Eur. Phys. J. E **15**, 83–93 (2004)
DOI 10.1140/epje/i2004-10039-7

THE EUROPEAN PHYSICAL JOURNAL E

Stepwise unfolding of collapsed polymers

D. Marenduzzo[1,a], A. Maritan[2,3], A. Rosa[4], and F. Seno[2]

[1] Department of Physics, Theoretical Physics, University of Oxford, 1 Keble Road, Oxford OX1 3NP, England, UK
[2] INFM and Dipartimento di Fisica "G. Galilei" dell' Università di Padova, via Marzolo 8, 35131 Padova, Italy
[3] The Abdus Salam Centre for Theoretical Physics, Strada Costiera 11, 34014 Trieste, Italy
[4] Institut de Mathématiques B, Faculté de Sciences de Base, École Polytechnique Fédérale de Lausanne, 1015 Lausanne, Switzerland

Received 7 January 2004 and Received in final form 5 May 2004 /
Published online: 1 October 2004 –

Abstract. Motivated by recent experimental data on DNA stretching in presence of polyvalent counterions, we study the force-induced unfolding of a homopolymer on and off lattice. In the fixed force ensemble the globule unravels via a series of steps due to surface effects which play an important role for finite-size chains. This holds both for flexible and stiff polymers. We discuss in a qualitative way how this result may impact on the interpretation of DNA stretching experiments showing peaks in the characteristic curves, by extracting from the raw data the corresponding elongation-*versus*-force characteristic curves. Furthermore, approximate analytical and numerical calculations, valid in a quasi-equilibrium fixed stretch ensemble, and if the initial low-temperature state is ordered in a spool, show that the average force *versus* elongation displays peaks related to the geometry of the initial configuration. We finally argue how the proposed mechanisms identified for the arising of peaks may couple in the experiments, and comment on the role of dynamic effects.

PACS. 82.35.Lr Physical properties of polymers – 87.15.-v Biomolecules: structure and physical properties – 36.20.Ey Conformation (statistics and dynamics)

1 Introduction

The behaviour of a single polymeric molecule under stretching has recently become a popular subject among experimental physicists thanks to the remarkable development of single-molecule techniques (see, *e.g.*, [1,2] for an updated review). Consequently, it has become possible to test thoroughly theoretical predictions made in the past few years. In particular, here we concentrate on a simple theoretical framework within which to rationalize the latest experimental observations on the stretching response of a polymer in a poor solvent [3–6] (see Fig. 1). We will analyze the data contained in references [3,4], which refer to a DNA molecule in presence of a significant concentration of condensing agents —polyvalent counterions. Our analysis also applies to the case of the stretching of a polymer below the θ point, *viz* in a poor solvent —where the preferred state, in the absence of a perturbing force, is an isotropic compact globule.

Until very recently, many of the theoretical predictions were refined versions of the original mean-field approach first proposed in reference [7] (see Refs. [8–10] for more recent related work on this subject). The mean-field approximation predicted that the force-*versus*-elongation curves should consist of a plateau followed by the steep increase which is also present in the case of a swollen polymer (*i.e.* DNA with no counterions) (see also Ref. [11]). The

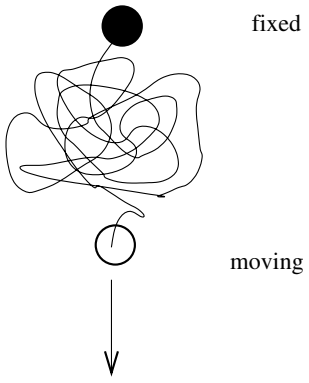


Fig. 1. Schematics of the experiment we want to compare our results with. In a typical experiment the two ends of a globule are controlled. One end is kept fixed and the other one is moved very slowly (to have a quasi-equilibrium situation). The average force is then measured, so that the experiments are done in the fixed stretch ensemble.

[a] e-mail: davide@thphys.ox.ac.uk

plateau is due to configurations where stretched and globular states coexist within the same chain, as is typical in a first-order transition. The instability of other configurations with respect to these coexisting conformations can be seen as a manifestation of the Rayleigh instability [6, 12]. The mean-field prediction of a force plateau successfully accounts for the curve recorded by Wang's group [3]. This is different from the results obtained by Bustamante's group and reproduced in the same reference [3]. We suggest that this behaviour may be due to the different contour lengths of the molecules used in the experiments. The λ phage molecule used in Wang's setup was relatively long, its contour length being ~ 16.5 μm, corresponding to roughly 48.5 kilobases or 330 persistence lengths (the base pair separation in double stranded DNA is ~ 0.34 nm and the DNA persistence length is usually agreed to be ~ 50 nm, though it has been reported to depend on the salt concentration in solution too). The length of the DNA plasmid treated by Bustamante's group is around 24 persistence lengths, so much shorter. Finally, the length of the DNA used in the experiment of reference [4] is intermediate, roughly 90 persistence length (~ 4.5 μm). It has to be noticed that the amount of spermidine —the condensing counterion— was in these cases comparable, between 100 and 700 μM. One of our main results is that a treatment of a commonly accepted model, namely that of a self-avoiding polymer in a poor solvent (below the theta point), not relying on the mean-field assumption of very long molecules, can account qualitatively for the different observations, as due to the different contour lengths of the samples used.

More recent experiments in reference [5] employed a higher concentration of spermidine. In this case peaks were more enhanced and, as noted by the authors themselves, their relative positioning was close to the length of one turn in a DNA spool of radius 50 nm. In this work we also show that if the initial configuration is ordered in a toroid as resulting from a high concentration of condensing agents, again peaks are expected in the experimental fixed stretch ensemble due to non-convexities of the free energy as a function of the elongation, with a peak-peak distance related to the geometrical dimensions of the initial configuration. These peaks would *not* in general result in step-wise unfolding in the conjugate ensemble and thus would be missed with the fixed force calculation.

In all experimental cases, there is a marked hysteresis between the stretching and the releasing of the DNA chain. This clearly means that dynamical effects are necessary for a quantitative understanding of the force-*versus*-elongation curves. However, the goal of this work is to show that peaks can be present even in the *absence* of dynamical effects. We do comment on the possible role of dynamics in our picture at the end of this work.

The outline of this work is the following. In Section 2 we will introduce the models to be studied in the fixed force equilibrium ensemble and discuss the appearance of steps. In Section 3 we analyze the data from the experiments, calculating the fixed force response from the fixed stretch raw data via a Legendre transform. In Section 4 we perform some analytical approximate calculations in the fixed stretch ensemble by assuming that the condensing agents are so strong that the initial state is highly ordered —*e.g.*, in a spool. We show that also this mechanism produces peaks. In Section 5 we then attempt to critically discuss the impact of our equilibrium results to the experiments, we argue on the role of dynamics and compare with related recent literature. Section 6 contains our conclusions.

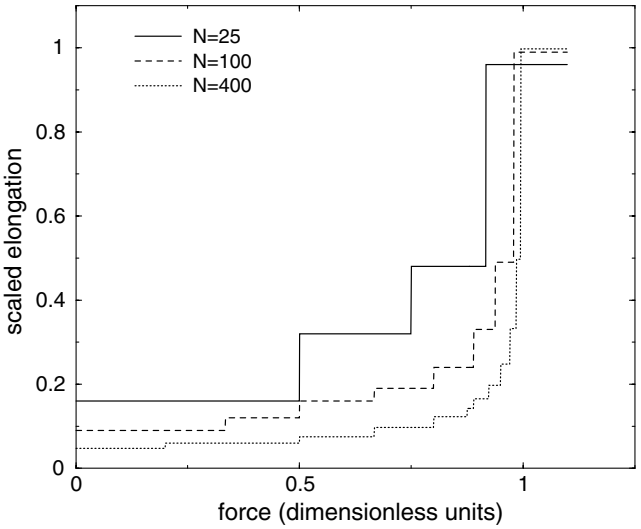


Fig. 2. Scaled elongation (defined as the average end-to-end distance, x, scaled by the length of the chain N), as a function of pulling force for a flexible polymer of different contour lengths in the ground-state approximation ($\epsilon_0 = 1$). As the contour length increases, the transition looks more and more like an all-or-none transition as will be the case for infinite length. The data have been obtained in the ground-state approximation, *i.e.* by considering only the configurations which minimizes the energy function defined in equation (1). The minimum has been searched among all rectangles of edges of sizes m and $\mathrm{int}(N/m)$, where $\mathrm{int}(a)$ denotes the integer closest to a and smaller than or equal to a.

2 Unfolding of globules in the fixed force ensemble: models and behaviour at low T

In this section, we will consider the unfolding of a collapsed polymer in the case in which the stretching force acting on the last bead is fixed (fixed force ensemble). This will be taken along the x-direction throughout this section. We will study four models, in order of increasing realism and hence complexity: first, a) a self-avoiding polymer and b) a necklace semiflexible polymer, in $d = 2$ on the lattice, and then c) a flexible and d) a semiflexible three-dimensional off-lattice self-avoiding polymer. In none of the cases will we attempt quantitative comparison with experimental data, since a lot of realistic details are omitted. Yet, we will see that the analysis of these models reveals a great amount of robust common features.

Let us then first consider a self-avoiding polymer under the action of a stretching force of modulus f, on the

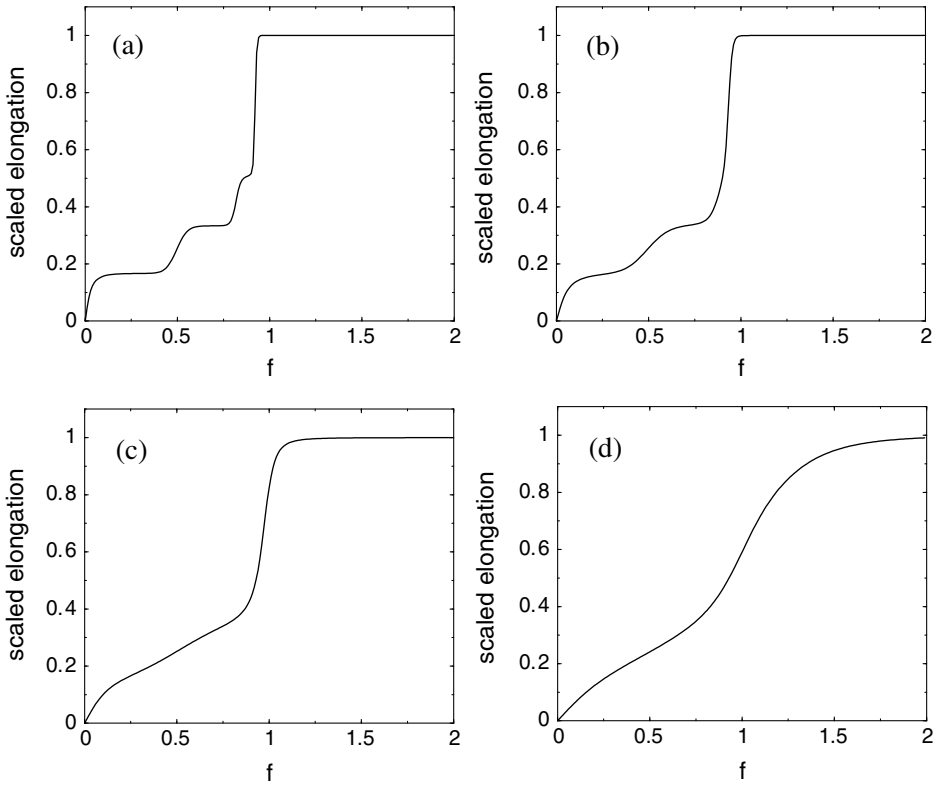


Fig. 3. Scaled elongation (defined as the average end-to-end distance, x, scaled by the length of the chain N), as a function of pulling force for a $N = 25$ chain, for (a) $T = 0.05$, (b) $T = 0.1$, (c) $T = 0.2$, and (d) $T = 0.4$. The multi-step character is washed out progressively by increasing T (for reference, the θ point is at $T = 1.5$).

two-dimensional lattice. The lattice spacing corresponds to the persistence length and the walk can change its direction after it has traveled one persistence length [13]. The polymer in this model is fully characterized by the attraction energy between non-consecutive beads, ϵ_0, by the variable force, f, and by the number of persistence lengths, N. Calculations will be performed in the ensemble where the force is fixed, which is related to the experimental ensemble (fixed stretch) by means of a usual Legendre transform [14]. When we deal with short polymers, it is actually possible to enumerate walks on the lattice in the fixed stretch ensemble. We will restrict to the case of low temperature, where the entropic fluctuations are negligible and ground state dominates. In Section 3 we will discuss how well this approximation is met in the experiments motivating our study. In this situation, the partition function is dominated by the ground state. In our simplified model the energy of a self-avoiding polymer chain is

$$H = -N_{\rm c}\epsilon_0 - fx, \qquad (1)$$

where $N_{\rm c}$ is the number of self-contacts of the polymer, f is the modulus of the pulling force and x is the elongation along the force direction. Since we are interested in the ground state only, we can restrict our analysis to rectangular walks elongated along the force axis which do not leave empty spaces in their interior (known also as rectangular Hamiltonian walks [13]). Let us consider the case of a chain consisting of 25 persistence lengths (see Fig. 2): as f increases we go from a configuration with extension 4 along the x-axis to one with elongation 8 and finally to the stretched coil. Thus here the unfolding occurs in steps, and each step is stable in a window of force values. In Figure 2 we plot the average (scaled) elongation *versus* force for different lengths of the chain (different values of N): as the chain gets longer each of the steps is stable only for a very small interval of forces, apart from the globule and the coil states, and indeed for infinitely long chains there would be an abrupt transition between these two states only. Finally, we calculated the same curves at non-zero T, obtained via an exact enumerations of the chains at $N = 25$ (Fig. 3): it is found that the multi-step character is progressively washed out by entropy, as expected.

What one learns from this simple exact calculation is that the steps are less and less pronounced as the chain gets longer, and that for short chains there is a number of intermediate states. These finite-size effects might be an important ingredient responsible of the observed peak pattern.

As we will see, while the flexible polymer model, usually employed in the interpretation of these experiments [3,6], has the virtue of qualitatively explaining the presence of peaks in terms of finite-size effects, in order to explain more quantitatively the experiments one needs a more complete model. Indeed, our simplified model would predict a series of steps progressively narrower as the unfolding proceeds, whereas the experimental steps do not

display this feature (see next section and the analysis of the experimental data).

In order to render the modelling more realistic, we can choose another level of coarse graining for a DNA molecule. Above, we considered DNA as a flexible polymer chain whose Kuhn length is one persistence length, $\sim$ 50 nm. A more detailed model for DNA is that of a chain with a non-zero stiffness, which applies to a different coarse graining level in which the beads are spaced less than one persistence length. A natural coarse graining is one in which neighbouring beads are 2–2.5 nm apart (DNA thickness) so that the persistence length is 20–25 beads. This introduces one more parameter, the stiffness or bending rigidity, $\epsilon_{\rm b}$, which is an energetic penalty paid for each turn of the chain on the lattice. This parameter has to be tuned in order to give the desired persistence length. At the same time, the ratio between the pair interaction energy and the stiffness will depend on the salt concentration of the solution DNA is in: roughly, the more collapsed DNA is, the higher the ratio $\epsilon_0/\epsilon_{\rm b}$ is. Furthermore, DNA is a polyelectrolyte and the inter-charge potential is believed to render the unperturbed configuration of a collapsed nucleotide more similar to a necklace of several blobs than to a single globule.

As a consequence, we have considered another lattice model, in which the $f \to 0$ structure is assumed to consist of a series of globules connected in series by straight portions of chain, oriented along the direction of the pulling force, $\hat{x}$. Different globules in the chain have negligible interaction with each other. As the force is switched on, each of the globule will behave like the single collapsed polymer treated above, the only difference being that now there is a penalty for each corner in the lattice chain. However, since the different globules have in general different sizes, different globules in the necklace will undergo different hierarchies of finite-size conformational transitions. The overall necklace extension-*versus*-force characteristic curve will be given by the superimposal of all the curves valid for the separate globules.

We have studied this more realistic model again with the same methods as discussed above (this is possible since we can treat each globule separately as they do not mutually interact). In Figure 4, we show an example of the *x-versus-f* curve obtained in the ground-state approximation upon stretching a necklace of 19 blobs of total contour length comparable to the one considered in reference [3], when the ratio $\epsilon_{\rm b}/\epsilon_0 = 2.8$ (the length of a blob varies from 50 to 300 nm). This is a typical extension-*versus*-force curve, and is of interest due to its generic qualitative features rather than in view of any quantitative comparison to experimental results. It can be seen that this time the width of the steps, which depends on the ratio $\epsilon_0/\epsilon_{\rm b}$, is not decreasing as the unfolding proceeds.

From what reported so far, it might not be clear if the multi-step effect depends on the artificial discreteness of the two-dimensional lattice. However, we observe that the multi-steps found with the preceding models are ultimately due to surface effects. Indeed, for a finite polymer, the gain in contact energy per monomer due to collapse into a compact conformation is smaller than the corresponding energetic gain in an infinite polymer. On the other hand, the extensional energy gain does not suffer from this. As a consequence, for intermediate forces and finite chains, it may become convenient to have an elongated rectangular shape rather than an isotropic globule. This already suggests that the multi-step behaviour is not an artifact of the lattice, but a common feature of all unfolding models, irrespective of dimension and discreteness of underlying space, and due to genuine surface finite-size effects. Indeed, in two dimensions the surface corrections and the intermediate steps have a role also in the thermodynamic limit, and they smoothen the transition, rendering it a second-order one [15].

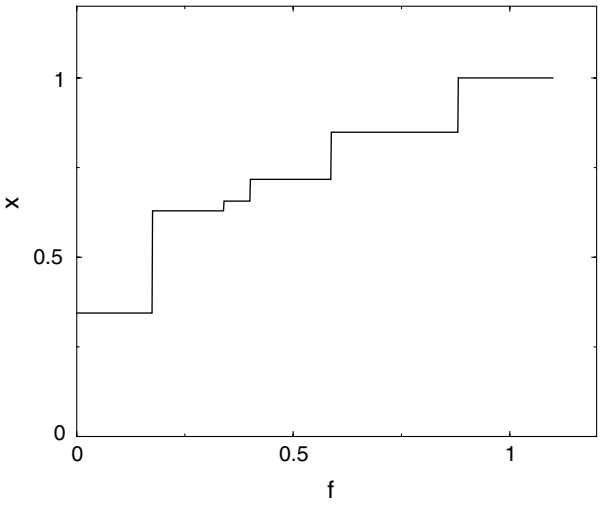


Fig. 4. This is an example, in $d = 2$, of the curve obtained numerically by considering a semiflexible polymer (with non-zero stiffness) and a ground-state configuration made up of several blobs of various lengths. It can be seen that this treatment gives a better qualitative comparison with experiments (see Fig. 7b). The qualitative reason underlying the multi-step behaviour is the same as the one discussed for the simplified model with no bending rigidity and no blobs (Fig. 1).

Quantitatively, we have also studied a continuum model in three dimensions, of a chain of equally spaced beads, with mutual distance taken equal to 1, in presence of a compacting potential and a stretching force of modulus f, acting along the x-direction. The Hamiltonian we have considered for this calculation is the following:

$$\mathcal{H} = \sum_{i<j;i,j=1}^{N} v(|\vec{r}_i - \vec{r}_j|) - f\hat{x} \cdot \ (\vec{r}_N - \vec{r}_0)\,, \qquad (2)$$

where $\{\vec{r}\}_{i=1,\ldots,N}$ refers to the position of the N beads in the three-dimensional space, $\hat{x}$ is the unit vector parallel to the x-axis, and the detailed shape of the two-body potential $v(|\vec{r}_i - \vec{r}_j|)$ is that of a square well. This square well has a hard core at $|\vec{r}_i - \vec{r}_j| = 1.2$, below which the potential is infinite, and extends up to $|\vec{r}_i - \vec{r}_j| = 1.6$, beyond which point the potential is zero. The value of the potential in the well was chosen to be -1. For a flexible polymer, we considered chain length up to 40 beads.

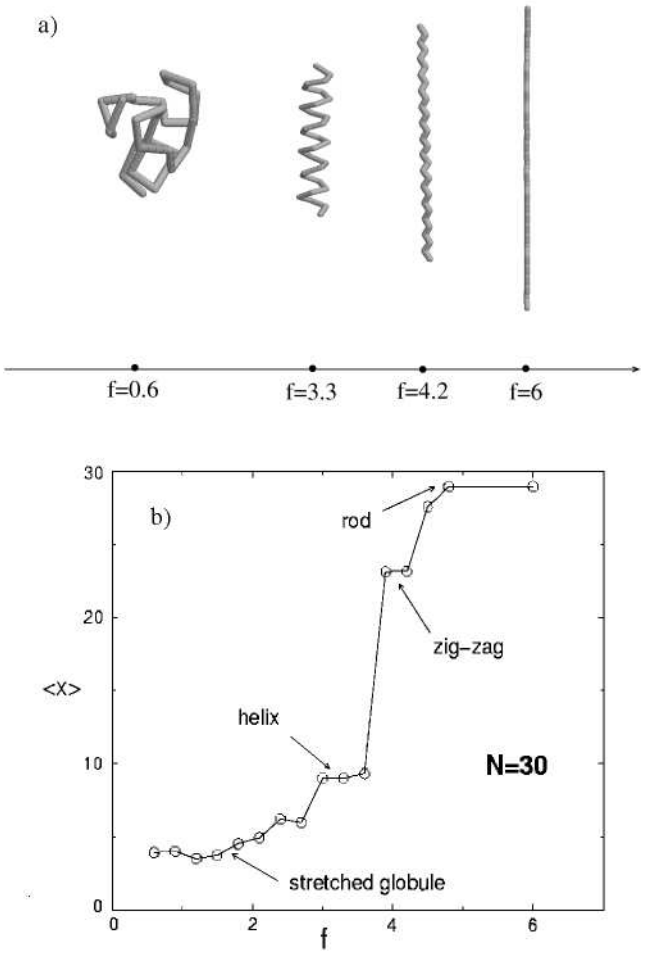


Fig. 5. a) Sequence of ground states for a three-dimensional flexible homopolymer with attractive self-interactions and variable force. b) Elongation-*versus*-force curve at low temperature.

The additional degrees of freedom now present make so that the calculation can now only be performed numerically. We have employed Monte Carlo simulations, parallel tempering [16] and standard simulated annealing techniques, to find the low-temperature configuration of the Hamiltonian $\mathcal{H}$, at a varying value of the stretching force modulus.

The resulting configurations and extension-force curves at low temperature are shown in Figure 5 for the flexible polymer unfolding. For small forces, the ground states is an elongated globule as expected from mean-field theories. For intermediate force, the selected structure is a regular helix, which can be thought as a cylinder, which always loses to the globule or the stretched state in the thermodynamic limit but is stabilized at finite length. It is notable to remark that the helix comes as a ground state with no other constraints or interactions apart from those in equation (2). For yet stronger forces one finds that the ground state is a zig-zag, followed only later by a stretched polymer. The presence of the intermediate states causes the elongation-*versus*-force curve to be multi-step as in the lattice models. We have also made calculations with a

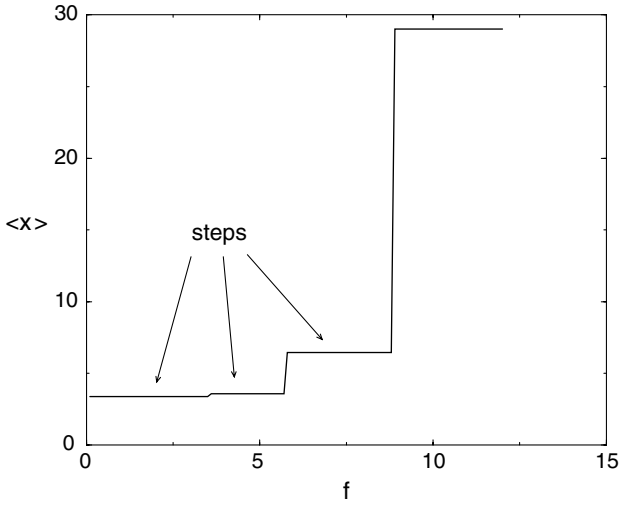


Fig. 6. Average elongation as a function of pulling force for a $N = 30$ chain, for low T. The bending rigidity was $\kappa = 2$, while the two-body attracting potential extended from $R = 1.1$ to $R = 1.6$.

semiflexible polymer and found that also in this case the unfolding occurred via steps for finite chains (see Fig. 6 for an example). The Hamiltonian describing a semiflexible polymer in a poor solvent with a stretching force is

$$\mathcal{H} = \sum_{i<j;i,j=1}^{N} v(|\vec{r}_i - \vec{r}_j|) - f\hat{x}\cdot(\vec{r}_N - \vec{r}_0) - \kappa \sum_{i=1}^{N-2} \vec{t}_i \cdot \vec{t}_{i+1}, \quad (3)$$

where κ is the bending rigidity and $\vec{t}_i \equiv \vec{r}_{i+1} - \vec{r}_i$ (which is a vector of unit length). In this case the selected configurations will depend on the values of the ratio of the bending to the attractive energy. The starting steps —for DNA-like parameters— at least for not too long polymers is expected to be a toroid, while for intermediate forces we expect that a fibril-like structure (see, *e.g.*, Ref. [17]) will be more stable.

In all the four models considered in this section, torsional rigidity has never been considered. It would be interesting to see quantitatively how this new term would affect the hierarchy of ground states encountered at increasing forces (the very presence of these steps, based on the previous analysis, is, on the other hand, most probably robust with respect to this new ingredient). Torsional rigidity would be introduced by adding to the previous three-dimensional Hamiltonians the extra term [18]:

$$\delta\mathcal{H} = -\kappa_{\rm t} \sum_{i=1}^{N-3} \vec{b}_i \cdot \vec{b}_{i+1}, \quad (4)$$

where $\vec{b}_i$ denotes the binormal vector associated with the i-th bead of the chain (see, *e.g.*, [19] for its definition) and $\kappa_{\rm t}$ is the torsional rigidity. In the case of a flexible polymer, this extra term is expected to destabilize consistently the zig-zag structure. On the other hand, the helix, as well as the toroid and fibril state which are expected for a semiflexible polymer should be much more robust.

In the next section we will argue that these steps in the conjugate fixed stretch ensemble, in thermal equilibrium, would give rise to peaks.

3 Analysis of the experimental data

While the calculations of Section 2 were performed in the fixed force ensemble, in the experiments the end-to-end distance is the controlled quantity. It is useful here to summarize a few facts about the relation between the two ensembles. In the fixed stretch ensemble, the partition function is given by

$$Z(x) = \sum_{\{\mathrm{conf}\}} \delta(x(\mathrm{conf}) - x) \exp(-\beta \mathcal{H}_{\mathrm{conf}}), \qquad (5)$$

where {conf} indicates the ensemble of chain configurations, while x(conf) and $\mathcal{H}_{\mathrm{conf}}$ denote, respectively, the end-to-end distance and energy of a given configuration (Eqs. (2) or (3) with $f = 0$). The partition function of the fixed force ensemble, $\hat{Z}(f)$, is related to $Z(x)$ through the relation

$$\hat{Z}(f) = \int \mathrm{d}x Z(x) \exp(\beta f x). \qquad (6)$$

From the partition functions one can define the corresponding free energies via

$$\mathcal{F}(x) = -k_{\mathrm{B}} T \log Z(x), \qquad (7)$$

$$\hat{\mathcal{F}}(f) = -k_{\mathrm{B}} T \log \hat{Z}(f). \qquad (8)$$

In turn, the average elongation in the fixed force ensemble is $\langle x \rangle = -\frac{\partial \hat{\mathcal{F}}(f)}{\partial f}$; while the average force in the fixed stretch ensemble $f(x) = \frac{\partial \mathcal{F}(x)}{\partial x}$. In general, the two ensemble are known to be inequivalent even when there are no self-interactions [14].

For the case we are interested in, we note that, in particular, the average elongation in the fixed force ensemble is always an increasing function of the force, since

$$\frac{\partial \langle x \rangle}{\partial f} = \beta \left[\langle x^2 \rangle - \langle x \rangle^2 \right] > 0, \qquad (9)$$

while

$$\frac{\partial^2 \mathcal{F}(x)}{\partial x^2} = \frac{\partial f(x)}{\partial x} = \beta \left[\left\langle \frac{\partial^2 \mathcal{H}_{\mathrm{conf}}}{\partial x(\mathrm{conf})^2} \right\rangle - \left\langle \frac{\partial \mathcal{H}_{\mathrm{conf}}}{\partial x(\mathrm{conf})} \right\rangle^2 \right], \qquad (10)$$

where $\langle \cdot \rangle$ is the average with the statistical weight $\delta(x - x(\mathrm{conf})) \exp(-\beta \mathcal{H}_{\mathrm{conf}})$. Thus, $f(x)$ might be non-monotonic corresponding to a free energy which is not convex as a function of x (see also Ref. [20], in which a related situation is found in the DNA unzipping problem).

It is also instructive to consider the two ensembles in the limit $T \to 0$. From equation (6) it is clear that only the ground state determines the average elongation in the fixed force ensemble, while the average force is given by the derivative of the internal energy (since for $T \to 0$ there is no entropy) with respect to x.

Let us first suppose that $f(x)$ is not monotonic, with both minima and maxima (the experimental situation). The average elongation in the fixed force ensemble is given by the global minimum of $\mathcal{F}(x) - fx$. This can be deduced by equation (6) in the $T \to 0$ limit using the saddle point approximation. The condition that $f(x)$ has a minimum implies that $\mathcal{F}(x) - fx$ has local minima. If the *global* minimum changes as a function of f, other steps in $\langle x \rangle$ *versus* f appear.

On the other hand, if there are steps, then the global minimum changes discontinuously as a function of f (and is neither 0 nor L). If these steps are "first-order"–like phase transition, then there are local minima in the free-energy landscape, hence peaks in the fixed stretch ensemble.

So, 1) peaks in the fixed stretch ensemble *do not* in general imply steps in the fixed force one, while 2) steps imply peaks if they are connected to "first-order"–like transitions. It is possible to refine this reasoning and prove a weaker statement than 1), namely that an *ascending* peak pattern like those recorded in reference [3] will translate into a stepwise pattern in the conjugate ensemble.

We then come to a semiquantitative analysis of the experimental data. This is shown in Figure 7. From the available force-*versus*-elongation characteristic curves, which are not easily obtainable in theoretical calculations, we can extract via a numerical integration the free energy as a function of the end-to-end elongation. In formulas, if $\mathcal{F}(x)$ is the above-defined free energy of a polymer with end-to-end elongation (along the stretching direction) equal to x, then it is related to $f(x)$ via (we choose $\mathcal{F}(0) = 0$):

$$\mathcal{F}(x) = \int_0^x f(x') \mathrm{d}x', \qquad (11)$$

so that from experimental data in references [3,4] we can find the experimental $\mathcal{F}(x)$. From this, we can find the average elongation-*versus*-force characteristic curves as explained above. To have a clearer comparison with the theoretical qualitative results, we do that in the ground-state approximation, which means assuming that the saddle point approximation in equation (6) is reliable. This is equivalent to finding how the minimum of the function $\mathcal{F}(x) - fx$ moves as the parameter f is varied. We used data coming from Figure 3B in reference [3] and from Figure 4, curve B from reference [4]. In the curve extracted from the measurements by Bustamante's group multi-step behaviour is very pronounced, while it gets less clear in the curve obtained from the data reported in reference [4], and finally disappears for data taken from the long-molecule studies performed by Wang's group (Fig. 1 in Ref. [3]). In the case of the data in reference [3], Figure 3B, we explicitly verified that even without making the ground-state approximation, the multi-step character was evident.

It is useful here to discuss the validity of the approximation of neglecting entropic effects in the calculation. The relative importance of energetic and entropic effects can be directly assessed from the experimental data in the following way. The entropic work performed in stretching a double-stranded DNA up to a relative extension

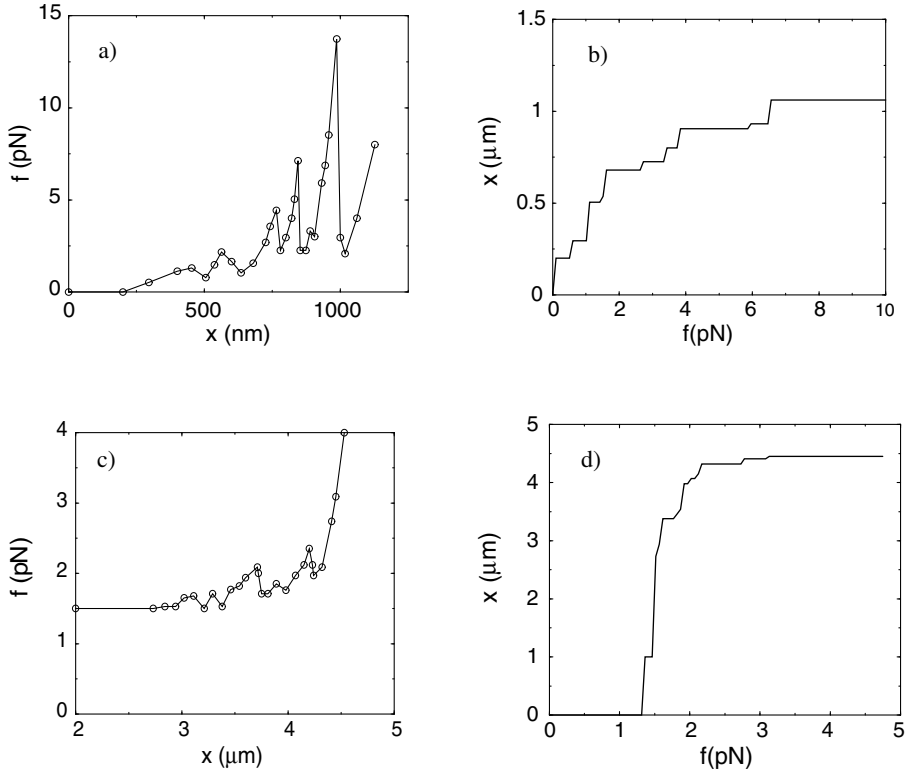


Fig. 7. Analysis of the data coming from references [3] and [4]. The raw data are reported in (a) (Fig. 3B of Ref. [3]) and (c) (Fig. 4B of Ref. [4]). The average end-to-end distance *versus* pulling force in the ground-state approximation is obtained for the data taken from reference [3] (b) and [4], (d). The multi-step character of the transition decreases with increasing contour length.

$x^*/N = \zeta^*$, with $0 < \zeta^* < 1$, can be estimated by considering the well-known Marko-Siggia formula for the stretching curve of a semiflexible polymer, namely

$$f(x) = \frac{k_{\rm B}T}{P}\left(\frac{1}{4(1-x/L)^2} - \frac{1}{4} + \frac{x}{L}\right), \qquad (12)$$

where P is the persistence length and L is the contour length of the polymer, and by integrating it in $x/L = \zeta$ from 0 to ζ^*. This yields a work of $\sim 0.01\ k_{\rm B}T$ per base pair to get up to $\zeta^* = 0.85$, by assuming a persistence length $P = 50$ nm. On the other hand, the extra work needed to stretch a collapsed polymer, which is imputed to energy, has been calculated in reference [5] and is equal to roughly $0.06\ k_{\rm B}T$ per base pair for a concentration of spermidine of 500 μM. So when we consider stretching experiments in which the DNA does not stretch more than ζ^* of its length, entropic effects can be neglected with a fair approximation.

4 Quasi-equilibrium fixed stretch ensemble: unfolding of a spool

In Sections 2 and 3 we proposed a mechanism for the appearance of an ascending series of peaks in the force-*versus*-elongation characteristic curves, due to conformational changes either in the whole polymer or in one of the "blobs" which constitute it if the polymer is modelled by a necklace due to electrostatic effective interactions. As observed in Section 1, however, there are also other patterns of peaks, not necessarily ascending. These were recorded in reference [5], and in that paper the authors observed that the patterns of peak-peak distances was reminiscent of the length of a turn in a DNA spool. Such peaks cannot be possibly captured within the fixed force calculation in Section 2. Indeed, the inverse Legendre transform of the experimental data would wash out these peaks so that in the fixed force ensemble one would not observe any steps. Here, we propose another mechanism which can cause peaks in the fixed stretch ensemble under the assumption that the molecule is initially wound into a toroid. Again, this is an equilibrium calculation in which we neglect dynamical effects: a qualitative discussion of these will be attempted in Section 5.

Let us assume that a DNA molecule, in the absence of external perturbations, attains, in a solution of condensing counterions, the shape of a toroid. This is due to the interplay between semiflexibility and attractive interactions. Let the radius of the spool be R, and let us neglect its vertical height (see Fig. 8). The positions of the two end beads are controlled externally. We will work in the ensemble where the relative distance along the x-axis is controlled: this means that we are going to find the x component of the average force. For symmetry, we require that the configuration of the overall system be symmetric

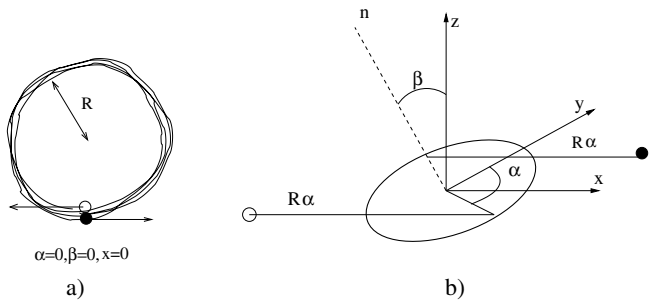


Fig. 8. Initial configuration (a) and definitions of the angles (b) for the calculation in Section 4. Initially, the chain is wound into a toroid of radius $\sim$ 50 nm, one persistence length of double-stranded DNA.

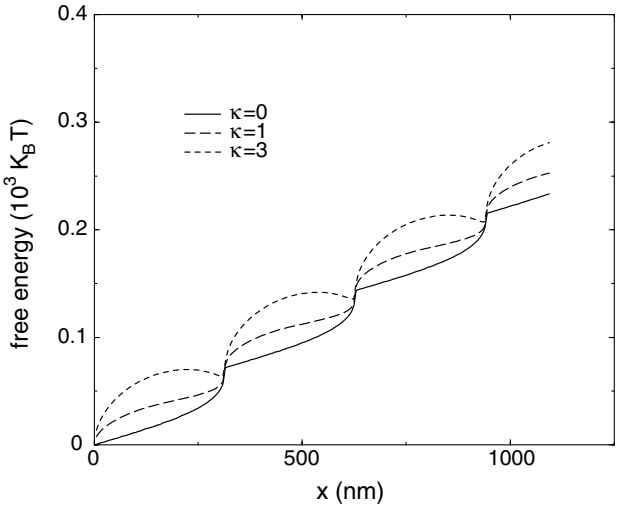


Fig. 9. Free energy of the toroid as a function of the projection of the end-to-end distance along x for different values of κ (the values of κ are measured in units of ϵR). The desorption energy has been taken equal to 0.075 $k_{\rm B}T$ per base pair (intermediate between the estimates given in Ref. [5]). The temperature is assumed to be 300 K, while the base pair spacing is assumed to be 0.34 nm as in B-DNA.

with respect to the toroid center (so the relative positions of the arms with respect to the toroid in Fig. 8 must be equivalent). We also require for simplicity that the spool axis is constrained to lie in the xz plane. Qualitatively, the results should apply to the experimental situation as well. If no conformational change occurs, due for example to high concentration of condensing agents, then as the DNA is stretched one can identify a region of DNA which remains in the spool and two equal sections which are desorbed from it, one on each side of the toroid. For simplicity, let us assume that no length is "wasted" and that the desorbed lines are straight and parallel to the x-axis. We call x the projection of the end-to-end distance along the x-axis and α the desorption angle, defined so that $2R\alpha$ is the total length of the chain detached from the spool (in the initial configuration at $x = 0$, which we consider, the two ends are coincident).

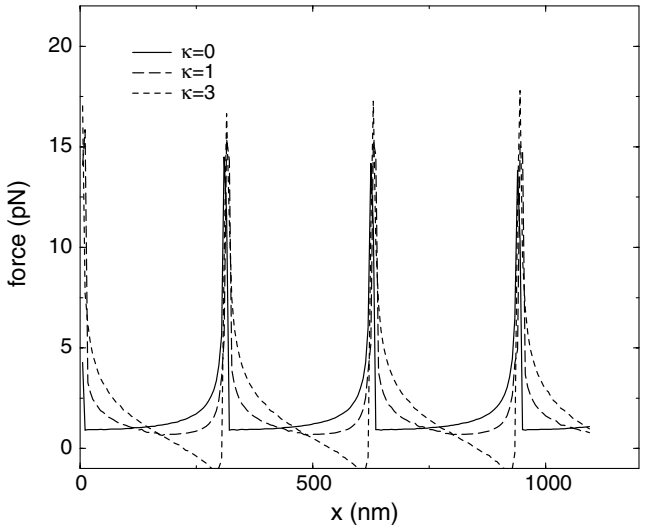


Fig. 10. Equilibrium force as a function of the projection of the end-to-end distance along x for different values of κ (measured in units of ϵR). There is a peak every turn. The desorption energy has been taken equal to 0.075 $k_{\rm B}T$ per base pair.

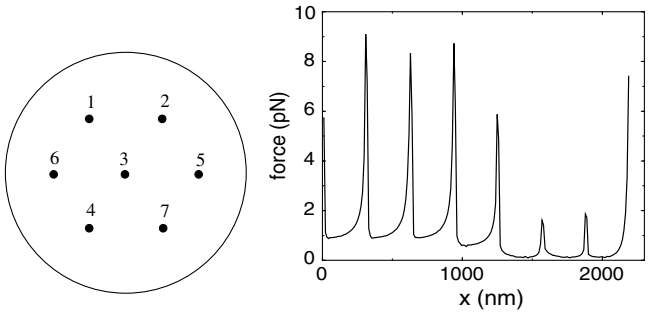
The free energy of the spool and the detached arms, if we neglect entropy, for a given spool-arm configuration, will be given by the loss of energy due to desorption and by the bending cost paid to curve the desorbed DNA in the point in which it joins the spool. We take this bending cost to be proportional to the dot product of the detached arm direction with the local tangent to the spool at the desorption point. In formulas the energy of the spool reads

$$\mathcal{F}(\alpha,\beta) = -2R\epsilon\alpha + \kappa\left(1 - \cos(\beta)\cos(\alpha)\right), \qquad (13)$$

where α and β are the desorption angle and the toroid tilt angle (with respect to the z-axis) characterizing completely the spool state, and $\epsilon > 0$ is the adsorption energy per unit length lost when the spool is opened (see Fig. 8 for an explanation of the definitions). The second term in equation (13) comes from assuming that the macroscopic change in orientation when the stretched arms join the spool must have an energetic cost in the case of a semiflexible polymer. Hence κ has the dimensions of an energy and will be given in units relative to ϵR in what follows. We thus do not address the question of what is the fine structure of the polymer at these junction points, and in particular of the local radius of curvature there, but rather introduce a phenomenological cost for the necessary bending associated with the change in direction. We now constrain all the spool + arms configurations to those which satisfy the constraint of having a projection of the end-to-end distance on the x-axis equal to x. By geometrical considerations, one can find that α and β are linked to the required x via the following simple relation (see Fig. 8):

$$x = 2R\alpha - 2R\sin(\alpha)\cos(\beta). \qquad (14)$$

Note that there are in general more than one possible pairs (α, β) which satisfy equation (14) for a given x. We



a) b)

Fig. 11. a) Example of a finite toroid: the strands are numbered according to the convention that 1 is attached to 2 via the chain, 2 to 3 etc. b) Force-*versus*-x curve: it can be seen the peaks are no longer equal in height. The last peak is due to the entropy of the chain. The desorption energy has been taken equal to 0.075 $k_{\rm B}T$ per base pair, while $\kappa = 0.15\ \epsilon R$.

now want to find the free energy as a function of x for low temperature. To do this, we specify the values of β and α among those which satisfy the constraint given by equation (14), for a given x, as the pair of angles which *minimizes* the free-energy equation (13). We solved this problem numerically by enforcing the constraint in a soft way via the introduction of a Lagrange multiplier. The numerical minimization yields the low-T free energy $\mathcal{F}(x)$ and by differentiation the average force along x, $f(x)$.

Three free energies and their corresponding force-*versus*-elongation curves, valid for different values of the ratio between bending cost and desorption energy, are shown in Figures 9 and 10. They display a peak each time a turn is unwound, since the free energy presents non-convexities due to a) the fact that the elongation x is non-linearly related to the desorption angle along the toroid, and b) to the different bending penalties experienced by the polymer as the toroid tilts and unwinds. From Figures 9 and 10 one can see that as the stiffness increases, the plateaus in the force-*versus*-elongation curves tilts down and for very high stiffness the free energy itself becomes non-monotonic, with pronounced local minima for the angular configurations corresponding to the minimum bending energy cost for the polymer. These peaks are equal in height in general and thus if we perform the Legendre transform of the characteristic curves, these do not display a multi-step behaviour: in other words in the conjugated fixed force ensemble the spool unravels in a cooperative way in an all-or-none transition.

Peaks of different height can be, on the other hand, found again if we include surface effect in the modelling. These can naturally come out if we think that different sections of the toroid which are unravelled have a different number of neighbours (see Fig. 11). Since in real experiments the number of turns in a toroid are somewhat limited, this effect cannot be realistically neglected and might give reasons for the irregular peak height patterns observed in reference [5].

Finally, we note that the present treatment does not consider topological obstacles upon the unravelling of successive turns in the spool. Phenomenologically, one can take this into account in principle by allowing for a non-homogeneous ϵ along the spool, since as one goes through the loop the cross-section may change and there might be entangled points which will cost energy to disentangle. It is not possible at this level to quantitatively discuss this effect. In the framework discussed here, this effect might add more peaks to the stretching curve, but the equispaced peaks we predicted would still be found.

5 Discussion and comparison with related works

In this section we attempt to critically assess to what extent our findings can explain the experimental phenomenology of DNA stretching in a poor solvent. We also discuss the relation of our calculation with recent literature of polymer stretching below the theta point aiming at explaining the same experimental facts (see Refs. [21–23]) invoking the importance of dynamical effects.

Within our equilibrium treatment, peaks may arise a) due to conformational changes of the whole polymer or part of it if the necklace model applies (see Sect. 1), or b) due to a non-convex free energy as a function of elongation if the initial state is ordered and configurational changes are excluded (see Sect. 4). In the former case, peaks should be ascending and are simply due to surface effects. In case b), the peaks are of constant height unless finite-size corrections are once more taken into account, and are dependent on to the geometry of the initial configuration of the collapsed state.

Coming to the experimental curves, let us first consider the responses recorded upon stretching the polymer chain. The stick-release curves reported in references [3, 4] show an ascending peak pattern. It appears reasonable that these are due to surface effects since the different events are always associated with different forces and the stick-release tendency is more enhanced for the shorter DNA. If this is the case, from the analysis in Section 3 of the Legendre transform of the raw data, it seems more likely that the initial configuration is a modular structure made up of globules of different lengths, which undergo complete stretching or different configurational changes at different values of the force depending on their size as shown in Section 1.

On the other hand, the curves reported in reference [5], do not display a regular pattern in their heights. The typical peak-peak distance is 300 nm, but the distance distribution is quite broad. Furthermore, the concentration of counterions is larger. It appears more likely that the initial configuration contains one or more toroids. Peaks should, in this case, predominantly be given by the mechanism pointed out in Section 4. Other peaks might be due to conformational changes of blobs or due to dynamical effects.

Of course a complete interpretation of experiments must take into account dynamical effects too. Indeed, the force-*versus*-elongation curves display hysteresis, namely the response force upon release is smaller than the

one found upon stretching. The only exception is the force-*versus*-extension curve obtained by Wang *et al.* in reference [1], and corresponds to the smallest concentration of condensing agent.

Regarding the release curves in references [3,4], these do not show peaks and might mean that the force basically comes back to 0 as the end-to-end distance is decreased. For example this happens if the polymer does not manage to re-collapse during the release. A similar effect occurs in titin stretching [24], too. However, the shape of the release response found in reference [5] is not the same: the force is not 0 for low extension and non-monotonicity is again observed, though much less pronounced. If the stretching occurs by unravelling toroids turn by turn, it appears unlikely that the re-folding also occurs turn by turn even if it might be energetically favourable. More likely, the refolding occurs via a different path, quite possibly via conformational changes to racquet-like states (see Ref. [17]), which would thus lead to non-monotonicity.

As discussed in references [21–23] dynamical effects may also lead to peaks. However, it is important to realize that, as shown here, the presence of peaks in the force-*versus*-elongation curve has to be expected, quite generally, also in equilibrium systems. Thus, our calculations present a definite prediction that peaks should be observed even in experiments in perfect thermal equilibrium, which are *in principle* obtainable if the stretching velocity becomes small enough.

The idea beyond the dynamical explanation given in reference [23] is similar in spirit to the surface effects postulated in Section 4 for the toroid unspooling. Reference [22] gives an interesting explanation of chromatin unspooling in terms of dynamical effects, proposing a calculation similar to the one reported in the previous section. It should be noted that from the model considered by the authors one would predict that the *fixed force* unspooling of a toroid proceeds turn by turn most probably because of the high free-energy barriers, a result different from our equilibrium calculation, which predicts an all-or-none transition in the fixed force ensemble but still the presence of peaks in the experimental fixed stretch ensemble.

A series of experiments can be devised which would help further clarify this situation. First, the behaviour of the stretching curves as a function of the stretching velocity could be analyzed. There are methods (see Ref. [25] and references therein) to extract from these curves the equilibrium free-energy profiles as a function of x. If the resulting equilibrium free energy is not convex, this would be an indication that the equilibrium mechanisms proposed here should be at work in the experiments. Second, fixed force unfolding experiments with flexible and semiflexible polymers could be performed to assess the importance of finite-size effects. If dynamic effects are the only reasons for steps, then the patterns should not differ between the two experiments. If the picture described in Section 2 is at work, however, there would be strong qualitative differences in the nature of steps.

We finally note that in the model of reference [26] one can identify an effect similar to the one pointed out here, namely when stretching a polyelectrolite, thought of as a modular arrangements of spherical globules, the force-*versus*-elongation curve displays an ascending peak pattern, corresponding to the unwinding of various globules, if the system is finite. This and our results can together explain also why multi-step transitions have been observed in constant force unfolding (with atomic force microscopy) of modular proteins such as titin as well [24].

6 Conclusions

In conclusion, we have studied the force-induced unfolding of a collapsed polymer. At low temperature the elongation-*versus*-force curve has a multi-step pattern due to finite-size effects. This is general and holds for models on and off lattice, for flexible and semiflexible polymers and any dimension. We have re-calculated from experimental raw data the elongation-*versus*-force curves in the fixed force ensemble, and shown they also display steps in general. We have discussed to which extent the modelling above can give reasons of the stick-release patterns. We also considered the situation in which a toroid is unfolded gradually and we showed that also in this case peaks appear (though in general not steps in the conjugated ensemble) due to the fact that the free energy is non-convex. Finally, we compared our approach with recent literature and argued on the role of dynamical effects leading to hysteresis in real situations.

This work was supported by FISR 2001, EPSRC and PRIN 2003.

References

1. C. Bustamante, J.C. Macosko, G.J.L. Wuite, Nat. Rev. Mol. Cell Biol. **1**, 130 (2000).
2. R. Lavery, A. Lebrun, J.F. Allemand, D. Bensimon, V. Croquette, J. Phys.: Condens. Matter **14**, R383 (2002).
3. C.G. Baumann *et al.*, Biophys. J. **78**, 1965 (2000).
4. Y. Murayama, M. Sano, J. Phys. Soc. Jpn. **70**, 345 (2001).
5. Y. Murayama, Y. Sakamaki, M. Sano, Phys. Rev. Lett. **90**, 018102 (2003).
6. B.J. Haupt, T.J. Senden, E.M. Sevick, Langmuir **18**, 2174 (2002).
7. A. Halperin, E. Zhulina, Europhys. Lett. **15**, 417 (1991).
8. I.S. Aranson, L.S. Tsimring, Europhys. Lett. **62**, 848 (2003).
9. T. Frisch, A. Verga, Phys. Rev. E **66**, 041807 (2002).
10. M. Cieplak, T.X. Hoang, M.O. Robbins, Proteins **49**, 104 (2002).
11. A. Rosa, T.X. Hoang, D. Marenduzzo, A. Maritan, Macromolecules **36**, 10095 (2003).
12. P. Grassberger, H.P. Hsu, Phys. Rev. E **65**, 031807 (2002).
13. C. Vanderzande, *Lattice Models of Polymers* (Cambridge University Press, 1998).
14. For the case with no interaction see, *e.g.*, R.M. Neumann, Phys. Rev. A **31**, 3516 (1985).
15. D. Marenduzzo, A. Maritan, A. Rosa, F. Seno, Phys. Rev. Lett. **90**, 088301 (2003).

16. M.C. Tesi, E.J. J. van Rensburg, E. Orlandini, S.G. Whittington, J. Stat. Phys. **82**, 155 (1996).
17. A. Montesi, M. Pasquali, F.C. MacKintosh, cond-mat/0308356.
18. C. Bouchiat, M. Mezard, Eur. Phys. J. E **2**, 377 (2000).
19. R.D. Kamien, Rev. Mod. Phys. **74**, 953 (2002).
20. S.M. Bhattacharjee, D. Marenduzzo, J. Phys. A **35**, L349 (2002).
21. H. Wada, Y. Murayama, M. Sano, Phys. Rev. E **66**, 061912 (2002).
22. I.M. Kulic, H. Schiessel, Phys. Rev. Lett. **92**, 228101 (2004).
23. I.R. Cooke, D.R. M. Williams, Europhys. Lett. **64**, 267 (2003).
24. A.F. Oberhauser *et al.*, Proc. Natl. Acad. Sci. USA **98**, 468 (2001).
25. O. Braun, A. Hanke, U. Seifert, cond-mat/0402496.
26. T.A. Vilgis, A. Johner, J.F. Joanny, Eur. Phys. J. E **2**, 289 (2000); M.N. Tamashiro, H. Schiessel, Macromolecules **33**, 5263 (2000).