

Ralf Metzler · Paul G. Dommersnes

Helical packaging of semiflexible polymers in bacteriophages

Received: 29 July 2003 / Revised: 3 November 2003 / Accepted: 15 December 2003 / Published online: 26 February 2004
© EBSA 2004

Abstract We investigate multilayered helical packaging of double-stranded DNA, or of a general polymer chain with persistence length l_b , into an ideal, inert cylindrical container, reaching densities slightly below close packaging. We calculate the free energy as a function of the packaged length, based on the energies for bending, twisting, the suffered entropy loss, and the electrostatic energy in a Debye–Hückel model. In the absence of charges on the packaged polymer, a critical packaging force can be determined, similar to the mechanism involved in DNA unzipping models. When charges are taken into consideration, in the final packaging state the charges which are chemically distant become geometrically close, and therefore a steep rise is seen in the free energy. We argue that due to the extremely ordered and almost closely packaged final state the actual packaging geometry does not influence the behaviour of the free energy, pointing towards a certain universality of this state of the polymer. Our findings are compared to a recent simulations study, showing that the model is sensitive to the screening length.

Keywords Bacteriophage λ · DNA and RNA · DNA packaging · Phage capsid

Introduction

The packaging of viral DNA into a small container, the capsid (Fig. 1), is a vital step in the assembly of phages, bacterial viruses, in whose lifecycle the capsid serves to transport the DNA from a lytic to a new host cell (Alberts et al. 1994; Catalano et al. 1995; Fujisawa and

Morita 1997; Snustad and Simmons 2003). The packaging corresponds to a highly optimized process, during which the viral DNA, which for many phages measures of the order of 5–10 μm , is squeezed into the capsid by a motor protein at forces of up to ~ 60 pN, such that pressures are built up inside the capsid hull which may reach values of 6 MPa (Catalano et al. 1995; Fujisawa and Morita 1997; Smith et al. 2001). Recently, considerable advances in the understanding of both the biochemical and biophysical processes involved in packaging have been achieved. Thus, the structure of the capsid container (see, for instance, Venkataram Passad et al. 1999), of the packaged DNA (Cerritelli et al. 1997; Hud and Downing 2001), as well as the packaging motor (Simpson et al. 2000), have been determined, and careful single molecule studies on bacteriophage $\phi 29$ capsids have revealed the biomechanical details of the packaging process, such as packaging rates and forces (Smith et al. 2001).

From a physics or engineering perspective, the biomechanical orders of magnitude involved in the packaging process are quite extraordinary. Thus, apart from the aforementioned pressure created in the capsid, the radius of this container is comparable to, or less than, the persistence length of the DNA chain to be packaged (Catalano et al. 1995; Hud 1995; Fujisawa and Morita 1997; Hud and Downing 2001; Smith et al. 2001). Therefore, fluctuation-based undulations are suppressed, and the chain can be approximately thought of as being wound up helically like thread on a bobbin (Fig. 1), or like a ball of yarn. Ultimately, a relatively highly ordered 3D configuration of the DNA inside the capsid is achieved, which under certain conditions may even lead to local crystallization of the DNA (Catalano et al. 1995; Hud 1995; Cerritelli et al. 1997; Fujisawa and Morita 1997; Kindt et al. 2001). It is generally argued that this ordered arrangement helps to avoid the creation of entanglements or even knots of the wound-up DNA, thus enabling easy ejection, i.e. release of the DNA once the phage docks to a new host cell; this ejection is not assisted by the packaging motor, but it

R. Metzler (✉)
NORDITA, Blegdamsvej 17, 2100 Copenhagen Ø, Denmark
E-mail: metz@nordita.dk

P. G. Dommersnes
Institut Curie, 11 rue Pierre et Marie Curie, 75231 Cedex 5 Paris, France

can be facilitated by host cellular DNA polymerase, which starts to transcribe the DNA and thereby pulls it out of the capsid (Alberts et al. 1994; Kindt et al. 2001; Arsuaga et al. 2002a; Snustad and Simmons 2003).

Based on the tuned ratio between capsid size and persistence length, due to which a highly ordered packaging is made possible, a number of analytical and numerical models for the packaging process for phage λ or similar phages with comparable properties have been proposed, such as a simple helical structure like the one shown in Fig. 1 (Riemer and Bloomfield 1978), a ball of yarn-type geometry (Richards et al. 1973), as well as hexagonal (Odijk 1998) and icosahedral packaging (Rudnick and Bruinsma 2003). In a detailed study of packaging and ejection under attractive and repulsive conditions (Kindt et al. 2001), it was shown that the typical force versus packaging ratio curves found experimentally (Smith et al. 2001) could be reproduced qualitatively. This qualitative agreement also supports the view that at high packaging densities the actual geometry of the packaged chain becomes irrelevant, indicating a universality of the free energy as a function of the packaging ratio $p \equiv l/L$ as it tends to 1 (here, l denotes the portion of the chain of total length L , which is already packaged into the capsid hull). Apart from these equilibrium or Brownian simulations studies, first steps towards a tilted ratchet modelling of the packaging kinetics have been proposed (Santo and Sebastian 2002). It should be noted that the suppression of entanglements and knots is much less efficient in mutant, tailless phages, in which DNA cyclization is not suppressed (Arsuaga et al. 2002b).

In what follows, we consider the general question of packaging a semiflexible polymer chain. To this end, we estimate the arising bending and twisting energies as well

as the entropy loss due to arranging the chain in a well-ordered fashion. These estimates are then compared to the electrostatic energy, which has to be overcome when the charged DNA is packaged. As it turns out that for higher packaging ratios this latter part is dominating, the actual configuration of the packaged DNA is of lesser significance. It is shown that the onset of a significant contribution from the electrostatic energy is close to packaging ratios p of 40–50%, and thus consistent with measured values. We also discuss the critical force needed to overcome the contributions for bending, twisting and entropy compensation. This latter case corresponds to the initially dominating contributions in the DNA packaging; however, it pertains to the packaging of electrically neutral chains with persistence length in more general terms, and is therefore of interest from a general polymer physics point of view, as well as for potential nanotechnological applications, such as drug delivery. The simplistic approach chosen in the following allows us to highlight the special physical features of this system.

Geometrical considerations

Before proceeding to the free energy computation, let us start by addressing some purely geometrical aspects of the packaging process. The shape of the capsid considered by Smith et al. (2001) is approximately cylindrical, and it possesses outer dimensions of $a_0 \approx 21$ nm times $b_0 \approx 54$ nm for radius and height (compare Fig. 1). We are going to employ these numbers in the following. Given a_0 and b_0 , the available (inner) volume is also cylindrical, with radius $a_{in} \approx 17$ –18 nm and height $b_{in} \approx 46$ nm, to accommodate a thickness and a certain ruggedness of the capsid hull (Wikoff et al. 2000). For the outer helix to be packaged, we choose the radius $a_1 = 16.5$ nm, due to the thickness of the DNA chain itself, and for the effective height we choose $b = 43$ nm (see below). Note that all relevant length scales are summarized in Table 1.

Double-stranded DNA (dsDNA) has a geometrical diameter of $d_g \approx 2$ nm, which is much less than the persistence length $l_b \approx 53$ nm, and for which dsDNA possesses a very small order parameter d_g/l_b , implying small excluded volume effects, i.e. dsDNA behaves almost like an ideal polymer (De Gennes 1979; Frank-Kamenetskii 1997). Owing to the existence of counterions, which form a layer around the DNA and whose concentration determines the screening length of the DNA, an effective diameter d_{eff} has to be assigned to the chain; usually, d_{eff} is larger than d_g (Frank-Kamenetskii 1997). However, when we compare the volume inside the capsid hull, which is $V_c \approx 4.6 \times 10^{-23}$ m³, to the volume occupied by the DNA of length $L = 6.6$ μ m [the DNA size occurring in Smith et al. (2001)], $V_{DNA-rod} \approx \frac{1}{4} \pi d_{eff}^2 L$, in a rod-like configuration, we obtain the estimate $d_{eff} \approx 3.0$ nm. Therefore, in the fully packaged state the DNA fills out the major portion of the accessible volume inside the

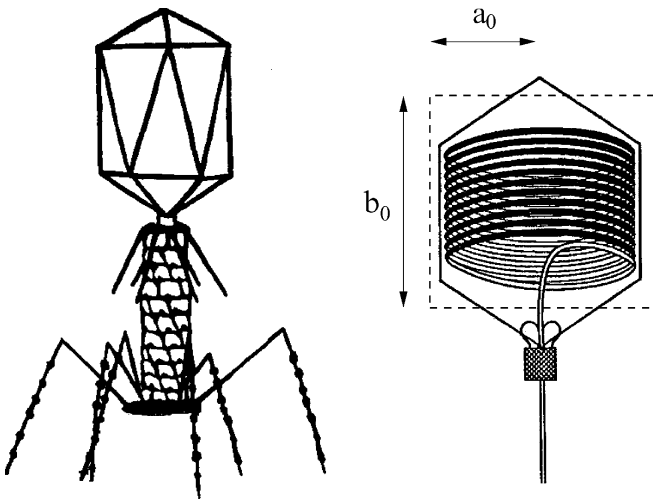


Fig. 1 *Left:* sketch of bacteriophage $\phi 29$, with the capsid sitting on the injection tube with attachment legs to the host cell wall. *Right:* stylized packaging of a DNA chain into a phage capsid through a motor consisting of an array of proteins. The chain is packaged into an almost crystal-like high-pressure configuration. In the model, the capsid is approximated by a cylinder of radius $a_0 \approx 21$ nm and height $b_0 \approx 54$ nm

Table 1 Relevant length scales of phage capsid and dsDNA

a_1	16.5 nm	Free radius of capsid ^a
b	43 nm	Free height of capsid
L	6.6 μm	Total DNA length
d_g	2 nm	Geometric (minimal) dsDNA diameter
d_{eff}	2.5 nm	Assumed effective dsDNA diameter ^b
l_b	53 nm	Bending persistence length
l_t	109 nm	Twist persistence length
λ	0.57/1 nm ^c	Screening length
l_B	0.7 nm	Bjerrum length

^aInner radius a_{in} minus $d_{\text{eff}}/2$ ^bCompare Frank-Kamenetskii (1997)^c1 nm at normal salt conditions; 0.57 nm is the effective value inside the capsid

capsid hull and the packaged configuration is expected to be highly ordered, and the effective diameter acquires a value lower than 3.0 nm (compare also Riemer and Bloomfield 1978; Kindt et al. 2001). We find that $d_{\text{eff}} \approx 2.5$ nm appears to be a reasonable value to model the experimental findings from Smith et al. (2001).

With $d_{\text{eff}} \approx 2.5$ nm, we can accommodate a maximum number $n_{\text{max}} \approx b/d_{\text{eff}} \approx 17.2$ of layers in the helix with radius 16.5 nm, corresponding to a packaged length of $l_1 \approx 2\pi a_1^2 n_{\text{max}} \approx 1.8 \mu\text{m}$. A monolayered helical structure therefore cannot account for the entire packaged length $L = 6.6 \mu\text{m}$. We therefore pursue the following scenario: at early stages of DNA injection through the packaging motor location, the DNA will be thought of as forming a single helix with a low number of windings. Subsequently, on further packaging, when the influence of charges, which are distant in respect to the chemical distance along the DNA chain but close geometrically, becomes appreciable, it will split up into a number of intertwined helices with different radii, i.e. successive layers will vary in their radii, to guarantee the maximum possible geometric distance between individual charges. We simplify this configuration by distributing the packaged length l into a number of separate, disconnected helices, inside one another and with an offset of half a winding to maximize the distance between the charges (see Fig. 2 for two nested helices). Thus, we assume a self-organizing, successive rearrangement of an initial single helix with a smaller number of windings towards a few nested helices. The exact turnover point from single to multiple nested helices, as will be argued, is of lesser importance. For a smaller number of windings per helix, the individual distances are still longer than the electrostatic screening length, and the influence of the different radii on the bending energy is appreciably small. In contrast, for high packaging ratios p , the nested helices structure enables both packaging of the entire chain length and a maximum possible distance between individual segments of the DNA. In the final stage of packaging, $p \rightarrow 1$, close packaging is approached, and therefore the actual geometry becomes irrelevant. The nested helices scenario therefore seems to be a reasonable model for the packaging geometry, and is in fact similar to the spooling structure found from

Brownian dynamics simulations by Kindt et al. (2001); compare also Richards et al. (1973).

The two stages of packaging can therefore be described as follows:

1. Mono-helical arrangement. In this case, we can package up to $1.8 \mu\text{m}$ in a helix of radius 16.5 nm. Although we expect a turnover to the multi-helical geometry before this packaging length is reached, we will use it to have at hand a control for the involved parameters. To be more specific, we assume that one chain end enters the cylinder through a hole in the centre of the cylinder face, corresponding to the location of the motor of the phage capsid. After the length l of the chain inside the container approaches the vertical extension b , it starts to bend and subsequently curls up helically alongside the hull. For $l > b$, the resulting curve $\mathbf{r}$ described by the chain can then be parametrized through $\mathbf{r}(s) = \{a_1 \cos(s/c), a_1 \sin(s/c), \xi s\}^T$, where we neglect the shortening portion from the motor to the upper end of the helix. This assumption is justified from more realistic shapes of capsid hulls. Here $\xi = b/l$, $c = a/\sqrt{1 - b^2/l^2}$ and the arc length s , with $0 \leq s \leq l$, parametrizes the space-curve $\mathbf{r}(s)$.
2. Multi-helix arrangement. This is the version providing enough storage capacity for the entire DNA of $L = 6.6 \mu\text{m}$ to be packaged. Assuming that in the fully packaged state the nested helices will just touch, i.e. the individual space-curves $\mathbf{r}_i(s)$ will be the effective diameter 2.5 nm apart from each other at the closest points, we obtain the difference in radii: $\Delta a_i \approx \sqrt{2.5^2 - 1.25^2} \text{ nm} \approx 2.2 \text{ nm}$. At the same time, we assign an offset along the symmetry axis z of the cylinder between successive helices, which equal half of the z -distance between points of equal phase within one helix. This offset ensures the maximum separation of successive helices from each other. For the first helix, we therefore have the parametrization:

$$\mathbf{r}_1(s) = \begin{pmatrix} a_1 \cos(s/c_1) \\ a_1 \sin(s/c_1) \\ bs/l_1 \end{pmatrix} \quad (1)$$

and for the second one, we use:

$$\mathbf{r}_2(s) = \begin{pmatrix} a_1 \cos(s/c_2) \\ a_1 \sin(s/c_2) \\ bs/l_2 + (\Delta z)_2 \end{pmatrix} \quad (2)$$

with $(\Delta z)_2 \approx \pi a_1 / \sqrt{l_1^2/b^2 - 1}$. We also find that if the length l_1 is stored in the first helix, the second one will contain:

$$l_2 \approx \sqrt{\frac{a_2^2}{a_1^2} l_1^2 + b^2 \left(1 - \frac{a_2^2}{a_1^2}\right)} \approx \frac{a_2}{a_1} l_1 \quad (3)$$

Fig. 2 *Upper*: two nested helices with offset of one half winding height along the symmetry axes such that the maximum distance between the two helices is achieved. *Lower*: top (perspective) and side views

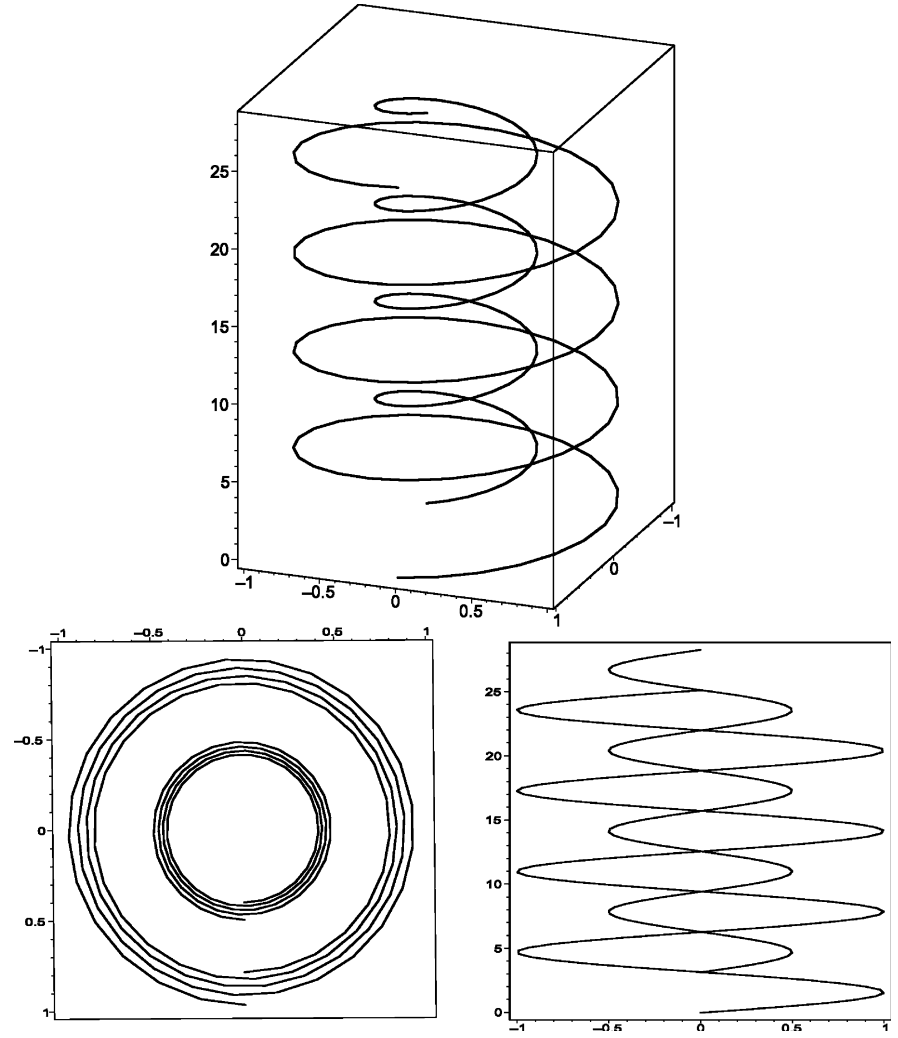


Table 2 Parameters for the five nested helices. The maximum overall length to be packaged in this arrangement is $l^{\max} = \sum l_i^{\max} = 6.6 \mu\text{m}$

i	a_i (nm)	$l_i^{\max}$ (nm)	$(\Delta z)_i$
1	16.5	1800	0
2	14.3	1560	$\pi a_1 / \sqrt{l_1^2/b^2 - 1}$
3	12.1	1320	0
4	9.9	1080	$\pi a_3 / \sqrt{l_3^2/b^2 - 1}$
5	7.7	840	0

where the last approximation becomes increasingly better in the course of the packaging process. Figure 2 shows a sketch of two nested helices with an offset along the symmetry axis of the helix. In order to package the required $L = 6.6 \mu\text{m}$, we find that five helices are needed whose parameters are summarized in Table 2. This result is also reasonable from the point of view that the sixth helix would already have a very small radius of 5.5 nm, which might already be close to the instability of dsDNA under bending. The “unused” volume left in the centre of the capsid hull amounts to $V_{\text{rest}} \approx 5.6 \times 10^{-24} \text{ m}^3$,

or 12% of the entire accessible capsid volume V_c . We note that the value of 2.5 nm for d_{eff} chosen above is consistent with the minimum distance reported by Kindt et al. (2001). In establishing the five helices in our self-organizing scenario, we envisage a turnover from one to multiple helices between 20–40% of packaging; see below.

Bend, twist and entropy loss

Let us at first regard the packaging problem for the case of an electrically neutral chain with (bending) persistence length l_b and an analogous quantity l_t for the twist. The associated free energy, $\mathcal{F}(l)$, consists of two contributions, the entropy loss, and the bending and twisting energies. We assume that bending and twisting are not coupled, and that twist, which is potentially induced by the packaging motor, can relax through the free end of the packaged DNA in the capsid. Moreover, internal bending and twisting due to the stacking of base pairs is neglected in our homogeneous dsDNA model. Ionic conditions in the capsid are not known in detail, and

therefore the standard salt conditions are assumed for the electrostatic energy below. Finally, we assume an inextensible chain model for the DNA.

The part of the effective Hamiltonian corresponding to the bending can be obtained through integration of the local curvature, i.e.:

$$\mathcal{H}_{\text{bend}}^i = \frac{\kappa}{2} \int_0^{l_i} ds \left(\frac{d^2 \mathbf{r}_i(s)}{ds^2} \right)^2 = \frac{\kappa}{2a_i^2} (1 - b^2/l_i^2)^2 l_i \quad (4)$$

for one helix i , where $\kappa = k_B T l_b$ is the bending modulus (Moroz and Nelson 1998; compare also Marko and Siggia 1995; Bouchiat and Mézard 1998), and the corresponding length scale is just the (bending) persistence length l_b , for which we take $l_b \approx 53$ nm (Smith et al. 2001), similar to the value of 49 nm given in Moroz and Nelson (1998). Already, for a relatively loosely packaged state, b^2 ($b = 43$ nm) becomes small in comparison to l_i^2 (which approach the $l_i^{\text{max}} \approx 1 \mu\text{m}$ listed in Table 2), and thus we can approximate the bending Hamiltonian through $\mathcal{H}_{\text{bend}}^i \approx \kappa l_i / (2a_i^2)$. The overall bending energy for the five nested helices is therefore:

$$\mathcal{H}_{\text{bend}}^i \approx \frac{\kappa l}{2} \frac{\sum a_i^{-1}}{\sum a_i} \approx \frac{3.7}{1000 \text{ nm}^2} \kappa l \quad (5)$$

The corresponding force $d\mathcal{H}_{\text{bend}}/dl$ to bend the DNA double helix is approximately constant at higher p , with a value of $F_{\text{bend}} \approx 0.8$ pN.

Similarly, the twist energy (solely due to the helical geometry, see above):

$$\mathcal{H}_{\text{twist}}^i = \frac{\rho}{2} \int_0^{l_i} ds \left(\mathbf{t} \times \mathbf{n} \frac{d\mathbf{n}}{ds} \right)^2 = \frac{\rho b^2}{2a_i^2} \left(1 - \frac{b^2}{l_i^2} \right) \frac{1}{l_i} \quad (6)$$

contributes to the effective Hamiltonian, where the twisting modulus is $\rho = k_B T l_t$ with $l_t \approx 109$ nm (Moroz and Nelson 1998). Here, $\mathbf{t}_i(s) = d\mathbf{r}_i(s)/ds$ is the normalized tangent vector, and $\mathbf{n}_i(s) = R_i d^2 \mathbf{r}_i(s)/ds^2$ is the normal vector, normalized via the radius of curvature, $R_i = a_i / (1 - b^2/l_i^2)$. The twist energy $\mathcal{H}_{\text{twist}}^i$ exhibits a maximum at $l_i = \sqrt{3}b$, corresponding to an initial packaging stage, and it then decays like $1/l_i$, and therefore $\mathcal{H}_{\text{twist}}^i \approx \rho b^2 / (2a_i^2 l_i)$. From all five helices, the cumulative twist energy becomes:

$$\mathcal{H}_{\text{twist}} \approx \frac{\rho b^2}{2l} \sum a_i^{-3} \sum a_i \approx 243 \frac{\rho}{l} \quad (7)$$

valid for larger values of l . For the entire range of packaging, the force to overcome the twist is negligible in comparison to the bending energy.

The entropy loss suffered during packaging can be estimated through the number of configurations in the unpackaged state and the assumption that the entropy in the packaged state is vanishingly small. The number of degrees of freedom of a linear chain of length $L-l$ (the not yet packaged portion), one end of which is attached to a wall, scales like $\omega \approx \mu^{(L-l)/(2l_b)} ([L-l]/[2l_b])^{\bar{\gamma}-1}$ measured in Kuhn lengths of order $\sim 2l_b$ (Frank-

Kamenetskii 1997). Here, $\bar{\gamma} \approx 0.69$ is the 3D configuration exponent for self-avoiding walks, one end of which is grafted (Eisenriegler 1993), and μ is the non-universal connectivity constant, which for the simple cubic lattice in 3D is $\mu \approx 4.69$ (De Gennes 1979); despite this non-universality of the μ value, we are going to adopt it for the following estimates. From $S = k_B \log \omega$, we then obtain:

$$S \approx k_B \left(\frac{L-l}{2l_b} \log \mu + (\bar{\gamma} - 1) \log \frac{L-l}{2l_b} \right) \quad (8)$$

which up to a logarithmic correction, that we are going to neglect, is proportional to the packaged length, $S \approx S_0 - 0.77 k_B (l/l_b)$, with $S_0 \approx 26.2 k_B$.

The total free energy $\mathcal{F} \approx \mathcal{H}_{\text{bend}} - TS$ for the packaging of an uncharged, semi-flexible chain therefore becomes:

$$\mathcal{F} \approx \left(\frac{3.7\kappa}{1000 \text{ nm}^2} + 0.77 \frac{k_B T}{l_b} \right) l - TS_0 \quad (9)$$

which is linear in the packaged length l . From this form for $\mathcal{F} = \mathcal{F}(l)$ we can determine the constant packaging force for the uncharged chain with persistence length $l_b \approx 53$ nm to be:

$$F_p^0 = \frac{d\mathcal{F}(l)}{dl} \approx \left(\frac{3.7l_b}{1000 \text{ nm}^2} + \frac{0.77}{l_b} \right) k_B T \approx 0.87 \text{ pN} \quad (10)$$

at $T = 300$ K, where the force contribution due to entropy loss is less than 10% of the bending value. Thus, in comparison to the orders of magnitude of the packaging force caused by the electrical charges calculated in the next section and measured in experiments, the mechanical and entropic contributions can be neglected. However, as they are relevant both for a charged chain towards a more advanced stage of ejection, when electrostatic repulsion becomes weak, and for an uncharged chain, we pursue the existence of an l -independent packaging force in the electrically neutral case somewhat further, before addressing the electrostatic problem in the next section.

The property $\mathcal{F}(l) \approx F_0 l$ gives rise to the existence of a critical packaging force F_c , in accordance with the findings from the modelling of DNA unzipping (Lubensky and Nelson 2000). There, owing to the linearity of the free energy in the number of unzipped monomers, an applied force F that comes close to the critical force F_c leads to the divergent behaviour $\langle m \rangle \approx (F_c - F)^{-1}$ of the mean number $\langle m \rangle$ of unzipped monomers, i.e. one observes a phase transition at F_c . In the case of DNA packaging, this translates to a divergence of the mean packaged length:

$$\langle l \rangle \approx (F_c - F)^{-1} \quad (11)$$

with critical force $F_c = F_0$, if we regard the system infinite. For a finite system, this packaging transition means that beyond the critical force, a constant motor force stores the entire chain in the container. We note that the

divergence at F_c due to the physical realizations is actually not a momentary one: in unzipping, the external force has to be adjusted within the single molecule setup with a finite response time, and in packaging the packaging motor limits the packaging speed. However, the statement (11) implies that if only a constant force F_c is provided, the entire chain will be processed.

The divergence at F_c becomes modified in the presence of disorder, e.g. provided by different DNA sequences in unzipping (Lubensky and Nelson 2000). A plausible cause for such disorder in chain packaging is the existence of fluctuations of the bending modulus along the chain, or fluctuations in the motor strength. In the former case, for instance, the modulus is expected to behave like $\kappa + \delta\kappa(s)$. This can be incorporated through a Langevin equation of the form $d\mathcal{F} = (F - F_c)dl + \Gamma(l)d$ where $\Gamma(l)$ represents a stochastic force. It is reasonable to assume that, above some given distance, the noise will be of the white Gaussian type with correlation function $\overline{\Gamma(l)\Gamma(l')} = K\delta(l - l')$. It can then be shown that the Langevin equation for $d\mathcal{F}(l)$ corresponds to a bivariate Fokker–Planck equation, from which it is possible to extract the disorder-averaged packaging length (Risken 1989; Lubensky and Nelson 2000):

$$\overline{\langle l \rangle} \approx (F_c - F)^{-2} \quad (12)$$

This is a remarkable finding: in the presence of Gaussian fluctuations of the modulus (and/or the motor power), the packaging transition is drastically enhanced.

Electrostatic free energy contribution

The above considerations are valid for the packaging of a neutral polymer chain with persistence length l_b , or for a charged chain for packaging ratios at which the influence of the electrostatic interaction is still negligible. We now address the influence of electrical charges distributed along the chain, leading to effects that were observed in the single molecule studies reported by Smith et al. (2001), as well as in the simulations study (Kindt et al. 2001). We note that in the presence of charges, an attractive force can come about when the distance between chains is small enough; see the discussion in Kindt et al. (2001) and references therein. This effect may give rise to further stabilization of the packaged chain, but we neglect this possibility as well as interactions with the capsid hull, which we consider chemically and electrostatically inert. The influence of the solvent is taken into account through the existence of a screening length in a Debye–Hückel approach according to which the electrostatic energy between two charges at a distance r falls off like $\propto e^{-r/\lambda}/r$; compare, for instance, Netz (2001) and Schiessel and Aranda-Espinoza (2001). More precisely, on a DNA strand, individual charges are an average distance $\delta \approx 0.17$ nm apart from each other and the screening length under standard conditions is $\lambda \approx 1$ nm. There is insufficient knowledge

about the counterion concentration inside the capsid, such that inside this capsid the screening length may differ from this value. Outside the capsid, we neglect charge–charge interactions.

The electrostatic energy within the Poisson–Boltzmann approximation is:

$$\mathcal{E}_{\text{el}} = \frac{k_B T l_B \tau^2}{2} \oint ds \oint ds' \frac{\exp(-|r(s) - r(s')|(\lambda))}{|r(s) - r(s')|} \quad (13)$$

where the line charge density $\tau = 1/\delta = 2/(0.34 \text{ nm})$ is used (Kunze and Netz 2002), corresponding to the rule of thumb that it corresponds to roughly 10% of its bare value (Grønbech-Jensen et al. 1997). Moreover $l_B = e^2/(4\pi\epsilon k_B T)$ is the Bjerrum length, which in water at room temperature is of size $l_B \approx 0.7$ nm. Finally, the screening length λ is given through $\lambda = 1/\sqrt{8\pi l_B c_s}$ in dependence on the salt concentration c_s ; a typical value at cellular salt conditions is $\lambda \approx 1.0$ nm (Netz 2001); see also below.

Before calculating the electrostatic energy of the multilayered DNA-helix configuration in the finite capsid, it is instructive to look at the simpler case of monolayered, helically packaged DNA in an infinite cylinder. Its space-curve is given by $r(s) = \{a \cos(s/c), a \sin(s/c), \zeta s\}^T$ with $\zeta = \sqrt{1 - (a/c)^2}$. The increase in electrostatic energy due to packaging per unit length after substitution to the relative coordinate, $S = s - s'$, and with $\sin^2(z/2) = (1 - \cos z)/2$, then becomes:

$$\frac{\Delta \mathcal{E}_{\text{el}}}{\kappa_B T L} = 2 l_B \tau^2 \int_{\epsilon}^{\infty} dS \left(\frac{\exp(-D(S)/\lambda)}{D(S)} - \frac{\exp(-S/\lambda)}{S} \right) \quad (14)$$

where $D(S)^2 = \zeta^2 S^2 + 4a^2 \sin^2(S/2c)$ and ϵ is a lower cut-off. If the helix is not too densely packaged and the screening length is short, i.e. if $D(S=c) \gg \lambda$, the main contribution in energy comes from the region $S < c$. Thus, for $S/c < 1$, $D(S) \approx S \{1 - a^2 S^2/(24c^4)\}$. Integration up to the upper limit Sc and expansion in orders of ϵ and λ produces, to leading order:

$$\frac{\Delta \mathcal{E}_{\text{el}}}{\kappa_B T L \tau^2 l_B} \approx \frac{a^2 \lambda^2}{8c^4} \left(1 + \frac{\epsilon}{\lambda} + \frac{1}{3} \left(\frac{\epsilon}{\lambda} \right)^2 \right) \exp(-\epsilon/\lambda) \quad (15)$$

For a cylinder of given height H and radius a , $(a/c)^2 = 1 - (H/L)^2$, where L is the length of packaged DNA. If $(H/L) < 1$, the prefactor from above expression behaves like $a^1 \lambda^2/(8c^4) \approx \lambda^2/(8a^2)$. Expression (15) can then in fact be considered as the electrostatic contribution to the bending energy at low packaging ratios, the corresponding force for $a = 16.5$ nm being $F_p^{\text{el}} \approx 0.27$ pN, which for $\epsilon \approx \delta$, and the other parameters chosen as above, compares well with the magnitude of the bending force. This is demonstrated in Fig. 3, in which we compare for $a = 16.5$ nm, $\lambda = 1$ nm and $\epsilon = 0.1$ nm the exact, numerically evaluated electrostatic energy with the approximation (3) as a function of the periodicity c , showing excellent agreement for values above 16.6 nm.

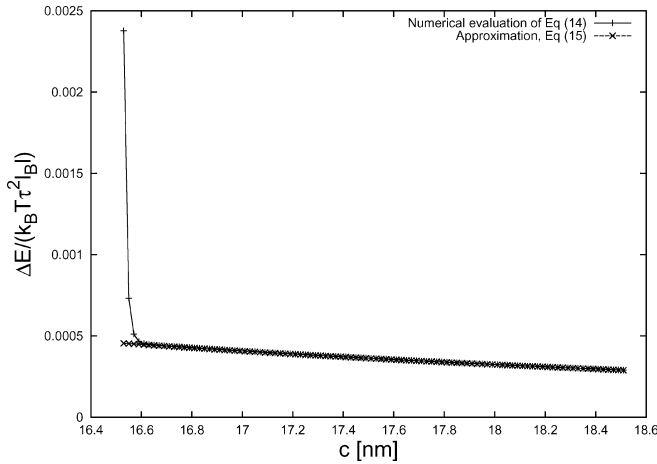


Fig. 3 Approximation (electro) of the electrostatic energy difference $\Delta\epsilon_{el}$ in comparison to numerical evaluation of equation (14) for the parameters $a = 16.5$ nm, $\lambda = 1$ nm and $\epsilon = 0.1$ nm, showing excellent agreement above periodicity $c \approx 16.6$ nm

For higher packaging ratios, we retreat to numerical evaluation of the electrostatic energy.

Let us come back to the packaging into the multihelical arrangement defined in the Geometrical considerations section (above). With the explicit parametrization $\mathbf{r}(s)$ of the space-curve described by the packaged chain, we can numerically integrate the electrostatic interactions for the monolayered helix:

$$\mathcal{E}_{el} = k_B T l_B \tau^2 \int_0^L ds \int_s^L ds' \frac{\exp(-|\mathbf{r}(s) - \mathbf{r}(s')|/\lambda)}{|\mathbf{r}(s) - \mathbf{r}(s')|} \quad (16)$$

Instead of this continuum form, however, we employ the discretized version:

$$\mathcal{E}_{el} = \frac{k_B T l_B}{2} \sum_{i,j:i \neq j}^{\text{int}(l/\delta)} \frac{\exp(-|\mathbf{r}(s_i) - \mathbf{r}(s_j)|/\lambda)}{|\mathbf{r}(s_i) - \mathbf{r}(s_j)|} \quad (17)$$

in which $s_i = i\delta$, by which choice we eliminate the prefactor τ . Note that given the parametrization $\mathbf{r}(s)$ of the space-curve(s), the functional behaviour of $\mathcal{E}_{el}$ up to a constant prefactor, depends solely on the screening length λ . The summation in expression (17) is carried out up to the number of charges we can put on the DNA helix, $n_{\max} \approx \text{int}(l/\delta)$, where $\text{int}(\cdot)$ denotes the integer value. In the case of the multiple nested helices, the summation extends over all individual helices, and the summation runs up to $n_{\max}$ for each individual helix.

In Fig. 4 we show the result for the electrostatic energy $\mathcal{E}_{el}(l)$ for a monolayered helix, normalized to the packaged length l , for three different screening lengths. The fully packaged length in this example corresponds to $L = 1.8$ μm , i.e. to the maximum packaging for the mono-helical geometry with $d_{\text{eff}} \approx 2.5$ nm. For the shorter screening length $\lambda = 0.5$ nm (lowest curve), the value of $\mathcal{E}_{el}(l)/l$ is almost constant over the entire packaging range, whereas it increases rather quickly for the larger value $\lambda = 2$ nm, due to which the interaction

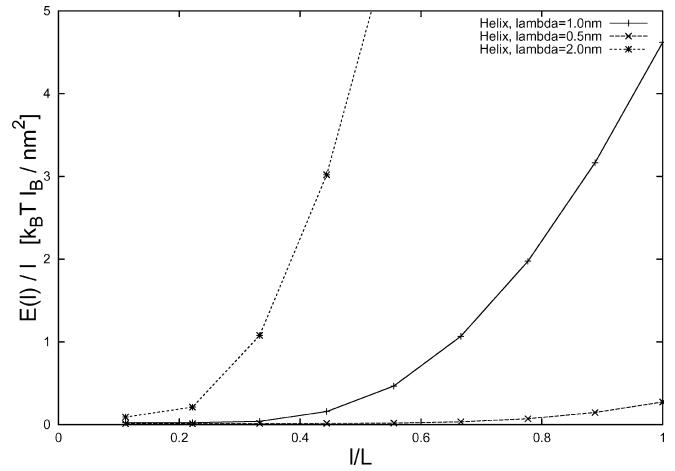


Fig. 4 Electrostatic energy $\epsilon_{el}(l)/l$ in units of $l_B k_B T / m^2$ as a function of the packaging density $p = l/L$, for $L = 1.8$ μm . The screening length λ is varied for the three graphs from 0.5 to 2 nm

between chemically remote charges is non-negligible already for rather low packaging densities. For the typically chosen value $\lambda = 1.0$ nm, the significant increase in the normalized electrostatic energy starts at packaging densities $p \approx 40\text{--}50\%$, and is therefore in agreement with the behaviour reported (Kindt et al. 2001; Smith et al. 2001).

To be able to accommodate the full $L = 6.6$ μm of dsDNA in the experiments of Smith et al. (2001), we suggested the multi-layered arrangement of five nested helices with the parameters as given in Table 2. The corresponding packaging energy is plotted in Fig. 5 for $\lambda = 0.57$ nm. Clearly, the turnover from the plateau-like regime to the strong increase occurs at around 40–50%, as required for our data. From the slope at $p = 1$, we can determine the maximum force needed to package the final portion of the chain: $F_{\max} \approx \Delta E / \Delta l$, so that we find $F_{\max} \approx 59$ pN, which is close to the experimentally observed value of 57 pN (Smith et al. 2001). This supports the view that the packaging force is indeed dominated by electrostatic repulsion even close to maximal packaging. In comparison, the maximum packaging force corresponding to Fig. 5 with $\lambda = 1.0$ nm is $F_{\max} \approx 248$ pN, and therefore overshoots the measured value by far, documenting the delicate sensitivity of the model on the screening length λ . Similarly, for $\lambda = 0.5$ nm, the value is already down to 40 pN.

This finding implies that, in the experimental system, the effective screening length appears smaller than the value of 1 nm generally used in experiments under physiological conditions. This may either be due to a higher counterion concentration within the capsid, to attractive interactions between chain and wall, or to reduced charge density τ due to enhanced counterion condensation on the DNA. As the latter, however, might counteract proper ejection once the phage docks to a new host cell, the low value for the effective λ seems a fair indication in favour of increased counterion presence, in comparison to the bulk value. Owing to the fact

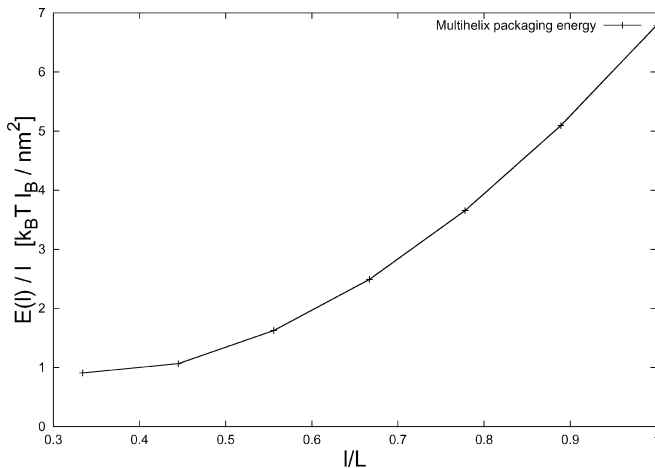


Fig. 5 Electrostatic energy $l_{el}/(\sum l_i)$ in units of the Bjerrum length l_B and $k_B T$, versus packaging density p for $\lambda = 0.57$ nm and $L = 6.6$ μm . A significant increase of $l_{el}/(\sum l_i)$ occurs at around 40–50%. The slope at $p \rightarrow 1$ corresponds to a maximum packaging force of approximately 59 pN, close to the experimentally measured value

that both the onset of a larger slope and the magnitude of the final slope (i.e., the maximum force) match the observed behaviour, together they seem good indicators for estimating the screening length. The analysis demonstrates that the system, due to the high packaging densities reached, is very sensitive to small variations of the screening length.

Conclusions

We have studied analytically a scenario for the packaging of a macromolecule with persistence length and charges into a small volume such that almost close packaging is reached. Based on some fundamental geometrical considerations, according to which the packaged chain occupies the major portion of the accessible volume, the effective diameter of the packaged dsDNA is necessarily small, and the packaged geometry is expected to be highly ordered. The latter is strengthened by the fact that the persistence length of dsDNA (53 nm) is of the order of half the circumference of the packaging cylinder, i.e. that fluctuation-based undulations of the chain become less and less likely when packaging proceeds. For these reasons, we chose a nested-helices arrangement for the packaging geometry, a fairly easy geometry to treat analytically in terms of the space-curve $r(s)$ described by the packaged chain. The different radii of the nested helices can be viewed as some kind of a self-organized structure, either due to layered packaging or as emanating from a single helix on increasing compression; in both cases, we neglect those portions of the chain which connect between the individual helices. Eventually, when the system approaches close packaging, the actual geometry becomes necessarily less and less important, until an approximately universal packaging energy is reached.

Whereas for uncharged polymers (and also for charged chains at lower values of the packaging ratio p) it is possible to determine a critical packaging force F_c , given the result that the packaging free energy grows linearly in the packaged length l , the packaging force grows significantly faster for charged chains when chemically distant charges come geometrically close, beyond the screening length λ . Conversely, on ejection, it is this critical (constant) force, dominated by the bending energy, which facilitates ejection below the packaging ratio at which electrostatic effects become all-dominating.

The parameters chosen in this study were tailored to the experimental system investigated by Smith et al. (2001). By choosing different geometrical parameters or packaging arrangements, as well as screening lengths, etc., different systems can be considered readily. The results were shown to be relatively sensitive to the screening length λ , which can be adjusted from both the onset of the significant growth of the packaging force and the maximum packaging force. In contrast, it can be checked that the radii of the nested helices do not provide much leeway for variations; in particular, we verified that a change in the difference between successive radii from 2.2 nm to 2.4 nm does not change the $\mathcal{E}_{el}/l$ behaviour significantly.

Our study is considerably more simplistic than, for instance, the detailed numerical investigation by Kindt et al. (2001), who use a Lennard-Jones potential and include the interaction with the capsid hull explicitly, considering different counterionic regimes, or the elaborate calculations by Odijk's group (Odijk 1998, 2003; Odijk and Slok 2003). We believe that this approach, combining an analytical space-curve with numerical evaluation of the electrostatic energy, helps shedding some more light on the most basic underlying, purely packaging-related phenomena. In particular, the high sensitivity to the screening length already at lower packaging ratios, reflecting the “volume interactions” in the multilayered arrangement chosen, and the surprisingly close reproduction of the maximum force once the screening length was adjusted to the turnover point, demonstrate that we actually capture the essential physics of the process. To stabilize the packaged capsid, DNA “condensation” effects come into play, but we decided to neglect such additional effects. Depending on the very nature of such an attractive interaction, it may be speculated whether at a given packaging ratio due to fluctuations a “condensation kernel” may spontaneously form, from which the DNA “zips” into the closely packaged final packaging state. We also note that given the experimental parameters available presently, it was possible to go beyond the more qualitative findings in the similar model of Riemer and Bloomfield (1978).

Expression (15) for the nested helices configuration and for a given packaged length can be evaluated through a Fortran77 program within a few minutes on a 2.0 GHz Linux platform, a reasonably efficient method to check the geometry for parameter variation, etc.

The packaging of semiflexible, charged chains as well as the converse ejection process, which we have disregarded in the current account, stretches into a fairly extreme regime of packaging densities and biomechanical properties of the elements involved: the tensile strength of the capsid hull, the maximum force of the packaging motor, as well as the bending strength of dsDNA itself. We therefore expect that the investigation of this parametric domain will continue to receive interest, in particular when future experimental observations of the process become available, such as the packaging parameters of different capsid geometries, the role of the capsid hull, dependences on salt concentrations, and the behaviour for varying capsid hull sizes and chain lengths.

Acknowledgements We acknowledge helpful discussions with Mehran Kardar. R.M. acknowledges partial funding by the DFG within the Emmy Noether programme.

References

- Alberts B, Roberts K, Bray D, Lewis J, Raff M, Watson JD (1994) The molecular biology of the cell. Garland, New York
- Arsuaga J, Tan RK-Z, Vazquez M, Sumners DW, Harvey SC (2002a) Investigation of viral DNA packaging using molecular mechanics models. *Biophys Chem* 101:475–484
- Arsuaga J, Vázquez M, Trigueros S, Sumners DW, Roca J (2002b) Knotting probability of DNA molecules confined in restricted volumes: DNA knotting in phage capsids. *Proc Natl Acad Sci USA* 99:5373–5377
- Bouchiat C, Mézard M (1998) Elasticity model of a supercoiled DNA molecule. *Phys Rev Lett* 80:1556–1559
- Catalano CE, Cue D, Feiss M (1995) Virus-DNA packaging – the strategy used by phage-lambda. *Mol Microbiol* 16:1075–1086
- Cerritelli ME, Cheng N, Rosenberg AH, McPherson CE, Booy FP, Steven AC (1997) Encapsidated conformation of bacteriophage T7 DNA. *Cell* 17:271–280
- De Gennes P-G (1979) Scaling concepts in polymer physics. Cornell University Press, Ithaca, New York
- Eisenriegler E (1993) Polymers near surfaces. World Scientific, Singapore
- Frank-Kamenetskii MD (1997) Biophysics of the DNA molecule. *Phys Rep* 288:13–60
- Fujisawa H, Morita M (1997) Phage DNA packaging. *Genes Cells* 2:537–545
- Gronbech-Jensen N, Mashl RJ, Bruinsma RF, Gelbart WM (1997) Counterion-induced attraction between rigid polyelectrolytes. *Phys Rev Lett* 78:2477–2480
- Hud NV (1995) Double-stranded DNA organization in bacteriophage heads – an alternative toroid-based model. *Biophys J* 69:1355–1362
- Hud NV, Downing KH (2001) Cryoelectron microscopy of λ phage DNA condensates in vitreous ice: the fine structure of DNA toroids. *Proc Natl Acad Sci USA* 98:14925–14930
- Kindt J, Tzili S, Ben-Shaul A, Gelbart WM (2001) DNA packaging and ejection forces in bacteriophage. *Proc Natl Acad Sci USA* 98:13671–13674
- Kunze K-K, Netz RR (2002) Complexes of semiflexible polyelectrolytes and charged spheres as models for salt-modulated nucleosomal structures. *Phys Rev E* 66:011918
- Lubensky DK, Nelson DR (2000) Pulling pinned polymers and unzipping DNA. *Phys Rev Lett* 85:1572–1575
- Marko JF, Siggia ED (1995) Statistical mechanics of supercoiled DNA. *Phys Rev E* 52:2912–2938
- Moroz JD, Nelson P (1998) Entropic elasticity of twist-storing polymers. *Macromolecules* 31:6333–6347
- Netz RR (2001) Strongly semiflexible extensible polyelectrolytes and DNA. *Macromolecules* 34:7522–7529
- Odijk T (1998) Hexagonally packed DNA within bacteriophage T7 stabilized by curvature stress. *Biophys J* 75:1223–1227
- Odijk T (2003) Statistics and dynamics of condensed DNA within phages and globules. eprint cond-mat/0305354
- Odijk T, Slok F (2003) Nonuniform Donnan equilibrium within bacteriophages packed with DNA. *J Phys Chem B* 107:8074–8077
- Richards KE, Williams RC, Calendar R (1973) Mode of DNA packing within bacteriophage heads. *J Mol Biol* 78:255
- Riener S, Bloomfield VA (1978) Packaging of DNA in bacteriophage heads – some considerations on energetics. *Biopolymers* 17:785–794
- Risken H (1989) The Fokker–Planck equation. Springer, Berlin Heidelberg New York
- Rudnick J, Bruinsma R (2003) Icosahedral packing of RNA viral genomes. eprint cond-mat/0301305
- Santo KP, Sebastian KL (2002) A simple model for the kinetics of packaging of DNA into a capsid against an external force. *Phys Rev E* 65:052902
- Schiessel H, Aranda-Espinoza H (2001) Electrostatically induced undulations of lamellar DNA-lipid complexes. *Euro Phys J E* 5:499–506
- Simpson AA, Tao YZ, Leiman PG, Badasso MO, He YN, Jardine PJ, Olson NH, Morais MC, Grimes S, Anderson DL, Baker TS, Rossmann MG (2000) Structure of the bacteriophage phi 29 DNA packaging motor. *Nature* 408:745–750
- Smith DE, Tans SJ, Smith SB, Grimes S, Anderson DL, Bustamante C (2001) The bacteriophage ϕ 29 motor can package DNA against a large internal force. *Nature* 413:748–752
- Snustad DP, Simmons MJ (2003) Principles of genetics. Wiley, New York
- Venkataram Passad BV, Hardy ME, Dokland T, Bella J, Rossmann MG, Estes MK (1999) X-ray crystallographic structure of the Norwalk virus capsid. *Science* 286:287–290
- Wikoff WR, Liljas L, Duda RL, Tsuruta H, Hendrix RW, Johnson JE (2000) Topologically linked protein rings in the bacteriophage HK97 capsid. *Science* 289:2129–2133