Theory and simulations of toroidal and rod-like structures in single-molecule DNA condensation

ARTICLE *in* THE JOURNAL OF CHEMICAL PHYSICS · MARCH 2015

Impact Factor: 2.95 · DOI: 10.1063/1.4914513 · Source: PubMed

4 AUTHORS:

Ruggero Cortini
Pierre and Marie Curie University - Paris 6
6 PUBLICATIONS **30** CITATIONS
SEE PROFILE

Bertrand R Caré
Pierre and Marie Curie University - Paris 6
12 PUBLICATIONS **72** CITATIONS
SEE PROFILE

Jean-Marc Victor
French National Centre for Scientific Resea…
59 PUBLICATIONS **1,107** CITATIONS
SEE PROFILE

Maria Barbi
Pierre and Marie Curie University - Paris 6
32 PUBLICATIONS **794** CITATIONS
SEE PROFILE

Theory and simulations of toroidal and rod-like structures in single-molecule DNA condensation

Ruggero Cortini,[1, a)] Bertrand R. Caré,[1, b)] Jean-Marc Victor,[1, c)] and Maria Barbi[1, d)]
Laboratoire de Physique Théorique de la Matière Condensée, UMR 7600, Université Pierre et Marie Curie, Sorbonne Université, 4 place Jussieu, 75252 Cedex 05, Paris, France

DNA condensation by multivalent cations plays a crucial role in genome packaging in viruses and sperm heads, and has been extensively studied using single-molecule experimental methods. In those experiments, the values of the critical condensation forces have been used to estimate the amplitude of the attractive DNA-DNA interactions. Here, to describe these experiments, we developed an analytical model and a rigid body Langevin dynamics assay to investigate the behavior of a polymer with self-interactions, in the presence of a traction force applied at its extremities. We model self-interactions using a pairwise attractive potential, thereby treating the counterions implicitly. The analytical model allows to accurately predict the equilibrium structures of toroidal and rod-like condensed structures, and the dependence of the critical condensation force on the DNA length. We find that the critical condensation force depends strongly on the length of the DNA, and finite-size effects are important for molecules of length up to $10^5 \mu$m. Our Langevin dynamics simulations show that the force-extension behavior of the rod-like structures is very different from the toroidal ones, so that their presence in experiments should be easily detectable. In double-stranded DNA condensation experiments, the signature of the presence of rod-like structures was not unambiguously detected, suggesting that the polyamines used to condense DNA may protect it from bending sharply as needed in the rod-like structures.

The following article appeared in R. Cortini *et al.*, The Journal of Chemical Physics **142**, 105102 (2015) and may be found at `http://dx.doi.org/10.1063/1.4914513`

I. INTRODUCTION

Single-molecule DNA micromanipulations provide a very powerful tool for the study of DNA mechanics[1,2], DNA-DNA interactions[3–6], and DNA-protein interactions[7]. In the case of DNA-DNA interactions, it was possible to study the reentrant behavior of cation-condensed DNA[3,8,9], and measure accurately the free energy of DNA condensation[5]. These studies are of great importance to determine the behavior of condensed DNA, which is crucial for DNA packaging and genomic ejection from viruses[10], and also to provide an accurate estimate of the forces involved.

In single-molecule DNA condensation experiments, a DNA molecule is tethered at one end to a surface, and at the other end to a micron-sized bead, which is trapped by magnetic or optical tweezers. The experimental extension-force curves show that there is a critical traction force, above which the system is fully extended and behaves as a worm-like chain semiflexible polymer, and below which the system progressively folds to a completely condensed state[3–5,11]. This critical condensation force allows to estimate the attractive DNA-DNA interaction per unit length. So far, this was done by neglecting finite size effects for long λ DNA molecules.

DNA condensation in single-molecule experiments was already studied theoretically[12,13], but the geometry of the condensed structure was always assumed *a priori* to be toroidal. It is possible however that toroids are not the only possible geometry for the condensate, as was already noticed in theoretical studies of condensation of semiflexible polymers with self-interactions[14,15]. It is therefore interesting to explore the possibility of the appearance of other geometrical configurations in single-molecule DNA condensation experiments.

In this paper we propose to study DNA condensation in single-molecule experiments by means of Langevin dynamics (LD) simulations, together with an analytical model. The scope of this work is two-fold: on the one hand, to better account for finite-size effects in the estimate of attractive DNA-DNA interactions using measured critical condensation forces; on the other hand, to evaluate the possibility of the appearance of the rod-like geometry of the condensate, alongside the well-known toroidal one.

Our LD simulation method is based on the work of Carrivain et al[16], which has been successfully applied to modeling single-molecule DNA supercoiling experiments[17]. The main advantage of using this technique is its speed and enhanced sampling efficiency, based on (a) efficient algorithms to simulate rigid bodies[18] and (b) the global thermostat scheme introduced by Bussi and Parrinello[19]. Using this method, we are able to observe nucleation of DNA condensation within our simula-

[a)] Electronic mail: cortini@lptl.jussieu.fr
[b)] Electronic mail: care@lptl.jussieu.fr
[c)] Electronic mail: victor@lptl.jussieu.fr
[d)] Electronic mail: barbi@lptl.jussieu.fr

tion windows. The well-known toroidal structure of DNA condensates[20] is not the only geometry that we observe: rod-like structures also appear.

The analytical model we propose is based on the theoretical ansatz proposed by Hoang et al.[21]. This was developed to study toroidal and rod-like structures of free DNA in solution, during crowding- or cation-induced condensation. We extended the model proposed by these authors to include the presence of a traction force applied at the extremities of the DNA. Using this model, we are able to evaluate the different contributions to the free energy of the system: the bending energy, the surface tension, the bulk attractive force, and the worm-like chain entropic free energy of the non-condensed region of the DNA. We are also able to calculate the critical condensation force at any DNA length.

In both our simulations and our analytical model we consistently model the counterions that are present in solution implicitly, using a pairwise attractive potential. The validity of this approximation is supported by a favorable comparison of our results with experimental data on DNA condensation.

II. METHODS

A. Geometry

The geometry of our system is described in figure 1. We describe a polymer of length L, tethered at one end

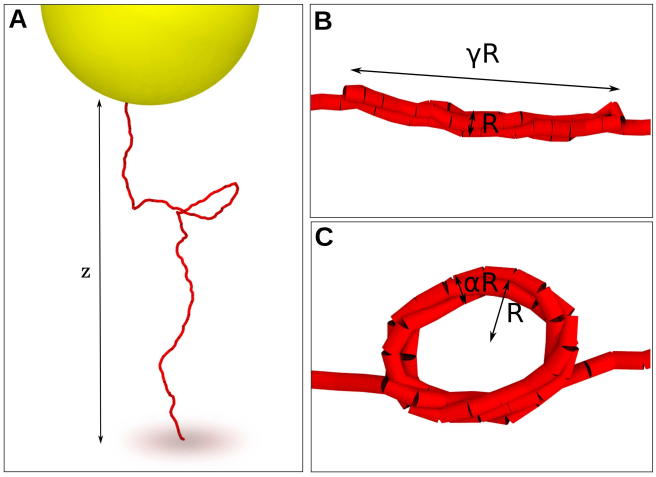

FIG. 1. Geometry of the single-molecule DNA manipulation experiment.(A) Bead and DNA molecule attached to it; (B) rod-like condensed structure: R is the rod diameter, γ is the aspect ratio of the rod; (C) toroid-like condensed structure: R is the toroid radius, α is the ratio between the radius and the thickness of the toroid.

to a surface and at another end to a bead. In magnetic tweezer experiments, the force F acting on the bead is imposed, and the end-to-end distance z is measured. In optical tweezers, the bead is trapped around a fixed height z_0, and the force is measured. These two experimental setups correspond to different statistical ensembles: at fixed force (where we use the Gibbs free energy), and at fixed end-to-end extension (where the Helmholtz free energy is used).

B. Analytical model

We describe the polymer as a made of two different phases: a condensed phase characterized by a length L_c and an extended phase of length $L - L_c$. We therefore have the following two expressions for the free energy, at fixed force and at fixed end-to-end extension respectively:

$$G = E_c + (L - L_c) g_{WLC}(F), \tag{1}$$

$$A = E_c + (L - L_c) a_{WLC}\left(\frac{z}{L - L_c}\right), \tag{2}$$

where E_c is the energy of the condensed phase, g_{WLC} and a_{WLC} are the free energies per unit length of the worm-like chain phase, in the fixed force and fixed extension ensembles, respectively. In our model, we neglect the entropic contribution due to fluctuations inside the condensed phase: we consider that the condensed phase has a definite conformational structure. We now discuss separately the contributions from the condensed phase and from the extended phase.

For the condensed phase, we base our model on the recent work by Hoang et al.[21]. Here, we report the main results of this study. More details on the derivation of the following formulas, as well as a discussion on their validity, is reported in the original study[21].

The authors model a polymer chain with self-interactions as a chain of spheres of radius b connected by rigid bonds, and propose an energy ansatz for the toroidal and rod-like geometry. In our model, we take their expressions, and assign the energy of the condensed phase E_c to one of the two formulas:

$$\begin{aligned} E_{toroid}(\eta, \alpha, L_c) = {} & 2 l_p \pi^{2/3} \eta^{2/3} \alpha^{4/3} b^{-4/3} L_c^{1/3} + \\ & - \phi(d) d^{-1} b^{1/3} \pi^{4/3} \alpha^{-1/3} \eta^{-2/3} L_c^{2/3} + \\ & + 3\phi(d) b^{-1} L_c \end{aligned} \tag{3}$$

$$\begin{aligned} E_{rod}(\eta, \gamma, L_c) = {} & \frac{32}{3} l_p \eta \left[4\eta\left(\gamma + 4/3\right)\right]^{-1/3} L_c^{1/3} + \\ & - 2\pi\phi(d) d^{-1} b^{1/3} \left(\gamma + 2\right) \left[4\eta\left(\gamma + 4/3\right)\right]^{-2/3} L_c^{2/3} + \\ & + 3\phi(d) b^{-1} L_c. \end{aligned} \tag{4}$$

Both the toroid and the rod geometries are characterized by η, the packing fraction in the condensate. The toroid is described also by α, which is the ratio between the toroid radius and the thickness of the toroid; the

rod is described by γ, the rod aspect ratio between the width and the length of the rod (see figure 1). The interaction of the spheres is given by the potential function $\phi(d)$. We chose the Lennard-Jones potential energy function, which is numerically very close to the more realistic Morse potential:

$$\phi(d) = \epsilon \left[\left(\frac{\sigma}{d}\right)^{12} - 2\left(\frac{\sigma}{d}\right)^{6} \right]. \quad (5)$$

Here, σ is the Lennard-Jones radius of the monomers. The variable d, which represents the inter-monomer lateral distance in the condensed phase, is related to η by the following expression:

$$d = b\sqrt{\eta_{hex}/\eta}, \quad (6)$$

where $\eta_{hex} = \pi/6\cot(\pi/6)$. Finally, l_p is the polymer bending persistence length. Notice that the precise form of the interaction potential was found to be of minor importance in determining the position and length dependence of the toroid/rod transition line in the work of Hoang et al[21]. We also checked that the Morse potential gave little or no difference compared to the the Lennard-Jones potential in our analysis.

Both the toroid and the rod energy functions are a sum of three terms: the bending energy to fold the polymer into its condensed state (which is $\sim L_c^{1/3}$); the surface tension (which is $\sim L_c^{2/3}$), and the bulk attractive energy, which is proportional to L_c. The surface tension and the bulk energy terms are proportional to the amplitude of the attractive potential, ε. These terms are derived by supposing that the packing of the polymer in the condensate is hexagonal, so that one can calculate the average number of neighbors in the bulk, and the average number of molecules at the surface. As we shall see, the dependence of these terms on L_c has important consequences on the equilibrium properties of the system.

The free energy of the extended phase may be expressed using the formula derived by Marko and Siggia[22] for the Gibbs free energy per unit length of a worm-like chain semi-flexible polymer of length $L - L_c$:

$$g_{WLC}(F) = \min_a \left\{ \left(\frac{a}{2l_p} - F \right) \left(\coth 2a - \frac{1}{2a} \right) \right\}. \quad (7)$$

This expression is used to calculate the relative extension $\rho = z/(L - L_c)$ at thermal equilibrium

$$\rho_{WLC}(F) = -\frac{\partial g_{WLC}}{\partial F}. \quad (8)$$

To express the free energy per unit length in the fixed end-to-end extension ensemble, we must perform the Legendre transform of g_{WLC}:

$$\tilde{a}_{WLC} = g_{WLC}(F) - F\rho. \quad (9)$$

To obtain the force as a function of the extension, we invert the $\rho_{WLC}(F)$ function and we obtain the $F_{WLC}(\rho)$ function. Once equation 9 is specified for $F = F_{WLC}(\rho)$, we finally obtain

$$a_{WLC} = g_{WLC}\left[F_{WLC}(\rho)\right] - F_{WLC}(\rho)\rho. \quad (10)$$

The equilibrium condition is obtained by minimizing the free energy (either equation 1 or 2) with respect to all the variables.

We may derive interesting predictions from minimization of the Helmholtz free energy (equation 2). The minimum condition for L_c is given by

$$\left.\frac{\partial E_c}{\partial L_c}\right|_{L_c = L_c^\star} - g_{WLC}\left[F_{WLC}\left(\frac{z}{L - L_c^\star} \right) \right] = 0, \quad (11)$$

where the starred symbols refer to the equilibrium condition. The above equation is derived in the Appendix. In the thermodynamic limit, $L_c \to \infty$, and equation 11 may be approximated by

$$-3\varepsilon = g_{WLC}\left[F_{WLC}\left(\rho^\star\right)\right]. \quad (12)$$

This equation depends only on $\rho^\star = z/(L - L_c^\star)$, and not on z and L separately. Therefore, in the limit of large L we expect $z/(L - L_c^\star)$ to be constant. This may be translated into the following useful formula, which we will use later:

$$L_c^\star = L - \frac{z}{\rho^\star} \quad (13)$$

At fixed end-to-end extension, the force acting on the bead at thermal equilibrium is given by:

$$F = F_{WLC}\left[\rho^\star\right]. \quad (14)$$

This equation is derived in the Appendix.

To obtain the equilibrium structures, we use numerical routines for minimization of multi-variable functions, as implemented in the GNU Scientific Library[23].

To calculate the critical condensation force, we switch back to the fixed force ensemble. The transition from the fully extended state to the fully condensed state is obtained by equating the chemical potentials of these two phases, at the critical force. We then obtain the following expression:

$$\frac{E_c^\star}{L} = g_{WLC}(F_c), \quad (15)$$

where in $E_c^\star$ we set $L_c = L$. This equation may be solved numerically to obtain F_c. Since the bending energy and surface tension terms vary as $\sim L_c^{1/3}$ and $\sim L_c^{2/3}$ respectively, in the thermodynamic limit of $L \to \infty$, we obtain

$$-3\varepsilon = g_{WLC}(F_c^\infty), \quad (16)$$

where F_c^∞ is the critical force for an infinitely long chain.

C. Molecular dynamics simulations

Our simulation assay is based on the recently published work by Carrivain et al[16]. We mention here only the most relevant features of this simulation method. First, it is a simulation of rigid bodies, where the angular velocity is taken into account, and the simulated bodies have an impenetrable volume. In our case, the DNA is composed by N cylinders of length l_s, with a hard-core radius equal to the DNA crystallographic radius, and is attached to a magnetical or optical bead, which is treated as a sphere (see figure 1A). Second, the DNA chain is modelled by an articulated system, where the joints act as holonomic contraints of a mechanical system. Forces and torques may be applied at each joint, so to model bending and twisting energies. In our case, we apply only a bending term, and neglect the twisting contribution, which is irrelevant because the DNA molecules are typically nicked in the experiments. Finally, the coupling to the thermal bath is treated using the global thermostat scheme, introduced by Bussi and Parrinello[19]. This allows for fast equilibration of the system and realistic treatment of thermal fluctuations.

We model the bending energy associated to two successive cylinders as:

$$E_b(\theta) = \frac{1}{2} g_b \theta^2, \tag{17}$$

where θ is the angle formed by the tangent vectors of the two cylinders, and g_b is a constant related to the bending persistence length (for full details, see reference[16]).

Here, we include the effect of intra-DNA interactions, as modelled by the Lennard-Jones potential equation 5. In our system, the potential acts between the centers of mass of the cylinders that compose the articulated system. The Lennard-Jones radius σ was chosen so that the minimum of the potential corresponds to 28Å. The full list of parameters used in our simulations is given in table I.

To simulate optical tweezer experiments, we add a potential energy term to the bead, expressed as an isotropic harmonic trap:

$$E_{trap} = \frac{1}{2} k_{trap} (|\boldsymbol{r} - \boldsymbol{r}_0|)^2. \tag{18}$$

Here, k_{trap} is the stiffness of the optical trap. Typical values of this parameter are of the order of $0.1\mathrm{pN/nm}^3$. The vectors $\boldsymbol{r}$ and $\boldsymbol{r}_0 \equiv (x_0, y_0, z_0)$ are the actual position of the bead, and the position of the center of the force (which would correspond to the laser focus), respectively.

We run an equilibration round of 10^6 steps, using a local Langevin thermostat and no inter-monomer potential. After this, we switch to the global thermostat and turn on the inter-monomer interactions, and perform 10^7 LD steps for production run.

D. Data analysis

We briefly describe here a few methods we used to analyze the data from our LD simulations.

Condensate length. To calculate the length of the condensed region from an LD trajectory, we proceed as follows. First, we calculate the distance matrix between all the DNA segments at a given time step. That is, calculate $M_{ij} = |\boldsymbol{r}_i - \boldsymbol{r}_j|$. Then we select all segments in which $|i-j| > 1$, and $M_{ij} < l_s$. The number of segments that satisfy this condition, multiplied by l_s, gives L_c.

Free energy. The free energy of a given state may be estimated by calculating the total energy of the condensate (Lennard-Jones potential plus bending energy), and adding the worm-like chain contribution of the uncondensed region. This can be calculated by averaging L_c over the trajectory, and using the expression for the worm-like chain free energy per unit length proposed by Marko and Siggia (see equations 10 and 7). As we have done in our analytical model, we neglect the entropic contribution of fluctuations of the segments in the condensed phase.

Critical force. It is challenging to estimate F_c with LD simulations directly. At the critical point, the time scale for the formation of an initial condensation loop diverges. Therefore, it is difficult to estimate F_c by starting from a high force value and decreasing it, because the time scale for nucleation is inaccessible in LD simulations. On the other hand, a similar problem is encountered when starting from a condensed structure and increasing progressively the applied force. Here, the kinetic barrier between the folded and unfolded states is significant, therefore presenting the same conceptual problem. However, one can estimate F_c by fitting the $\langle L_c \rangle$ ($\langle z \rangle$) curves to a line, and estimating the intercept of the line with the $L_c = 0$ axis (see equation 13 and figure 6). The value of z at which $L_c = 0$ gives an estimate of $\rho^\star$, from which we can estimate the critical force using the $F_{WLC}(\rho)$ function.

Distinguishing between toroids and rods. To automatically detect whether a condensed state is toroidal or rod-like, we developed an heuristic method outlined below. Once the cylinders that are part of the condensed phase are identified, for each pair of cylinders i and j in contact, we evaluate $\cos\gamma_{ij} = \boldsymbol{t}_i \cdot \boldsymbol{t}_j$. If $\cos\gamma_{ij} > 0$, we say that those segments are in parallel contact, otherwise they are in antiparallel contact. We can then define an order parameter as $s = (N_p - N_a)/N_c$, where N_p and N_a are the number of parallel and antiparallel contacts, respectively, and N_c is the total number of contacts. For a toroid, most segments will be in parallel contact, so we expect $s \approx 1$. On the other hand, for a rod-like configuration we expect $s \approx 0$, since $N_p \approx N_a$.

III. RESULTS

We performed Langevin dynamics simulations of a 3 kb DNA molecule ($N = 150$, $L = 1\mu m$), as described in

Methods, using parameters shown in table I, in an optical tweezer-like setup. We compared the simulation results with the predictions of our analytical model (equations 1 and 2) using the same parameters, for the toroidal and rod-like geometries. We present here the results of these two approaches.

A. Fixed z_0/L simulations

First, we performed 20 independent simulations of the DNA chain, at fixed $z_0/L = 0.6$ and at $\varepsilon = 0.7 k_B T/\text{nm}$. We observed that after 10^7 time steps, 9 chains adopted the rod-like condensed state, 10 adopted the toroidal state, and one remained uncondensed. Snapshots of the rod-like and toroidal states are shown in figure 1B and 1C.

Figure 2 shows two example traces of $(L - L_c(t))/L$ for a toroid and a rod. The condensation of the chain occurs as discrete jumps, both for the toroid and rod-like geometries. The steps correspond to adhesion of successive portions of the DNA chain onto an initially formed condensation loop for toroids and a condensation stretch for rods. In the time trace of the toroid condensation, a more progressive, linear condensation step is also observed.

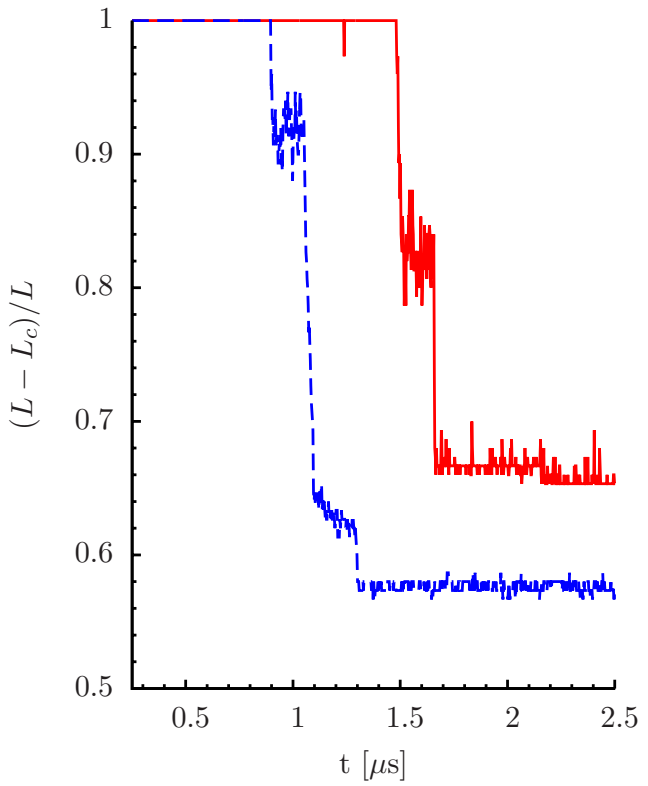

FIG. 2. Variation of the uncondensed length $(L - L_c)/L$ (L_c is calculated as explained in Methods, as a function of time for two specific simulations, one which nucleates into a rod-like condensate (solid red) and one which nucleates into a toroid (dashed blue).

B. Force-extension curves

Next, we used the final configurations of the simulations described in the previous section to perform stretching simulations. Here, every 10^6 time steps the value of z_0/L was incremented by 5%. The resulting force-extension curves are depicted in figure 3. The toroid- and

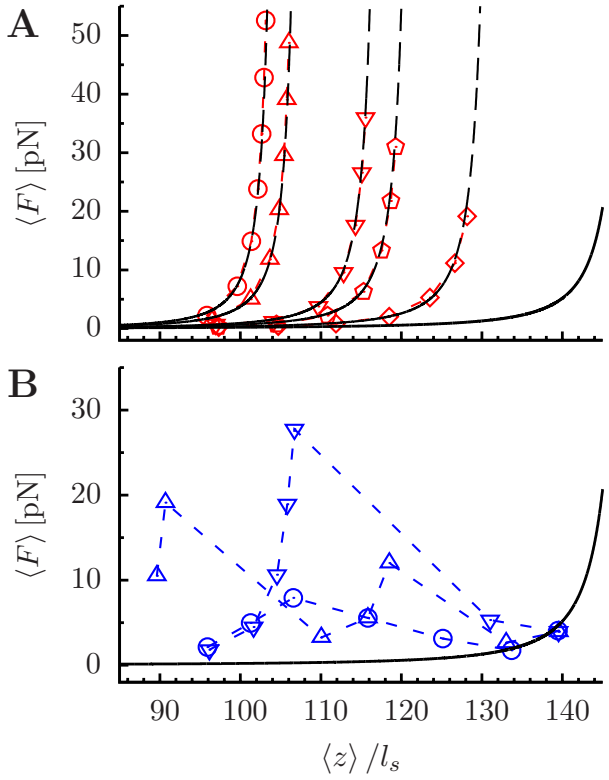

FIG. 3. Average force versus average extension of single stretched rods (A) and toroids (B). The different symbols correspond to simulations starting from independent configurations (see Main Text). The worm-like chain behavior for a completely uncondensed DNA is shown as a black continuous line. In A, black dashed lines correspond to the fit to a worm-like chain behavior for, from left to right, $L/l_s = 105.5$, $L/l_s = 108.5, 118.5, 122.5, 132.5$. The simulation protocol is described in Methods, and the parameters used are shown in table I.

rod-like structures have a markedly different response under stretching force. The toroidal structures unfold progressively, exhibiting a saw-tooth-like force-extension behavior, eventually becoming fully extended (see final blue points in figure 3). The peaks in the force value can go up to ∼30pN. In contrast, rod-like structures never unfold, even at the highest forces probed here (∼50pN). In all our simulations, we could not unfold rod-like structures, whereas toroidal structures all eventually unfold.

The difference in the force-extension behavior of toroids and rods under tension can be understood in terms of the direction of the applied force. In a rod-like condensate, the rod axis is in general approximately par-

Parameter	Description	Value
l_p	Bending persistence length	45 nm
T	Absolute temperature	300 K
N	Number of DNA segments	150
n_s	Number of base pairs per DNA segment	20
Ω	Thermostat coupling frequency	3 $10^{10} s^{-1}$
k_{trap}	Optical trap stiffness	0.2 pN nm
ε	Lennard-Jones potential well depth	0.2–0.7 k_BT/nm
σ	Lennard-Jones radius	25.08 Å

TABLE I. Summary of the parameters used in Langevin dynamics simulations and in our analytical model (where applicable).

allel to the direction of the applied force (see figure 1B). As the force increases, the direction of the force becomes more and more orthogonal to the direction of the DNA-DNA contacts. This makes it difficult for the force to unfold the condensate. On the other hand, the toroidal structures have a larger degree of lateral mobility, that results in fluctuations aligning the lateral contacts with the direction of the applied force.

C. Free energy of toroidal and rod-like structures

To further gain insight on the system, and to test the results of our analytical model (equations 1 and 2), we performed simulations at different values of z_0/L. For each value of z_0/L, we obtained results for 10 independent starting configurations. We then calculated the free energy of the final configurations as outlined in Methods, averaging over the 10 simulations. Figure 4 shows the comparison of the numerical minimization of our analytical model, for the toroid and rod geometries, and the free energies obtained from Langevin dynamics simulations as outlined above.

For each value of z_0/L that we studied, the toroidal states are the ones having the lowest free energy. The free energy minima for toroidal and rod-like configurations, however, are remarkably close (within $\sim 0.1 k_BT$/nm difference). This fact is also manifest in the observation of simulations in which a toroidal and a rod-like condensed phase coexist within the same DNA chain (not shown).

Our simulations show that at $\langle z \rangle / L > 0.7$ for toroids and $\langle z \rangle / L > 0.6$ for rods, the condensed state ceases to exist. At values of $\langle z \rangle / L$ higher than this, the free energy is equal to the worm-like chain free energy.

We notice that there is a significant spread between the values of the free energies calculated in our simulations. However, the lowest values of the free energy for a given conformation are remarkably close to the predictions of our analytical model. This means that in the other cases, the conformation adopted by the chain is not an equilibrium one, but a local minimum. This is a sign that the final conformation is strongly dependent on the nucleation of a first condensation loop.

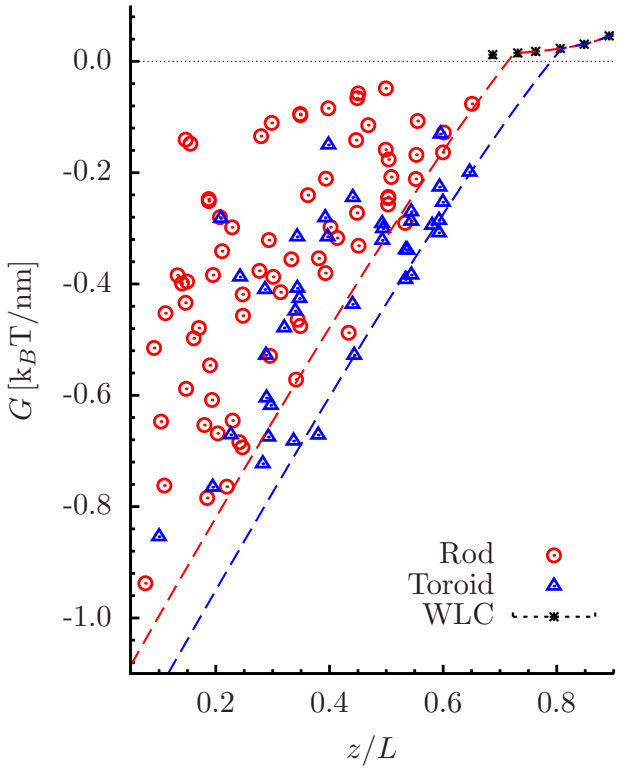

FIG. 4. Minimum free energy of rod-like and toroidal geometries (blue and red dashed lines, respectively), as calculated from numerical minimization of equation 2, and average free energies of final states in Langevin dynamics simulations, calculated as explained in Methods, for rods (red circles) and toroids (blue triangles).

D. Critical condensation force

Figure 5 shows the values of the critical condensation force F_c as estimated from the analytical model (see equation 15), as a function of the DNA length, at fixed ε. The two values chosen for ε correspond to the case of spermidine (0.20 k_BT/nm) and spermine (0.33 k_BT/nm) given in reference[4]. As we explained in the

Methods section, for a very long chain the critical condensation force reaches a finite value F_c^∞, and becomes independent of the geometrical structure of the condensate. Here, we show that this thermodynamic limit is reached only for very long molecules. As an example, for $\varepsilon = 0.33 k_B T/\text{nm}$, the value of F_c reaches F_c^∞ only for $L = 10^8$nm, the size of a genomic DNA molecule. The experimental values of the maximum condensation force, as extracted from the work of Todd et al[4] and Murayama et al[3] are also depicted. There is good agreement between the theoretical and experimental values.

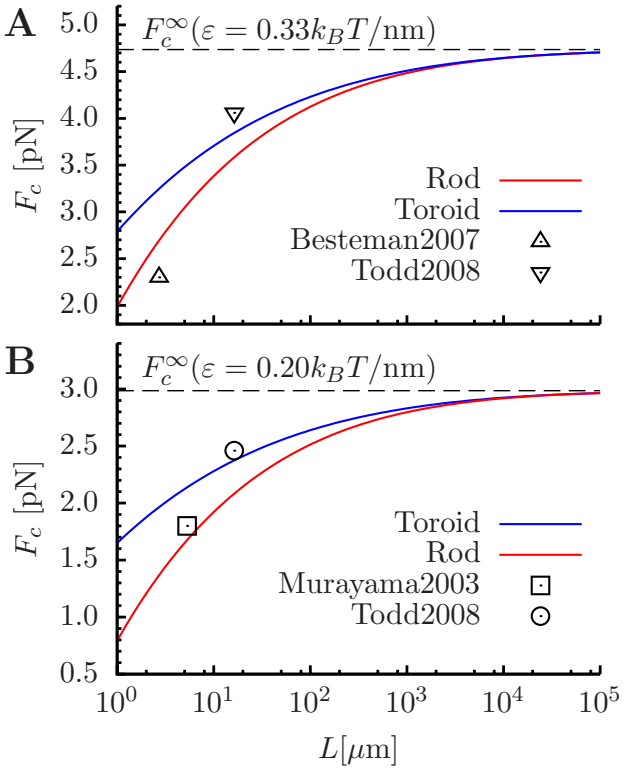

FIG. 5. Dependence of the critical condensation force F_c on the length of the DNA, for $\varepsilon = 0.33 k_B T/\text{nm}$ (A) and $\varepsilon = 0.20 k_B T/\text{nm}$ (B). These two values correspond to the case of spermidine and spermine, respectively. The points are extracted from the experimental data of Murayama et al[3], Besteman et al[9], and Todd et al[4].

The reason for the very slow convergence of F_c to F_c^∞ is to be found in the surface tension term. In fact, this term is proportional to ε (see equations 3 and 4, which is typically a very significant value, and depends on $L_c^{2/3}$, which is not far from linear dependence. Therefore, this term represents a significant energy penalty even for a very long DNA.

As outlined in the Methods section, one can estimate the critical condensation force from LD simulations by looking at the dependence of the condensed length L_c on z. Figure 6 shows the variation of the average L_c, for toroids and rods, at fixed z. The data clearly shows a linear dependence of $< L_c >$ on $< z >$, which is one of the key predictions of our analytical model (see Methods). A linear fit of this data provides a value of the intercept

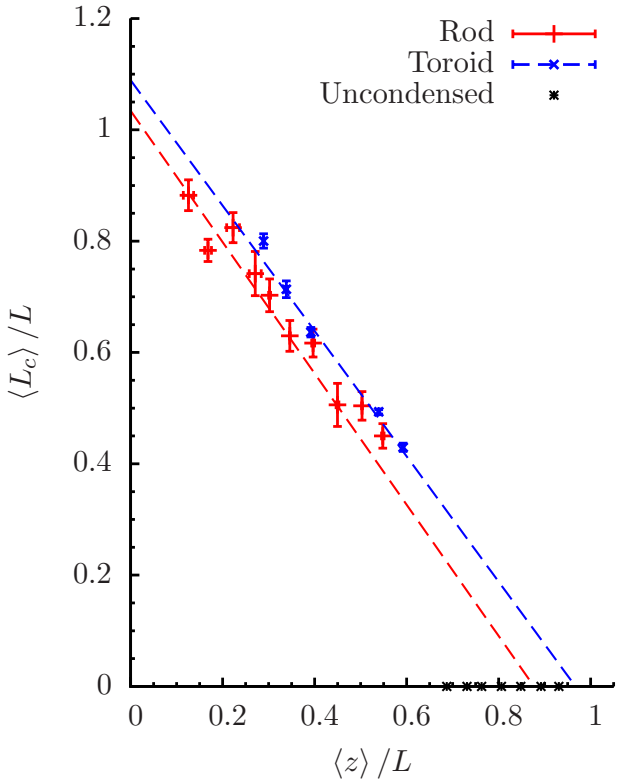

FIG. 6. Condensed length in simulations, as a function of average $\langle z \rangle / L$. Condensed length was calculated as described in Methods. Error bars are standard deviations. The dashed lines are linear fits to the data.

with the $z = 0$ axis which is close to 1, meaning that in the absence of an external force, the whole molecule is in the condensed state. The value of $\rho^\star$ estimated from our data is 0.96 and 0.87 for toroids and rods, respectively. Unfortunately, small errors in $\rho^\star$ result in large errors on the estimate of F_c, as for $\rho > 0.85$ the $F_{WLC}(\rho)$ function has a very steep increase (see figure 3 for comparison). Therefore, we estimate F_c to be between ≈1.5 and 20pN, at $\varepsilon = 0.7 k_B T/\text{nm}$.

IV. DISCUSSION

In this study, we developed an analytical theory and a simulation method to investigate the behavior of a DNA under traction and in the presence of self-attraction. This study was primarily aimed at the investigation of single-molecule DNA condensation in optical and magnetic tweezers[3–6,9,11,24]. Whereas our simulations allow to study the kinetic aspects of DNA condensation, our analytical theory is aimed at the study of the equilibrium properties of the system. The latter are experimentally

accessible, using a very slow unloading rate of a magnetic bead (see e. g. Ref.[6]). In this section we discuss our results in relation to the experimental data.

A. General features of single-molecule DNA condensation experiments are reproduced by our simulations

Firstly, we notice that several experimental features of DNA condensation are well captured by our model. The condensation of the chain proceeds in discrete steps, and each step was shown to correspond to folding of a portion of the molecule onto an initial condensation loop[6,24]. When the condensate is toroidal, the force-extension curve has the characteristic saw-tooth shape, which was seen in experiments[3] and in previous simulations of DNA condensation[13]. The force peaks of our simulations are quantitatively in agreement with the experimental ones, being in the 10-30 pN range (see figure 3). Our values are somewhat higher than the experimental ones, and this may due to the limited time window accessible by LD simulations. Moreover, we find that the conformation of the condensed structures appearing in the LD simulations is strongly dependent on the size of the initial condensation loop, which is a well-known feature of DNA toroids observed in electron microscopy experiments[25–27].

Despite its simplicity, our simple assumption of a pairwise potential between the monomers is able to capture some of the principal features of the system.

B. Evaluating the amplitude of DNA-DNA interactions using values of critical condensation force

The analysis of the dependence of the critical condensation force on the length of the DNA gives an interesting prediction: the limit of infinitely long chains is reached very slowly, as shown in figure 5. In the work of Todd et al[4,5], the authors accurately measured the condensation force at different ionic concentrations. The condensation force has a maximum value as a function of the concentration of multivalent ions. For this value, one can show that the entropic contribution to the free energy due to ion mixing vanishes, and one can estimate the amplitude of the attractive DNA-DNA interactions using the value of the maximum critical force. This estimate is based on the idea that for a very long DNA molecule, the total free energy of the condensate may be approximated by its bulk value. Our study provides a quantitative way of veryfing this hypothesis. Figure 7 shows the relationship between the estimated values of ε and the condensation force. The dashed line corresponds to the hypothesis of Todd et al, i. e. assuming direct proportionality between ε and F_c, independently of the molecular length. Remarkably, the curve for $L = 16.3\mu$m (corresponding to the λ DNA molecules used in the experiments of Todd et al) falls almost exactly on the dashed line. Therefore, even though ε should have been estimated more rigorously, the values supplied by Todd et al for ε are correct. Using the fact that those values are correct, we also set other values of experimental critical forces in the figure, as extracted from the works of Murayama et al[3] and Besteman et al[9], along with the theoretical estimates corresponding to the molecular lengths used in those experiments. The theoretical and experimental values are in reasonable agreement.

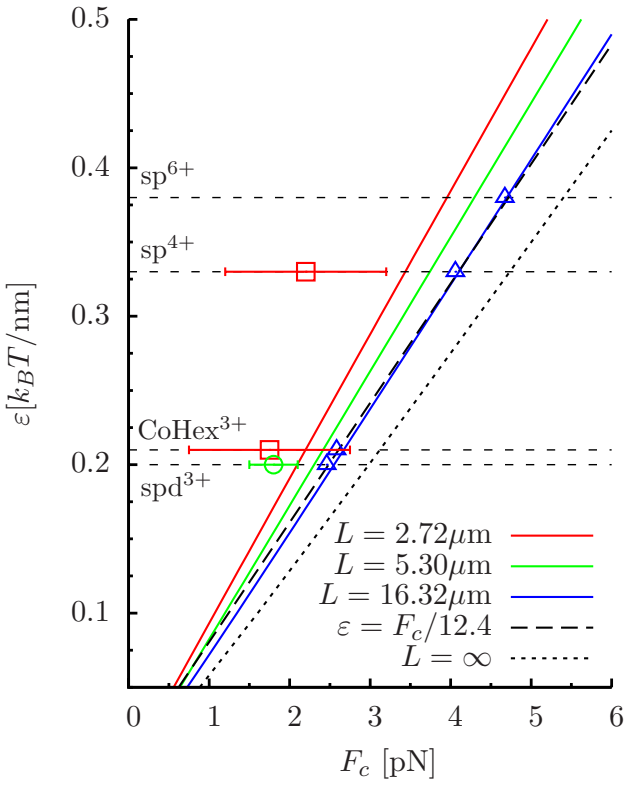

FIG. 7. Estimated value of the amplitude of the DNA-DNA attraction ε, from a given measured value of the maximum critical condensation force F_c. Solid colored lines correspond to the theoretical prediction of our analytical model, as extracted by numerically solving equation 15. Dotted line corresponds to the case of $L = \infty$. Dashed line corresponds to the estimate done by Todd et al[4], assuming direct proportionality between ε and F_c. Experimental points were assigned by assuming that the values of ε estimated by Todd et al[4](blue triangles) were correct (see main text for further details) at a given ionic condition, and assigning the points extracted from the studies of Murayama et al[3](green circle) and by Besteman et al[9](red squares) to the corresponding ionic conditions.

C. Force-extension curves allow to discriminate between toroids and rods

This study also showed that rod-like condensates have a different force-extension behavior compared to toroids. In the paper by Baumann et al.[11], it was reported that

some of the plasmid-length DNA molecules could not be stretched to the full original extension. Our results may provide a possible explanation for the two different force-extension behaviors seen experimentally. To our knowledge, the results of Baumann et al. are the only ones that show two different force-extension curve classes. In other experimental works[3,6] on the same system, this was not reported.

The geometrical configuration of the condensate is highly dependent on the model chosen for the DNA bending. If one allows for formation of kinks in the chain, with low energetic cost, the rod-like structure becomes the most favorable geometry for the condensate. In contrast, adding a high energy cost for the hairpin turns strongly favors the toroidal geometry (data not shown). These results are in line with previous theoretical studies[14,28]. This, together with the fact that the force-extension signature of rods was not clearly reported in the literature, suggests that the high bending angles required in the rod-like structures have a higher energetic cost than that expected by a simple harmonic model as we used here (see equation 17). One possibility is that the polyamines used in DNA condensation experiments protect the secondary structure of DNA from kinking or bending sharply. It is in fact known that the polyamines stabilize base-pairing[29] and do not disrupt the B-DNA structural form[30]. It is therefore unlikely that DNA kinks are favored by inhomogeneous patterns of adsorption of counterions on DNA.

The structure of condensed DNA is a subject of great interest[10], because of the importance, for instance, of DNA ejection from viral capsids[31]. Our results show that under the influence of an external traction force, the toroidal condensed structures are the only ones that can readily unfold. When a DNA molecule becomes rod-like, the unfolding is strongly impeded. This could be an important element to take into account when envisaging therapeutic strategies against DNA viruses.

ACKNOWLEDGMENTS

The authors wish to thank D. J. Lee, A. Kornyshev, A. Giacometti, G. Wuite, and A. Grosberg for useful discussions. This work has been funded by the French Institut National du Cancer, grant INCa 5960 and by the French Agence Nationale de la Recherche, grant ANR-13-BSV5-0010-03.

Appendix: Derivation of equations 11 and 14

The force acting on the bead at thermal equilibrium is obtained by taking the total derivative of the minimum Helmholtz free energy with respect to z:

$$F = \frac{\mathrm{d}A^\star}{\mathrm{d}z} = \sum_i \left.\frac{\partial E_c}{\partial X_i}\right|_{X_i=X_i^\star} + \\ + \left.\frac{\partial}{\partial z}\left[(L-L_c)a_{WLC}\left(\frac{z}{L-L_c}\right)\right]\right|_{L_c=L_c^\star}, \quad \text{(A.1)}$$

where the starred symbols indicate the equilibrium values of the variables. In the above expression, the first term on the right-hand side vanishes, because the gradient of the condensed energy is zero at equilibrium. The second term is the only one that contributes to the force, and may be calculated by straightforward differentiation. The first step is to calculate the derivative of a_{WLC} with respect to ρ:

$$\frac{\partial}{\partial\rho}a_{WLC}(\rho) = \frac{\partial}{\partial\rho}\left\{g_{WLC}\left[F_{WLC}(\rho)\right] + F_{WLC}(\rho)\rho\right\} = \\ = \frac{\partial F_{WLC}}{\partial\rho}\frac{\partial g_{WLC}}{\partial F_{WLC}} + \rho\frac{\partial F_{WLC}}{\partial\rho} + F_{WLC}(\rho) = \\ = F_{WLC}(\rho), \quad \text{(A.2)}$$

where we have taken into account equation 8. Using this result we may now calculate easily the force acting on the molecule at equilibrium by inserting equation A.2 into equation A.1:

$$F = \left.\frac{\partial}{\partial z}\left[(L-L_c)\,a_{WLC}\left(\frac{z}{L-L_c}\right)\right]\right|_{L_c=L_c^\star} = F_{WLC}(\rho^\star), \quad \text{(A.3)}$$

which is equation 14.

We may now explicitly calculate the derivative of the Helmholtz free energy equation 2 with respect to L_c:

$$\left.\frac{\partial E_c}{\partial L_c}\right|_{L_c=L_c^\star} - \left.\frac{\partial}{\partial L_c}\left[(L-L_c)\,a_{WLC}\left(\frac{z}{L-L_c}\right)\right]\right|_{L_c=L_c^\star} = 0. \quad \text{(A.4)}$$

The second term on the left-hand side of this equation can be calculated using again equation A.2:

$$\frac{\partial}{\partial L_c}\left[(L-L_c)\,a_{WLC}\left(\frac{z}{L-L_c}\right)\right] = \\ = -a_{WLC}(\rho) + (L-L_c)\frac{\partial\rho}{\partial L_c}\frac{\partial}{\partial\rho}a_{WLC}(\rho) = \\ = -a_{WLC}(\rho) + (L-L_c)\frac{z}{(L-L_c)^2}F_{WLC}\left(\rho\right) = \\ = -g_{WLC}\left[F_{WLC}\left(\frac{z}{L-L_c}\right)\right]. \quad \text{(A.5)}$$

Using this result, equation A.4 becomes

$$\left.\frac{\partial E_c}{\partial L_c}\right|_{L_c=L_c^\star} - g_{WLC}\left[F_{WLC}\left(\rho^\star\right)\right] = 0, \quad \text{(A.6)}$$

which is equation 11.

[1] S.B. Smith, L. Finzi, and C. Bustamante. Direct mechanical measurements of the elasticity of single DNA molecules by using magnetic beads. *Science*, 258(5085):1122–1126, 1992.

[2] TR Strick, J.F. Allemand, D. Bensimon, A. Bensimon, and V. Croquette. The elasticity of a single supercoiled DNA molecule. *Science*, 271(5257):1835, 1996.

[3] Y. Murayama, Y. Sakamaki, and M. Sano. Elastic Response of Single DNA Molecules Exhibits a Reentrant Collapsing Transition. *Physical Review Letters*, 90(1):018102, 2003.

[4] B.A. Todd, V. Adrian Parsegian, A. Shirahata, TJ Thomas, and D.C. Rau. Attractive forces between cation condensed DNA double helices. *Biophysical journal*, 94(12):4775–4782, 2008. ISSN 0006-3495.

[5] B.A. Todd and D.C. Rau. Interplay of ion binding and attraction in DNA condensed by multivalent cations. *Nucleic acids research*, 36(2):501–510, 2008.

[6] B. Van den Broek, M.C. Noom, J. Van Mameren, C. Battle, F.C. MacKintosh, and G.J.L. Wuite. Visualizing the formation and collapse of DNA toroids. *Biophysical journal*, 98(9):1902–1910, 2010.

[7] M.C. Noom, B. van den Broek, J. van Mameren, and G.J.L. Wuite. Visualizing single DNA-bound proteins using DNA as a scanning probe. *Nature Methods*, 4(12):1031–1036, 2007.

[8] K. Besteman, S. Hage, NH Dekker, and SG Lemay. Role of tension and twist in single-molecule DNA condensation. *Physical review letters*, 98(5):58103, 2007.

[9] K. Besteman, K. Van Eijk, and SG Lemay. Charge inversion accompanies DNA condensation by multivalent ions. *Nature Physics*, 3(9):641–644, 2007.

[10] Pascal Carrivain, Axel Cournac, Christophe Lavelle, Annick Lesne, Julien Mozziconacci, Fabien Paillusson, Laurence Signon, Jean-Marc Victor, and Maria Barbi. Electrostatics of DNA compaction in viruses, bacteria and eukaryotes: functional insights and evolutionary perspective. *Soft Matter*, 8(36):9285–9301, 2012.

[11] C.G. Baumann, V.A. Bloomfield, S.B. Smith, C. Bustamante, M.D. Wang, and S.M. Block. Stretching of single collapsed DNA molecules. *Biophysical journal*, 78(4):1965–1978, 2000.

[12] C. Battle, B. Van Den Broek, MC Noom, J. Van Mameren, GJL Wuite, and FC MacKintosh. Unraveling DNA tori under tension. *Physical Review E*, 80(3):031917, 2009.

[13] Paul Cárdenas-Lizana and Pai-Yi Hsiao. Stick-release pattern in stretching single condensed polyelectrolyte toroids. *Macromolecules*, 42(8):3211–3214, 2009.

[14] Mark J Stevens. Simple simulations of DNA condensation. *Biophysical journal*, 80(1):130–139, 2001.

[15] T Sakaue and K Yoshikawa. Folding/unfolding kinetics on a semiflexible polymer chain. *The Journal of chemical physics*, 117(13):6323–6330, 2002.

[16] Pascal Carrivain, Maria Barbi, and Jean-Marc Victor. In Silico Single-Molecule Manipulation of DNA with Rigid Body Dynamics. *PLoS computational biology*, 10(2):e1003456, 2014.

[17] F. Mosconi, J.F. Allemand, D. Bensimon, and V. Croquette. Measurement of the torque on a single stretched and twisted DNA using magnetic tweezers. *Physical review letters*, 102(7):78301, 2009. ISSN 1079-7114.

[18] R. Smith. Open dynamics engine, 2008. `http://www.ode.org`.

[19] Giovanni Bussi and Michele Parrinello. Stochastic thermostats: comparison of local and global schemes. *Computer Physics Communications*, 179(1):26–29, 2008.

[20] N.V. Hud and I.D. Vilfan. Toroidal DNA condensates: unraveling the fine structure and the role of nucleation in determining size. *Annual Review of Biophysics and Biomolecular Structure*, 34:295–318, 2005.

[21] Trinh Xuan Hoang, Achille Giacometti, Rudolf Podgornik, Nhung TT Nguyen, Jayanth R Banavar, and Amos Maritan. From toroidal to rod-like condensates of semiflexible polymers. *The Journal of chemical physics*, 140(6):064902, 2014.

[22] J.F. Marko and E.D. Siggia. Stretching DNA. *Macromolecules*, 28(26):8759–8770, 1995.

[23] M. Galassi, B. Gough, G. Jungman, J. Theiler, J. Davies, M. Booth, and F. Rossi. *GNU Scientific Library Reference Manual - Third Edition*. Network Theory Ltd., 2009. ISBN 0954612078. URL `http://www.gnu.org/software/gsl`.

[24] W.B. Fu, X.L. Wang, X.H. Zhang, S.Y. Ran, J. Yan, and M. Li. Compaction dynamics of single DNA molecules under tension. *Journal of the American Chemical Society*, 128(47):15040–15041, 2006.

[25] M.R. Shen, K.H. Downing, R. Balhorn, and N.V. Hud. Nucleation of DNA condensation by static loops: formation of DNA toroids with reduced dimensions. *Journal of the Americal Chemical Society*, 122(19):4833–4834, 2000.

[26] N.V. Hud and K.H. Downing. Cryoelectron microscopy of λ phage DNA condensates in vitreous ice: the fine structure of DNA toroids. *Proceedings of the National Academy of Sciences of the United States of America*, 98(26):14925, 2001.

[27] C.C. Conwell, I.D. Vilfan, and N.V. Hud. Controlling the size of nanoscale toroidal DNA condensates with static curvature and ionic strength. *Proceedings of the National Academy of Sciences of the United States of America*, 100(16):9296, 2003.

[28] H Noguchi and K Yoshikawa. Morphological variation in a collapsed single homopolymer chain. *The Journal of chemical physics*, 109(12):5070–5077, 1998.

[29] A.A. Ouameur and H.A. Tajmir-Riahi. Structural analysis of DNA interactions with biogenic polyamines and cobalt (III) hexamine studied by Fourier transform infrared and capillary electrophoresis. *Journal of Biological Chemistry*, 279(40):42041–42054, 2004.

[30] Hong Deng, Victor A Bloomfield, James M Benevides, and George J Thomas Jr. Structural basis of polyamine–DNA recognition: spermidine and spermine interactions with genomic B-DNAs of different GC content probed by Raman spectroscopy. *Nucleic acids research*, 28(17):3379–3385, 2000.

[31] A. Leforestier and F. Livolant. Structure of toroidal DNA collapsed inside the phage capsid. *Proceedings of the National Academy of Sciences*, 106(23):9157, 2009. ISSN 0027-8424.