SCIENTIFIC REP⚙RTS

OPEN

Shapes of minimal-energy DNA ropes condensed in confinement

Antonio Šiber

Shapes of a single, long DNA molecule condensed in a confinement of a virus capsid are described as conformations optimizing a model free energy functional accounting for the interplay between the bending energy of the DNA and the surface energy of the DNA bundled in a "rope". The rope is formed by bundled DNA brought together by (self-)attractive interactions. The conformations predicted by the model depend on the shape of the confinement, the total amount of the packed DNA but also on the relative contributions of the bending and surface energies. Some of the conformations found were not predicted previously, but many previously proposed DNA conformations, some of which are seemingly contradictory, were found as the solutions of the model. The results show that there are many possible packing conformations of the DNA and that the one which realizes in a particular virus depends on the capsid geometry and the nature of condensing agents.

Organisms have developed mechanisms to physically compress the DNA strand(s) in order to fit in the crowded space of the cell - the use of proteins to condense the DNA in chromosomes[1] is typical both in the eukaryotic and prokaryotic cells, but also in viruses which sometimes use specific proteins to confine their DNA in a capsid (e.g. adenoviruses, see refs 2 and 3). The DNA molecule carries a large negative charge, so that its confinement requires screening agents[4] - that is why the "condensing" proteins typically have a pronounced positive charge.

The role of screening of (double stranded) DNA self-interaction is in some bacteriophages, e.g. T4, played also by multivalent counterions, such as polyamines found in bacteria where bacteriophages assemble[5]. Multivalent counterions are known to induce *condensation* of DNA, i.e. the effective self-attraction of the DNA[6]. The viral DNA is thus expected to be condensed in cases when the virus capsid is impermeable to condensing counterions, so that they, once packed in the bacteria/cell, cannot leave the virus once it enters the inter-cellular space. The role of the condensing agents may be very different in the process of DNA packing[7,8] and in the condensation of the DNA once it enters the virus - the later is of primary interest to this work.

Shapes of many virus capsids are experimentally determined to high precision, but the conformation of DNA in viruses (bacteriophages in particular) is less well known and many different shapes and orderings of the viral DNA have been proposed. Condensation of DNA in free space has been investigated *in vitro*. The shape of the DNA condensate is typically toroidal[9]. In such a configuration a single, long DNA strand is wound in a toroidal "rope" so that it forms locally parallel DNA streamlines ("bundle" of DNA) enabling thus the DNA self-attraction. The shapes and sizes of the DNA toroids have been theoretically explained[10] in a model of the condensate that accounts for the costs of the DNA bending and the formation of the condensate surface. In order to apply the model to viruses one needs to consider the DNA condensation in the confinement of the capsid. Such calculations have been performed[11,12], but the toroidal symmetry of the condensate was *assumed* from the start. The obtained theoretical predictions were found to be adequate for T5 bacteriophages which were experimentally modified so to eject about 60% of their DNA with the remaining part condensed using a buffer containing spermidine or PEG. A question still unresolved is[13] whether the *fully packed* bacteriophages and other dsDNA viruses (e.g. herpes) contain the condensate with toroidal (axial), the so-called *inverse spool* symmetry, and whether such conformation is universal. Non-toroidal conformations of bundled DNA have been proposed previously, most notably by Hud[14], Petrov *et al.*[15] and Earnshaw *et al.*[16] (for the partially disrupted capsid of giant T4 bacteriophage). Hud proposed a conformation where the DNA is packaged in a toroid that is "folded" forming thus a highly compact structure. Similar conformations may have been detected in numerical simulations[15,17,18]. However, the energetic, dynamic and possibly entropic reasons for the formation of the "folded toroid" configuration are still not obvious and additional investigation is warranted. Furthermore, it is not clear whether the folded toroid conformation is somehow special or it may be only a member of a set of possible conformations of a circularized DNA bundle, favorable energetically or dynamically in given circumstances. In this respect, it is important to investigate the

Center of Excellence for Advanced Materials and Sensing Devices, Institute of Physics, Bijenička cesta 46, 10000 Zagreb, Croatia. Correspondence and requests for materials should be addressed to A.Š. (email: asiber@ifs.hr)

interplay of surface and bending energies and how their competition influences the formation of the confined DNA structure - this it the aim of the present paper. Brownian dynamics simulations of the DNA packing[15] revealed that the DNA conformation may depend on the shape of the virus capsid, i.e. whether it is elongated or not. The influence of the geometry of the confinement on the shape adopted by the DNA condensate is thus of interest and will also be investigated here.

Previous investigaions indicate that a unique DNA conformation is not likely to be representative for all the DNA phages, irrespectively of the capsid geometry (which may be quite elongated in some bacteriophages and can have large internal portal structures[19]) and of the condensing agent (dispersed proteins, proteins bound to capsid, or multivalent counterions)[20] - the present study aims to clarify all of this factors and their influence on the conformation assumed by the condensed DNA.

The DNA conformations are investigated here within a framework of the model used to explain formation of DNA toroids in free space[10]. The shape of the DNA condensate is thus assumed to minimize the free energy but the model in ref. 10 is extended suitably to allow for investigation of non-toroidal conformations of condensed DNA. Such an approach is advantageous with respect to the simplicity of modeling as it requires only two parameters for competing energies: the persistence length (i.e. the bending rigidity) of DNA, L_p, and the surface energy per unit area of the DNA rope, σ. The shapes of interest can be most easily visualized as solids generated by motion of a planar domain $\mathcal{D}$ of area S bounded by curve $\mathcal{C}$ (cross-section of the bundle) along a *closed* three-dimensional curve Γ of length $\mathcal{L}$ so that the geometric centroid of $\mathcal{C}$ lies on Γ. This produces a shape (rope) of volume $V = S\mathcal{L}$ and area $A = \mathcal{O}\mathcal{L}$ where $\mathcal{O}$ is the length of $\mathcal{C}$. Physically, the rope is realized by DNA wounding many times around Γ, forming locally a bundle of parallel DNA streamlines. The curve Γ needs thus to be closed, which makes the problem importantly different from and more difficult than those investigated in the previous studies of shapes of open strings in confinement[17].

The DNA ropes contain the bundled DNA, together with condensing agents. It is easy to conceptualize the distribution of small multivalent couterions in a DNA rope, but the model in principle applies to condensing proteins also - in this case, the proteins may disturb the DNA streamlines around them, but the notion of the rope may still make sense. For example, when the capsid of adenovirus is disrupted, the DNA is detected in a form of a thick fiber, most likely containing several DNA strands ("string") complexed with condensing proteins ("bead")[2], so that the notion of a "nucleoprotein filament"[21] (the rope) within a capsid may make sense.

Concerning the model, one could allow for the variation of cross-section bounding curve along Γ, as long as its area remains the same - this may simulate rearrangements of DNA strands in a cross-section, depending on its position on Γ. Such rearrangements may be required for tighter packing of DNA in the capsid. In the model adopted here, the cross-section of the bundle is assumed to be a circle of radius r, for all points on Γ. This simplifies the analysis, as for given r, one needs to determine only the optimal path Γ. Note, however, that for the given *volume* (or length) of packed DNA, the shapes differ depending on r - larger cross-sectional areas imply shorter paths Γ, as V is preserved. DNA ropes dominated by surface energy (σA, where σ is the DNA surface tension) are expected to be stout and compact, i.e. their cross-sectional areas should be as large as possible so that their exposed surface is minimal. On the other hand, ropes dominated by elastic energy are expected to be slim and long so to maximally reduce the curvatures of individual DNA strands - this implies smaller cross-sectional areas[10].

The total (free) energy of the condensate in the described approach is

$$F = \gamma V + 2\pi\sigma r\mathcal{L} + \frac{\pi k_B T L_p \mathcal{T}^2}{V} \int_{\Gamma; s=0}^{s=1} ds \left[1 - \sqrt{1 - \left(\frac{r}{R_s}\right)^2} \right],$$

(1)

where the first term represents condensation energy of the rope (γ is the condensation energy per unit volume and may be positive if the DNA-DNA separation, d, is smaller than the optimal), the second is the surface free energy (including also the entropic contribution of the DNA strands at the bundle surface), and the third is the bending energy. The model does not include the DNA twist explicitly. As in ref. 10, the surface energy is assumed to be proportional to the exposed surface of the bundle, $A = 2r\pi\mathcal{L}$, with the surface tension, σ as the proportionality coefficient i.e. the energy scale. The bending energy is obtained by integrating the squared curvature over the whole length of DNA, $\mathcal{T}$. This can be reduced to the single integral once the integration over the circular cross-section of the rope/bundle is performed. The remaining integral in Eq. 1 is performed over the centroid curve Γ, where s is the arc-length coordinate of Γ ($s \in [0, 1]$), and R_s is the radius of curvature at the center of the circular cross-section. Bulk and surface energy parameters, γ and σ, depend on d and are, in the simplest approximation, proportional[22].

For known energy parameters σ and L_p, and fixed DNA length (or volume - these are related through d[22]) of DNA, the minimization of F reduces to finding optimal r and Γ which are mutually dependent as the total DNA volume, $V = r^2\pi\mathcal{L}$, is fixed. The minimization of F with respect to centroid curve Γ and radius r needs to account for (*i*) the confinement and (*ii*) possible self-collisions of the bundle in cases when the curve Γ approaches itself. Both physical effects are essentially important for the problem in question and both can be treated as penalizations of the energy functional in Eq. (1). It should be noted that in the limit of a thin bundle the approach enables one to follow essentially a *single* DNA strand - in this limit it is thus complementary to numerical simulations of packed DNA[8,15] and other models of confined semiflexible polymers[23].

For given cross-sectional radius r, infinitely many closed curved Γ can be constructed, and finding the one which minimizes the free energy is the aim of the numerical procedure described in Methods. The minimization is performed somewhat asymmetrically, due to the fact that the surface energy depends only on r and not on Γ. For given volume V, specification of r uniquely determines the surface (free) energy, and the minimization of *total*

free energy reduces to finding Γ which, for given r, yields minimal bending energy of the rope. The optimal shape of the rope depends on the magnitudes of elastic and surface energy parameters, L_p and σ, respectively. However, whatever their magnitudes may be, the rope will eventually have certain cross-sectional radius r and centroid curve Γ. Instead of considering three-parameter space (V, L_p, σ), it suffices to consider two-parameter space (V, r), as the rope radius "encodes" the relative importance of surface vs. bending energy of the rope. There is a bonus to this representation, because it should enable a simpler comparison between the theory and experiments - whereas the bending and surface energy parameters are inaccessible from the experiments, the rope radius should not be. The problem is thus reduced to *geometry*. For given V and r, Γ is found such that the bending energy *only* is minimal. The rope radius r (and the corresponding surface free energy) is systematically modified and the minimization of bending energy performed separately for each value of r. Note that in this representation the results do not depend on the value of persistence length L_p, as it appears only as the scale (multiplying factor) of the bending energy. One should remember, however, that small persistence lengths diminish the contribution of the bending energy in the total free energy balance, emphasizing importance of the surface energy contribution - in such case, one may expect formation of ropes with large values of cross-sectional radius[10].

The bacteriophage capsids are either icosahedral or elongated along an axis connecting two opposite icosahedron vertices (prolate), such as e.g. T4[16,24]. In the calculations, the confinement (capsid) is represented by a cylinder of base radius r_B and height $2r_B f$, where f is the "elongation factor" - $f = 1$ approximately represents nearly isometric virus particles, and $f > 1$ prolate viruses. The two-parameter space (V, r) is appropriately scaled so that the two relevant parameters are ratio of radii, $p_1 = r/r_B$ ($p_1 < 0.5$), and the volume ratio, $p_2 = V/(2\pi f r_B^3)$. In a wild-type virus, volume ratio is determined by the length of the viral genome. It may be useful to consider it as a parameter in the minimization as such calculations may give insight into the structures of DNA condensates in partially filled viruses, possibly also during the packaging or ejection process[25], although one should always keep in mind that the structures obtained here are thermodynamical minima and may not be realized dynamically[26]. For a certain volume of the DNA (fixed p_2) there are solutions with different bundle radii (p_1). For given σ and L_p, only one of these radii minimizes the total free energy. The values of σ and L_p depend on the ionic content of the bacteriophage interior and presence of the condensing proteins.

Results

The applicability of the model to viruses depends on the range of parameters examined. If the model is understood in its most restricted sense - to represent the condensed DNA "rope" - the ratio of radii needs to be sufficiently large. This is due to the *continuum* concepts of bulk and surface DNA in the model, so that the number of DNA streamlines in the cross-section $\mathcal{C}$ needs to be large enough in order that the two concepts make sense. In the condensed DNA medium, the center-to-center distance between the DNA strands/streamlines is about 3 nm[9,12]. Assuming a hexagonally packed DNA[12], the notion of "bulk" and "surface" may make some sense for at least 7 DNA helices, close-packed so that the central helix is surrounded with six equally distanced neighbors. The effective radius of such a bundle would be about 4 nm ($3 + 1.5$ nm). The size of the capsid depends on the bacteriophage type, but taking, e.g. $\phi 29$[25] with a base radius of about 25 nm, the minimum ratio of radii that could be simulated by the model would be about $p_1 = 4.5/25 \sim 0.2$. For larger bacteriophages, such as T4[24], the minimum ratio is about $p_1 = 4.5/45 \sim 0.1$. In the simulations presented below (Fig. 1), $p_1 > 0.275$. As demonstrated, smaller values of p_1 (~ 0.1) may also be of interest - these shapes with small cross-sectional radii occur for sufficiently low surface tensions. The numerical method adopted here, however, becomes increasingly slower as p_1 is diminished. This is due to the fact that, for given volume V, slim ropes (small p_1) are longer, requiring effectively more degrees of freedom for description of centroid curve Γ (see Methods). This increases the number of configurations which need to be examined by the minimization procedure.

The minimal energy shapes can, of course, be plotted once found, yet a quantitative measure is required to classify them. An adequate measure was found to be the *writhe* of Γ (Wr) which quantifies the topological complexness of the rope - in two dimensions, writhe is the total number of self-crossings of Γ, which, in general, contains both positive and negative contributions, depending on whether the curve crosses itself from "above" or "below"[27]. In three dimensions (3D), the writhe can be evaluated by performing a double integral over the closed curve[27]. For our purposes it suffices to consider the absolute value of writhe, |Wr|, as shapes with the same absolute value of Wr have the same free energies. Writhes obtained in 3D are real numbers, yet the shapes are found to belong to *classes* represented by territories in parameter space, in which their writhe is typically within ~ 0.25 around the integer value - there are some exceptions, specified below. Writhes of minimal free energy shapes in a (p_1, p_2) parameter space are shown in Fig. 1 for elongation factors $f = 1, 1.5, 2$ (from left to right). The elongation factors chosen can be compared to real bacteriophages. Elongation factor $f = 1$ roughly corresponds to bacteriophages with isometric icosahedral capsid, such as T12 and λ phages. T4 phage has an elongated icosahedral capsid whose height to width ratio is about 1.35[24], similar to $\phi 29$ (1.35[25]) and T2 (1.45[14]) bacteriophages. In giant vibriophages[28], the elongation ratios are between 1.7 and 2.1. Values of $f = 1.5$ and $f = 2$ are thus relevant to real bacteriophages. Much larger elongation ratios are found in giant headed T4 phages[16] which have a variable head morphology (elongation ratios from ~ 2.5 to ~ 12) due to a mutated coat protein.

The territories with different writhes have sharply defined borders so that the solutions from the two sides of the border differ by $|\Delta Wr| = 1$ or more. The exceptions were found for $f = 1.5$ in neighboring regions having Wr $= 0$ and 2, and for $f = 2$ in neighboring regions Wr $= 0$ and 1 and Wr $= 2$ and 0 - there the transitions between different conformations are continuous and the borders are fuzzy. A particularly densely packed shapes (large p_2 parameters) with different writhes were selected for graphical representation. These are indicated by circles and letters which correspond to panel labels in Fig. 2.

In case of isometric capsids ($f = 1$), a large portion of the parameter space is occupied by shapes with Wr $= 0$. These shapes are "squeezed" and folded tori, which become maximally folded for large packing ratios p_2, as shown in Fig. 2a ($f = 1$). Exactly such shapes were predicted for organization of double-stranded DNA in the
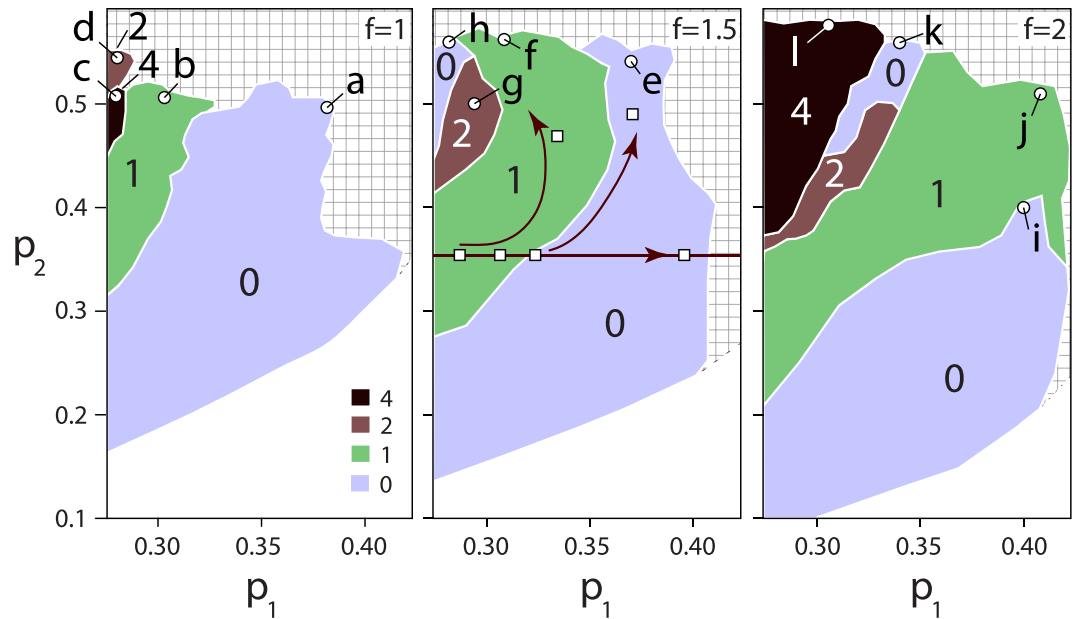
Figure 1. Regions of (p_1, p_2) space with different absolute values of writhe as indicated. In white regions the shapes are not influenced by the confinement - the (non-selfintersecting) solutions there are tori. In cross-hatched regions, the solutions were not found - all the shapes in these regions either penetrate the confinement or self-intersect. Shapes denoted by circles (squares) are represented in Fig. 2 (Fig. 3).

bacteriophage heads by Hud[14] (see Fig. 2b in ref. 14). The results obtained here show that the folded toroid proposed by Hud is the shape with minimum free energy for sufficiently thick DNA ropes/condensates ($p_1 \gtrsim 0.33$). The maximum packing fractions that can be obtained with such a shape are about $p_2 \sim 0.5$. As also explained by Hud[14], further increase in the packing fraction would be possible by allowing for a change in the cross-section of the DNA condensate along the centroid curve. Intriguingly, Hud's shape seems to be quite generic as it appears also in different contexts and has been found in simulations of growth of thin, frictionless and non-attracting elastic filaments in spherical confinement[17,18].

Shapes with $\mathrm{W}r = 1$ in isometric case can be conceptualized as an additional coil turn of the DNA rope sandwiched between the folded toroid (Fig. 2b, $f = 1$). For $f = 1$, this packing pattern produces a maximum packing fraction of about $p_2 = 0.51$, but more efficient packings with the same overall pattern are possible in case of elongated capsids ($p_2 = 0.54$), as shown in Fig. 2f for $f = 1.5$. The shapes could also be understood as pieces of tightly packed helix found in open tubes[17], but in this case their appearance is complicated by the requirement on Γ curve being closed. A surprising conformation was found in a narrow range of parameters for isometric capsids where $\mathrm{W}r \approx 4$ (Fig. 2c, $f = 1$). The conformation of the DNA rope there is *knotted*, which can be most easily detected by removing the confinement once the minimal shape is found and letting it relax - a prototypical trefoil knot appears, with three-fold symmetry and maximally extended loops so that their curvature is minimal. The writhe of the relaxed shape drops by one, to ≈ 3. The knotted shapes cannot thus be unfolded to a toroid once the capsid is destroyed - this is a situation quite different from the one discussed by Hud[14] and applies also to the shape shown in Fig. 2l, representative for the knotted shapes in the large region of parameter space when $f = 2$ (see below).

In addition to reproducing Hud's shape, the calculations also reproduce the "supertwisted" shapes proposed by Earnshaw *et al.* for partially disrupted capsids of giant T4 phages[16] (the "supertwisting" of a thick DNA rope found here should not be mistaken for supercoiling of a *single* DNA strand - see e.g. ref. 29). The shape which Earnshaw *et al.* describe as "twisted skeins of yarn" is quite similar to the one found here (twice twisted closed loop - "double helix") for minimal energy condensates in $f = 2$ capsids having $\mathrm{W}r \simeq 1$ (Fig. 2j). The elongation ratio of capsids containing supertwisted DNA in experiments[16] is about $f \sim 4$. A similar shape, but twisted only once, has been found by Petrov *et al.*, also in elongated capsids[15]. The model elaborated here does not explicitly include the twisting contribution to elastic energy of the condensed DNA rope. The DNA twist may be important for the condensation, especially for its dynamics, as shown e.g. in ref. 30. However, it is interesting to note in this context that the DNA twist is *not* required to obtain some of the rope conformations whose shape might suggest otherwise. For example, the appearance of a shape similar to the one shown in Fig. 2j (but twisted only once) has been previously proposed to result from the twisting contribution to the DNA elastic energy[31]. In the model presented here, such a shape occurs due to a particular interplay of the DNA rope bending energy and the packing constraints effected by the rope confinement in the cylinder and its self-collisions.

In many densely packed conformations of bundles in elongated cylinders, DNA strands are mostly parallel to the cylinder axis (see Fig. 2f,h,k,l), a feature found also by Earnshaw *et al.* in their experiments[16] which led them to propose the model of coiled DNA with the axis of coiling perpendicular to the elongated axis. Some of the structures found could also be loosely interpreted as a *concentric* spool[15] ("ball of yarn"), without axial symmetry,
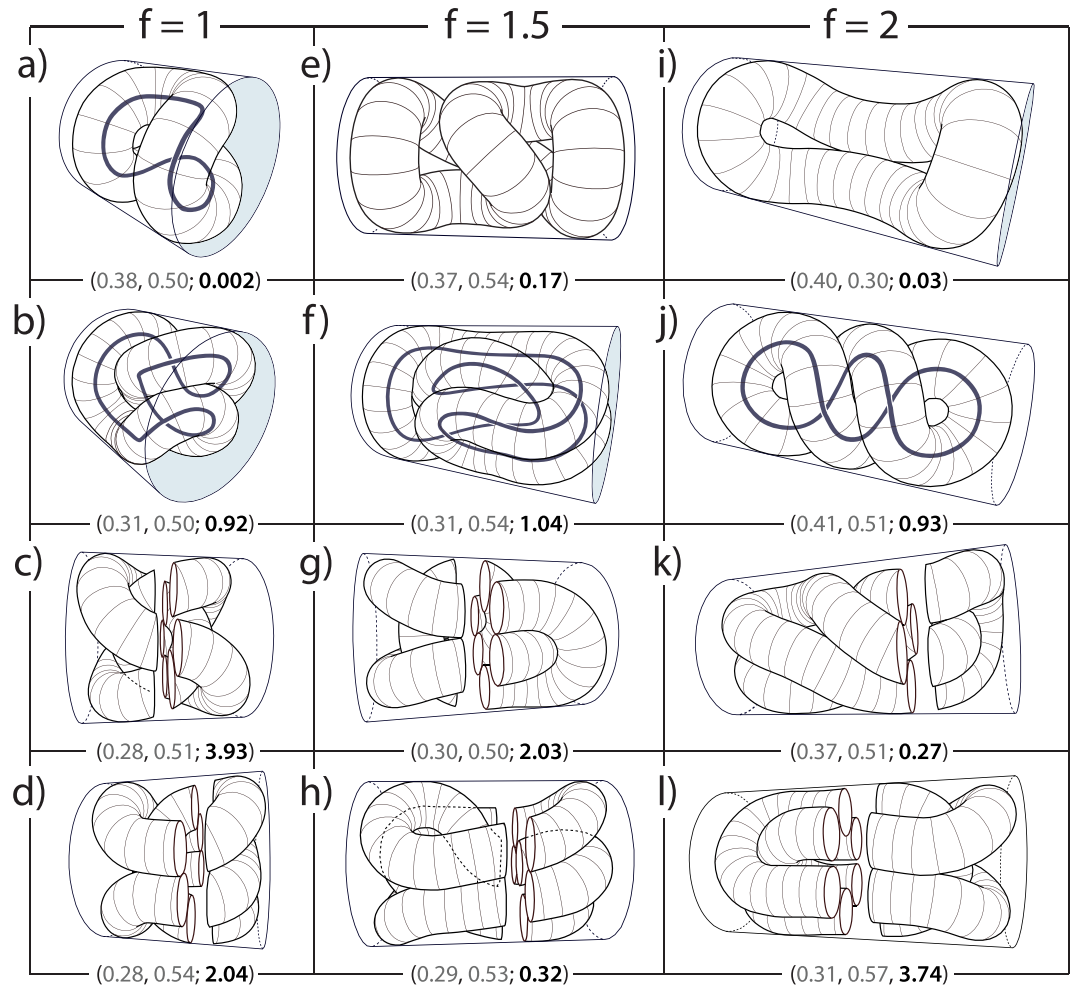
Figure 2. Minimum free energy shapes for $f = 1$, $f = 1.5$, and $f = 2$ (left, middle, and right columns). Below each of the shapes displayed, the corresponding p_1, p_2 parameters and the writhe are indicated as (p_1, p_2; Wr). To better represent the three-dimensional nature of the shape, in panels a,b,f,j, the centroid curves of the shapes, Γ's, are shown. In panels c,d,g,h,k,l cylindrical slab is cut out, so that the cross-sections of the shapes, parallel to the confinement cylinder bases can be seen.

see e.g. Fig. 2c - there the DNA strands spool along different axes, leaving a hole in the middle of the capsid. When viewed from a fixed perspective many of the shapes would appear as seemingly "random" configurations of short, liquid-crystalline like segments of DNA - this is due to complicated connectivity of the shapes (see e.g. Fig. 2c,h) and the large part of their volume being obscured from view. In that respect, some of the shapes obtained are consistent with the "liquid crystalline drop" model[13,32] and the cross-sectional size of the hexagonally crystallized domains of DNA in that model corresponds to the rope diameter $2r$. The message here is that different models of DNA packing in bacteriophages proposed previously can be reproduced in the present model for different values of p_2 and p_1, i.e. depending on the total length of the DNA and relative importance of surface vs. bending energy of the DNA condensate - this depends on the chemical environment in a capsid which determines the bulk and surface energy of the DNA bundle but also influences the persistence length of DNA. The predictions of model are, however, not consistent with the previously proposed conformations which require sharp bends of the DNA, such as the so-called spiral fold model proposed by Black *et al.*[33].

Figure 3 illustrates how the conformations of the minimal energy ropes vary throughout the (p_1, p_2) space. The shapes chosen for representation are denoted by squares in Fig. 1 for $f = 1.5$. The figure illustrates how the shapes with the same volume ratio ($p_2 = 0.35$) change as the rope radius is increased (from left to right; first, second, third and sixth shape). Note how the conformation of the rope changes as the border between the two writhe region is crossed (between the second and the third shape, the writhe changes from ∼1 to ∼0). The two additional shapes displayed (for $p_2 = 0.47$ and 0.50) enable one to visualize how the shapes vary in the Wr = 1 and Wr = 0 regions, all until the shapes denoted by f and e in Fig. 1 are reached, as indicated by arrows in Figs 1 and 3.

A specific *motif* can be detected in both knotted and unknotted shapes, such as those shown in Fig. 2d,f,g. Roughly, the motif consist of three half-tori, two of them lying in parallel planes and placed one above the other, and the third one situated in the void created by such an arrangement, lying in a plane perpendicular to those of the two parallel half-tori - an efficient packing of such sort is possible when the major radii of the tori are about
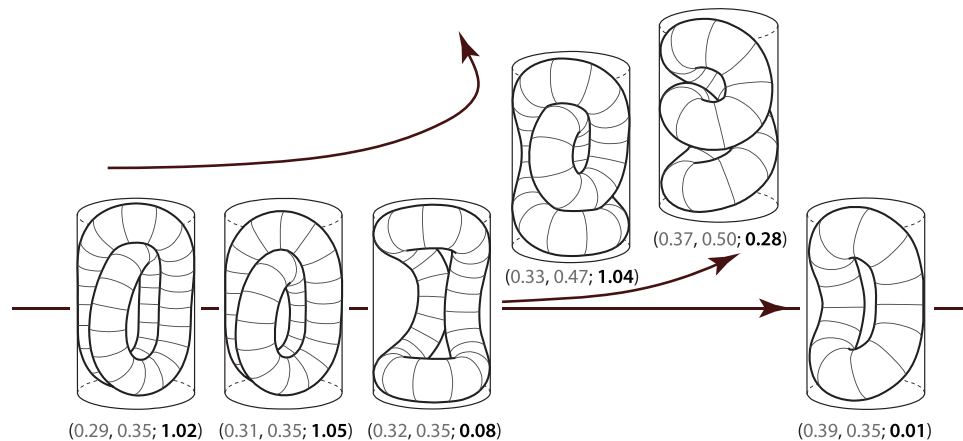
Figure 3. Minimum free energy shapes for $f = 1.5$ denoted by squares in Fig. 1. Below each of the shapes displayed, the corresponding p_1, p_2 parameters and the writhe are indicated as $(p_1, p_2; \text{Wr})$. The arrows suggest the path through the (p_1, p_2) space, all up to shapes denoted by f and e in Fig. 1 in $\text{Wr} = 1$ and $\text{Wr} = 0$ regions, respectively.
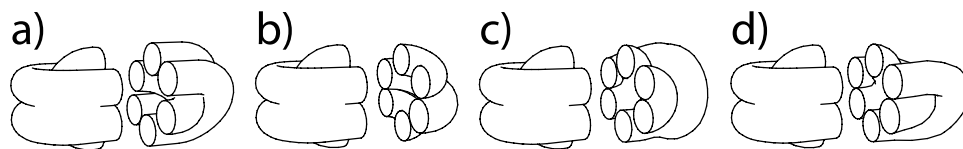

Figure 4. Four different closures of the three-toroids motif resulting in shapes having a single, closed centroid curve.

two times larger than the tube radius r ($p_1 \sim 0.3$). Often, the two parallel half-tori are found to be slightly folded, upwards and downwards for a better packing (see left parts of the shapes in Fig. 2d,g, and right part of the shape in Fig. 2l). Some of the shapes with large packing fractions can be understood different closures of the three-toroid motif, as illustrated in Fig. 4. A closure of the motif shown in Fig. 4a corresponds to minimal energy shapes in Fig. 2c,l - this case of closure is particularly interesting as the conformation of the closing shape is the same as the motif it closes, only rotated by ~60 degrees around the cylinder axis. Closures in Fig. 4b–d correspond to minimal energy shapes in Fig. 2d,f,g, respectively. All of the shapes with the characteristic motif shown have large packing fractions, $p_2 > 0.5$.

A finite thickness of the confined DNA rope is what makes the appearance of certain conformations possible only in some regions of parameter space. For example, the minimal length of the rope required to tie a trefoil (3_1) knot depends on the rope radius, so that $\mathcal{L}/r|_{min} \approx 32.74$[34] - the trefoil knots formed by ropes with ratios $\mathcal{L}/r$ larger than this (loose knots) can be tied, while those with smaller can not. Combining parameters, one finds that $\mathcal{L}/r = 2f\, p_2/p_1^3$. The trefoil knot is the simplest of all the knots and the $\mathcal{L}/r$ ratio (rope length) required to tie it is the smallest[34]. The next shortest length knot is 4_1 knot, requiring $\mathcal{L}/r|_{min} \approx 42.09$, but as the rope length increases, the space of knots that can be tied rapidly expands. All of these knots are candidates for the solutions of the problem, and are interesting in particular when stiffly tied (densely packed), i.e. as soon as the rope length $\mathcal{L}/r$ becomes larger than a critical value (32.74 and 42.09 for 3_1 and 4_1 knots, respectively). The trefoil knots were found by the SA procedure starting from random and torus conformations, so no specific initial configurations were required, yet one may wonder whether more complicated knots may evade detection. To investigate this, knots whose coordinates were taken from ref. 34, with simply scaled centroid curves so to account for ratios $\mathcal{L}/r$ larger than the minimal, were used as initial conformations in the SA procedure with different initial temperatures. More complicated knots were not found to be the minimal energy solutions for the f factors investigated when $p_1 > 0.275$. One should also note that the knotted conformation when $f = 1$ appears for rope lengths significantly larger ($\mathcal{L}/r \sim 50$) than in the case of tight knot, so that the centroid curve of the observed conformation is significantly deformed with respect to the centroid curve of the tight knot in order to fit in the cylinder (see Fig. 2c).

Discussion

All the findings indicate that there may not be a unique model of the DNA conformation appropriate for all bacteriophages. The conformation with minimal free energy depends sensitively on the total amount of the DNA, the nature of the condensing agents, and the shape of the bacteriophage capsid. The phase space of conformations is divided in territories of different writhe, separated mostly by sharp borders, which do not allow for a continuous change of the conformation. Some of the conformations proposed previously were also found in the present work, in different positions in the parameter space, further corroborating the idea of the non-uniqueness of the DNA

packing in bacteriophages. The results indicate that various conformations proposed previously may be understood within the same theoretical framework, as minimal-free-energy states of the DNA condensate, obtained in particular condensation conditions, for given amount of the DNA and in a geometrical confinement provided by the shape of the capsid. They thus interlink the previous studies (e.g. Hud's folded toroid[14], Earnshaw's *et al.* twisted skeins of yarn[16] and Lepault's *et al.* liquid crystalline drop[13,32]), showing that different shapes of confined DNA condensates may be understood in a unifying model.

Methods

The DNA rope was represented by a string of intersecting spheres, all with the same radius r, with centers distributed on centroid curve Γ and separated by a fixed amount qr, $q \sim 1$[35]. The total volume of such a system is proportional to the number of spheres, N. The degrees of freedom of the rope are related to the positions of sphere centers on Γ. The self-collisions of such a shape are easily checked for by calculating the distances between all the spheres. Once detected, the energy functional are penalized to essentially forbid the sphere-sphere penetration for spheres that are not nearest neighbors along Γ. Collisions of the shape with the container (capsid) are detected by calculating the distances between all the spheres and the cylinder surface (including cylinder bases) and by penalizing the functional to forbid the penetration of the rope outside the container. The closedness of Γ is imposed by penalizing the energy functional when the distances between the first and the last (N-th) sphere are larger or smaller than qr. The elastic energy is calculated from the positions of all the spheres by interpolating a cubic spline through all the points and integrating its curvature along Γ, as specified in Eq. (1). The minima of free energy functional, penalized as explained, are found by the simulating annealing procedure. The numerical procedure was run many times, with different initial conditions (i.e. initial shapes at high temperatures of the algorithm) for Γ, including random curves, circles, spirals/coils, (trefoil) knots, and appropriately modified curves found by the algorithm for values of p_1 and p_2 parameters different from the one considered in the minimization. The solutions were carefully manually sifted, some of them reselected for repeated minimization, all until no better solutions could be found.

References

1. Bednar, J. *et al.* Nucleosomes, linker DNA, and linker histone form a unique structural motif that directs the higher-order folding and compaction of chromatin. *Proc. Natl. Acad. Sci. USA* **95**, 14173–14178 (1998).
2. San Martin, C. Latest Insights on Adenovirus Structure and Assembly. *Viruses* **4**, 847–877 (2012).
3. Perez-Berna, A. J. *et al.* Distribution of DNA-condensing protein complexes in the adenovirus core. *Nucleic Acids Res.* **43**, 4274–4283 (2015).
4. Šiber, A., Lošdorfer - Božič, A. & Podgornik, R. Energies and pressures in viruses: contribution of nonspecific electrostatic interactions. *Phys. Chem. Chem. Phys.* **14**, 3746–3765 (2012).
5. White Tabor, C. & Tabor, H. Polyamines in Microorganisms. *Microbiol. Rev.* **41**, 81–99 (1985).
6. Teif, V. B. & Bohinc, K. Condensed DNA: Condensing the concepts. *Prog. Biophys. Mol. Biol.* **105**, 208–222 (2010).
7. Keller, N., delToro, D., Grimes, S., Jardine, P. J. & Smith, D. E. Repulsive DNA-DNA Interactions Accelerate Viral DNA Packaging in Phage Phi29. *Phys. Rev. Lett.* **112**, 248101-1–248101-5 (2014).
8. Petrov, A. S. & Harvey, S. C. Role of DNA-DNA Interactions on the Structure and Thermodynamics of Bacteriophages Lambda and P4. *J. Struct. Biol.* **174**, 137–146 (2011).
9. Hud, N. V. & Downing, K. H. Cryoelectron microscopy of λ phage DNA condensates in vitreous ice: The fine structure of DNA toroids. *Proc. Natl. Acad. Sci. USA* **98**, 14925–14930 (2001).
10. Ubbink, J. & Odijk, T. Deformation of toroidal DNA condensates under surface stress. *Euorophys. Lett.* **33**, 353–358 (1996).
11. Tzlil, S., Kindt, J. T., Gelbart, W. M. & Ben-Shaul, A. Forces and Pressures in DNA Packaging and Release from Viral Capsids. *Biophys. J.* **84**, 1616–1627 (2003).
12. Leforestier, A., Šiber, A., Livolant, F. & Podgornik, R. Protein-DNA Interactions Determine the Shapes of DNA Toroids Condensed in Virus Capsids. *Biophys. J.* **100**, 2209–2216 (2011).
13. Leforestier, A. Polymorphism of DNA conformation inside the bacteriophage capsid. *J. Biol. Phys.* **39**, 201–213 (2013).
14. Hud, N. V. Double-Stranded DNA Organization in Bacteriophage Heads: An Alternative Toroid-Based Model. *Biophys. J.* **69**, 1355–1362 (1995).
15. Petrov, A. S., Boz, M. B. & Harvey, S. C. The Conformation of Double-Stranded DNA Inside Bacteriophages Depends on Capsid Size and Shape. *J. Struct. Biol.* **160**, 241–248 (2007).
16. Earnshaw, W. C., King, J., Harrison, S. C. & Eiserling, F. A. The Structural Organization of DNA Packaged within the Heads of T4 Wild-Type, Isometric and Giant Bacteriophages. *Cell* **14**, 559–568 (1978).
17. Maritan, A., Micheletti, C., Trovato, A. & Banavar, J. R. Optimal shapes of compact strings. *Nature* **406**, 287–290 (2000).
18. Vetter, R., Wittel, F. K. & Herrmann, H. J. Morphogenesis of filaments growing in flexible confinements. *Nat. Commun.* **5**, 4437-1–4437-8 (2014).
19. Guo, F. *et al.* Visualization of uncorrelated, tandem symmetry mismatches in the internal genome packaging apparatus of bacteriophage T7. *Proc. Natl. Acad. Sci. USA* **110**, 6811–6816 (2013).
20. Black, L. W. & Thomas, J. A. Condensed Genome Structure. *Adv. Exp. Med. Biol.* **726**, 469–487 (2012).
21. Nermut, M. V. The Architecture of Adenoviruses: Recent Views and Problems. *Archives of Virology* **64**, 175–196 (1980).
22. Marion, S. & Šiber. A. Ejecting Phage DNA against Cellular Turgor Pressure. *Biophys. J.* **107**, 1924–1929 (2014).
23. Guven, J. & Vazquez-Montejo, P. Confinement of semiflexible polymers. *Phys. Rev. E* **85**, 026603-1–026603-16 (2012).
24. Fokine, A. *et al.* Molecular architecture of the prolate head of bacteriophage T4. *Proc. Natl. Acad. Sci. USA* **101**, 6003–6008 (2004).
25. Comolli, L. R. *et al.* Three-dimensional architecture of the bacteriophage ϕ 29 packaged genome and elucidation of its packaging process. *Virology* **371**, 267–277 (2008).
26. Catalano, C. E. ed *Viral Genome Packaging Machines: Genetics, Structure, and Mechanism* (Landes Bioscience, 2005).
27. Berger, M. A. in *Lectures on Topological Fluid Mechanics* (ed. Ricca, R. L.) 75–98 (Springer-Verlag, 2009).
28. Lal, T. M., Sano, M., Hatai, K. & Ransangan, J. Complete genome sequence of a giant Vibrio phage ValKK3 infecting Vibrio alginolyticus. *Genomics Data* **8**, 37–38 (2016).
29. Irobalieva, R. N. *et al.* Structural diversity of supercoiled DNA. *Nat. Commun.* **6**, 8440-1–8440-10 (2015).
30. Li, W. *et al.* Complex kinetics of DNA condensation revealed through DNA twist tracing. *Phys. Rev. E* **92**, 022707-1–022707-8 (2015).
31. Kulić, I. M., Andrienko, D. & Deserno, M. Twist-bend instability for toroidal DNA condensates. *Europhys. Lett.* **67**, 418–424 (2004).

32. Lepault, J., Dubochet, J., Baschong, W. & Kellenberger, E. Organization of double-stranded DNA in bacteriophages: a study by cryo-electron microscopy of vitrified samples. *The EMBO Journal* **6,** 1507–1512 (1987).

33. Black, L. W., Newcomb, W. W., Boring, J. W. & Brown J. C. Ion etching of bacteriophage T4: Support for a spiral-fold model of packaged DNA. *Proc. Natl. Acad. Sci. USA* **82,** 7960–7964 (1985).

34. Ashton, T., Cantarella, J., Piatek, M. & Rawdon, E. Knot Tightening by Constrained Gradient Descent. *Exp. Math.* **20,** 57–90 (2011). The datafiles of configurations of tight knots can be downloaded from the web page of J. Cantarella, www.jasoncantarella.com (Accessed: 2nd June 2016).

35. Baranska, J., Pieranski, P. & Przybyl, S. Length of the tightest trefoil knot. *Phys. Rev. E* **70,** 051810-1–051810-9 (2004).

Additional Information