Simulations of Knotting in Confined Circular DNA

C. Micheletti,* D. Marenduzzo,† E. Orlandini,‡ and D. W. Sumners§

*International School for Advanced Studies (SISSA), Consiglio Nazionale delle Ricerche e Istituto Nazionale di Fisica della Materia (CNR-INFM) Democritos unit and Italian Institute of Technology, Trieste, Italy; †Scottish Universities Physics Alliance (SUPA), School of Physics, University of Edinburgh, Edinburgh, Scotland; ‡Dipartimento di Fisica, Consorzio Nazionale Interuniversitario per le Scienze Fisiche della Materia (CNISM) and Sezione Istituto Nazionale di Fisica Nucleare (INFN), Universita' di Padova, Padova, Italy; and §Department of Mathematics, Florida State University, Tallahassee, Florida

ABSTRACT The packing of DNA inside bacteriophages arguably yields the simplest example of genome organization in living organisms. As an assay of packing geometry, the DNA knot spectrum produced upon release of viral DNA from the P4 phage capsid has been analyzed, and compared to results of simulation of knots in confined volumes. We present new results from extensive stochastic sampling of confined self-avoiding and semiflexible circular chains with volume exclusion. The physical parameters of the chains (contour length, cross section, and bending rigidity) have been set to match those of P4 bacteriophage DNA. By using advanced sampling techniques, involving multiple Markov chain pressure-driven confinement combined with a thermodynamic reweighting technique, we establish the knot spectrum of the circular chains for increasing confinement up to the highest densities for which available algorithms can exactly classify the knots. Compactified configurations have an enclosing hull diameter ~2.5 times larger than the P4 caliper size. The results are discussed in relation to the recent experiments on DNA knotting inside the capsid of a P4 tailless mutant. Our investigation indicates that confinement favors chiral knots over achiral ones, as found in the experiments. However, no significant bias of torus over twist knots is found, contrary to the P4 results. The result poses a crucial question for future studies of DNA packaging in P4: is the discrepancy due to the insufficient confinement of the equilibrium simulation or does it indicate that out-of-equilibrium mechanisms (such as rotation by packaging motors) affect the genome organization, hence its knot spectrum in P4?

INTRODUCTION

The packing of DNA inside bacteriophages arguably yields the simplest example of genome organization in living organisms. The phage genome occupies a substantial fraction of the capsid volume and this leads to a buildup of large forces exerted by DNA on the capsid walls. It is believed that phages harness this high internal force to initiate the infection cycle by a pressure-driven ejection of the viral DNA into the cytosol of the bacterial host (1–4).

Genome organization inside the phage is usually assumed to be spool-like, although many recent numerical simulations have quite seriously questioned this assumption (2,5–10). Instead, they have proposed a scenario in which the phage DNA possesses a rather weak liquid crystalline order, at least when DNA packaging is modeled by an elastic self-avoiding rod (DNA) being fed adiabatically slowly inside a spherical rigid cavity (the phage head) (see, e.g., (6–8)). The extent and kind of DNA ordering also significantly depends on capsid shape (2,6) and on solvent quality (1,4).

In a recent series of experiments, the knotting probability and knot spectrum of DNA extracted from P4 phage capsids have been characterized (11,12). The P4 phage capsid is ~40 nm in diameter, and contains ~10 kilobases (kbases) of DNA. The experiments considered both wild-type P4, in which one of the DNA ends is rooted at the capsid surface, and a tailless mutant, in which the two cohesive ends were often able to reptate inside the capsid and stick to each other. Upon extraction from the P4 capsid, the remainder of the DNA cyclize, producing a complex knot spectrum that was characterized by gel electrophoresis (11–13). Ninety-five percent of the DNA was knotted, with at least 48% of the tailless mutant knots coming from DNA which cyclized inside the P4 capsid (11).

The quantitative characterization of the spectrum of DNA knots provided a powerful indirect experimental probe for the overall organization of DNA inside the P4 capsids (12). In fact, it was found that there is a higher percentage of chiral knots than achiral ones, and that torus knots greatly outnumber twist knots (12). What is responsible for the observed bias? Is the tendency to form chiral and torus knots driven by confinement alone?

These questions have stimulated several theoretical and computational investigations of the spatial organization and/or knotting of biomolecules both in general (14–16), and specifically to compare the experimental knot spectrum with the simulated one for circular DNA subject to confinement. In this context, the simplified models for circular DNA adopted so far were based on infinitesimally thin polymers (11,12,17–19). (The only simulations of knotting in confined self-avoiding walks, of which we are aware, are in the on-lattice study in (20).) Infinitely thin polymer models have proved extremely useful to our current understanding of the fundamental properties of knots (21,22). With the addition of suitable bending rigidity terms, they have also been shown to be capable of accurately reproducing the knot spectrum of unconstrained circular DNA (23). Extending these models to the case of

Submitted May 16, 2008, and accepted for publication June 18, 2008.

Address reprint requests to Davide Marenduzzo, Tel.: 44-01-31-650-5293; E-mail: dmarendu@ph.ed.ac.uk.

Editor: Ruth Nussinov.

0006-3495/08/10/3591/09 $2.00

doi: 10.1529/biophysj.108.137653

spatial confinement represents a major challenge due to the difficulty in conformational sampling (11) which requires the introduction of advanced sampling techniques (19).

The results obtained with confined circular chains with no excluded volume qualitatively reproduced the observations that the probability of DNA knotting and the complexity of the detected knots increased dramatically upon confinement (11,19). However, neither the presence of a bias toward chiral knots or toward torus knots were observed (19). Intriguingly, these observations could be reproduced by a writhe-biased sampling of the infinitely thin circular polymers (12). But does the writhe-bias represent a genuine biophysical aspect of how DNA is packaged in the P4 viral capsid? It might, for instance, reflect nonequilibrium effects, such as a spooling enhancement induced by a rotary packaging motor at the capsid portal. Or is this procedure an ad hoc correction which is only needed in view of the excessive simplifications employed? In other words, do highly confined self-avoiding circular chains in thermodynamic equilibrium lead to the observed statistical biases to chiral and torus knots? (Note that, as we consider circular chains, like in previous modeling work, we are focusing on the tailless mutant case in which ends are assumed to meet either inside the capsid or while still partially confined.)

This work aims at tackling these questions by providing a reference equilibrium study for the occurrence of knots in semiflexible self-avoiding circular polymers. To this end, we generate ensembles of semiflexible circular chains with excluded volume and characterize their knot spectrum inside a sphere of radius R. As the reference experimental system is constituted by the P4 bacteriophage, DNA is modeled as a 3.4 μm-long semiflexible self-avoiding chain of cylinders (23) with thickness equal to 2.5 nm (the hydrated dsDNA diameter) bending rigidity appropriate to yield a persistence length equal of 50 nm in the wormlike chain model (24). Two possible coarse-graining schemes were considered where the 10 kbases-long P4 circular genome was modeled as a circular chain of 200 and 400 cylinders, respectively. Such chains were subjected to progressive confinement into spheres with radius R down to approximately four cylinders' length. Although this represents a threefold reduction of the confining radii with respect to previous studies, the P4 phage has a diameter which is still 2–3 times smaller than our smallest confining sphere. While we could equilibrate conformations inside a smaller volume, the bottleneck is given by the computational routines used to precisely determine the knot type, which are unable to characterize the majority of knots if the compactification exceeds the one we focus on here.

We find that, in agreement with the experiments, confinement leads to a very high proportion of complex knots, with high crossing number. For the range of confining radii R for which we can obtain reliable results, confinement favors chiral over achiral knots as well, again in qualitative agreement with experiments. For instance the ratio between six crossing chiral knots over six crossing achiral knots increases with $1/R$. However, in the explored range of spatial confinement, our simulations indicate the preferential formation of twist knots (type 5_2) as opposed to torus ones (type 5_1), opposite to the results of the P4 experiments. Thus the torus/twist bias is the most prominent feature of the knot spectrum where a qualitative discrepancy is found between experiments and simplified models for DNA packaging. Its further clarification should probably advance significantly the understanding/modeling of the DNA packaging inside the P4 phage.

This work is organized as follows. In the next section, we motivate and introduce the coarse-grained model for DNA and briefly review the techniques used to generate, sample, and classify the knotted configurations that a circular DNA molecule 3.4 μm-long (the length of the P4 genome) attains in equilibrium when spatially confined. The results are finally presented, discussed, and compared in relation to the available experimental data on DNA knotting inside the P4 phage.

MODEL AND METHODS

The coarse-grained DNA description used in this study is based on the model adopted by Rybenkov et al. (23), which has been proved valuable for reproducing salient physical properties of DNA molecules as long as the P4 genome (23,25). Specifically, a circular DNA molecule is described as a closed linear chain made of N equal cylinders which, with the exception of nearest neighbors along the chain, are subjected to steric hindrance, in that no overlap of the cylinders is allowed (26). The thickness (the diameter of the circular cross section) and length (the long axis) of each cylinder will be indicated with d and l, respectively. We considered a thickness $d = 2.5$ nm, equal to the hydration diameter of double-stranded DNA. We employed two discretization levels, corresponding to a coarse-graining of the circular chain into $N = 200$ and $N = 400$ cylinders. For a more direct comparison with the experimental context, all lengths will be expressed in nanometers. Accordingly, as the P4 genome experimentally studied in the literature (11,12) is 10 kb-long (corresponding to 3.4 μm), the lengths, l, of the cylinders in the two coarse-graining levels are, respectively, $l_{200} = 16.7$ nm and $l_{400} = 8.3$ nm.

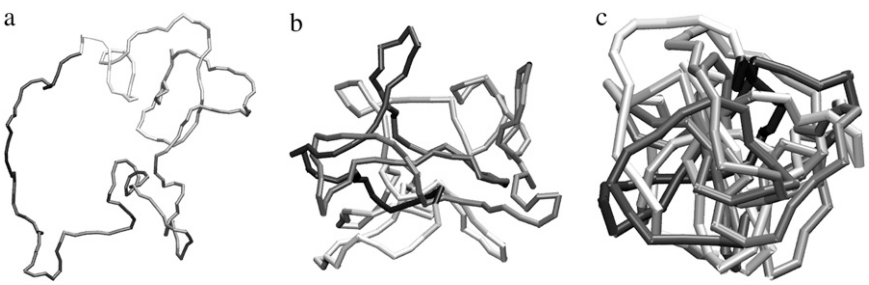

FIGURE 1 Typical configurations of circular chains of $N = 200$ cylinders generated at (*a*) low, (*b*) medium, and (*c*) high pressure. While different magnifications are used for the three panels (reflecting the different radii of gyration), the ratio of the cylinders' diameter to the total chains contour length is kept fixed to the nominal value, given by 2.5 nm/3400 nm, used in the simulation.

The elasticity of the polymer is accounted for by a bending rigidity term with strength appropriate for yielding a persistence length of $l_p = 50$ nm in a free (spatially unconstrained) wormlike chain (24),

$$V_{\text{bending}} = -\kappa \sum_{i=1}^{N-1} \vec{t}_i \cdot \vec{t}_{i+1}, \quad (1)$$

where $\kappa = k_B T l_p / l$ is the elastic modulus of the polymer. We denote the coordinates of the vertices of the circular chain by $\vec{r}_1, \vec{r}_2, \vec{r}_N, \vec{r}_{N+1} \equiv \vec{r}_1$, and we define the tangent vectors $\vec{t}_i \equiv \vec{r}_{i+1} - \vec{r}_i$ ($i = 1, \ldots, N$). Configurations in which two cylinders are interpenetrating are disallowed by penalizing them with a very large (infinite for practical purposes) energy cost.

The method used to generate uncorrelated configurations of semiflexible self-avoiding circular polymers closely parallels the one we adopted in Micheletti et al. (19) for the case of random circular chains with no excluded volume.

Briefly, to each conformation, Γ ($\Gamma \equiv \{\vec{r}_1, \vec{r}_2, \vec{r}_N, \vec{r}_{N+1} \equiv \vec{r}_1\}$) we first associate a hull radius, $R(\Gamma)$, defined as the distance of the farthest vertex from the center of mass of Γ. We note that this definition may not necessarily yield the radius of the smallest sphere (the miniball) containing all vertices of Γ (27). This is because the miniball center need not coincide with the chain center of mass. However, under increasing confinement the polymer adopts more and more compact conformations so that the adopted radius and the miniball radius practically coincide (9). We also note that the definition of R is based on the location of the vertices of Γ, which therefore lie on the centerline of the chain. Therefore it does not account for the extra room required to accommodate the thick cylinders entirely in the sphere. Finally we point out that, in Ali et al. (8), a slightly different definition of the enclosing radius R was used.

To generate the confined circular chains, a series of multiple Markov chains (28) are run in parallel by evolving 16 different stochastic trajectories of closed circular chains. At each time step each configuration is modified by an unrestricted crankshaft move of the vertices. Such moves preserve the chain connectivity and cylinders lengths and ensure an efficient exploration of the configuration space. The 16 Markov processes are subjected to various confining pressures, P, in that the reduced energy (i.e., in units of thermal energy) of a given configuration, Γ, is defined as

$$E(\Gamma) = PR(\Gamma) + V_{\text{bending}}(\Gamma) \quad (2)$$

if no overlap occurs between nonconsecutive cylinders, otherwise the energy is treated as infinite. The newly generated configurations are accepted with the standard Metropolis criterion. Occasional swaps of configurations in different Markov chains are also proposed and accepted/rejected with a suitable generalization of the Metropolis scheme (28). Examples of configurations generated at different pressures are provided in Fig. 1.

This stochastic scheme ensures that configurations explored by each of the Markov processes are sampled with probability proportional to $\exp(-PR(\Gamma) - V_{\text{bending}}(\Gamma))$, where P is the pressurelike parameter of the given Markov chain (replica). Notice that, while the hull radius is subjected to different pressures across the replicas, the elastic energy term is controlled by the same bending rigidity parameter, κ, for all replicas. The ensemble of conformations explored by the various replicas is next sampled at intervals greater than the autocorrelation time of the hull radius. From the sampled conformations at pressure P one then computes the histogram, $n_P(K, 1/R)$, of the number of independent circular chain configurations with knot type K and hull radius R. The use of the inverse confining radius, $1/R$, as variable, is a natural choice that captures how various geometrical/topological properties change with increasing confinement. The entries of the recorded histogram result from the combination of 1), the statistical weight of elastic circular chains of knot type K and hull radius R, $W(K, 1/R)$; and 2), the pressure bias, $\exp(-PR)$,

$$n_P(K, 1/R) \propto W(K, 1/R) \exp(-PR). \quad (3)$$

Notice that W depends implicitly on the chain length, N, and the bending rigidity, κ. The proportionality constant between the two sides of the above relationships can be fixed through suitable normalization of the histogram (29). From the above relationship it is clear that the statistical weight of unconstrained knotted elastic chains can be recovered from the recorded histograms by ''undoing'' the pressure-bias (where ''undoing'' means solving Eq. 3 to obtain W up to a multiplicative constant). Data obtained by the various Markov processes can be optimally merged (29) to return the desired statistical weight $W(K, 1/R)$ for a wide range of $1/R$.

The addition of a confining bias via P is necessary because, due to the rapid decrease of the chain entropy upon confinement (30), it would be impossible to efficiently sample compact configurations in ordinary, unconstrained, Monte Carlo simulations.

Once $W(K, 1/R)$ is known (it clearly must be calculated separately for each chain length, N) various quantities of interest can be estimated. For example, the expected fraction of circular polymers of knot type K that can fit inside a sphere of radius R is given by the weighted contribution of all circular chains having hull radius up to R,

$$P_N(K, 1/R) = \frac{\sum_j W_N(K, 1/R_j)}{\sum_{K'} \sum_j W_N(K', 1/R_j)}, \quad (4)$$

where K' spans all knot types and the sum over the discrete equispaced values of inverse hull radii is restricted to the cases where $1/R_j > 1/R$. Notice that, due to the rapid decrease of W for increasing confinement, the translational entropy of configurations with size smaller than R is neglected.

The errors of the quantities in equation 4 are estimated from the semidispersion observed by applying the weighted histogram method separately to the first and second half of the circular chains generated stochastically.

The reweighting procedure we use depends on the knowledge of the knot type of each sampled conformation. The knot identification procedure is carried out in the same spirit of Micheletti et al. (19). First a smoothing and shrinking procedure (31,32) is employed to simplify the configuration as far as possible, by reducing, without changing topology, the average number of crossings of its two-dimensional projections. After the conformational simplification, the Dowker code (33) of one of the simplified two-dimensional projections is computed, simplified with suitable algebraic manipulations (equivalent to Reidmeister moves) and finally processed by the Knotfind program, a routine implemented in the KnotScape package (34), which can identify prime components of up to 16 crossings for a given knot.

It should be noted that while all positive knot identifications of Knotfind are exact, it is not true that all simple knots of up to 16 crossings can be identified from an arbitrarily complicated projection. In such cases, though the knot identity might be tentatively assigned on the basis of available polynomials, we shall conservatively term the knot type as unclassified or unknown. From the above discussion it is clear that the circular chain simplification procedure is a critical step in our analysis scheme since the degree of achieved simplification impacts on the efficiency with which simple knots (or knot components) can be precisely identified. Owing to the large number of processed conformations for every value of P ($2.5 - 3 \times 10^4$ for $N = 200$ and 4.5×10^3 for $N = 400$), the simplification follows a fast, greedy, algorithm (see (19) for details) which becomes rather inefficient at very small values of R (equivalently large values of P). This is the main factor limiting the compactification level for which we can report results in this study.

As anticipated, when Knotfind is unable to characterize the knot type, we term the knot as unclassified. The number of unclassified knots and of unknots is plotted in Fig. 2. In what follows, solid lines are used for the range of R where the fraction of unclassified knots is $<10\%$. When the latter is between 10 and 50% dashed lines will be used, while dotted lines are employed beyond the 50% threshold.

RESULTS

The knot spectrum of the simplest prime knots and of some of the composite (35) knots for confined circular chains of $N = 200$ and $N = 400$ cylinders is shown in Fig. 3. The two dif-

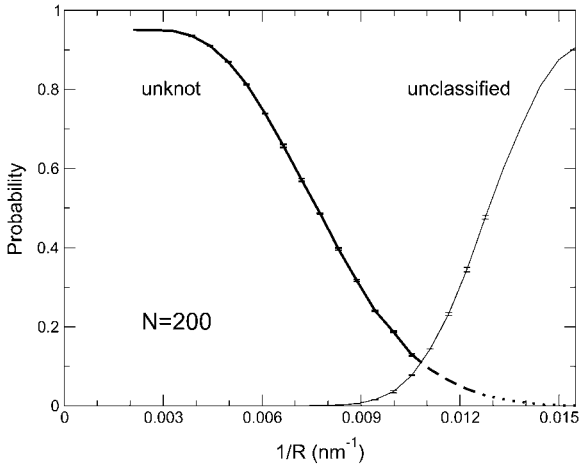

FIGURE 2 Unknotting probability and probability of unclassified knots as a function of the inverse radius of the confining sphere, $1/R$, for circular chains of $N = 200$ cylinders.

ferent discretization levels we employed lead to consistent results with high accuracy. In particular, the probabilities of forming a given knot type is independent of discretization in the unconfined case, $1/R \rightarrow 0$, suggesting, a posteriori, a dependence only on the persistence length, l_p, of the system (and not on the level of coarse graining). In the absence of excluded volume, our model would reduce to the Kratky-Porod model (discrete wormlike chain) characterized by a single-exponential decay of the tangent-tangent correlation, with decay length equal to $\langle \vec{t}_i \cdot \vec{t}_{i+x} \rangle = \exp^{-(x\,l/\xi)}$, where $\xi = -l/[\log(\coth(\kappa/K_BT) - K_BT/\kappa)]$(notice that, as κ is inversely proportional to l, ξ reduces to the nominal persistence length

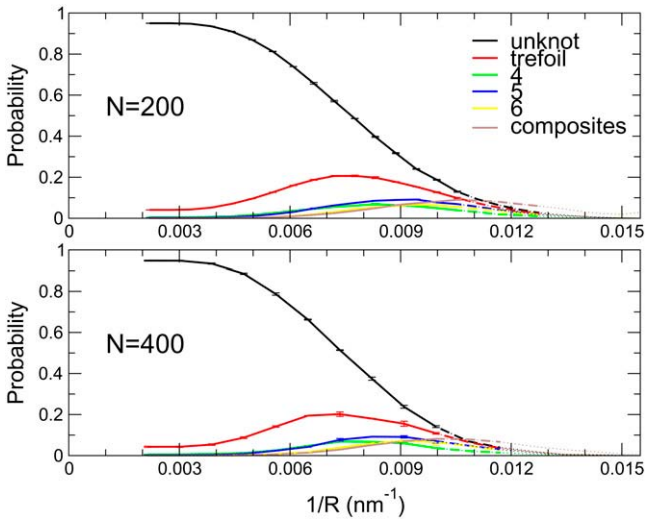

FIGURE 3 Probabilities of simple prime knots and composite knots for circular chains with $N = 200$ and $N = 400$ cylinders as a function of the inverse radius of the confining sphere, $1/R$. Dashed and dotted lines refer to statistics in which the fraction of unclassified knots is, respectively, between 10 and 50% and above 50%.

$l_p = 50$ nm in the continuum limit $l \rightarrow 0$). We calculated the tangent-tangent autocorrelations for our model with excluded volume and fitted them with a single-exponential, persistence lengths of 42 nm and 46 nm for the $N = 200$ and $N = 400$ cases, respectively. These are in good agreement with the values obtained for the Kratky-Porod model for infinitesimally thin chains with the expression provided above. The fact that l is significantly smaller than l_p for both $N = 200$ and $N = 400$, provides an intuitive a posteriori explanation of why the probability curves of Fig. 3 are insensitive to the discretization level.

Slight differences in the two discretizations are seen, instead, in the compact regime owing to the appearance of a further length scale, the radius of the confining sphere. For the levels of compactification reached in this study, differences due to the coarse-graining employed are limited, as indicated by the consistency of the two panels in Fig. 3. Owing to the better statistics collected for $N = 200$, we shall henceforth mostly focus on these data.

We first discuss the knot spectrum in Fig. 3 in the limit of unconstrained semiflexible self-avoiding circular chains, $1/R \rightarrow 0$. This constitutes a first validation of our model and algorithms, as the predicted knotting probabilities may be compared against the experimental knotting probabilities of free, unconfined, circular DNA in a 1 M NaCl buffer (23). This is the buffer for which our calculations are most appropriate, because the high concentration of counterions makes it so that electrostatic interactions between different DNA segments, neglected here, are efficiently screened. For $1/R \rightarrow 0$ trefoils, 4_1 and five crossing knots occur with probabilities $4.0 \pm 0.15\%$, $0.46 \pm 0.01\%$, and $0.28 \pm 0.01\%$, respectively. These predictions compare well with the experimental results by Rybenkov et al. (23), who report values of 3.5%, 0.44%, and 0.25%, for 10-kb-long circular DNAs in a 1 M NaCl buffer.

The first notable qualitative feature of the probability distributions of Fig. 3 is their nonmonotonic behavior as a function of the inverse confining radius. Clearly, the succession of the knot distributions is such that the progressive confinement of the semiflexible self-avoiding circular chains leads to the appearance of knots with larger and larger number of minimal crossings. This intuitive trend (which is shared with simulations of circular polymer chains with no excluded volume (19)) is compatible with the experimental findings on DNA knotting in P4 capsids where, at compactifications still higher than those explored here ($R = 20$ nm), the knot population is dominated by knots with ~30 crossings (the unknot and trefoil accounting for only 2% and 0.36% of the total population, respectively).

Before examining the distributions of Fig. 3 in further detail, we quantify how confinement impacts on the minimal number (over all possible projections) of crossings $\mathcal{C}$. It is important to estimate $\mathcal{C}$, as this correlates well with the electrophoretic migration velocity of the knots within the gel, which is the quantity that experimentalists measure (36,37).

The computation of the minimal crossing number of a strongly confined configuration presents challenges similar to the problem of finding its knot type. Therefore, also in this case, there is a large fraction of the knots for which the minimal crossing number cannot be exactly determined. For these, we can, however, compute an upper and lower bound on $\mathcal{C}$. The upper bound is simply given by the minimal number of crossings one gets after the simplification of the Dowker code based on the Knotfind program (34). The lower bound is obtained by combining the bound

$$\mathcal{C} \geq 2(i-1), \tag{5}$$

where i is the braid index of the knot (38) with the Morton-Franks-Williams inequality (39,40)

$$i \geq \frac{1}{2}(E-e)+1, \tag{6}$$

where E and e are, respectively, the largest and the smallest power of the HOMFLY polynomial of the knot. This gives

$$\mathcal{C} \geq E-e. \tag{7}$$

Fig. 4 shows the upper and lower bounds on the minimal number of crossings as a function of the inverse radius of the confining sphere. For relatively mild confinement all the knot types are classified and the two bounds coincide with the exact value $\mathcal{C}$. As the confinement increases, however, the difference between the two bounds broadens due to the increasing number of unclassified configurations. Nevertheless our results are enough to provide an interesting estimate of the growing rate of $\mathcal{C}$ with $1/R$. It would be interesting to compare this prediction with data obtained with gel electrophoresis on experiments with different double-stranded DNA bacteriophage (whose internal DNA density may differ).

Coming back to knot types, Fig. 5 shows the whole knot spectrum inside a confining sphere of $R \sim 90$ nm (so still significantly larger than P4, for which R would be $\sim$20 nm).

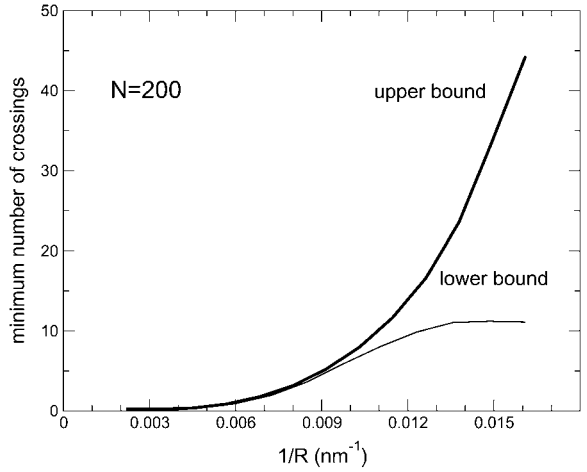

FIGURE 4 Upper and lower bounds on the minimal crossing number as a function of the inverse radius $1/R$ for circular chains with $N = 200$ cylinders.

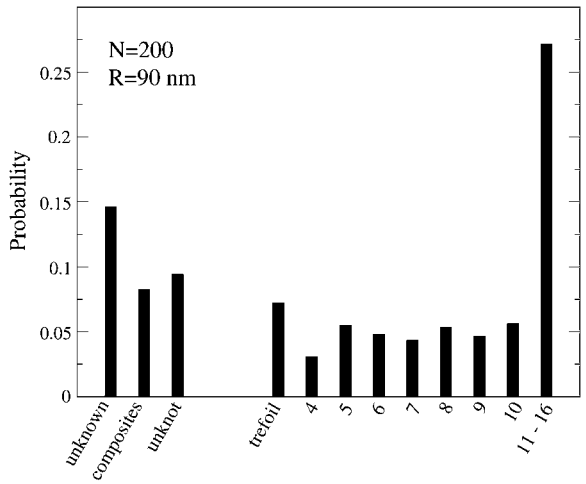

FIGURE 5 Knot spectrum for $N = 200$ and $R = 90$ nm.

The most populated knot, excluding the unknot, is the trefoil, followed by composite knots. The detailed breakup of knot types in the spectrum illustrates that the population of prime knots at fixed degree of compactification does not present obvious trends with the increase in complexity (minimal crossing number).

To make contact with the experimental observation, we shall now analyze the obtained spectrum to detect any biases in the population of chiral versus achiral knots and of twist versus torus knots. In our model any given chiral knot can exist in either handedness with the same probability, by symmetry, hence we do not specify whether a chiral knot is right- or left-handed. (Note that, on the other hand, we will have to detect handedness for the prime knot components of a composite knot, to determine whether it is globally chiral or achiral; see the discussion on granny versus square knots later on.)

We start by noting that the trends in Fig. 3 indicate that chiral five crossing knots outnumber the achiral 4_1 knot even for quite moderate values of compactification. This result is, qualitatively, in line with the P4 experimental knot spectrum although the five crossing knots outnumber the four crossing knot to a much greater degree in the experimental spectrum.

This result is, however, not sufficient to indicate unambiguously an effect of confinement on chirality since all five crossing knots are chiral but the five crossing knot is more complex than 4_1. As progressive confinement shifts the distribution of knot types toward complex knots, it can be expected that for an appropriate degree of compactification the five crossing population will outnumber the 4_1 one.

The simplest context in which the chirality bias can be analyzed within knots of the same complexity is the family of six crossing knots which comprises the chiral 6_1 and 6_2 knots and the achiral 6_3 knot. We thus wish to probe whether there is any trend in the ratio between the probabilities of occurrences of these prime knots. These data are plotted in Fig. 6. It emerges, that the achiral knot is less populated at all values of

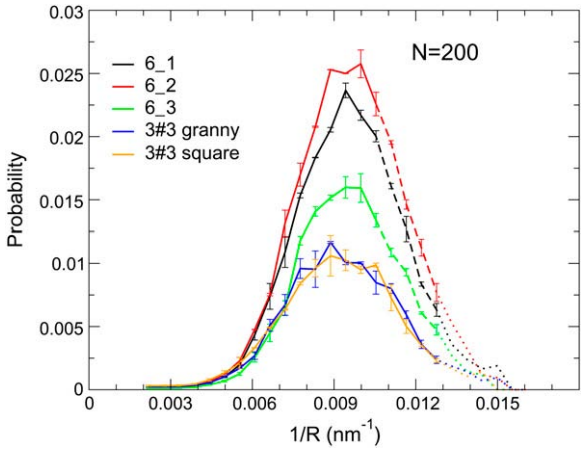

FIGURE 6 Probabilities of occurrence of 6_1, 6_2, and 6_3 knots and of square and granny knots as a function of the inverse radius $1/R$ for circular chains of $N = 200$ cylinders.

$1/R$. It should be noted that the same feature is found for infinitely thin unconfined circular polymer chains as a function of N. Therefore, by itself this result again does not prove chiral knots are favored by confinement.

A more robust criterion to assess the existence of a chirality bias can be found by computing the ratio between globally chiral versus achiral knots on the family of all, both prime and composite, six crossing knots. A global chirality bias should lead to a higher number of chiral knots with respect to achiral ones.

To find this ratio, we have to identify the chirality of composite knots as well. A composite six crossing knot is made up by two trefoils, each of which can come in either handedness. If the two handednesses are different, then we have a globally achiral square knot, otherwise we end up with a globally chiral granny knot.

The identification of granny and square knots is nontrivial in our framework, given that the computed Dowker code does not account for the handedness of a given chiral knot. We followed a heuristic classification scheme inspired by the properties of distribution of the writhe of the composite six crossing knots after the smoothing and shrinking simplification procedure. As illustrated by Fig. 7, the writhe distribution is very broad and featureless before smoothing, but acquires a distinctive trimodal character after the conformational simplification. The location of the maxima at writhe values of 0 and ± 6.8 correspond to the values of the writhe that would be expected in ideal configurations of square and granny knots, respectively (41). This observation strongly suggests that the smoothing and shrinking procedure is sufficiently effective in simplifying knots that the resulting writhe acquires the corresponding ideal value. We use this observation to separate the six crossing composite knots into square and granny knots according to whether the modulus of the writhe after smoothing is, respectively, $<$ or >3.4.

The resulting ratio between chiral and achiral six crossing knots, as a function of $1/R$, is visible in Fig. 7 *b*, and indicates

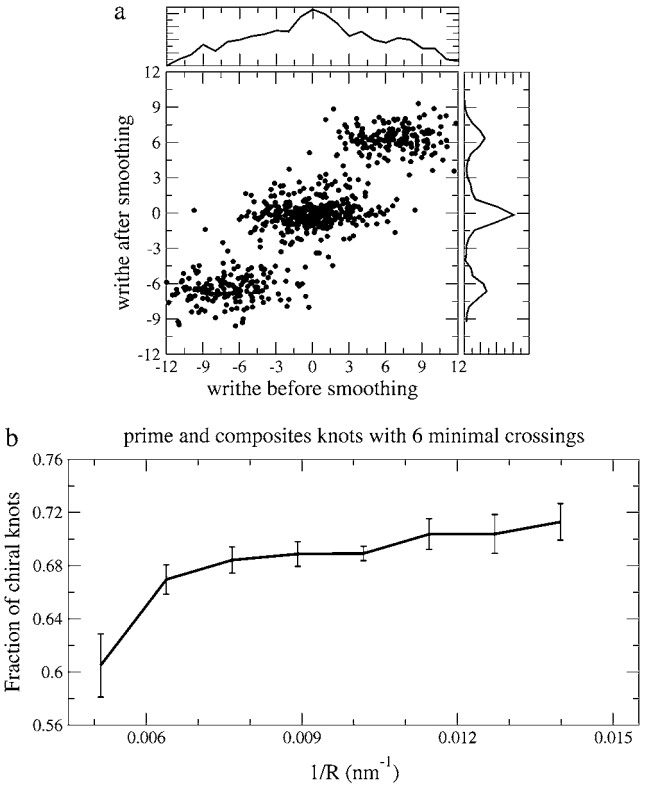

FIGURE 7 (*a*) Writhe of the composite six crossings knots conformations before and after smoothing. The side graphs represent the histogram of the writhe before and after smoothing. (*b*) Fraction of chiral knots (6_1, 6_2, and granny knots) within the class of knots with six minimal crossings (6_1, 6_2, 6_3, granny, and square knots). All results are for circular chains of $N = 200$ cylinders.

a mild, but persistent, bias toward chiral conformations upon confinement. Restricting to square and granny knots only, we find a small imbalance favoring the latter, again pointing to a confinement induced chirality bias.

We now investigate the second type of topological bias observed in P4 experiments, namely the fact that torus knots outnumber twist knots. The simplest instance of such bias occurs within the family of five crossing knots: in P4 experiments the 5_2 twist knot is found to occur much less frequent that the 5_1 torus knots (11). We have accordingly undertaken an analogous investigation of the knot spectrum of semiflexible self-avoiding circular chains. The top panel of Fig. 8 shows our data for the breakup of the five crossing knots into 5_1 and 5_2. The data indicate a bias opposite to the one found in the P4 experiments; in fact, throughout the reliable $1/R$ window, twist knots occur more frequently than torus ones. An alternative way to check whether confinement gives a bias to torus knots is to consider the ratio between the probabilities of occurrences of 5_1 and 5_2 knots as a function of $1/R$ (*bottom panel* of Fig. 8). Again, this ratio does not appear to display a significant increasing trend, so that we conclude that our data at these confinements are compatible with no torus knot bias.

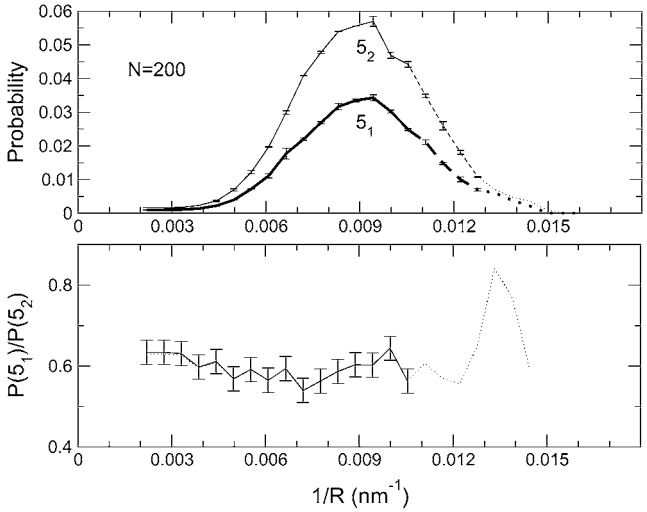

FIGURE 8 (*Top*) Probability of occurrences of 5_1 and 5_2 knots as a function of the inverse radius $1/R$ and for circular chains with $N = 200$ cylinders. (*Bottom*) Ratio between the probability of having a 5_1 torus knot and a 5_2 twist knot.

This qualitative incongruity with experiments poses the question of whether equilibrium properties of elastic self-avoiding chains are inadequate to account for the P4 data. While this is a possibility, it is not necessarily so as the degree of compactification probed by the experiments and the equilibrium simulations are, by necessity, different. In the P4 experiment, the number of the five crossing knots total to $\sim$0.26%, so that this regime falls definitely beyond our reliable $1/R$ window. Therefore, it cannot be ruled out that the tails of the 5_1 and 5_2 distributions intersect at the appropriate $1/R$ value.

DISCUSSION AND CONCLUSIONS

In conclusion we have performed Monte Carlo simulations to find the knotting probabilities and knot spectrum of a DNA molecule inside a confined region. The motivation for this study comes from the characterization of the knots which are formed inside the P4 bacteriophage. We have modeled P4 DNA as a semiflexible self-avoiding polymer, with nominal persistence length and thickness equal to 50 and 2.5 nm, respectively. The interaction potential which we have used corresponds to screened electrostatic interactions so it can be realized in practice with a monovalent salt buffer with a large enough salt concentration (e.g., 0.1–1 M) (23). The values of the confining diameter for which we manage to generate statistically uncorrelated configurations is $\sim$2.5 times larger than the diameter of the P4 phage, albeit significantly before then our algorithms fail to reliably detect knot types due to the complexity of the structure in terms of number of crossings. The different packing fraction achieved by simulated and real P4 DNA should be kept in mind when comparing the results: one should expect qualitative or at best semiquantitative agreement with experiments.

Even at this level there are still several questions to ask to assess whether sampling conformations of self-avoiding circular DNA in thermodynamic equilibrium leads to realistic trends in knotting probabilities as observed in the bacteriophage experiments. We asked, does confinement lead to a preference of chiral versus achiral knots? Does it also lead to a bias toward torus knots as opposed to twist knots? Both these trends were apparent from the observations, and they were speculated to be related to the onset of spooling order in the confined DNAs.

Our results show that the ratio between chiral and achiral six crossing knots does reproduce a chirality bias, as the fraction of chiral knots is $>$50% and increases with $1/R$, i.e., with confinement. Coming to composite knots, still with six crossings, we also computed the relative fractions of chiral granny knots and of (globally) achiral square knots, finding a small imbalance favoring the former at least inside the window of $1/R$ values for which our knot identifier can be proved to work reliably. There is therefore a promising qualitative agreement on the observation of a chirality bias which is present, although small, in confined knots. We believe this result is consistent with some studies in the literature on compactified linear self-avoiding polymers or copolymers (42,43), which were shown to display a bias toward locally helical or chiral configurations (but with no systematic preferred handedness). Compactification was ensured in those studies by means of a two-body attractive potential and not via the constraint of being contained inside a sphere of prescribed radii. A similar phenomenon leads to the formation of regular helical structures, again with no preference for a specific handedness, for self-avoiding tubes subject to compactification via locally or globally hydrophobic or solvophobic interactions (44–48). In Banavar et al. (47) it was also proposed from a geometric point of view that helices naturally arise when maximizing the surface area of a tube which is accessible to solvent molecules.

On the other hand, we did not find an appreciable tendency to favor torus over twist knots with confinement. In particular, in our reliable $1/R$ window the 5_2 twist knot probability is larger than the 5_1 torus knot probability, while in the experiments the opposite is found, and, more importantly, the ratio between torus and twist knots does not appreciably increase with $1/R$. Would the observed bias toward torus knots still be found at P4 levels of DNA packing? Or does one have to look for nonthermodynamic explanations of this feature of the experiments, e.g., kinetic effects which favor spooling and are due to the way the DNA is fed inside the capsid? (Cyclization of the DNA inside the capsid may occur with the two ends being either parallel or antiparallel when joining, with the latter possibility being likely disfavored as it leads to the formation of a kink. As the annealing into a circle is a nonequilibrium process, there is no guarantee that the fraction of parallel versus antiparallel ends found in the experiment is close to the thermodynamic equilibrium value. This might be another kinetic factor affecting knot spectrum.) These are both interesting questions and our present view is

that one should attempt to answer them just in the order in which we have posed them. To this end we are presently trying to enhance both our sampling and our knot finding algorithms to extend the value of $1/R$ for which we can confidently assess knot type. This is a significant computational challenge and it might well be that, to reach reliable results for smaller values of R, one will have to resort to a finer-grained modeling to resolve kinks and bends in the DNA filament at a scale which is shorter than the length of the cylindrical segments currently employed. Still, it is very important to do a reference calculation to quantitatively assess how far one can go in explaining the observed data in a thermodynamic framework, as this will also lead to a clear quantification of any nonequilibrium effect.

It would also be of interest to study theoretically the effect of salt concentration and electrostatic interactions on the knot spectrum, as well as of the persistence length of the confined semiflexible polymers as these studies may suggest new experiments either with phage DNA in different buffers or with synthetic semiflexible polymers in vesicles.

We acknowledge financial support from the Italian Ministry for Education (grant No. PRIN-2006025255) and from Regione Friuli Venezia Giulia (Biocheck, grant No. 200501977001).

REFERENCES

1. Kindt, J., S. Tzlil, A. Ben-Shaul, and W. M. Gelbart. 2001. DNA packaging and ejection forces in bacteriophage. *Proc. Natl. Acad. Sci. USA*. 98:13671–13674.
2. Ali, I., D. Marenduzzo, and J. M. Yeomans. 2006. Polymer packaging and ejection in viral capsids: shape matters. *Phys. Rev. Lett*. 96:208102.
3. Lof, D., K. Schillen, B. Jonsson, and A. Evilevitch. 2007. Forces controlling the rate of DNA ejection from phage lambda. *J. Mol. Biol*. 368:55–65.
4. Ali, I., D. Marenduzzo, and J. M. Yeomans. 2008. Ejection dynamics of polymeric chains from viral capsids: effect of solvent quality. *Biophys. J*. 94:4159–4164.
5. Arsuaga, J., R. K. Z. Tan, M. Vazquez, D. W. Sumners, and S. C. Harvey. 2002. Investigation of viral DNA packaging using molecular mechanics models. *Biophys. Chem*. 101:475–484.
6. Petrov, A. S., M. B. Doz, and S. C. Harvey. 2007. The conformation of double-stranded DNA inside bacteriophages depends on capsid size and shape. *J. Struct. Biol*. 160:241–248.
7. Petrov, A. S., and S. C. Harvey. 2008. Packaging double-helical DNA into viral capsids: structures, forces and energetics. *Biophys. J*. In press.
8. Ali, I., D. Marenduzzo, and J. M. Yeomans. 2004. Dynamics of polymer packaging. *J. Chem. Phys*. 121:8635–8641.
9. Marenduzzo, D., and C. Micheletti. 2003. Thermodynamics of DNA packaging inside a viral capsid: the role of DNA intrinsic thickness. *J. Mol. Biol*. 330:485–492.
10. Siber, A., M. Dragar, V. A. Parsegian, and R. Podgornik. 2008. Packaging nanomechanics of viral genomes. *Eur. Phys. J. E*. In press.
11. Arsuaga, J., M. Vazquez, S. Trigueros, D. W. Sumners, and J. Roca. 2002. Knotting probability of DNA molecules confined in restricted volumes: DNA knotting in phage capsids. *Proc. Natl. Acad. Sci. USA*. 99:5373–5377.
12. Arsuaga, J., M. Vazquez, P. McGuirk, S. Trigueros, D. W. Sumners, and J. Roca. 2005. DNA knots reveal a chiral organization of DNA in phage capsids. *Proc. Natl. Acad. Sci. USA*. 102:9165–9169.
13. Trigueros, S., J. Arsuaga, M. E. Vazquez, D. W. Sumners, and J. Roca. 2001. Novel display of knotted DNA molecules by two dimensional gel electrophoresis. *Nucleic Acids Res*. 29:e67.
14. Liu, Z. R., J. K. Mann, E. L. Zechiedrich, and H. S. Chan. 2006. Topological information embodied in local juxtaposition geometry provides a statistical mechanical basis for unknotting by type-2 DNA topoisomerases. *J. Mol. Biol*. 361:268–285.
15. Rollins, G. C., A. S. Petrov, and S. C. Harvey. 2008. The role of DNA twist in the packaging of viral genomes. *Biophys. J*. 94:L38–L40.
16. Liu, Z. R., and H. S. Chan. 2008. Efficient chain moves for Monte Carlo simulations of a wormlike DNA model: excluded volume, supercoils, site juxtapositions, knots, and comparisons with random-flight and lattice models. *J. Chem. Phys*. 128:145104.
17. Michels, J. P. J., and F. W. Wiegel. 1986. On the topology of a polymer ring. *Proc. Roy. Soc. A*. 403:269–284.
18. Arsuaga, J., T. Blackstone, Y. Diao, K. Hinson, E. Karadayi, and M. Saito. 2007. Sampling large random knots in a confined space. *J. Phys. A Math. Theor*. 40:11697–11711.
19. Micheletti, C., D. Marenduzzo, E. Orlandini, and D. W. Sumners. 2006. Knotting of random ring polymers in confined spaces. *J. Chem. Phys*. 124:064903.
20. Tesi, M. C., E. J. J. van Rensburg, E. Orlandini, and S. G. Whittington. 1994. Knot probability for lattice polygons in confined geometries. *J. Phys. A*. 27:347–360.
21. Katrich, V., W. K. Olson, A. V. Vologodskii, J. Dubochet, and A. Stasiak. 2000. Tightness of random knotting. *Phys. Rev. E Stat. Phys. Plasmas Fluids Relat. Interdiscip. Topics*. 61:5545–5549.
22. Dobay, A., J. Dubochet, K. Millett, P. E. Sottas, and A. Stasiak. 2003. Scaling behavior of random knots. *Proc. Natl. Acad. Sci. USA*. 100: 5611–5615.
23. Rybenkov, V. V., N. R. Cozzarelli, and A. V. Vologodskii. 1993. Probability of DNA knotting and the effective diameter of the DNA double helix. *Proc. Natl. Acad. Sci. USA*. 90:5307–5311.
24. Marko, J. F., and E. D. Siggia. 1994. Entropic elasticity of λ-phage DNA. *Science*. 265:1599.
25. Vologodskii, A. V., and N. R. Cozzarelli. 1995. Modeling of long-range electrostatic interactions in DNA. *Biopolymers*. 35:289–296.
26. Shimamura, M. K., and T. Deguchi. 2000. Characteristic length of random knotting for cylindrical self-avoiding polygons. *Phys. Lett. A*. 274:184–191.
27. Calvo, J. A., K. C. Millett, and E. J. Rawdon. 2005. Physical and Numerical Models in Knot Theory. World Scientific, Singapore.
28. Tesi, M. C., E. J. J. Janse van Rensburg, E. Orlandini, and S. G. Whittington. 1996. Monte Carlo study of the interacting self-avoiding walk model in three dimensions. *J. Stat. Phys*. 82:155–181.
29. Ferrenberg, A. M., and R. H. Swendsen. 1989. Optimized Monte Carlo data-analysis. *Phys. Rev. Lett*. 63:1195–1198.
30. Calvo, J. A., K. C. Millett, and E. J. Rawdon, editors. 2001. Physical Knots: Knotting, Linking and Folding Geometric Objects in R^3. AMS Special Session on Physical Knotting and Unknotting, Las Vegas, Nevada, April 21–22, 2001 (Contemporary Mathematics).
31. Koniaris, K., and M. Muthukumar. 1991. Self-entanglement in ring polymers. *J. Chem. Phys*. 95:2873–2881.
32. Taylor, W. R. 2000. A deeply knotted protein structure and how it might fold. *Nature*. 406:916–919.
33. Dowker, C. H., and M. B. Thistlethwaite. 1983. Classification of knot projections. *Topology Appl*. 16:19–31.
34. Hoste, J., and M. Thistlethwaite. 1999. *Knotfind* routine. KnotScape 101 Software Package. www.math.utk.edu/morwen/knotscape.html.
35. Adams, C. C. 1994. The Knot Book. W. H. Freeman and Company, New York.
36. Katrich, V., J. Bednar, D. Michoud, J. Dubochet, and A. Stasiak. 1996. Geometry and physics of knots. *Nature*. 384:142–145.

37. Vologodskii, A. V., N. J. Crisona, B. Laurie, P. Pieranski, V. Katritch, J. Dubochet, and A. Stasiak. 2002. Sedimentation and electrophoretic migration of DNA knots and catenanes. *J. Mol. Biol.* 278:1–3.

38. Ohyama, Y. 1993. Crossing number and the braid index of links. *Canad. J. Math.* 45:117–131.

39. Franks, J., and R. F. Williams. 1987. Braids and the Jones polynomial. *Trans. Am. Math. Soc.* 303:97–108.

40. Morton, H. R. 1986. Seifert circles and knot polynomials. *Math. Proc. Cambridge Press Soc.* 99:107–109.

41. Cerf, C., and A. Stasiak. 2000. A topological invariant to predict the three-dimensional writhe of ideal configurations of knots and links. *Proc. Natl. Acad. Sci. USA.* 97:3795–3798.

42. Toan, N. M., D. Marenduzzo, P. R. Cook, and C. Micheletti. 2006. Depletion effects and loop formation in self-avoiding polymers. *Phys. Rev. Lett.* 97:178302.

43. Magee, J. E., V. R. Vasquez, and L. Lue. 2006. Helical structures from an isotropic homopolymer model. *Phys. Rev. Lett.* 96:207802.

44. Maritan, A., C. Micheletti, A. Trovato, and J. R. Banavar. 2000. Optimal shapes of compact strings. *Nature.* 406:287–290.

45. Stasiak, A., and J. H. Maddocks. 2000. Mathematics—best packing in proteins and DNA. *Nature.* 406:251–253.

46. Marenduzzo, D., A. Flammini, A. Trovato, J. R. Banavar, and A. Maritan. 2005. Physics of thick polymers. *J. Polym. Sci. [B].* 43:650–679.

47. Banavar, J. R., T. X. Hoang, J. H. Maddocks, A. Maritan, C. Poletto, A. Stasiak, and A. Trovato. 2007. Structural motifs of biomolecules. *Proc. Natl. Acad. Sci. USA.* 104:17283–17286.

48. Yu, B., Q. Jin, D. Ding, B. Li, and A.-C. Shi. 2008. Confinement-induced morphologies of cylinder-forming asymmetric diblock copolymers. *Macromolecules.* 41:4042–4054.