Article

Thermodynamics of Long Supercoiled Molecules: Insights from Highly Efficient Monte Carlo Simulations

Thibaut Lepage,[1,2] François Képès,[1,3] and Ivan Junier[2,4,5,*]

[1]Institute of Systems and Synthetic Biology, Genopole, CNRS, University of Évry, Évry, France; [2]Laboratoire Adaptation et Pathogénie des Micro-organismes-UMR 5163, Université Grenoble 1, CNRS, Grenoble, France; [3]Department of BioEngineering, Imperial College London, London, United Kingdom; [4]Centre for Genomic Regulation (CRG), Barcelona, Spain; and [5]Universitat Pompeu Fabra (UPF), Barcelona, Spain

ABSTRACT Supercoiled DNA polymer models for which the torsional energy depends on the total twist of molecules (Tw) are a priori well suited for thermodynamic analysis of long molecules. So far, nevertheless, the exact determination of Tw in these models has been based on a computation of the writhe of the molecules (Wr) by exploiting the conservation of the linking number, Lk = Tw + Wr, which reflects topological constraints coming from the helical nature of DNA. Because Wr is equal to the number of times the main axis of a DNA molecule winds around itself, current Monte Carlo algorithms have a quadratic time complexity, $O(L^2)$, with respect to the contour length (L) of the molecules. Here, we present an efficient method to compute Tw exactly, leading in principle to algorithms with a linear complexity, which in practice is $O(L^{1.2})$. Specifically, we use a discrete wormlike chain that includes the explicit double-helix structure of DNA and where the linking number is conserved by continuously preventing the generation of twist between any two consecutive cylinders of the discretized chain. As an application, we show that long (up to 21 kbp) linear molecules stretched by mechanical forces akin to magnetic tweezers contain, in the buckling regime, multiple and branched plectonemes that often coexist with curls and helices, and whose length and number are in good agreement with experiments. By attaching the ends of the molecules to a reservoir of twists with which these can exchange helix turns, we also show how to compute the torques in these models. As an example, we report values that are in good agreement with experiments and that concern the longest molecules that have been studied so far (16 kbp).

INTRODUCTION

The folding properties of DNA are strongly constrained by the physical intricacy of its strands. In a situation of torsional stress, this induces the polymer to curl up for complementary bases to properly face each other. As a result, when DNA molecules are underwound (negative supercoiling) or overwound (positive supercoiling), so-called plectonemes form (Fig. 1). Although these properties have important implications both for the physics of single molecules (1,2) and for our understanding of DNA biology (3), several aspects have remained elusive, especially for long molecules that exceed 10 kilobasepairs (kbp). A proper modeling of DNA that can handle multiple scales, from the kbp scale (gene scale) to a scale of hundreds of kbp, is nevertheless a strong requirement for improving our capacity to model biological systems (4), more particularly to rationalize the impact of chromosome structuring on the functioning of cells (5).

The fact that local properties of double-stranded DNA can be accommodated only via a global folding of the molecule stems from the existence of a topological invariant, the linking number (Lk), which occurs both for circular molecules and for linear molecules whose ends cannot rotate (Fig. 1). Namely, the conservation of Lk imposes that the variation of the twist (Tw) be equal to the opposite variation of the writhe (Wr):

$$\mathrm{Lk} = \mathrm{Tw} + \mathrm{Wr}, \tag{1}$$

with Tw reflecting the torsional properties of the molecule and Wr its global geometry (see below) (6).

Direct information about the conformations of supercoiled molecules was first provided by numerical simulations of circular molecules (1). Deep insights have then been provided at the single-molecule level thanks to magnetic tweezers experiments (7–9). In these experiments, one end of the molecule is maintained fixed, whereas the other end is tethered to a superparamagnetic bead that can be rotated using small magnets (see Figs. 1 and 4). The end-to-end extension of the molecule is measured as a function of the stretching force, f, applied to the molecule and of the number of turns of the magnetic bead, that is, as a function of the supercoiling density, $\sigma = (\mathrm{Lk} - \mathrm{Lk}_0)/\mathrm{Lk}_0$, where Lk_0 is the linking number of the molecule at rest. The formation of plectonemes is then monitored indirectly by a distinctive diminution of the extension as σ varies (10–12) (see Fig. 4).

Several theoretical models have been able to reproduce with good accuracy the mean extension of these molecules (2,13–15). Nevertheless, predictions differ in the numbers

Submitted December 23, 2014, and accepted for publication June 2, 2015.

*Correspondence: ivan.junier@ujf-grenoble.fr

Editor: Keir Neuman.

0006-3495/15/07/0135/9 $2.00

http://dx.doi.org/10.1016/j.bpj.2015.06.005

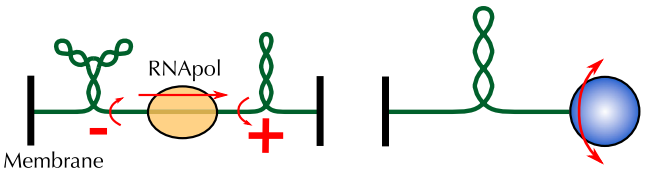

FIGURE 1 Torsionally induced super-structuring of DNA molecules. (*Left*) In vivo, plectonemes, which can be branched, occur because, e.g., of the activity of RNA polymerases and of the anchoring of DNA to the membrane (52). (*Right*) Topological constraints are exerted in vitro by keeping one end of a molecule fixed and adding helix turns at the other end, thanks to magnetic tweezers (9). To see this figure in color, go online.

of plectonemes and in the types of structures that are able to form (14,15). Apparently incompatible reports have also raised important questions about the proper way to measure the torques these molecules can exert (15). Recently, structural predictions have been compared to experiments that combine magnetic tweezers with fluorescent staining of DNA, allowing direct visualization of the formation and dynamics of plectonemes (16). In these experiments, multiple plectonemes were observed to form in a wide range of forces and supercoiling densities, putting some important constraints on the outcomes of models.

Polymer simulations are expected to provide deep insights into the problem. In this regard, Monte Carlo (MC) methods are particularly well adapted to investigate thermodynamic properties of supercoiled molecules and, in particular, to rationalize single-molecule experiments (17–19). To this end, DNA is modeled as a wormlike chain (WLC) and the WLC is simulated by discretizing the chain into a succession of rigid segments that include bending and torsional energies (1). In the spirit of Brownian dynamics (20–22), a torsional energy can then be defined locally using Euler angles between local frames associated to the segments, with, in particular, the possibility of measuring the torques these molecules can exert (18,19). In this case, a local twist can be defined between any two consecutive segments as the sum of the local Euler angles, providing an approximation of the exact twist that is all the better when torsional angles are small (18,20,21).

The torsional energy can also be defined globally, as a function of the total twist of the molecule (17,23,24), with the advantage of leading to much smaller relaxation times (25) (see below). In this case, the exact twist can be computed by using the topological relationship $\mathrm{Tw} = \mathrm{Lk} - \mathrm{Wr}$ (17,23,25); Lk is then constrained to remain constant during the simulation (topological constraint), and Wr, which is a geometrical property solely of the main axis of the molecule, is computed at every iteration. As Wr is a double integral over this main axis (26), the corresponding algorithms have nevertheless a quadratic time with respect to the contour length (L) of the polymer, making them poorly adapted for investigating thermodynamic properties of long (e.g., >10 kbp) molecules. Moreover, in contrast to models with local torsional energies (18,19), torque computation has remained an open issue in these models.

Here, we propose, to our knowledge, a novel treatment of the topological properties of WLC-based DNA models by including an explicit representation of the double-helix, which allows an easy exact computation of Tw. As a result, we obtain a highly efficient MC algorithm that both relies on a global torsional energy and features a theoretical linear time complexity, $O(L)$, which in practice is $O(L^{1.2})$. Our approach also provides a basis for computing torques in models with global torsional energies. As an application, we study the thermodynamic properties of 16 kbp (11) and 21 kbp long molecules (16), which are the longest molecules that have been exhaustively investigated so far.

MATERIALS AND METHODS

The self-avoiding supercoiled wormlike chain with a global torsional energy

Our model is based on the self-avoiding supercoiled wormlike chain (sWLC) introduced by Vologodskii and collaborators (23,27), which has been successfully used in the context of magnetic tweezers experiments to rationalize extension curves and to provide structural aspects of small molecules (17). In this model, self-avoidance of the chain simplifies the treatment of the short-range electrostatic repulsions of DNA at the cost, nevertheless, of introducing an effective radius that cannot be determined directly experimentally and that is related to the second virial coefficient of the sWLC (25). Here, in accord with recent Brownian dynamics studies (28), best results were obtained over a wide range of experimental conditions by considering an effective radius, r_e, that is given by $r_e = r_0 + \lambda_D$; $r_0 = 1$ nm is the radius of the DNA double-helix (B-form) and $\lambda \simeq 10/\sqrt{[\mathrm{NaCl}]}$, with [NaCl] in mM, and λ in nm corresponding to the Debye screening length for an aqueous solution of NaCl at room temperature (29). In the sWLC model, the conformational energy of the chain reads

$$E = E_\mathrm{b} + E_\mathrm{t} + E_\mathrm{s}, \tag{2}$$

with E_b, E_t, and E_s the energies of bending, torsion, and stretching, respectively. The bending energy is given by $E_\mathrm{b} = K/2 \int_0^L (\partial_s \vec{t})^2 ds$, where $\vec{t}(s)$ is the vector tangent to the chain at the curvilinear abscissa s and K is the bending modulus that defines the persistence length: $l_\mathrm{p} = K/k_\mathrm{B}T$. The stretching energy is given by $E_\mathrm{s} = -fz$, where z is the molecular extension along the direction of the stretching force, f.

Strictly speaking, the torsional energy should be given by $E_\mathrm{t} = 2\pi^2 k_\mathrm{B} TC \int_0^L (\tau(s) - \tau_0)^2 ds$, with C the torsional modulus and $\tau(s)$ the local twist of the molecule, that is, the number of times per unit of contour length that the DNA helix turns around its main axis, τ_0 being the corresponding value at rest (0.29 turns/nm for the B-form of DNA). Nevertheless, E_t can be decomposed as the sum of two terms: $2\pi^2 k_\mathrm{B} TC(\mathrm{Tw} - \mathrm{Tw}_0)^2/L$, which reflects the cost of having a total twist, Tw, different from that at rest (Tw_0), plus a term accounting for the fluctuations around Tw. The latter term can then be dropped without affecting the folding properties of DNA, with the advantage of drastically reducing the sWLC equilibration time (17,24). As a result, we use a sWLC for which the torsional energy is a function of the total twist (Tw) only:

$$E_\mathrm{t} = 2\pi^2 k_\mathrm{B} TC(\mathrm{Tw} - \mathrm{Tw}_0)^2/L. \tag{3}$$