A similar regression for $\Phi = 150°$ produces $\overline{N} \approx (0.170\,72 \pm 0.000\,05)\ln n$ for $n \gg 4$.

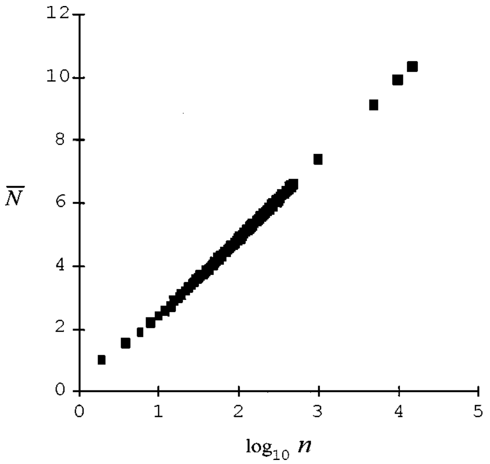

FIG. 3. Logarithmic scaling of the mean overcrossing number as a function of the number of monomers in an all-*trans* linear chain. (Results are shown for n even, $n \gg 4$ and $\Phi = 60°$.)

The analysis presented in this section leads us to two important conclusions.

(a) The scaling behavior of $\overline{N}$ depends on the spatial dimension D. Whereas $\overline{N}$ increases as $\sim n^{\beta}$ for random and regular 3D conformers, the same need not be true in 2D space.

(b) The scaling law in Eq. (12) is well defined even for *rather short chains*. The correlation observed in Fig. 3 can be obtained from chains even as short as $n=20$. This observation suggests that it might be sufficient to use medium sized chains ($n<100$) to obtain a good representation of scaling behavior for the entanglements in 2D polymers.

The properties of $\overline{N}$ in regular conformers are contrasted in the next section with the scaling behavior of its configurational average, $\langle\overline{N}\rangle$, in random chain polymers.

IV. SCALING BEHAVIOR OF $\langle\overline{N}\rangle$ FOR SELF-AVOIDING WALKS IN THE CONTINUUM

We have computed the configurational average of the mean number of overcrossings, $\langle\overline{N}\rangle$, using off-lattice 2D self-avoiding walks with variable excluded volume and constant bond length. The chains are specified by two parameters, the number of monomers n and the reduced excluded volume $y=r_{\mathrm{ex}}/2b$. For every (n,y) pair, shape descriptors were evaluated on an ensemble of 10^3 uncorrelated conformers, collected by the Monte Carlo search discussed in Sec. II. This procedure yields reliable independent structures, though it suffers from strong conformational attrition as n or y increases [20]. Thus, evaluating mean properties becomes prohibitive for chains with $n>100$, and for short chains with $y>0.5$. For example, at $y=\frac{2}{3}$ and $n=15$ there are $\sim 7\times 10^3$ rejected conformers for every one accepted. At $y=\frac{2}{3}$ and $n=30$, the yield diminishes to $\sim 3\times 10^8$ rejected conformers for each one accepted. The detailed behavior of the mean attrition rate as a function of y and n is given elsewhere [27,28].

We have computed $\langle\overline{N}\rangle$ for a wide series of y values ($y=\frac{1}{3000}$, $\frac{1}{6}$, $\frac{1}{3}$, $\frac{1}{2}$, $\frac{2}{3}$, $\frac{5}{6}$, and $\frac{11}{12}$). At low excluded volume ($y=\frac{1}{3000}$), simulations were carried out for polymer lengths of $n=10,15,20,\ldots,80$. At high excluded volume ($y=\frac{11}{12}$), only results with $n=5,6,7,\ldots,15$ were possible. In addition to the descriptor of folding, $\overline{N}$, we have computed a number of descriptors of molecular size (the mean radius of gyration, $\langle R_G^2\rangle^{1/2}$, the span, and the end-to-end distance in the chain). The well-known scaling behavior of these descriptors ($\langle R_G^2\rangle^{1/2}\sim n^{\nu}$) can be used as a benchmark to test the reliability of our sampling. With the same configurations used to calculate $\langle\overline{N}\rangle$, we obtained $\nu\approx 0.75\pm 0.02$, for all $y<0.5$. This result, obtained consistently for all molecular size descriptors, is in good agreement with the accepted exact result, $\nu=\frac{3}{4}$ [23]. In addition, our results for $y>0.5$ (where chains are too short) can be well described by a leading asymptotic term $\nu=\frac{3}{4}$ and a subdominant correction-to-scaling term, with exponent $\Delta\approx\nu$. This value is also consistent with results in the literature, which indicate $0.66\leqslant\Delta\leqslant 1.2$ [20]. These tests suggest that our simulations for $y<0.5$ could be already in the asymptotic regime for $\langle\overline{N}\rangle$. As well, these observations agree with the results in Sec. III, which indicate that $n>20$ may be enough to obtain the leading asymptotic behavior in 2D polymers.

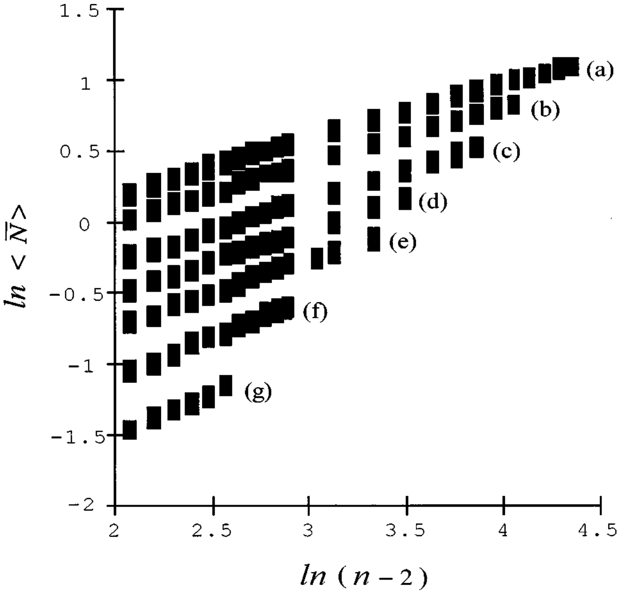

FIG. 4. Power-law scaling of the configurationally averaged mean overcrossing number $\langle\overline{N}\rangle$ as a function of the number of monomers in self-avoiding walks with variable excluded volume. [The letters indicate different values of the excluded volume variable $y=r_{\mathrm{ex}}/2b$, where $y=\frac{1}{3000}$ (a), $\frac{1}{6}$ (b), $\frac{1}{3}$ (c), $\frac{1}{2}$ (d), $\frac{2}{3}$ (e), $\frac{5}{6}$ (f), and $\frac{11}{12}$ (g).]

Using the above sampling, we have studied the asymptotic behavior of $\langle\overline{N}\rangle$. Since $\overline{N}$ vanishes for $n=2$, a reasonable scaling law for 2D chains is [cf. Eq. (1)]

$$\langle\overline{N}\rangle\sim a(y)(n-2)^{\beta}, \tag{13}$$

where the $a(y)$ must vanish for $y\to 1^{-}$ (i.e., when the chain becomes a rod).

Figure 4 tests Eq. (13) with a logarithmic plot $\ln\langle\overline{N}\rangle$ versus $\ln(n-2)$. The results in Fig. 4 are consistent with those for molecular size descriptors. First, the asymptotic regime appears to have been reached for $y=\frac{1}{3000}$, $\frac{1}{6}$, and $\frac{1}{3}$, precisely the cases where the mean radius of gyration shows the correct scaling behavior. Second, the results for $y>0.5$ also appear to be approaching the same asymptotic regime, despite the short chains used.

With the results in Fig. 4, we have estimated the scaling exponent β (with 95% confidence intervals) as follows.

(a) Linear correlations for chains with $n>25$ give $\beta=0.37\pm 0.02$, 0.36 ± 0.03, and 0.43 ± 0.04, for $y=\frac{1}{3000}$, $\frac{1}{6}$, and $\frac{1}{3}$, respectively. Note that each of these fittings has a different number of points.

(b) Linear correlations using the eight longest computed chains give exponents $\beta=0.39\pm 0.04$, 0.41 ± 0.03, and 0.46 ± 0.02, for $y=\frac{1}{3000}$, $\frac{1}{6}$, and $\frac{1}{3}$, respectively.

These results do not rule out a dependence of the exponent β with y. However, considering the trends in Fig. 4 for $y>0.5$, such a dependence appears to be small, as is also the case for 3D chains [15,16]. From the above values, we can make a conservative estimate:

$$\beta = 0.40 \pm 0.05 \tag{14}$$

in 2D self-avoiding walks with excluded volume, which is in accord with the bound in Eq. (5).

The estimation of the preexponential function $a(y)$ is more difficult. Our simulations appear to be represented semiquantitatively by adopting $a(y) \approx \exp\{-y^2 P(y)/(1-y)\}$, where $P(y)$ is a smooth, bound function of the reduced excluded volume.

V. FURTHER COMMENTS AND CONCLUSIONS

In this work, we have introduced the concept of folding complexity for 2D polymers, and provided a quantitative measure in terms of a geometrical property of planar curves. The resulting ''line-intersection descriptor,'' $\overline{N}$, conveys the folding features in a 2D polymer. By using an analogy with 3D curves, we could state that $\overline{N}$ measures the equivalent to a ''degree of self-entanglement'' for a polymer adsorbed onto a surface. The present approach is general, and it can be applied without major modification to different topologies, e.g., branched polymers, 2D cellular decompositions of the plane, or 2D networks. The latter case is particularly interesting, because the size and anisometry of a network configuration do not change as its connectivity is modified. Accordingly, descriptors of folding complexity (e.g., the mean overcrossing number) are especially convenient for analyzing networks, since they describe shape features related to both connectivity *and* geometry. Descriptors relying only on molecular geometry would not provide much useful information on these systems.

We have also established the qualitative scaling behavior of the descriptor $\overline{N}$ in some regular and random 2D linear polymers. In random polymers with excluded volume, we have shown the occurrence of power-law scaling. As a result, a picture emerges where the descriptor $\langle \overline{N} \rangle$ shares a key property with molecular size descriptors such as the mean radius of gyration: the scaling exponent depends on the dimensionality of space. Results up to the present indicate that (i) the mean overcrossing number in 1D chains is independent of the length (thus suggesting that $\beta = 0$), (ii) the ''entanglement'' scaling exponent for 2D chains is $\beta = 0.40 \pm 0.05$, and (iii) the exponent for 3D chains is $\beta = 1.2 \pm 0.1$ [15,16]. Based on these values, we can conjecture that β increases faster than linearly with the dimensionality of space D. From the results above, a rough fitting would indicate $\beta \sim O(D(D-1))$.

Regular and random conformations provide a benchmark for the analysis of realistic 2D polymers with attractive and repulsive monomer-monomer interactions. An interesting recent development is the possibility to study the shape of protein backbones using 2D projections [29]. The Sammon projection algorithm [30] represents an appealing choice, because it minimizes the difference between the distance matrices in two and three dimensions. Consequently, the resulting 2D projection conserves most of the original 3D folding features of a protein. By performing a systematic survey of 2D protein shape, our method should provide the type of quantitative measures of homology between protein folds needed in molecular engineering.

In closing, it should be noted that our formulation can also be adapted to the *analytical* computation of mean overcrossing numbers in selected conformers. To this end, one can use a path-integral representation of $\overline{N}$ for 2D polymers that is similar to the one used in 3D chains [18,19]. Details of this approach will be discussed elsewhere.

ACKNOWLEDGMENTS

G.A.A. would like to thank J. Cantarella (Pennsylvania) for sending a copy of Ref. [18] prior to publication. This work was supported by FRUL (Laurentian) and by grants from NSERC (Canada).

---

[1] J. M. Hammersley, G. M. Torrie, and S. G. Whittington, J. Phys. A **15**, 539 (1982); S. G. Whittington, Adv. Chem. Phys. **51**, 1 (1982); J. Tobochnik, I. Webman, J. L. Lebowitz, and M. H. Kalos, Macromolecules **15**, 549 (1983); F. W. Wiegel, in *Phase Transitions and Critical Phenomena*, edited by C. Domb and J. L. Lebowitz (Academic, London, 1983), Vol. 7; K. De'Bell and T. Lookman, Rev. Mod. Phys. **65**, 87 (1993); A. K. Vuppu, A. A. García, and C. Vernia, Biopolymers **42**, 89 (1997); I. Golding and Y. Kantor, Phys. Rev. E **56**, R1318 (1997); T. Hjelt, S. Herminghaus, T. Ala-Nissila, and S. C. Ying, *ibid.* **57**, 1864 (1998).

[2] W. Korneta, S. K. Mendiratta, and J. Menteiro, Phys. Rev. E **57**, 3142 (1998); G. Caldarelli, R. Cafiero, and A. Gabrielle, *ibid.* **57**, 3878 (1998).

[3] J. Stavans, Rep. Prog. Phys. **56**, 733 (1993); B. Levitan, Phys. Rev. E **55**, 1205 (1997); O. Herrera, A. Olguín, M. Ortíz, and C. H. Wörner, *ibid.* **55**, 4475 (1997); T. Aste, *ibid.* **55**, 6233 (1997).

[4] D. Stauffer and A. Aharony, *Introduction to Percolation Theory* (Taylor and Francis, London, 1992).

[5] R. F. W. Bader, *Atoms in Molecules: A Quantum Theory* (Clarendon, Oxford, 1990).

[6] P. G. Mezey, *Shape in Chemistry: An Introduction to Molecular Shape and Topology* (VCH, New York, 1993).

[7] G. A. Arteca, in *Reviews in Computational Chemistry*, edited by K. B. Lipkowitz and D. B. Boyd (VCH, New York, 1996), Vol. 9.

[8] P. J. Flory, *Statistical Mechanics of Chain Molecules* (Interscience, New York, 1969).

[9] P.-G. de Gennes, *Scaling Concepts in Polymer Physics* (Cornell University Press, Ithaca, 1985).

[10] K. Solč, Macromolecules **6**, 378 (1973); J. Rudnick and G. Gaspari, J. Phys. A **19**, L191 (1986); J. Rudnick and G. Gaspari, Science **237**, 384 (1987); H. W. Diehl and E. Eisenriegler, J. Phys. A **22**, L87 (1989); A. Baumgärtner, J. Chem. Phys. **98**, 7496 (1993); S. J. Sciutto, J. Phys. A **27**, 7015

(1994); G. Zifferer, J. Chem. Phys. **102**, 3720 (1995).

[11] G. A. Arteca and P. G. Mezey, Biopolymers **32**, 1609 (1992).

[12] G. A. Arteca, Biopolymers **33**, 1829 (1993).

[13] E. J. Janse van Rensburg, D. W. Sumners, E. Wasserman, and S. G. Whittington, J. Phys. A **25**, 6557 (1992).

[14] A. Stasiak, V. Katritch, J. Bednar, D. Michoud, and J. Dubochet, Nature (London) **384**, 122 (1996); V. Katritch, J. Bednar, D. Michoud, R. G. Scharein, J. Dubochet, and A. Stasiak, *ibid.* **384**, 142 (1996); A. Yu. Grosberg, A. Feigel, and Y. Rabin, Phys. Rev. E **54**, 6618 (1996); V. Katritch, W. K. Olson, P. Pieranski, J. Dubochet, and A. Stasiak, Nature (London) **388**, 148 (1997); A. V. Vologodskii, N. J. Crisona, B. Laurie, P. Pieranski, V. Katitch, J. Dubochet, and A. Stasiak, J. Mol. Biol **278**, 1 (1998).

[15] E. Orlandini, M. C. Tesi, S. G. Whittington, D. W. Sumners, and E. J. Janse van Rensburg, J. Phys. A **27**, L333 (1994).

[16] G. A. Arteca, Phys. Rev. E **49**, 2417 (1994); **51**, 2600 (1995).

[17] G. A. Arteca, Phys. Rev. E **56**, 4516 (1997).

[18] J. Cantarella, R. B. Kusher, and J. M. Sullivan, Nature (London) **392**, 237 (1998); G. Buck, *ibid.* **392**, 238 (1998); J. Cantarella, D. DeTurk, and H. Gluck (unpublished).

[19] A. L. Kholodenko and D. P. Rolfsen, J. Phys. A **29**, 5677 (1996); A. L. Kholodenko and T. A. Vilgis, Phys. Rep. **298**, 251 (1998).

[20] A. Sokal, in *Monte Carlo and Molecular Dynamics Simulations in Polymer Science*, edited by K. Binder (Oxford University Press, New York, 1995); N. Madras and G. Slade, *The Self-Avoiding Walk* (Birkhäuser, Boston, 1996).

[21] G. A. Baker, Jr., B. G. Nickel, and D. I. Meiron, Phys. Rev. B **17**, 1365 (1978); J. C. LeGuillou and J. Zinn-Justin, *ibid.* **21**, 3976 (1980); J. Phys. (France) Lett. **46**, L137 (1985); J. Phys. (Paris) **50**, 1365 (1989); A. J. Guttmann, J. Phys. A **22**, 2807 (1989); D. C. Rapaport, *ibid.* **18**, 113 (1985); N. Madras and A. Sokal, J. Stat. Phys. **50**, 109 (1988); B. Li, N. Madras, and A. Sokal, *ibid.* **80**, 661 (1995); J. S. Pedersen, M. Laso, and P. Schurtenberger, Phys. Rev. E **54**, R5917 (1996).

[22] B. Duplantier and H. Saleur, Phys. Rev. Lett. **59**, 539 (1987); **62**, 1368 (1989); R. M. Bradley, Phys. Rev. A **39**, 3738 (1989); M. Wittkop, S. Kreitmeier, and D. Göritz, J. Chem. Phys. **104**, 3373 (1996).

[23] B. Nienhuis, Phys. Rev. Lett. **49**, 1062 (1982).

[24] G. A. Arteca, Biopolymers **35**, 393 (1995); G. A. Arteca, Macromolecules **29**, 7594 (1996); G. A. Arteca, J. Phys. Chem. B **101**, 4097 (1997).

[25] C. Brändén and J. Tooze, *Introduction to Protein Structure* (Garland, New York, 1991). (A 3_{10} helix corresponds to an elongated helix, with closer hydrogen bonds per turn as compared with an α helix.)

[26] G. A. Arteca, Can. J. Chem. **73**, 241 (1995).

[27] G. A. Arteca and S. Zhang, Phys. Rev. E **58**, 6817 (1998).

[28] In addition to the limitations in the length of ''treatable'' polymers, the procedure poses a heavy demand on the random number generator. In our case, we employ a random number generator that combines two sequences plus shuffling, for an effective period of $\sim 2\times 10^{18}$. For a discussion on the performance of this random number generator in the context of the present type of simulations, see G. A. Arteca, Int. J. Quantum Chem. **65**, 519 (1997).

[29] T. W. Barlow and W. G. Richards, J. Mol. Graphics **13**, 373 (1995); **14**, 232 (1996).

[30] J. W. Sammon, Jr., IEEE Trans. Comput. **18**, 401 (1969).