Reactive and Functional Polymers 131 (2018) 243–250

Contents lists available at ScienceDirect

Reactive and Functional Polymers

journal homepage: www.elsevier.com/locate/react

Coarse-grained models of double-stranded DNA based on experimentally determined knotting probabilities

Florian C. Rieger[a,b,*], Peter Virnau[a]

[a] Institute of Physics, Johannes Gutenberg University Mainz, Staudingerweg 9, D-55128 Mainz, Germany
[b] Graduate School Materials Science in Mainz, Staudingerweg 9, D-55128 Mainz, Germany

ARTICLE INFO

Keywords:
DNA modelling
DNA knots
Monte Carlo simulation
Model fitting
Kratky-Porod wormlike chain model

ABSTRACT

To accurately model double-stranded DNA in a manner that is computationally efficient, coarse-grained models of DNA are introduced, where model parameters are selected by fitting the spectrum of observable DNA knots: We develop a general method to fit free parameters of coarse-grained chain models by comparing experimentally obtained knotting probabilities of short DNA chains to knotting probabilities that are computed in Monte Carlo simulations, resulting in coarse-grained DNA models which are tailored to reflect DNA topology in the best possible way. The method is exemplified by fitting ideal chain models as well as a bead-spring model with excluded volume interactions, to model double-stranded DNA for physiological as well as for high salt concentrations. The resulting coarse-grained DNA models predict the correct persistence length and effective diameter of double-stranded DNA, and can in principle be used for dynamical investigations using Molecular Dynamics. Our modelling ansatz thus provides a blueprint for building coarse-grained models of polymers, which are solely based on knotting spectra.

1. Introduction

Mathematical polymer models, designed for large-scale computational studies of double-stranded DNA (dsDNA), are coarsened descriptions of the DNA molecule, tailored to model its most important physical properties. With time and growing computational power, DNA models which include more details of DNA structure have become tractable, meeting the demands of more sophisticated theoretical studies [1]. If physical interactions and the geometry of dsDNA are modelled in detail, the range of possible dsDNA models is broadened, and an increased number of physical and geometric parameters have to be known to specify the computational model. On the other hand, for coarser models, the relation between model parameters and observable physical quantities is obscured.

Choosing proper values of DNA model parameters is a fundamental problem in coarse-grained modelling of DNA. The aim of this work is to introduce and illustrate a general fitting procedure which selects parameters of coarse-grained dsDNA models to mimic global topological properties, i.e. polymer self-entanglements, of dsDNA: Parameters are chosen so that experimentally observed knotting probabilities of short dsDNA chains are in agreement with model predictions. The fitting procedure is widely applicable as it is solely based on treating dsDNA as a space curve, and does not depend on the details of a specific polymer model.

Knots in polymers [2, 3] are known to be more or less likely to occur, depending on overall physical conditions [4–9]. The probability of knot formation, seen as a fingerprint of overall system conditions, may therefore be used to gauge DNA model parameters. More specifically, the likelihood of knots in dsDNA [3, 10] sensitively depends on salt conditions: Knotting probabilities of short dsDNA strands were first measured by gel electrophoresis [11, 12], finding that, due to screening of electrostatic interactions, for high salt concentrations, the fraction of knotted chain conformations is increased. More recently, knotting probabilities of significantly longer DNA chains for high salt concentrations have been obtained by studying translocation events in solid-state nanopores [13]. The first theoretical estimate of dsDNA knotting probability was obtained in [14], simulating lattice random walks to model knot formation on chain closure, and assuming a segment length of $b = 100$nm. In the 1990s, the first seminal attempts to model DNA [11] based on topological information were undertaken using experimental knotting probabilities from [11] in conjunction with the known persistence length of DNA. The resulting model consists of a chain of cylinders, whose diameter and stiffness was obtained from matching knotting spectra and persistence length, respectively. This work builds upon earlier investigations [15, 16], which study the effect of excluded volume on knotting probabilities. In [17], the topological

* Corresponding author at: Institute of Physics, Johannes Gutenberg University Mainz, Staudingerweg 9, D-55128 Mainz, Germany.
E-mail address: riegerf@uni-mainz.de (F.C. Rieger).

https://doi.org/10.1016/j.reactfunctpolym.2018.08.002
Received 14 January 2018; Received in revised form 29 June 2018; Accepted 2 August 2018
Available online 09 August 2018
1381-5148/

approach has been extended to a bead-stick model. In contrast to [11], however, the knotting spectrum from [11, 12] was the only input, and self-consistently the correct persistence length of DNA (50 nm) was recovered for physiological salt conditions. In [17], simulations corresponding to dsDNA strands of up to half a million base pairs were undertaken and are seen to be in good agreement with experimental results from [13] (see Fig. 2 below).

In [17], the outlined fitting procedure had been specialized to model dsDNA for physiological salt concentrations by a bead-stick chain with hard-core excluded volume interactions. In this work, ideal chain models as well as a bead-spring model with excluded volume interactions (which can in principle be used for dynamical simulations) are fitted to model dsDNA for physiological as well as for high salt concentrations of $c = 1.0$M NaCl:

We first illustrate the method by fitting the most simple ideal chain model, the random walk or freely-jointed chain, as well as a semi-flexible chain model, the Kratky-Porod model [18], which, in case that the polymer is described as a continuous space curve of fixed length, is known as the wormlike chain model [19–21]: For sufficiently long dsDNA strands, experiments show that the wormlike chain, much more so than the freely-jointed chain, accurately describes the stretching elasticity of dsDNA, with measured force-extension curves of dsDNA and predictions of the wormlike chain model in close agreement for low and intermediate forces [22]. For physiological salt concentrations, the persistence length l_p of dsDNA is known to be roughly $l_p \approx 50$nm [23–25]. Furthermore, among the numerous experimental methods that have been employed to estimate l_p, inference of l_p from the measured rate of formation of DNA circles [26, 27] has been instrumental in estimating how intrinsic curvature of dsDNA contributes to l_p [28], and has also been used to study the sequence dependence of DNA rigidity [29]. In (3) it is shown that the fitted Kratky-Porod model as well as the fitted bead-spring chain predict a persistence length of dsDNA in close agreement with experimental findings, for physiological as well as for high salt concentrations.

Ideal chain models of dsDNA do not contain any effects of DNA self-interaction. Arguably, the simplest approach to define a real chain model of dsDNA is to model the combined effect of excluded volume and electrostatic interactions by introducing an effective dsDNA diameter [11, 30]: In [11], dsDNA has been modelled as a sequence of impenetrable cylinders of fixed length and diameter, using a fixed value for the dsDNA Kuhn length as input to model chain bending, with the length of cylinders given by the Kuhn length. The effective diameter of dsDNA sensitively depends on salt concentration, reflecting the screening of repulsive polyelectrolyte self-interactions: In [11], experimentally measured and simulated knotting probabilities have been used to estimate the salt dependent effective dsDNA diameter, which for the first time demonstrated the use of knotting probabilities to predict model parameters. In (3.2) it is shown that the fitted bead-spring model predicts an effective dsDNA diameter that is consistent with the results in [11], for physiological as well as for high salt concentrations.

2. Theory and methods

To illustrate how fitting of knotting probabilities can be utilized to introduce coarse-grained models of dsDNA, in (3), the method is applied to fit model parameters of a bead-spring model, the Kratky-Porod model with quadratic bending potential, as well as the random walk model. In [17], the method has been used to introduce a bead-stick model of dsDNA. These models, as well as the implementation of computer simulations to derive knotting probabilities, are discussed in the following sections. Technical details of the fitting procedure are introduced as well.

2.1. Mathematical models

While the well-known freely-jointed chain (or random walk) model, which describes a polymer chain as a sequence of N jointed segments of fixed length b and arbitrary orientation, requires no additional model parameters to be fully defined, the Kratky-Porod model [18] also models the bending rigidity of the polymer, and therefore requires the stiffness parameter g as additional input: Starting from the Hamiltonian H_{WLC} of the wormlike chain model [19–21], where chain conformations are taken to be continuous space curves of fixed length L in natural parametrization, discretization of the space curves as sequences of N jointed segments of fixed length b, and subsequent discretization of the integral defining H_{WLC}, gives the Hamiltonian H_{KP} of the Kratky-Porod model:

$$\begin{aligned}\beta H_{WLC} &= \beta H_{WLC}[\vec{r}(t)] = \frac{\beta\kappa}{2}\int_0^L ds\left(\frac{\partial^2\vec{r}(s)}{\partial s^2}\right)^2 \\ &\approx \frac{\beta\kappa b}{2}\sum_{k=1}^{N-1}\left(\frac{\overrightarrow{t_{k+1}}-\overrightarrow{t_k}}{b}\right)^2\end{aligned} \tag{1}$$

with $\beta = 1/k_BT$, $\kappa = \epsilon b$, the bending modulus of the chain [31], and normalized vectors $\overrightarrow{t_k}$, $k = 1, ..,N$, tangent to the segments of the discretized chain. Further expanding, we have.

$$\begin{aligned}\beta H_{WLC} &\approx \frac{\beta\kappa b}{2}\sum_{k=1}^{N-1} b^{-2}(\overrightarrow{t_{k+1}}^2 + \overrightarrow{t_k}^2 - 2\,\overrightarrow{t_{k+1}}\cdot\overrightarrow{t_k}) \\ &= \frac{\beta\kappa b^{-1}}{2}\sum_{k=1}^{N-1} 2(1-\cos(\theta_k)) =: -g\sum_{k=1}^{N-1}\cos(\theta_k) + \beta E_0 \\ &= \beta(H_{KP} + E_0)\end{aligned} \tag{2}$$

where $g = \beta\kappa b^{-1}$ is the dimensionless stiffness parameter, $\cos(\theta_k) = \overrightarrow{t_{k+1}}\cdot\overrightarrow{t_k}$, and E_0 is a constant independent of chain conformation. The persistence length $l_p = l_p(g)$ of the Kratky-Porod chain can be obtained analytically [32], giving.

$$l_p(g) = -b/ln(coth(g) - 1/g) \tag{3}$$

For large g the persistence length is approximately given by $l_p(g) \approx bg = \beta\kappa = \beta\epsilon b$.

With the approximation $2(1 - \cos(\theta)) \approx \theta^2$ for angles $\theta = \theta_k$ not too far from zero, an alternative implementation of bending rigidity is based on the quadratic bending potential U_{QB}, giving.

$$\beta H_{WLC} \approx \beta U_{QB} := (g/2)\sum_{k=1}^{N-1}{\theta_k}^2 \tag{4}$$

To introduce excluded volume interactions, in a simple extension of the Kratky-Porod model, the chain beads are modelled as impenetrable spheres of diameter $d = b$, resulting in a bead-stick model, which has been employed in [17] to model dsDNA, choosing model parameters by application of a specialized version of the fitting procedure discussed in (2.3).

As a basis for molecular dynamics (MD) simulations, the bead-spring model introduced in [33] is more suitable than a bead-stick model: In this model, the angle potential of the Kratky-Porod chain is combined with a Weeks-Chandler-Anderson (WCA) potential to model excluded volume interactions, and the distance of adjacent beads is not fixed, but is kept finite by introducing a finitely extensible nonlinear elastic (FENE) potential.

In this work, we also employ this bead-spring model, with all choices of constants as in [33], but chain stiffness is modelled in terms of the quadratic bending potential (4) instead. The total Hamiltonian H is therefore given by $H = U_{WCA} + U_{FENE} + U_{QB}$, with.

$$U_{WCA}(r_{ij}) = 4\varepsilon((\sigma/r_{ij})^{12} - (\sigma/r_{ij})^6) + \varepsilon \tag{5}$$

for two beads at a distance $r_{ij} \le 2^{1/6}\sigma$, and $U_{WCA}(r_{ij}) = 0$ in case that $r_{ij} > 2^{1/6}\sigma$. For adjacent beads at distance r_{ij}, the potential U_{FENE} adds a non-vanishing contribution of.