PAPER

Quantifying supercoiling-induced denaturation bubbles in DNA

Jozef Adamcik,[a] Jae-Hyung Jeon,[b] Konrad J. Karczewski,[a] Ralf Metzler*[bc] and Giovanni Dietler[a]

In both eukaryotic and prokaryotic DNA sequences of 30–100 base-pairs rich in AT base-pairs have been identified at which the double helix preferentially unwinds. Such DNA unwinding elements are commonly associated with origins for DNA replication and transcription, and with chromosomal matrix attachment regions. Here we present a quantitative study of local DNA unwinding based on extensive single DNA plasmid imaging. We demonstrate that long-lived single-stranded denaturation bubbles exist in negatively supercoiled DNA, at the expense of partial twist release. Remarkably, we observe a linear relation between the degree of supercoiling and the bubble size, in excellent agreement with statistical modelling. Furthermore, we obtain the full distribution of bubble sizes and the opening probabilities at varying salt and temperature conditions. The results presented herein underline the important role of denaturation bubbles in negatively supercoiled DNA for biological processes such as transcription and replication initiation *in vivo*.

1. Introduction

It is becoming increasingly clear that the intrinsic and protein-induced topological properties of the DNA double helix, such as chromosomal packing, DNA looping, DNA knots, or local DNA denaturation zones, influence virtually every cellular reaction involving DNA. This pertains particularly to gene regulation and DNA replication, as well as to chromosomal matrix attachment.[1–3] These biochemical processes are intimately connected with the propensity to unwind of specific local DNA regions, the DNA unwinding elements of typical size of 30–100 base-pairs.[1,4] Such unwinding elements give rise to locally unpaired regions forming a single-stranded denaturation bubble. DNA bubbles are a primary requirement for the initiation of DNA transcription and replication in both prokaryotic and eukaryotic cells: DNA unwinding provides access to the core of the base-pairs and thus the reactive groups of the nucleobases, which are otherwise shielded deeply inside the double helix.

Local unwinding of the DNA double strand may be spontaneously effected by thermal fluctuations. The resulting denaturation bubbles are sequence-specific, located in regions rich in stacks of thermodynamically weaker AT base-pairs.[5,6] The bubbles correlate with promoter transcriptional start sites and other major regulatory sites.[7–14] Pure thermal denaturation bubbles typically encountered in linear DNA molecules often

used in *in vitro* studies are statistically rare and relatively short-lived at temperatures well below the thermal melting temperature.[15–17] Long-lived denaturation bubbles relevant for biochemical processes such as protein binding are promoted by mechanical rather than entropic forces. Thus using magnetic and optical tweezers setups it was demonstrated that sufficiently long DNA denaturation occurs under mechanical torsion[18] or by overstretching of the DNA chain.[19,20] These observations are consistent with mechanical effects in statistical DNA models without[21,22] and with superhelical stress.[23–28] The stress-induced duplex destabilisation (SIDD) in mechanically stressed DNA has indeed been shown to be strongly correlated with DNA-regulatory elements.[7–10] The existence of supercoiling-induced DNA unwinding was demonstrated in an imaging study roughly a decade ago.[29]

Native DNA is typically subject to topological constraints.[30,31] Thus prokaryotic DNA predominantly occurs in circular form, while eukaryotic DNA is packaged in complex chromatin structures. When such topologically constrained DNA is not nicked, the total linking number Lk = Tw + Wr is conserved according to the Călugăreanu–White–Fuller theorem.[32] The twist Tw is equal to the number of times that the two DNA strands wind around the curvilinear central axis of the molecule, and the writhe Wr is the number of superhelical turns. Fully relaxed B-DNA carries approximately one turn per 10.4 base-pairs, and no supercoiling is present. That means that relaxed B-DNA with N base-pairs has the linking number $Lk_0 = Tw_0 = N/10.4$. Naturally occurring DNA is not fully relaxed but characterised by the superhelical density $\sigma = (Lk - Lk_0)/Lk_0$ of magnitude $\sigma \approx -0.06$, *i.e.*, it is negatively supercoiled (Fig. 1).

In native DNA twist and writhe are constantly being exchanged under the constraint that the linking number is

[a]Laboratoire de Physique de la Matière Vivante, Institute of Physics of Biological Systems (IPSB), Ecole Polytechnique Fédérale de Lausanne (EPFL), CH-1015 Lausanne, Switzerland
[b]Physics Department, Tampere University of Technology, FI-33101 Tampere, Finland
[c]Institute for Physics & Astronomy, Potsdam University, 14476 Potsdam-Golm, Germany. E-mail: rmetzler@uni-potsdam.de
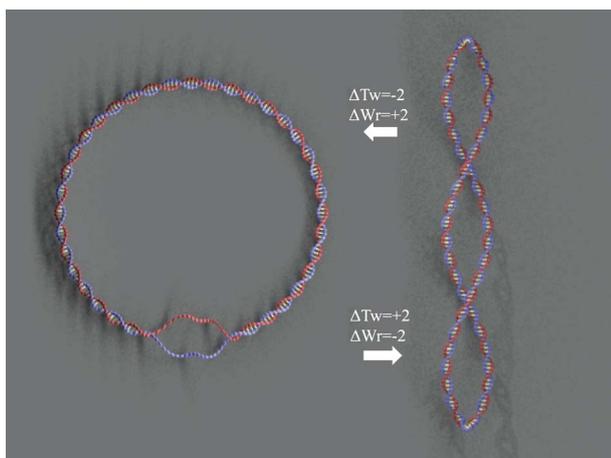
Fig. 1 Negatively supercoiled circular DNA plasmid (right). The twist may be exchanged for writhe by unwinding part of the double helix, producing a plasmid with lower twist and a denaturation bubble (left).

conserved: if the writhe is increased the twist must decrease. This is achieved by unwinding of the DNA double helix, eventually producing one or several single-stranded denaturation bubbles (Fig. 1). Based on extensive single DNA plasmid imaging by atomic force microscopy (AFM) we here present a statistical analysis of DNA bubble formation in circular DNA plasmids. We demonstrate that in negatively supercoiled DNA long-lived
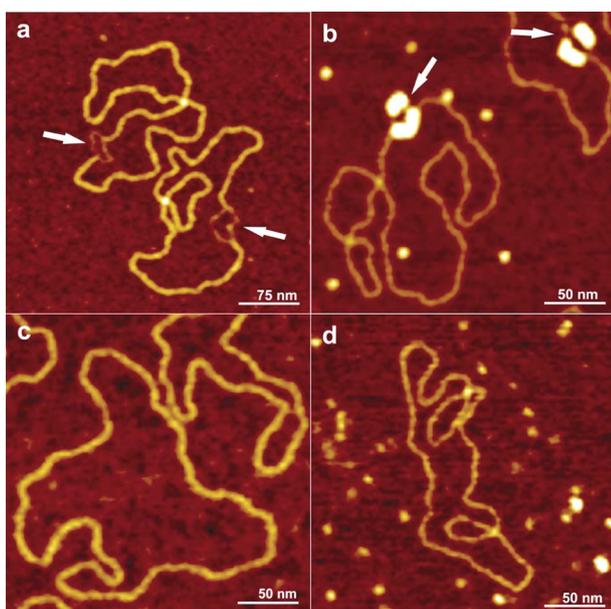

Fig. 2 AFM images of pUC19 plasmid at different conditions deposited on APTES-treated mica. (a) Supercoiled pUC19 plasmid in 1 mM Tris–HCl buffer (pH 7.8) at 25 °C. The appearance of a single denaturation bubble per plasmid is marked by the arrows. (b) Supercoiled pUC19 plasmid in 1 mM Tris–HCl buffer (pH 7.8) incubated with SSB at 25 °C for 15 minutes. One can easily discern the SSB bound to the ssDNA bubbles, visible as bright spots (arrows). (c) Relaxed pUC19 plasmid in 1 mM Tris–HCl buffer (pH 7.8) at 25 °C without a bubble. (d) Relaxed pUC19 plasmid in 1 mM Tris–HCl buffer (pH 7.8) incubated with SSB at 25 °C for 15 minutes; no SSB are observed to be bound to the relaxed plasmid.

denaturation bubbles come about by partial relaxation of the twist. Such sustained DNA bubbles are likely to play an active role in cellular processes, by facilitating access to the core of the base-pairs. In particular we find a linear relation between the degree of plasmid supercoiling and the bubble size. Moreover, we analyse the full distribution of bubble sizes, uncovering a distinct quantisation of bubble sizes consistent with the discreteness of DNA twist and writhe. We also present statistical results for the propensity to bubble formation at different salt and temperature conditions. Good agreement with predictions from a statistical model is demonstrated.

2. Results and discussion

We studied three different circular double-stranded DNA cloning vectors: pUC19, PhiX174, and pBR322. Throughout the paper we focus on results for pUC19, which consists of 2686 base-pairs corresponding to a length of 0.92 μm. The pUC19 plasmids were irreversibly deposited on a flat, APTES-treated mica surface from liquid solution (see Section 4). By AFM imaging, the conformations of the deposited plasmids were imaged at varying salt concentrations, ranging from almost 0 to 150 mM NaCl, and at different temperatures. Examples for the pUC19 plasmid configurations at zero NaCl concentration are shown in Fig. 2a. The AFM resolution indeed allows us to recognise single-stranded DNA bubble domains, as indicated by the arrows. Under these experimental conditions nearly 100% of the plasmids have *exactly* one bubble, consistent with predictions of the Benham model, as discussed below. In the AFM images double-stranded DNA crossings are located at the small bright spots on the DNA contour in Fig. 2a. These crossings on average reflect the degree of supercoiling in the plasmid, each crossing corresponding to one writhe unit. Bubbles absorb part of the twist, so that negatively supercoiled plasmids with larger bubbles have a smaller number of crossings, as quantified below. To demonstrate the generality of our results, we also studied other supercoiled plasmids such as pBR322 (Fig. 3a), PhiX174 (not shown) and supercoiled DNA ladder (Fig. 3c) under identical conditions as for pUC19. The supercoiled ladder contains a mixture of supercoiled plasmids with different sizes ranging from 2 to 10 kbp. As indicated by the arrows in Fig. 3, in all cases exactly one bubble per plasmid molecule was observed. In addition, plasmids deposited on chemically modified HOPG also feature exactly one bubble, as for the case of pBR322 plasmids (Fig. 3a and b). These results suggest that the deposition of DNA molecules on a substrate does not influence the formation of single-stranded bubbles.

To verify that the bubbles are not merely artefacts from the deposition process we performed various tests. First, we incubated the DNA plasmids in solution with selectively single-stranded DNA binding proteins (SSB) and then deposited them on the mica substrate for AFM imaging. SSB bind specifically to previously denatured, single-stranded regions of the DNA molecule. As shown in Fig. 2b indeed a bubble domain covered with SSB is visible, of a size comparable to the bubble domain in Fig. 2a. On the AFM images, SSB appear as relatively large spots, significantly brighter than the DNA crossings. The deposition process is quite fast, approximately 30 seconds, and we verified that incubation of already deposited DNA with SSB does
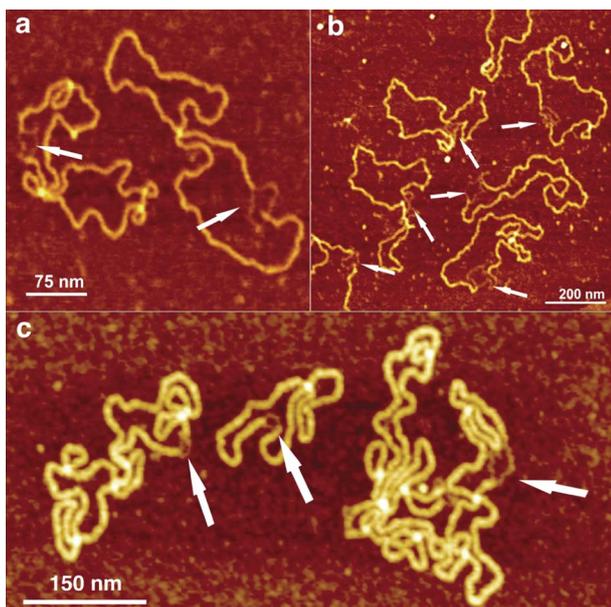
Fig. 3 AFM images of different supercoiled plasmids in 1 mM Tris–HCl buffer (pH 7.8) at 25 °C. (a) Supercoiled pBR322 plasmids deposited on APTES-treated mica. (b) Supercoiled pBR322 plasmids deposited on chemically modified HOPG. (c) Supercoiled DNA ladder (supercoiled plasmids with 2 to 10 kbp) deposited on APTES-treated mica. In all cases each DNA molecule contains exactly one bubble.
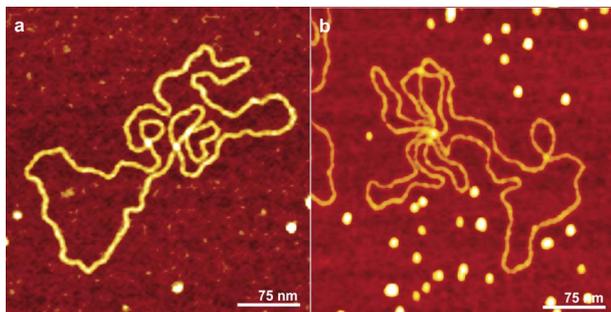

Fig. 4 AFM images of nicked PhiX174 plasmids with 7 nicking sites in 1 mM Tris–HCl buffer (pH 7.8) incubated at 37 °C for 15 minutes in the presence of (a) 2 μg ml^{-1} of SSB and (b) 10 μg ml^{-1} of SSB.

not lead to formation of SSB–DNA complexes. Therefore the experiment conclusively suggests that a single long-lived bubble already exists in solution, i.e., the SSB–bubble complex is not an effect of the deposition onto the mica substrate.† Incubation of DNA plasmids with SSB in solution offers the additional advantage that even small bubbles can be detected, even though they would hardly be discernible in the AFM images without SSB. Thus, Sucato et al.[34] have shown that supercoiled DNA molecules in 20 mM ionic strength form a small bubble. The

---

† While it was recently observed that SSB may lose the specificity to single-stranded DNA at very low salt concentrations,[33] we observed a matching percentage of bubble occurrence and SSB binding complexes in AFM images with and without SSB at zero salt concentration. Moreover we do not observe SSB–DNA complex formation for topologically relaxed DNA.
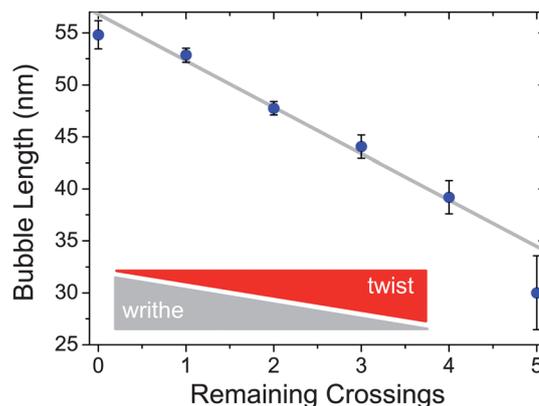

Fig. 5 Bubble contour length *versus* the number of remaining crossings extracted from approx. 1000 AFM images of DNA plasmids. The contour length of a bubble is defined as half of the entire circumference of a single-stranded open region. pUC19 plasmids were treated at experimental conditions of 1 mM Tris–HCl buffer, pH 7.8, and 25 °C. The grey line represents the fit-free theoretical expectation from our statistical model for a 2686-base-pair topoisomer with the following parameters: superhelical density $\sigma = -0.06$, nucleotide–nucleotide distance in single-stranded region $l_s = 0.43$ nm, base-pair–base-pair distance in B-DNA of 0.34 nm, average energy cost of base-pair opening of $1.1 k_B T$, and twist persistence length of B-DNA of 95 nm. For details see Appendix A.

AFM investigation of supercoiled pUC19 plasmids in 20 mM NaCl did not show any bubbles; however, after incubation with SSB almost all molecules showed the formation of SSB–DNA complexes.‡

To further corroborate that the long-lived denaturation bubbles are primarily due to superhelical stress inherent in topologically constrained DNA, we treated the plasmids with topoisomerase I before deposition. This enzyme cuts one strand and, after the DNA is fully relaxed (Wr ≈ 0), reseals the cut.[30] Indeed, the images of plasmids incubated with topoisomerase I show hardly any supercoiling (Fig. 2c). Unlike Fig. 2b, no SSB attachment was observed for the relaxed plasmids shown in Fig. 2d, i.e., no sufficiently long-lived denaturation bubbles existed in solution. Treatment of supercoiled plasmids with nicking enzymes produces fully relaxed plasmids with one or more cuts of one DNA strand. The incubation of nicked PhiX174 plasmids containing 7 nicking sites with different concentrations of SSB (Fig. 4) proved that SSB can only bind to sufficiently long single-stranded DNA.

We analysed approximately 1000 imaged DNA plasmid configurations such as those shown in Fig. 2a. Each plasmid contains exactly one bubble. From the individual plasmid configurations we extracted the contour length of the single-stranded domain as well as the number of crossings. Fig. 5 shows the experimentally obtained relation between the bubble contour length and the number #RC of remaining crossings. From the linking number conservation Lk = Tw + Wr for closed circular, negatively supercoiled DNA, we expect a negative-linear relation of the exchange between twist and writhe: the

---

‡ The minimum bubble size for SSB binding is expected to be a few tens of nucleotides. For instance, two tetramers of *E. coli* SSB occlude approximately 35 nucleotides.[35]
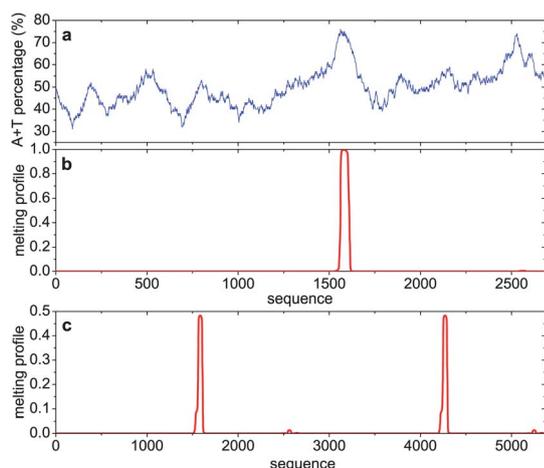
Fig. 6 (a) Percentage of AT base-pairs along the genome. (b) Opening profile for the plasmid pUC19. (c) Opening profile for the double plasmid consisting of two pUC19 fused together into one circular DNA (bottom). Parameters are $\sigma = -0.055$, $T = 37\,°C$, and 10 mM NaCl. Data obtained from the Benham model.[36]

remaining writhe of the deposited plasmids equals the negative number of the remaining crossings, $Wr_d = -\#RC$. Concurrently, the twist change mostly converts to bubble opening, which absorbs $\Delta Tw = (\text{bubble length})/(10.4 l_s)$, where l_s is the nucleotide–nucleotide distance in the single-stranded region. Linking number conservation then yields the linear relation between bubble size and remaining twist. Over the observed range from our AFM images we indeed find an approximately linear dependence between bubble contour length and superhelical crossings, as demonstrated in Fig. 5. The observed behaviour is quantitatively well explained by a simple statistical model which considers twisting and base-stacking energies of a topologically constrained DNA (see Appendix A), from which a linear relation between bubble contour length and writhe is derived. Thus, while the bubble contour length is determined by the superhelical density and the energetics of base stacking and DNA elasticity, interestingly, the linear slope Nl_s/Lk_0 (equal to $10.4 l_s$) is *independent* of these quantities. The grey line in Fig. 5 shows the theoretical prediction from our model with typical parameter values (Appendix A). Without fitting the theory exhibits excellent agreement with the experiment. The average gain of the bubble contour length by a unit change of the number of remaining crossings corresponds to ≈ 4.5 nm. The fact that the last data point for 5 remaining crossings deviates from linear behaviour is likely due to insufficient sampling statistics.

The fact that we always observe exactly one bubble per plasmid is intimately related to the stability landscape along the genome. From analysis with the Benham model[36] it can be seen that the probability of finding an open bubble is almost one at base-pairs 1525 to 1633. All other sites on the genome stay closed with probability close to one (see Fig. 6b). Fig. 6a shows that the designated stress-induced duplex destabilisation (SIDD) is strongly correlated with the long stretch in the pUC19 genome with relatively high content of weaker AT base-pair stacks. Interestingly we observed a small number of plasmids with two bubbles (see Fig. 7). The contour length of these plasmids is found to be twice the length of a single pUC19 plasmid, so that
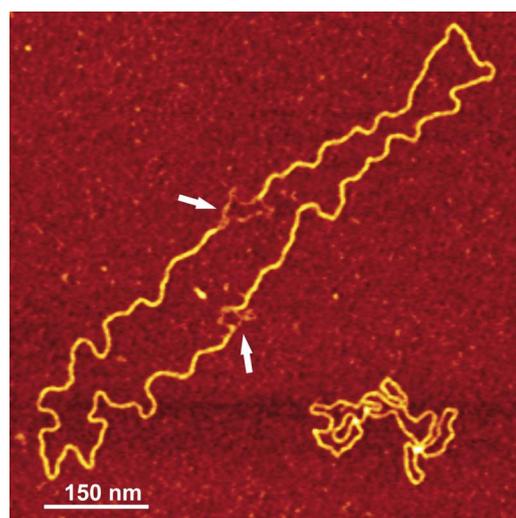

Fig. 7 AFM image of a dimerised form of pUC19 plasmid in 1 mM Tris–HCl buffer (pH 7.8) at 25 °C deposited on APTES-treated mica. The appearance of two bubbles is detected (arrows). Tracing out of the contour indeed yields twice the pUC19 plasmid length. Statistical analysis of such a dimer confirms the existence of two bubble zones in such a dimerised plasmid (Fig. 6c).

we conclude that plasmids with two bubbles correspond to rare cases of dimerised pUC19. Statistical analysis indeed corroborates the existence of two very distinct denaturation zones (Fig. 6c). The locality of the denaturation is thus preserved also in longer plasmids.

From our extensive data it is also possible to deduce the effect of the DNA topology on the distribution of bubble lengths induced by superhelical stress. As the linking number change ΔLk occurs only with integer values, the length of bubbles attributed to such a change should attain certain discrete values. This is indeed shown in Fig. 8 depicting the bubble contour length histogram for plasmids with zero crossings (corresponding to the first data point in Fig. 5). Indeed, only preferred

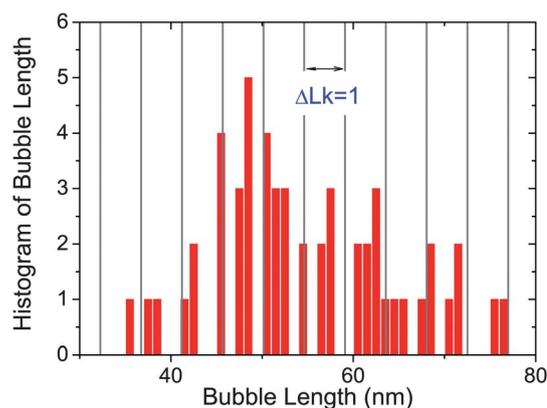Fig. 8 Histogram of bubble lengths in plasmids with zero remaining crossings ($Wr_d = 0$). The distribution reflects the discrete values of the bubble contour length, permitted by the discrete changes of the linking number. Vertical lines are the theoretical spectrum of bubble contour lengths corresponding to $\Delta Lk = 1$. Similar distributions are observed for non-zero remaining crossings (see Fig. 9).
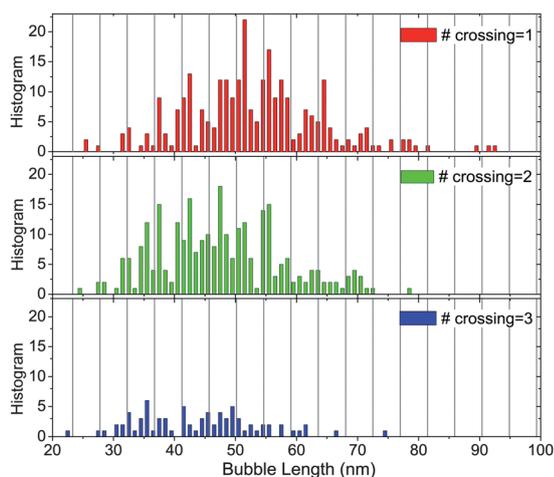
Fig. 9 Experimental histogram of bubble contour lengths of pUC19 plasmids with number of remaining crossings = 1 (red), 2 (green), and 3 (blue). Vertical lines are the theoretical spectrum of bubble contour lengths corresponding to $\Delta Lk = 1$, with the same parameter values as in Fig. 5 and 6.

lengths for the bubbles appear to occur. Consistent with our picture, the frequently observed lengths reasonably agree with our theoretical spectrum of bubble lengths with the separation $\Delta Lk = 1$ (vertical grey lines in Fig. 8). A similar behaviour is also found for bubble length distributions of plasmids at non-zero
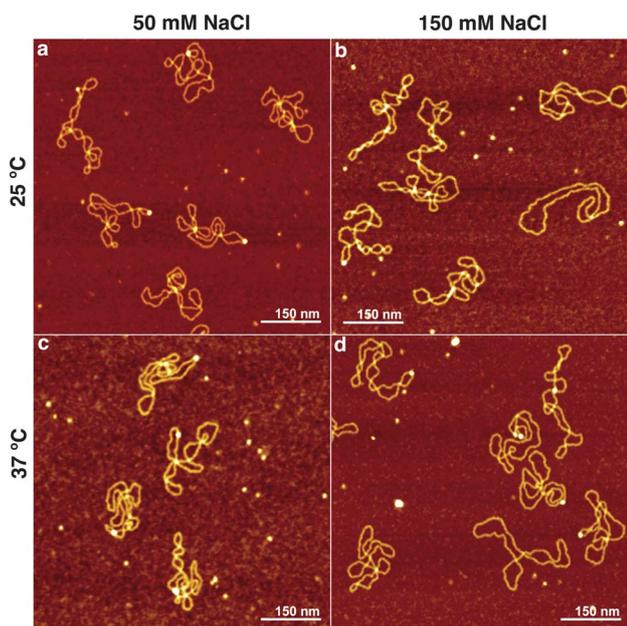

Fig. 10 AFM images of supercoiled pUC19 plasmids in 1 mM Tris–HCl buffer (pH 7.8) incubated with SSB at different concentrations of monovalent salt (NaCl) and at different temperatures, and deposited on APTES-treated mica at (a) 25 °C and 50 mM NaCl, (b) 25 °C and 150 mM NaCl, (c) 37 °C and 50 mM NaCl, and (d) 37 °C and 150 mM NaCl. The large bright spots along the DNA contours indicate SSB–DNA complexes, while crossings appear fainter. A statistical analysis of approximately 300 images per salt/temperature condition leads to the percentage of bubble occurrence at given salt/temperature condition indicated in Table 1.

Table 1 Bubble opening probabilities as a function of incubation temperature and ambient NaCl concentration. Lower salt concentration and/or higher temperature foster bubble formation. The lower part of the table shows results of our statistical model for bubble occurrence, with and without inclusion of the free energy gain upon SSB binding (see Appendix A and ref. 27)

		Salt concentration [NaCl]			
		20 mM	50 mM	100 mM	150 mM
Experiment	25 °C	97%	76%	45%	21%
	37 °C	99%	98%	78%	49%
Theory results at 37 °C	With SSB	100%	99%	78%	50%
	W/o SSB	100%	95%	74%	59%

crossing numbers, although less clear (Fig. 9). The deviation from the expected steps of $\Delta Lk = 1$ is presumably caused by an uncertainty in identifying the bubble size (in bps) from measuring the contour length of the single-stranded region. Moreover during deposition small local stresses may be induced, to which the bubble size is sensitive. Note that Fig. 8 also reflects the writhe distribution of pUC19 topoisomers in solution since, at zero crossing, ΔLk converts to ΔTw predominantly *via* bubble formation. Fig. 8 and 9 further demonstrate the unique quantitative potential of our AFM imaging technique, allowing access to previously unexplored, relevant aspects of DNA behaviour.

We also studied the effect of ambient salt concentration and temperature on the propensity towards bubble formation. Typical results are displayed in Fig. 10. Bubbles are identified by the bright spots caused by complexation with SSB. Analysing approximately 300 plasmids at each condition, we find the percentage of bubbles shown in Table 1. The effect of salt concentration and temperature is appreciable: lower salt concentration and/or higher temperature increase the occurrence of long-lived denaturation bubbles. In Table 1 we also compare the results at 37 °C with a theoretical model.[27] Taking the additional free energy gain upon SSB binding into consideration, the agreement between model and experimental results is surprisingly good. Even without accounting for SSB binding the agreement is reasonable, and represents well the observed trends.

Our findings are in line with previous analyses of DNA conformations and their complexes. Various single DNA molecule studies have shown the effects of different ionic conditions on the conformation of supercoiled DNA.[37–40] Thus, at high ionic strength (near physiological conditions, 160 mM NaCl) DNA molecules have a plectonemic form, while at low salt concentrations a considerably less interwound, open configuration of supercoiled DNA is observed, similar to Fig. 10. In the absence of salt or at very low salt concentrations (below 1 mM NaCl) a single denatured region was detected.[39]

3.   Conclusions

Biologically, the superhelical tension adopted by chromosomal or plasmid DNA is a fundamental structural property of DNA that plays a fundamental role in many genetic processes such as replication, transcription, recombination, and chromosomal matrix attachment. It is actively regulated by the balancing activities of two topological enzymes, topoisomerase I and DNA

gyrase (topoisomerase II). Physical–chemical studies of super-coiled DNA have revealed unique structural and thermodynamic features that can explain many biological effects of DNA supercoiling. The torsional energy stored in supercoiled DNA can lead to alternative DNA structures. Local alternative DNA structures stabilised by negative DNA supercoiling include left-handed Z-DNA, triple-helical H-DNA, cruciforms, and locally denatured regions.[30] Supercoiling, and therefore the relative proportion of twist and writhe, is very sensitive to environmental conditions, such as temperature and the nature and concentrations of cations. Quantitative information on the degree of supercoiling *versus* unwound regions of the DNA is therefore vital towards a better understanding of the biochemistry of DNA-related processes in cellular regulation.

Here we demonstrated that negative supercoiling in topologically constrained DNA is a necessary prerequisite for long-lived, stable denaturation bubbles even at physiological salt concentrations. These bubbles are formed by partial exchange between twist and writhe. Due to their stability they explicitly allow the attachment of single-stranded DNA binding proteins. From extensive AFM imaging we found a linear relation between the degree of supercoiling and the length of the unwound region. Moreover, we obtained the full distribution of bubble sizes for a given degree of supercoiling, mirroring the discreteness of the topological parameters twist and writhe. We also determined the bubble opening probability as a function of ambient salt concentration and temperature. Good agreement with statistical models is observed. It will be interesting to obtain similar information on more complex DNA topologies, such as packaged prokaryotic or eukaryotic DNA, or knotted DNA,[30,32,41] as well as the quantitative behaviour under molecular crowding conditions.

The fundamental advantage of our method lies in the high-throughput capability, plus the fact that the imaged DNA provides a quite truthful representation of the bubble size in solution. We note that the experimental determination and its quantitative theoretical description of the relation between residual topology and bubble size, as well as the distribution of bubble sizes to the best of our knowledge have not been previously reported.

4. Experimental procedures

Plasmids pUC19, PhiX174 and pBR322, topoisomerase I, and nicking enzymes were purchased from Fermentas. Supercoiled DNA ladders (2–10 kbp) and SSB were purchased from Promega. Plasmid pUC19 was treated with topoisomerase I to obtain relaxed molecules. In order to prepare nicked DNA molecules all plasmids were treated by nicking enzymes. DNA from all enzymatic reactions was purified with phenol–chloroform extraction. All DNAs (supercoiled, relaxed and nicked) were diluted in 1 mM Tris–HCl buffer (pH 7.8) to a final DNA concentration of 1 μg ml^{-1}. For DNA–SSB reactions, SSB were added to a final concentration of 2 or 10 μg ml^{-1}. DNA–SSB solutions with different concentrations of NaCl (20 mM, 50 mM, 100 mM and 150 mM) were incubated at 25 °C or 37 °C for 15 minutes.

Mica modified with APTES (APTES-treated mica) was prepared according to ref. 42 and 43, and chemically modified HOPG was prepared according to ref. 44. A 10 μl aliquot of the DNA solution or DNA–SSB protein solution was deposited onto APTES-treated mica and incubated for 30 seconds at room temperature. The sample was then rinsed with nanopure (Ultra High Quality) water (USF Elga, High Wycombe, England) and blow-dried with air. The adsorption of pBR322 plasmid on HOPG was performed according to the protocol of ref. 44.

Images were collected using a Nanoscope IIIa (Veeco Inc., Woodbury, NY, USA) operated in tapping mode in air. Ultra-sharp non-contact silicon cantilevers (NT-MDT Co., Zelenograd, Moscow, Russia) with a nominal tip radius <10 nm were used and were driven at oscillation frequencies in the range of 150 to 300 kHz. During imaging, the surface was scanned at a rate of one line per second. Images were flattened using the Nanoscope III software without further image processing. The plasmids' contour length was measured with Ellipse.[45]

Appendix

A. Modelling supercoil–bubble exchange

The experimental result in Fig. 5 manifests a linear relation between bubble length and the number of supercoiling turns for closed circular DNA. We here demonstrate that this behaviour can be successfully understood by a simple DNA model. Suppose that a supercoiled circular DNA of N base-pairs (bps) allows a bubble of n bps at certain conditions of superhelical density σ, temperature T and salt concentration c. The (free) energy $\mathscr{E}$ corresponding to this state will be the sum of various energies including base stacking, twisting, bending, *etc.* We assume that no twist energy is stored in denaturation bubbles, corroborated by the excellent fit of our model to Fig. 5. As a minimal model capturing the essentials of the phenomena, we consider base-stacking and twisting energies in the free energy, $\mathscr{E}(n) = \mathscr{E}_{\text{bp}}(n) + \mathscr{E}_{\text{tw}}(n)$, with the topological constraint

$$\Delta \text{Lk} = \text{Wr} + \Delta \text{Tw} = \text{Wr} - \frac{n}{10.4} + \Delta \text{Tw}_{\text{r}}, \qquad (1)$$

where $\Delta \text{Tw}_{\text{r}}$ is the residual twist in the double-stranded part. Using the 10 different nearest-neighbour energies of base stacking $\varepsilon_{j,j+1}$ (compare ref. 6), the base-stacking energy is

$$\mathscr{E}_{\text{bp}}(n) = \varepsilon_{\text{I}} + \sum_{j=1}^{i+n} \varepsilon_{j,j+1}(c, T) \qquad (2)$$

where $\varepsilon_{\text{I}} \approx 11 k_{\text{B}} T$ is the energy cost for bubble initiation.

For a homogeneous sequence of average strength of base stacking $\bar{\varepsilon}$, the above energy cost is simply $\mathscr{E}_{\text{bp}}(n) = \varepsilon_{\text{I}} + n\bar{\varepsilon}(c, T)$. Due to the residual twist each base-pair in the double-strand part, on average, has a twist strain $2\pi \Delta \text{Tw}_{\text{r}}/(N-n)$, so that the twisting energy yields

$$\begin{aligned} \mathscr{E}_{\text{tw}}(n) &= \frac{1}{2} \frac{C_{\text{tw}}}{l_{\text{d}}} \left(\frac{2\pi \Delta \text{Tw}_{\text{r}}}{N-n} \right)^2 (N-n), \\ &= \frac{2\pi^2 C_{\text{tw}}}{l_{\text{d}}(N-n)} \left\{ \left(\Delta \text{Lk} - \text{Wr} \right) + \frac{n}{10.4} \right\}^2. \end{aligned} \qquad (3)$$

Thus the twisting energy is a function of bubble size n and supercoiling turns Wr. In this expression $C_{\text{tw}}/k_{\text{B}} T$ is the twist persistence length of B-DNA (C_{tw} is called the twist modulus). $l_{\text{d}} = 0.34$ nm is the base-pair–base-pair distance in B-DNA. At

equilibrium the bubble size for a plasmid of given supercoiling turns is determined by free energy minimization:

$$\left.\frac{\partial \mathscr{E}}{\partial n}\right|_{n=n^*} = 0. \tag{4}$$

We found that the above equation leads to a simple analytic relation between bubble size and supercoiling, with the approximation that $1/(N-n)$ in eqn (3) is replaced by $(1/N)(1 + n/N) \approx 1/N$, which is reasonable for typical bubble size ≈ 100 bps for pUC19 plasmids ($N = 2686$). Converting the bubble size into its contour length b by $b = l_s n^*$, we find that

$$b(\mathrm{Wr}) = -\frac{Nl_s}{\mathrm{Lk}_0}|\mathrm{Wr}| + |\sigma|Nl_s - \frac{l_d l_s N^3}{4\pi^2 C_{\mathrm{tw}} \mathrm{Lk}_0^2}\bar{\varepsilon}. \tag{5}$$

Consistent with our experiment (displayed in Fig. 5), eqn (5) shows that the bubble contour length linearly decreases with increasing number of supercoiling turns $|\mathrm{Wr}|$. Note that the linear relation is not a consequence of the approximation. In the experimental range of $|\mathrm{Wr}|$ considered, hardly any difference is seen between eqn (5) and the exact numerical solution of eqn (4). Importantly the above shows that the slope is solely determined by the unconstrained twist repeat $N/\mathrm{Lk}_0 = 10.4$ of B-DNA, independent of parameters such as superhelical density and base-stacking and elastic energies of DNA. Meanwhile, the bubble length itself does depend on such parameters. For a given topological state σ, the bubble length decreases with increasing stacking energy $\bar{\varepsilon}$, and then the unrelaxed twist is distributed uniformly along the double-stranded chain. Conversely, for increased twist persistence length C_{tw}, the bubble length should be increased to reduce the twisting energy stored in the double-stranded part. In the limiting case of large C_{tw}, the linking number difference $\sigma \mathrm{Lk}_0 - \mathrm{Wr}$ all should go to the twist release by bubble formation, so $b = -Nl_s|\mathrm{Wr}|/\mathrm{Lk}_0 + l_s N|\sigma|$. In Fig. 5, we plot eqn (5) (grey line) with the following parameter values: superhelical density $\sigma = -0.06$, DNA size $N = 2686$ bp, twist persistence length $C_{\mathrm{tw}}/k_{\mathrm{B}}T = 95$ nm, average stacking energy of AT region $\bar{\varepsilon} = 1.1 k_{\mathrm{B}}T$, nucleotide–nucleotide distance in single-stranded DNA $l_s = 0.43$ nm, and base-pair–base-pair distance in B-DNA $l_d = 0.34$ nm. Without fitting, the above model shows surprisingly good agreement with the experimental data.

B. Distribution of bubble contour length

Based on the relation between bubble length and supercoiling turns, eqn (5), we anticipate that bubble contour lengths should be discretised for topology-fixed DNA. Especially this is obvious for the case of zero (supercoiling) crossings as displayed in Fig. 6. Here, the writhe vanishes,

$$\mathrm{Wr} = \frac{1}{4\pi}\iint \frac{\mathrm{d}\boldsymbol{r}_1 \times \mathrm{d}\boldsymbol{r}_2 \cdot (\boldsymbol{r}_1 - \boldsymbol{r}_2)}{|\boldsymbol{r}_1 - \boldsymbol{r}_2|^3} = 0, \tag{6}$$

since $\mathrm{d}\boldsymbol{r}_1 \times \mathrm{d}\boldsymbol{r}_2 (\boldsymbol{r}_1 - \boldsymbol{r}_2) = 0$ for the plasmids on the plane without crossings. Thus the above bubble contour length is rewritten as

$$b(\mathrm{Wr} = 0) = \left|\Delta\mathrm{Lk}\right|\frac{Nl_s}{\mathrm{Lk}_0} - \frac{l_d l_s N^3}{4\pi^2 C_{\mathrm{tw}} \mathrm{Lk}_0^2}\bar{\varepsilon}. \tag{7}$$

As the linking number deficit $\Delta\mathrm{Lk}$ occurs only with integer values such as $-1, -2, -3, \ldots,$ etc., the bubble has a discrete

spectrum of length distribution with increment Nl_s/Lk_0. The vertical lines in Fig. 6 indicate this spectrum, which is in reasonably good agreement with the experimentally obtained distribution (corresponding to the point at zero crossing in Fig. 5).

For non-zero superhelical crossings, similar characteristics are observed. Fig. 9 presents the distributions of bubble contour lengths at remaining crossings = 1, 2, and 3. Less sharply, the frequently observed bubble contour lengths overlap with the calculated values.

References

1 R. R. Sinden, *DNA Structure and Function*, Academic Press, 1994.
2 A. Kornberg and T. A. Baker, *DNA Replication*, ed. W. H. Freeman, 1974.
3 M. Ptashne, *A Genetic Switch: Phage λ and Higher Organisms*, Cell Press and Blackwell, 2nd edn, 1992.
4 C. J. Benham, *J. Mol. Biol.*, 1996, **255**, 425.
5 J. SantaLucia Jr, *Proc. Natl. Acad. Sci. U. S. A.*, 1998, **95**, 1460.
6 A. Krueger, E. Protozanova and M. D. Frank-Kamenetskii, *Biophys. J.*, 2006, **90**, 3091.
7 H. Wang, M. Noordewier and C. J. Benham, *Genome Res.*, 2004, **14**, 1575.
8 C. J. Benham, *J. Mol. Biol.*, 1992, **225**, 835.
9 C. J. Benham, *Proc. Natl. Acad. Sci. U. S. A.*, 1993, **90**, 2999.
10 P. Ak and C. J. Benham, *PLoS Comput. Biol.*, 2005, **1**, 41.
11 C. H. Choi, G. Kalosakas, K. Ø. Rasmussen, M. Hiromura, A. R. Bishop and A. Usheva, *Nucleic Acids Res.*, 2004, **32**, 1584.
12 C. H. Choi, *et al.*, *Biophys. J.*, 2008, **95**, 597.
13 T. Ambjörnsson, T. S. K. Banik, O. Krichevsky and R. Metzler, *Phys. Rev. Lett.*, 2006, **97**, 128105.
14 T. Ambjörnsson, S. K. Banik, O. Krichevsky and R. Metzler, *Biophys. J.*, 2007, **92**, 2674.
15 T. S. van Erp, S. Cuesta-Lopez, J. G. Hagman and M. Peyrard, *Phys. Rev. Lett.*, 2005, **95**, 218104.
16 A. Bar, Y. Kafri and D. Mukamel, *Phys. Rev. Lett.*, 2007, **98**, 038103.
17 H. C. Fogedby and R. Metzler, *Phys. Rev. Lett.*, 2007, **98**, 070601.
18 T. R. Strick, V. Croquette and D. Bensimon, *Biophys. J.*, 1998, **74**, 2016.
19 K. Pant, R. L. Karpel and M. C. Williams, *J. Mol. Biol.*, 2003, **327**, 571.
20 K. R. Chaurasiya, T. Paramanathan, M. J. McCauley and M. C. Williams, *Phys. Life Rev.*, 2010, **7**, 299.
21 T. Hwa, E. Marinari, K. Sneppen and L.-H. Tang, *Proc. Natl. Acad. Sci. U. S. A.*, 2003, **100**, 4411.
22 A. Hanke, M. G. Ochoa and R. Metzler, *Phys. Rev. Lett.*, 2008, **100**, 018106.
23 T. B. Liverpool, S. A. Harris and C. A. Laughton, *Phys. Rev. Lett.*, 2008, **100**, 238103.
24 A. Bar, A. Kabakcioglu and D. Mukamel, *Phys. Rev. E: Stat., Nonlinear, Soft Matter Phys.*, 2011, **84**, 041935.
25 A. Kabakcioglu, E. Orlandini and D. Mukamel, *Phys. Rev. E: Stat., Nonlinear, Soft Matter Phys.*, 2009, **80**, 010903(R).
26 D. Jost, Z. Asif and R. Everaers, *Phys. Rev. E: Stat., Nonlinear, Soft Matter Phys.*, 2011, **84**, 031912.
27 J.-H. Jeon, J. Adamcik, G. Dietler and R. Metzler, *Phys. Rev. Lett.*, 2010, **105**, 208101.
28 J.-H. Jeon and W. Sung, *Biophys. J.*, 2008, **95**, 3600.
29 V. N. Potaman, *et al.*, *J. Mol. Biol.*, 2003, **326**, 1095.
30 A. D. Bates and A. Maxwell, *DNA Topology*, Oxford University Press, 2005.
31 D. P. Snustad and M. J. Simmons, *Principles of Genetics*, John Wiley & Sons, 2003.

32 C. C. Adams, *The Knot Book*, W. H. Freeman, 1994.
33 L. S. Shlyakhtenko, A. Y. Lushnikov, A. Miyagi and Y. L. Lyubchenko, *Biochemistry*, 2012, **51**, 1500.
34 C. A. Sucato, D. P. Rangel, D. Aspleaf, B. S. Fujimoto and J. M. Schurr, *Biophys. J.*, 2004, **86**, 3079.
35 M. E. Ferrari, W. Bujakowski and T. M. Lohman, *J. Mol. Biol.*, 1994, **236**, 106.
36 C.-P. Bi and C. J. Benham, *Bioinformatics*, 2004, **20**, 1477.
37 C. Brack, T. A. Bickle and R. Yuan, *J. Mol. Biol.*, 1975, **96**, 693.
38 S. Dasgupta, D. P. Allison, C. E. Snyder and S. Mitra, *J. Biol. Chem.*, 1977, **252**, 5916.
39 D. I. Cherny and T. M. Jovin, *J. Mol. Biol.*, 2001, **313**, 295.

40 H. Tanaka, S. P. Mielke, C. J. Benham and T. Kawai, *J. Phys. Chem. B*, 2009, **112**, 16788.
41 E. Ercolini, F. Valle, J. Adamcik, R. Metzler, P. de los Rios, J. Roca and G. Dietler, *Phys. Rev. Lett.*, 2007, **98**, 058102.
42 J. Adamcik, V. Víglaský, F. Valle, M. Antalík, D. Podhradský and G. Dietler, *Electrophoresis*, 2002, **23**, 3300.
43 J. Adamcik, F. Valle, G. Witz, K. Rechendorff and G. Dietler, *Nanotechnology*, 2008, **19**, 384016.
44 J. Adamcik, F. Valle, G. Witz, S. Sekatskii and G. Dietler, *FEBS Lett.*, 2006, **580**, 5671.
45 J. Marek, E. Demjénová, Z. Tomori, J. Janáček, I. Zolotová, F. Valle, M. Favre and G. Dietler, *Cytometry, Part A*, 2005, **63A**, 87.